4-2024

Channel of exposure and its effect on purchase decision: Evidence from an online retailer

Lei LI

Singapore Management University, lei.li.2017@sjdba.smu.edu.sg

CHANNEL OF EXPOSURE AND ITS EFFECT ON PURCHASE DECISION: EVIDENCE FROM AN ONLINE RETAILER

LI LEI

SINGAPORE MANAGEMENT UNIVERSITY

2024

Channel of Exposure and its Effect on Purchase Decision: Evidence from an Online Retailer

Li Lei

Submitted to Lee Kong Chian School of Business

in partial fulfillment of the requirements for the Degree of

Doctor of Business Administration

Dissertation Committee:

Ma Dan (Chair)

Associate Professor of Information Systems

Singapore Management University

Jiang Ming (Co-supervisor)

Associate Professor of Economics

Shanghai Jiao Tong University

Geng Xuesong

Associate Professor of Strategic Management

Singapore Management University

SINGAPORE MANAGEMENT UNIVERSITY

2024

I hereby declare that this dissertation is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in this dissertation.

This dissertation has also not been submitted for any degree in any university previously.

Li Lei

Li Lei

2 April 2024

Channel of Exposure and its Effect on Purchase Decision: Evidence from an Online Retailer

Li Lei

Abstract

In the digital era, online retailers engage consumers through a variety of channels, making it critical to understand how multi-channel exposure impacts consumer behavior, purchase conversion and firm outcomes. Drawing upon information overload and attention theories as well as extant research on conversion models, this study examines how channel-specific browsing time and channel diversity influence online users' conversion and revenue contribution. By utilizing a multi-touch attribution approach, this study analyzes user-level visit data from an online eyewear retailer including millions of visits over a three-month period and finds that cumulative browsing time has a positive impact on both conversion and revenue, with search engine emerging as the most effective channel for driving purchase conversions and organic social media yielding higher-value orders. Recency is considered to reflect the effect decay over time.

Importantly, channel diversity acts as a double-edged sword - while it has a positive main effect on conversion probability, it negatively moderates the impact of browsing time within each channel on both outcomes, suggesting potential information overload and curvilinear relationship. These findings underscore the complex interplay between channel effectiveness and consumer attention allocation in a multi-channel environment.

This research advances theories of information processing and multi-channel marketing while providing actionable insights for online retailers to optimize their marketing strategies by aligning budgets with channel effectiveness to achieve an optimal channel mix yileding optimal conversion and revenue.

Keywords: multi-channel marketing, online conversion, attribution modeling, information overload, attention allocation, channel diversity

Table of Contents

List of Figures

List of Tables

List of Equations

Acknowledgement

The last 6 years has been an exciting, challenging, fruitful and memoriable period. I always remember the great coaching of our top professors, the inspiring discussions in and after the classes, the huge volume of papers we read, the hard time in each course assignment and in this dissertation, and finally I come to this great achievement. Thank you all, my teachers, classmates, my family and myself!

Sincere gratitude goes to my esteemed supervisors Prof. Ma Dan, Prof. Jiang Ming and Prof. Geng Xuesong, for their profound insight, knowledgable guidance and rigorous yet realistic research attitude from which I learnt and benefited a lot. I am especially grateful to Prof. Jiang Ming and Elaine Gao, for their selfless help and patience in the so-many discussions from topic selection to the final paper.

Family is the source of energy whenever I get frustrated or exhausted. Without their understanding and support I couldn't have finished this joury. I would like to thank all my family members, for their unwavering support, encouragement, and love all the time.

Chapter 1 Introduction

With the development of information and communication technologies (ICT) such as mobile internet, social media and 5G network, more and more retail business and communication happen online. As a result, e-commerce has grown impressively in the last decades. As an indispensable part of global business, the global e-commerce market size was estimated at 5.7 trillion US dollars in 2022 (Statista, 2023). Online retailers have benefited a lot from e-commerce: they can tap into the tremendous global online user base, interact actively and bidirectionally with them through multiple digital channels, and access granular record of every visit along the purchase journey to provide more personalized products or services with personalized marketing methods (Molla, 2007; Li & Kannan, 2014). Studies also show that e-commerce provides a basis for online retailers to increase productivity and provide better customer service (Abou-Shouk et al., 2013).

However, online retailers consistently face the significant challenge of how to convert a higher portion of users from visitors into buyers or consumers, or the visit-to-purchase conversion problem. For one thing, the global potential user base is quite huge – in 2021 the number of global Internet users has increased from 361 million in 2000 to 5.4 billion, a little more than two thirds of the world's population [1]. For another, the overall visit-to-purchase conversion rate remains notably low in the global e-Commerce market – throughout 2021 and 2022, the online conversion rate is always below 3% in the United States (Statista, 2022). One reason for the low conversion rate is that with the ease and flexibility provided by e-commerce, users have plenty of choices regarding what brand and model to buy – considering the very large and growing number of online shops plus the low cost and convenience of visiting the websites to

[1] Refer to Internet World Stats (https://www.internetworldstats.com/stats.htm) for most up-to-date statistics of Internet usage and growth by regions across the world.

shop around. In addition, online users can retrieve plenty of information before making a purchase decision from various sources, such as online shops and social media platforms. Information can be objective (about the product itself and its competitors) and subjective (other people's comments or ratings) (Kuhlthau, 2005 & 2008). In addition, the information can either be actively searched by the user from channels like the search engine or be passively received from social media or display advertisements. As a result, the purchase journey is full of complication and uncertainty (McKinsey, 2009). For online retailers to increase revenue and expand share in the fiercely competitive e-commerce market, it is vital importance to make clear factors that impact the purchase decision and take corresponding measures to increase the conversion rate and generate more sales.

Another challenge online retailers face is the efficient allocation of substantial marketing budget across various marketing channels. In this digital era, e-commerce is filled by the explosion of product choices, relevant information and digital touch points, coupled with the increasingly well-informed users and firecely competing market environment. To attract users to their online shops and to motivate them to make purchases, online retailers spend plenty of money (often 10~20% of their revenue) on multiple online marketing channels (including paid search advertising, social platform and display networks, etc.), and the channels of exposure need be available on desktop and laptop computers and mobile devices. In 2011, worldwide online advertising spending was reported at 522.5 billion US dollars (Statista, 2023a). However, the effectiveness of such spending, or the conversion effectiveness each channel remains unclear and hard to measure. To optimize the utilization of the substantial marketing budget and allocate resources effectively across various channels, it is crucial for online retailers to measure the influence power of each channel of exposure (from which the user is engaged) and the impact of channel diversity.

A lot of researchers have studied such problems from different perspectives. According to Kuhlthau (2005 & 2008)'s six-stage model of the information search process, a user may resort to various sources to seek information, and his affective feelings and cognitive thoughts may increase or decrease the uncertainty towards a decision during the process (Kuhlthau, 2005 & 2008). Additionally, research on user conversion suggests that the effects on purchase decision are accumulated over the user's visit-to-purchase journey (awareness to order) journey (Moe & Fadar, 2004b). Furthermore, research on channel attribution proposes several models to measure the contribution of each channel and assign credits of a conversion accordingly, indicating different influencing power of each channel on the effect of visit. (Shao & Li, 2011; Ji & Wang, 2016).

This paper follows the research stream on purchase conversion and channel attribution and intends to address the two afore-mentioned challenges of increasing purchase conversion rate and better allocate resource among various marketing channels from the perspective of channel effectiveness. First, this paper intends to shed light on the cummulative visit effects (taking into consideration factors of visit duration and recency) and the influence of channel of exposure on the users' purchase decisions. The channel of exposure refers to the specific incoming channel associated with each visit to the online retailer's website, such as paid search on Google (*cpc_google*) or paid social on Facebook (*psoc_facebook*). The assumption is that given the same browsing time or attention paid to the website, visits from certain channels can add more affective and cognitive confidence to the user's purchase decision making thus more effectivess than those from other channels.

Second, this article introduces the concept of channel diversity, measured by number of unique channels of each user along his visit history to the website.

This research seeks to examine its direct impact on conversion as well as its indirect impact on the influence power of each channel of exposure. The diversity of a user indicates his purchasing behavior and decision-making style. For example, for users with a higher level of channel diversity, additional channels may bring more information about the products or promotions to them, while in the meantime may add the user's cognitive fatigue and purchase uncertainty because of information overload. In addition, channel diversity can indirectly strengthen or weaken the influence power of each channel of exposure. The research will be based on about millions of visit records from users of an online eyewear retailer website.

This research intends to add some theoretical contributions to online users' purchase conversion and channel effectiveness theories in the following aspects. First, instead of treating each visit as a homogeneous point (as some researchers did), we include visit's detailed information such as its duration (how long the visit lasts, or the browsing time on the website) and recency of the visit in the purchase conversion model. Second, to measure the different effect on user conversion and revenue of visits from different channels of exposure, the influence power of each channel is investigated. Third, the direct and indirect impact of channel diversity is studied on the final purchase conversion and the influence power of each channel. To our knowledge, this research is among the first to introduce the concept of channel diversity and to consider channel diversity in the purchase decision and incorporate it into the conversion model to examine its influence on users purchase conversion. Fourth, aside from conversion which is the focus of most extant research, this research also checks the impact on revenue of converted users, considering the high importance of revenue to actual business. It also provides empirical evidence and support in favor of the information overload and attention theory in consumer decisions.

This research also intend to carry several managerial implications for business professionals by providing valuable insights into the impact of each channel of exposure on user purchase decisions - for example, whether *affiliate* drives a higher conversion rate than *retargeting* and whether the effect of *paid search* is overestimated or not. Online retailers can then make informed decisions about the allocation of marketing resource (money, personnel, promotions, etc.) across various digital marketing channels and platforms. In addition, by understanding what channels are strengthened by channel diversity and what are weakened, online retailers can decide whether to engage proactively with a user through other channels.

The content structure of this paper is organized as follows. Chapter 1 introduces the current problems in online retail of low conversion rate and high marketing spending with unclear effectiveness. It then introduces our research questions about the impact of visit details and channel of exposure on purchase conversion and the impact of channel diversity on each channel. Chapter 2 reviews the existing literature related to this research, typically theories on information retrieval, attention theory, online user conversion models, different impact on purchase of each channel of exposure, and the channel attribution model. The hypotheses and research model are explained in Chapter 3. Chapter 4 introduces the research design, including the structure of the data set, independent and dependent variables and how the raw data is processed to fit the models. Chapter 5 makes descriptive analysis of the data with explanatory statistics and examines and analyzes the regression results. After checking the models' robustness with visit time discounted by recency, Chapter 5 ends with a conclusion of the regression models. Chapter 6 summarizes the contributions and limitations of this research and future research directions.

Chapter 2 Literature Review

In this chapter, we first review literature about users' information retrieval which supports their decision-making in Section 2.1. Section 2.2 reviews the development and contributions of attention theory about how users spend and allocate their attention. Section 2.3 reviews the literature about online users' conversion model from the perspectives of website design, visit history, and past visit details etc.. Section 2.4 summarizes the typical online channels of exposure and reviews the literature on the impact on users' purchase decision of each individual channel. Section 2.5 focuses on channel attribution models on how to assign credits across different channels of a conversion. Finally, the last section concludes the related literature as the theoretical foundation of our research model in this paper.

2.1 Information Retrieval Theory

It is natural that people seek to get "enough" information before making uptheir decisions. In the context of online retail or e-commerce, users often go to several different kinds of websites to check product information, review other users' comments, look for promotions or coupons and compare among brands or products etc. Generally speaking, there are two kinds of information retrieval scenarios: internal and external (Gretzel et al., 2007; Engel et al., 1990). Internal retrieval refers to the behavior of retrieving information from existing knowledge (such as a person's memory), while external retrieval refers to information from external sources such as the Internet (blogs, videos, shopping websites, social networks, etc.). There are three types of external retrieval of online information: search (typing in search terms in a search box), navigation (clicking links) and information organization (printing down information on a sheet of paper). Navigation can be further divided into goal-oriented navigation and browsing without a clear purpose.

2.1.1 Information Search and Retrieval Process

Kuhlthau (2005) proposes a six-stage model of the user's information search process: initiation, selection, exploration, formulation, collection, and presentation. The model highlights the affective, cognitive and physical aspects of users' information search behavior and claims that user experience during the process can be regarded as a series of affective feelings, cognitive thoughts, and physical actions with varying uncertainty. The user's uncertainty in both feeling and thoughts increases or decreases during the process depending on the information they get that influences their confidence, interest and action. The authors revisited the model in 2008 and suggest that this model is still useful for navigating the complex and dynamic information landscape today (Kuhlthau et al., 2008). In the context of e-Commerce, it is natural to understand that people can seek information over the Internet from various sources, including search engines, social platforms, and retailers' online shops before making a purchase decision. This implies the online retailers that each exposure to the user adds or subtracts the user's intention to buy, depending on factors like the exposed time, website design and the affective feeling, etc.

McKinsey (2009) summarizes the user decision journey in a funnel metaphor of five stages: awareness, familiarity, consideration, purchase, and loyalty. This essentially echoes the Information Search Process (ISP) model in explaining users' information-seeking behavior before making a purchase decision. Figure 2-1 illustrates the essential similarity between the ISP model and the user decision journey. In the context of e-commerce, the decision journey and information search process of users remain the same, though the illustrated steps can be in one or multiple visits to the online retailers' website.

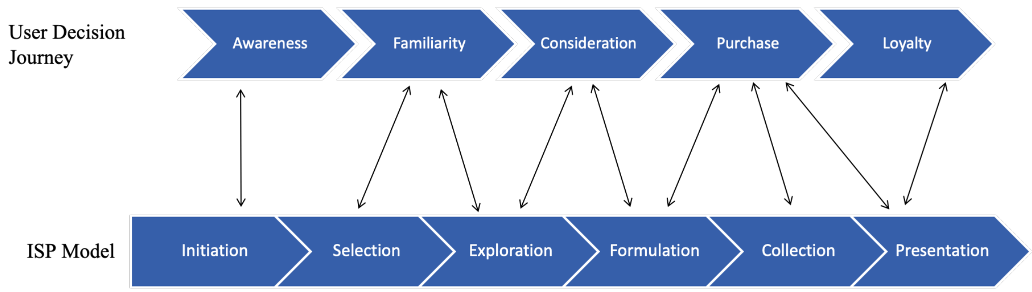

Figure 2-1 Relationship between ISP model and the Purchase Journey

2.1.2 Information Overload Problem

Researchers have reported several challenges or problems to information retrieval over the Internet. For example, users may be not satisfied with the slow retrieval speed, communication delay, poor quality of retrieved result, and inability to find relevant information (Kobayashi, 2000). Another problem is the information overload, which is a kind of 'paradoxical situation': the Internet is abundant in information of different forms and from different providers, but it is often hard, time-consuming, and confusing to obtain useful and relevant information. For centuries, there has been concern about the problem of “too much to read’, and it is becoming more urgent in the digital era considering the ubiquitous and explosive digital information spreading the Internet in a variety of forms and from several sources, especially the social media platforms with user-generated content (UGC) and AI-generate content (AIGC). As a result, content from different sources of various formats can be complementary and helpful, but can also be inconsistent, conflicting and misguiding or confusing.

For individual users, too much ‘potentially’ relevant and useful information available has become a hindrance rather than help, with one consequence that it damages the decision-making process by delays and poor decisions (Bawden, 2009). Additionally, people with a high level of information overload may experience lowered well-being, accompanied by a perceived loss of control over the situation (Bawdem, 2020). In the context of e-Commerce, information overload adds complexity to the user decision-making process with reduced confidence and can result in a negative impact on the quality of the final decision-making.

However, it’s worthy of note that the information overload problem doesn’t always happen whenever the user is retrieving information. For example, Reiley and Li (2010) perform a controlled field experiment by varying number of “north” sponsored ads displayed on the result page after an organic search on their search engine (Yahoo) and report that adding more northern ads in search results can unexpectedly increase the click-through rates of existing ads. This positive externality suggests that displaying multiple north ads can lead to higher overall clicks, though more ads provide more information (which are sometimes conflicting as the ads may come from different brands) to the user.

2.2 Attention Theory

Attention is a critical factor in influencing consumers' decision-making processes. We review attention theory in this paper because it provide valuable insights into the cognitive processes underlying online users' purchase decisions. In our digital era, consumers can assess (either actively or passively) quite a lot of information and their attention is attracted by numerous products, advertisements and reviews etc. That's why understanding how attention operates in online environments is crucial for our research. Extant literature reviews the impact of advertising strategies, consumer knowledge, and attention economies on consumers' choices to synthesize the role of attention and cognitive factors in shaping consumers' purchase intentions and behaviors.

Dai and Sheng (2022) conduct a study to investigate the effects of advertising strategies, particularly green advertising appeals, and subjective busyness on green purchase intention. The research highlights the significance of attention and cognitive factors in shaping consumers' intentions to purchase sustainable products. Their findings underscore the importance of attention and cognitive processes in influencing the user’s purchasing behavior and decision, especially in the context of sustainable and environmentally responsible consumption.

Chen et al. (2017) focuse on the impact of consumer knowledge (gained from multiple sources as well as online channels such as product information from search and e-commerce websites as well as users comments etc.) on search behaviors, information processing, decision-making, and purchase intentions. Their research provides valuable insights into the cognitive aspects of attention and information processing in the context of purchasing a specific product, and emphasizes the role of attention in information acquisition and decision-making processes, shedding light on how consumer knowledge gained from multiple sources influences consumer behavior.

Pieters et al. (2002) examine the benefits of advertisement originality and familiarity for brand attention and memory are investigated and highlight the challenges of attracting and holding consumers’ attention in the face of increasing advertising competition. This study provides valuable insights into the role of attention in advertising effectiveness, emphasizing the importance of capturing consumers' attention in highly competitive markets.

Falkinger (2007) discusses attention economies and their implications for consumer choices in complex markets. The research explores the effects of international integration and information technologies on global diversity. Understanding attention economies is relevant to comprehending the impact of attention on consumer choices, particularly in markets characterized by information overload and complexity.

Armstrong and Chen (2009) present a model in which some consumers make purchase decisions based solely on price, disregarding product quality. This study offers a perspective on the influence of attention on consumer choices in markets characterized by information congestion and emphasizes the importance of attention in guiding consumer decision-making, even when factors such as product quality may be overlooked.

In conclusion, these studies demonstrate the significance of attention and cognitive factors in shaping consumers' purchase intentions and behaviors. The role of attention in advertising effectiveness, consumer knowledge, attention economies, and decision-making processes highlights its importance in understanding and predicting consumer choices of products. However, the extent literature mainly focuses on studying attention theory under a single advertisement, and to our knowledge, there is no relevant theoretical research on the attentional shift of consumers in the context of multi-channel situations and its impact on purchase behavior. Future research should further explore the complex interaction between attention and other factors that influence consumer decision-making in various contexts.

2.3 User Conversion Models

The conversion rate is defined as the ratio of visitors with a desired action to all visitors to a website. For example, the registration conversion rate is defined as the percentage of visitors who register as a member to all visitors. The mainstream research focuses on the purchase conversion rate, which is defined as the ratio of visitors who have made at least one purchase in a given period to all visitors.

Plenty of research has investigated the probability of conversion of the online purchase conversion probability in the current visit or future visits at the individual level in order to shed light on the key factors impacting users' purchase decision making. Based on the user's visit history data in a given period. Most research treats each visit as an equal black-box event without looking into its details, such as Moe and Fader (2004a), Moe and Fader (2004b), and Park and Park (2016), while some also take into consideration the detailed information such as the number of pages viewed (page views), what categories of pages have been viewed (Home, Category, Detail etc.) along with what activities have been done (Add to Favorite, Add to Cart, Remove from Cart, Order etc.) in each visit (Montgomery, 2002).

2.3.1 Research on Website Quality

Some research on users' purchase decisions is done from the perspective of the online store (the retailer's website). The assumption is that website quality (such as the layout, color schema, interaction styles, and performance) is key to convert users and generate sales as the website is the primary user interface for online users to search for information and make purchases.

In this regard, Kuan et al. (2009) investigate the impact of perceived website quality in increasing customer's purchase conversion and retention. The authors defined website quality as a combination of three categories, namely System quality which is about the usability (ease of navigation, consistency of layout, website speed, etc.), Information quality which is about the content delivered in the website (accuracy, completeness, understandability, etc.) and Service quality which is about how the user feels about the services provided (responsiveness, interactivity, clarity on security and privacy, etc.). Based on a survey completed by 50 students where each category of quality is measured using seven questions, the researchers reported that the three perceived qualities of the website significantly influence the user's intention to first purchase and continued purchase and demonstrated the importance of website quality. The limitation is that the authors just did a test about users' intention to buy instead of actual purchase conversion, and the sample includes only 50 students, while there are millions of visitors on real e-commerce websites. Future research in this area can expand the sample size to fit the real business.

Gudigantala et al. (2016) furthur examine the relationship among purchase intention, conversion, and website satisfaction at website (online shop) level. Based on such data from 85 online retailers' websites, their research reported that the user's perceived website satisfaction positively influences his purchase intention, while both website satisfaction and purchase intention positively influence website conversion rate. However, this research measures purchase intention rather than the actual purchase decision.

Form another perspective, Hauser et al. (2009) extend the research on website quality by reporting that dynamically adjusting website design (content, style, look and feel, etc.) to match the user's cognitive style or preference can increase purchase intention by 21%. The authors mentioned that a visitor's cognitive style can be inferred from his clickstream data, since each click is a decision point that reveals the visitor's cognitive preference. However, in this research, the authors adopted a simple way to determine a user's cognitive style by asking him a list of questions when he first enters a website. For example, some people may prefer vivid images with big characters and numbers, while some others may prefer rich detailed information at one page. Some people may be impulsive (to make decisions quickly), while some are deliberative (to get information in depth and make a lot of comparisons before making a purchase decision). The authors suggest that online retailers dynamically change the website content and look and feel or send personalized promotion messages to match user's cognitive preference of the user.

2.3.2 Research on Visit History

Aside from website design, researchers investigate visit-to-purchase conversion from the perspective of users. With the advancement of information technologies, which is also a benefit of e-commerce, online retailers can access granular record of every visit to the website along the purchase journey (Li & Kannan, 2014). Based on the user's visit history data (at what time a visit occurs), Moe and Fader (2004a) examine the relationship between visit frequency (number of visits in a selected period) and purchase at individual level. Their research reports that more active shoppers (more visits and higher frequency of visits) tend to be more likely buyers.

Park and Park (2016) propose a model to identify user's visit behavior patterns and report that such patterns can be clustered and act as an important source of predictive information about his future visit – when the user will visit the website again and whether there will be a purchase in his future visits. In this research, each visit is considered a homogeneous point and a user's visit can then be characterized as a one-dimensional point process. Visits that are in close temporal proximity and followed by a period of no visits are considered as a cluster of visit events. A user's visit history consists of several visit clusters of different size, frequency, and distance. When a new visit occurs, the model determines what cluster it belongs to and then predicts what behavior might happen based on previous patterns. With an experiment from visit history data of 1,000 randomly selected users from an online retailer website, the authors claim that there are higher visit rates within clusters and lower visit rates between clusters. Consequently, conversion rates are relatively higher at subsequent visits compared with earlier visits within a cluster. One problem with the model is that it considers each visit equally as a point of no size, but in fact, visits are not equal – some visits may be triggered by the user clicking on an advertisement by mistake and are terminated immediately, while some visits last longer when the user is actively searching for information or making comparisons. As a result, treating these different visits equally may cause the clustering to be inaccurate.

Park's model predicts the user's purchase decision in a visit by fitting the visit into a prior generated group without considering the influence of previous visits. In another research done by Moe and Fader (Moe & Fader, 2004b), the authors model the accumulated visit effects that increase or decrease over visits and compare it against a dynamic purchase threshold to predict whether the user will make a purchase decision at a given visit. The purpose in this research is only to separate visits that are more likely to result in a purchase from others so that these visits can then be redirected to a higher-performance server that provides a better shopping experience to users. According to this research, each visit may be for different purpose (for example, some are out of planned purchase while others are simply browsing) and has a different effect on purchase decision. Similar to information retrieval theories, the authors suggest that although some visits do not have a purchase, they can be part of information search or comparison along the user decision journey and have important (positive or negative) effects on the final purchase decision. Although the paper does not explicitly state that more visits lead to a higher probability of purchase, the authors suggest that the net effect of past visits is an important factor in determining the probability of purchase at a given visit. They modeled the net effect of past visits as a gamma-distributed random variable and showed that it has a significant impact on the purchase decision or purchase conversion. The implication of this research is that accumulated visits, instead of a single visit, should be taken into account for online users' purchase decision. However, though this research assigns different effects (random value) for each visit, it still treats each visit as an equal point and does not look into the details of a visit – such as the channel of exposure, the browsing sequence, what types of pages have been viewed, the duration on each page and the time of the entire visit session – which are of great importance to evaluate the visit effect to the final purchase.

2.3.3 Research on Visit History From Multiple Retailers

Considering that users may search information and make comparison over other websites of the same category (the competitors' websites), some researchers also examined visit history data from multiple websites. Park and Fader (2004) conduct a research by combining users' clickstream data from multiple online retailers of the same category (eg, books) provided by Media Metrix. This research reported better performance compared to models that use visit data only from a single website. It is true that adding visit data from other similar websites makes the data source more comprehensive and the prediction more accurate, but in real business, it is not easy and up-to-date to get such data. Also, considering the difference and difficulty in identifying a user across different retailers who will not share customer data, only a small percentage of users can be matched from different websites. That's why in our research we still focus on visit history data on a single website.

2.3.4 Research on Visit Details

The research mentioned above treats each visit as a homogeneous point without looking at its content (for example, what pages have been viewed for how long), which is, in fact, of great importance in determining a user's behavior and visit intention.

Montgomery et al. (2002) study the information of visit details and examine the type (such as Home, Category, Detail, and Order, etc.) of pages viewed in a visit session and categories path sequence to predict the user's future moves on the website in the same session. The primary target of their model is to predict whether the user will make a purchase (by visiting the Order pages) in a visit session given a sequence of pages viewed. For example, given that the user has browsed pages of Category (C) and Shopping Cart (S) thus the page sequence may be CCSCCC, whether the user will visit an Order (O) page in the following moves within the current visit session. Based on visit data of 1160 users in 30 days, the authors trained the model with path sequence data of previous sessions. Research reports that the user's state (likely or unlikely to purchase) can change in the midst of a session, indicating that it is inaccurate to simply describe a user or a whole session as purchase- or nonpurchase- oriented. The implication of this research is that the content of a visit session is informative to predict the probability of purchasing by the user. The problem with this research is that it is too computing intensive in that it requires dynamic calculation after every page view – when the user has viewed another page, the visit path is updated and the model needs to calculate the next step and calculate the purchase probability in this session in real time. To solve this, we suggest making the calculation of purchase probability at visit level – that's, after each visit. However, such research has highlighted the importance of detailed information of each visit on modeling users' purchase conversion. So in our research, the details of each visit including its duration (the user's browsing time on the website), channel of exposure, recency compared to the user's last visit is considered. We also suggests future research based on user visit history check visit details instead of treating each visit equally as a homogeneous point.

2.4 Channel Attribution Models

2.4.1 Research on Digital Marketing Channels

In the digital era, online retailers can reach, interact with and engage its audience through various digital channels. Each channel has its pros and cons. Major digital marketing channels are explained as follows.

- Search engine marketing, including search engine optimization (SEO, orgnanic search) and pay-per-click (PPC, paid search or sponsored search) advertising on search engines like Google and Bing. Some giant social media platforms also provide the search functionality and apply a cost-per-click model, same as search engines.
- Social media marketing, where merchants place advertisements on social media platforms like Facebook, Twitter, Instagram, and Tiktok to promote their brand or products. The advertising can be of a variety of forms like text, image or video and targets to users of specific demographics and interests.
- Display ads, where merchants place visual ads, banners, or videos on the top or side of multiple websites or mobile apps. There is often a call-to-action (CTA) message that links to a landing page on display ads. Typical platforms

providing display ads are Google display network (GDN), Facebook audience network and LinkedIn advertising etc.

- Affiliate marketing, where merchants partner with affiliates who promote products and earn a commission on sales generated through their referral links across their ads network. Some well-known platforms for affiliate marketing are Amazon associate, CJ affiliate, Impact and Rakuten LinkShare etc.
- Retargeting (or remarketing), which re-engages users who have previously interacted with a website or brand. It leverages data tracking techniques (e.g., cookie) to display targeted ads to users who have shown interest in a product or brand on other websites. Typical platforms for retargeting includes Criteo, Steelhouse and Outbrain etc.
- Email marketing, where merchants work with partners to send emails to subscribers or potential customers to inform, engage, or promote products or services. In some scenarios, sms (short message) is also used.
- Content marketing, where retailers post relevant content (such as blogs, articles videos) to attract and engage potential customers on forums or user groups. Some well-known platforms for content marketing include Instagram, Youtube and Quora etc..

Li and Kannan (2014) conduct a research to categorize these online channels into user-initiated (such as organic search and direct visit) and firm initiated (such as display ads and email). From the perspective of information retrieval, users may actively seek information from the user-initiated channels or passively receive information (such as an email or sms) from the firm-initiated channels. It is natural to understand that different channels of exposure reflect different user intention.

When selecting one or more to place advertisements among the various channels, online retailers need understand the pros and cons of each channel to fit their business goal. Extant research has examined the different effects on user purchase conversion of different channels. Considering the overwhelming popularity of search engine in the Internet, search engine marketing (SEM) is currently the most widely used online marketing form. It allows advertisers to place ads that depend on the keywords the user has entered in a search engine and can be divided into two categories: organic search, which leverages search engine optimization (SEO) techniques to optimize the website to improve its ranking in the search engine platform so that it appears on the top of an organic search result, and paid search, which is about buying keywords from search engines to show their product ads in the top of the search result page. Paid search is also known as pay-per-click (PPC) or cost-per-click (CPC). The advantage of search engine marketing is that it can reach a large number of users in the right time of the user's information search process and take effect immediately, while the disadvantage is that improper selection of keywords can cause mismatch between products being displayed and the users' expectation, resulting in a negative impact on user's purchase intention (Sun and Spears, 2012). In a study conducted by Jansen and Schuster (2011), the researchers investigate the effects of keywords on the user throughout the buying funnel from research to purchase and report the importance of the search engine as an important channel for web advertising to promote the products of online retailers. Skiera et al. (2010) examine the effect of keywords by studying the effect of keywords on search engine marketing campaigns and reported the importance of selecting few but appropriate keywords in impressing more users and attract them to websites. Oliver and Michael (2011) introduce a two-stage model to evaluate the effect of paid search (such as Google Ads) on its performance (such as number of clicks and conversions) and reported that the position (e.g., top or right side) where the ad was displayed on the website has a significant impact.

Compared to search engine marketing where ads appear only to users who are actively searching for the required products, display ads are paid placement and can appear anywhere on various website that a user may be navigating. Display ads are helpful when the retailer wants to move users from the discovery or awareness stages of the marketing funnel to consideration and boost brand awareness among shoppers with low purchase intent. Chatterjee et al. (2003) study the effect of display ads on websites and reported that repeated or too-often ad exposure is, in fact, negative to conversion. In other words, display ads should be displayed at the right time of the user's information search process.

Ewijk et al. (2020) assess the effects of display ads for 154 CPG (Consumer Packaged Goods) brands and reported that display ads are ineffective for low-involvement utilitarian products, but they can significantly enhance sales for other CPG product types. Research also reported that if the display ads are spread more evenly, its long-term effectiveness can increase significantly.

Another important marketing channel is social media because of its increasing popularity and frequency of use. In 2022 over 4.59 billion people were using social media with an average of 2.5 hours spent on social media every day (Statista, 2003c). With the popularity of mobile Internet and user-genenrated content, social media platforms are becoming more and more important in the retrieval of user information and purchase decision. The advantages of social media marketing (SMM) are that online retailers can easily and effectively reach a targeted audience and get instant feedback, while the disadvantages are that it's time-intensive and may generate negative or controversial posts (Arthsa, 2018). The significant influence of perceived social media marketing activities on customer loyalty is also found in previous research (Ismail, 2017). Zhang and Duan (2014) reported on the importance of social media marketing from a different perspective. Their research examined the spillover effect of referrals on search and reported that social media referrals to competing online stores have a significant negative impact on conversion. However, too much content on various social media platforms also introduces the problem of low-quality, controversial, and even contradictive information. Such an information overload problem on social media platforms, called social media fatigue, adds complexity to research on users' purchase decision.

2.4.2 Research on Channel Attribution

Channel attribution refers to the process of interpreting the influences of and assigning credits of conversion among the various marketing channels that contributed to a conversion event, such as a purchase or a registration (Shao and Li, 2011). Research on channel attribution is of great managerial importance in that it can help online retailers to better understand the effectiveness of different marketing channels and optimize their marketing spend accordingly in an informed and data-driven way.

Kannan et al. (2016) summarized the state-of-the-art in channel attribution research and called for more work in this area. In general, there are two kinds of attribution model: rule-based or data-driven. Table 2-1 lists the category of channel attribution models.

Table 2-1 Category of Channel Attribution Models

Category	Model	Description
Rule-based	Last-Click	The last view or click channel gets 100% credit. Easy to implement and interpret but may overestimate channels closer to conversion while underestimating others on the path.
	First-Click	The first-view or click channel gets all 100% credit. May overestimate the first channel touching the customer
	Average / Even	Each channel that the user has touched gets the same credit (100% divided by the number of touch points on the conversion path).
	Position Based	The first and the last channels get the most (say 40%) and all other channels in between share the rest.
Data-driven	Time Decay	Channels of later visits receive more credits than those from earlier visits.
	Multi-Touch	Use algorithms to calculate the attribution of each channel based on historical data of engagements and purchases. All touch points along the path are considered.

The rule-based attribution models are easy to understand and implemented in business, but are likely to be biased with limited effectiveness caused by their underlying assumptions. For example, the Last-Click Attribution, also known as Last-Touch Attribution (LTA), is used in Google Analytics by default. This model attributes every conversion to the last touch point and ignores any other channels along the conversion path. As a result, it fails to reflect the effects of previous channels, which may, in fact, contribute more to the final conversion. Its limitations have been thoroughly discussed in previous research (Anderl et al., 2016; Li and Kankan, 2014). To address such limitations, both academic research and managerial practice call for more sophisticated data-driven attribution models.

2.4.3 Research on The Effect of Single Channel

Some research studies the impact of individual channel such as search, social, or email on conversion. Section 2.3 has reviewed research on channels of search engine and social media. Chatterjee et al. (2003) study the effect of display ads on websites and reported that repeated or too-often ad exposure is, in fact, negative to conversion. Oliver and Michael (2011) introduce a two-stage model to evaluate the effect of paid search (such as Google Ads) on its performance (such as number of clicks and conversions) and reported that the position (e.g., top or right side) where the ad was displayed on the website has a significant impact. Reiley et al. (2012) conduct a field experiment to investigate the impact of added northern ads on click-through rates (CTRs) in search engine advertising and reveal a positive externality that increasing the number of north ads displayed actually benefit existing ads by increasing the CTR and boosting overall clicks.

The limitation is that in the current digital era, both online retailers and consumers are in a multiple-channel market environment. Such research contributes to the research on the effectiveness of each channel on conversion, providing a basis for researchers to furthur compare and measure the relative effectiveness in a multi-channel context.

2.4.4 Research on the effect of multiple channels

In the digital era, online retailers are faced with a multi-channel environment and need more sophisticated models to optimize their marketing strategy across channels. The straightforward last-click attribution has many limitations, such as underestimating the effects of email and referral, but inflating that of search. Shao and Li (2011) propose a bagged logistic regression model to quantify the attribution of different channels, based on data on what and how many ads from which channel have been viewed. This research improves traditional models by looking at the whole journey and including historic visit data from nonconverted users as well. One limitation is that this research only contains data in one month, which may not be long enough to cover the complete visit-to-purchase journey of users.

Ji and Wang (2016) extend this research by adding the time factor (visit time, which is about the exact time when a visit happens) of each visit to their model. They make an analogy between biological death and conversion and introduce a probabilistic multitouch attribution model to measure the influence of each visit. The research reports that some ads that lead to final conversions are, in fact, a consequence of previous interactions from other channels. This research is among the first to take into account the visit time and to measure the impact decay over time. One limitation of this research is that it treats each visit homogeneously without considering the duration of a visit.

Li and Kannan (2014) furthur propose a three-level measurement model based on individual-level visit path data of all channels along the visit-to-purchase journey. Their research reports that customers vary in their consideration of channels through which to visit a firm's website, given people's diverse habits for seeking information in the online shopping. Some customers may visit the online retailer's website directly, while others may consider the search channel for better prices and options, and some may consider both. From the perspective of the opportunity cost and cognitive cost in buying decision making, this research investigates the carryover effect in the same channel and the spillover effect between different channels with a Bayesian model. They report a significant spillover effect of firm-initiated channels such as display ads and emails on customer-initiated channels such as search. The paper reveals that the incremental contribution of each channel to purchase conversions varies significantly and the simple rule-based attribution models (last click and the seven-day average) may not accurately reflect the true contribution of each channel. As a result, the authors suggest that firms should use a more comprehensive approach to measure the effectiveness of their marketing channels and allocate their marketing budget accordingly. Zhang and Duan (2014) also investigate the spillover effect of referrals on search and report that social media referrals to other competing online stores have a significant negative impact on conversion. Though this research doesn’t directly address whether social media referral has a positive impact on conversion or not on the retailer’s own website, its finding reveals the importance of social media marketing in driving conversion.

2.5 Summary

Extant research on purchase conversion models mostly treats consumer search behavior as homogeneous points, lacking consideration for the details of each visit which have an important impact on assessing consumer’s purchase intention, behavior and decision. Among them, the duration (browsing time on website) of visits and the recency of visits can reflect the level of consumer interest and purchase intention towards a product or service or brand. Longer visit durations and more recent visit times may indicate greater consumer interest, thereby increasing the likelihood of conversion.

Secondly, most literature focuses on studying attention theory under the single marketing effect, and there is no relevant theoretical research on the attentional shift of consumers in the context of multi-channel situations and its impact on purchase behavior. Filling this research gap will contribute to a better understanding of the consumers’ purchase decision-making process in a multi-channel context and provide strong support for practitioners to develop more effective marketing strategies. This has become increasing important as in the digital age and in the context of online purchasing, there're abundant sources of information and it's worth investigating what kinds of sources the users pay their attention to and allocate their attention among.

Furthermore, channel exposures also have an important influence on assessing consumer behavior and intent. Different channel exposures, such as search engines, social media, and advertisements, may attract different types of consumers, thus having varying effects on purchase conversion. Similarly, the effectiveness of the visited page can also impact consumer purchase behavior. For example, an intuitive and easy-to-navigate page may increase consumer purchase intent, while a slow-loading or poorly designed page may decrease conversion rates.

Moreover, the extent literature has not fully considered the impact of channel diversity on the final purchase decision. In the actual purchase decision-making process, consumers often obtain information through multiple channels, including multiple online and offline channels. However, most research either treats these channels as independent factors without considering their interactions and influences, or the time span is insufficient to account for consumer purchase decisions. Therefore, this study aims to fill this research gap by incorporating channel diversity into the conversion model for the first time, to explore its impact on purchase decisions.

Research on user conversion models reveals that past visits have cummulative effects on the user's purchase decision. Therefore, when studying the purchase decision of online users, the cummulative visit effects should be considered instead of the current single visit. Previous research also reports that when measuring the effect of a visit on the final purchase, the detailed information such as what has been viewed for how long and the recency, is meaningful.

Both information retrieval theories and channel attribution models reveal that channel exposure, or the incoming channel of a visit, reveals a user's visit intention and has different influence over the visit effect. For example, visits from a user-initiated channel (such as search engine) indicate that the user is actively seeking information and is more likely to make a purchase decision than visits from a firm-initiated channel (such as display ads) where the user passively receives information. Research on channel attribution has reported the characteristics and different influence power of each channel on purchase. For example, display ads are good at moving low-intent users from the discovery or awareness stages to consideration, while social media referral to the competitor's website has a significant negative impact on the user's purchase decision.

Based on information retrieval theories, the channel diversity of a user, or how many channels from which the user visited the website, reflects his decision-making style. In addition, visits from other channels can be a supplement of the user's information retrieval or become a barrier (information overload) to the user's purchase decision. Therefore, it is worth considering the positive or negative impacts either directly on user conversion or indirectly on each channel of exposure.

Chapter 3 Hypothesis Development

Research on the user purchase journey has implied that the user's visit-to-purchase conversion could be a long journey from awareness to purchase (McKinsey, 2009). Some visits are out of planned purchase intention or price comparison, while some are just for information retrieval or from clicking a link shared by others on social media platforms. According to the theories of information retrieval, users may spend much time seeking information before making a decision from initiazation and information exploration to formulation and presentation (Kuhlthau, 2008).

Extant research on conversion has revealed that a user's purchase intention and purchase probability can accumulate over visits (Moe and Fader, 2004; Park and Park, 2016). In addition, the recency of each visit also impacts the cummulative effect on purchase decision as events happen closer have higher influences. In other words, the impact of a visit can decay over time and more recent visits will have higher weight on conversion than earlier ones (Ji and Wang, 2016). Related research also demonstrates the importance of visit details (such as pages viewed, time spent on browsing) in the purchase decision making of the user. The duration of the visit can reflect a user's intention to buy and attention allocated to the website's brand or product - a visit lasting 10 minutes should bring more information to the user and is more important in influencing his purchase decision compared to a visit lasting only 5 seconds. As a result, the duration (time spent in browsing the website) of each visit will play a key role in users' purchase and in modeling purchase conversion. To be specific, visit time has a positive impact on purchase conversion.

Research on channel attribution models indicates that different channels will have different influencing power or impact on the user's final purchase decision or conversion. For example, some research argues that paid search has a higher ROI (return of investment) or ROAS (return of advertising spend) than other channels such as affiliates, while some claims that the "search" channel is overestimated while the "retargeting" channel is underestimated (Nottorf, 2014). Despite multiple attribution models, it is generally accepted that the effect on conversion of each visit is influenced by the channel of exposure from which this visit comes, and can be accumulated with time decay depending on its recency.

Putting it all together, we propose that the effect of each visit is accumulated towards the user's purchase conversion, and that the effect of each visit is determined by its details including its the duration or browsing time on the website as well as the channel of exposure which has different influencing power on the visit effect. What's more, when accumulated, the effect of each visit will decay over time which is determined be its recency compared to the user's last. Thus, Hypothesis 1 is proposed as follows with 3 items:

Hypothesis 1.1: *The browsing time has a positive impact on conversion. More browsing time spent on the website leads to a higher purchase conversion probability.*

Hypothesis 1.2: *The browsing time (discoutned by recency) has a positive impact on conversion. With recency taken into consideration, more browsing time spent on the website also leads to a higher purchase conversion probability.*

Hypothesis 1.3: *The effect of each visit is influenced by its channel of exposure. Different channels of exposure show different influencing power on.*

Channel diversity refers to the distinct number of channels through which a user has gone through during his visit journey. It measures the variety of sources where users seek information or are exposed to different types of information. From the perspective of information retrieval for decision making, channel diversity reflects a user's decision-making style and can be used as an indicator of purchase conversion probability. In addition, according to the theory on information search process, a user can obtain different affective feelings or cognitive thoughts about the retailer and the product on different channels, which influences the general influence power of each channel (Kuhlthau, 2005). Furthurmore, when users are exposed to various information from different sources, the information may not always be helpful or informative but sometimes inconsistent or conflicting, causing the problems of information overload or cognitive fatigue. From the perspective of online retailers, users can be attracted and engaged from multiple channels, such as search engines, social media platforms, advertising networks and online stores.On the one hand, getting more information from more channels of exposure adds confidence and interest to purchase, as the users becomre more aware of the product, brand and promotions. On the other hand, too much information which sometimes is conflicting from different channels may increase the user's level of uncertainty or cognitive fatigue, making the information a hindrance rather than a helping hand in the decision-making of users. As a result, the user may experience a decreased well-being accompanied by a perceived loss of control (Bawden, 2020) and consequently a decreased willingness to purchase. Thus, Hypothesis 2 is proposed as follows,

Hypothesis 2: *Channel diversity has a direct positive effect on conversion and moderates the relationship between visit effect and purchase probability.*

As consumers engage in more extensive information search, they may encounter conflicting, ambiguous, or overwhelming information that leads to hesitation and choice deferral. Jacoby et al. (1974) find that too much information can lead to poorer decision making and less confidence in choices. In the context of online shopping, Wan et al. (2012) suggest that the abundance of information and alternatives can lead to information overload and decision paralysis, reducing purchase likelihood. Thus, Hypothesis 3 is proposed as follows,

Hypothesis 3: *The more times a user visits the website of a product, the less likely he will make the purchase decision.*

Some empirical studies have found results consistent with the hypothesis. For example, Moe (2003) analyzed clickstream data from an online store and found that visitors who engaged in more search-related activities (which could involve multiple visits) were less likely to make a purchase compared to those who exhibited more goal-directed browsing. Similarly, Bucklin and Sismeiro (2003) found that increased exposure to a website (in terms of total page views) was associated with a decrease in purchase probability, suggesting that more extensive browsing may reflect greater uncertainty or hesitation.

To sum up, the theoretical framework is illustrated in Figure 3-1.

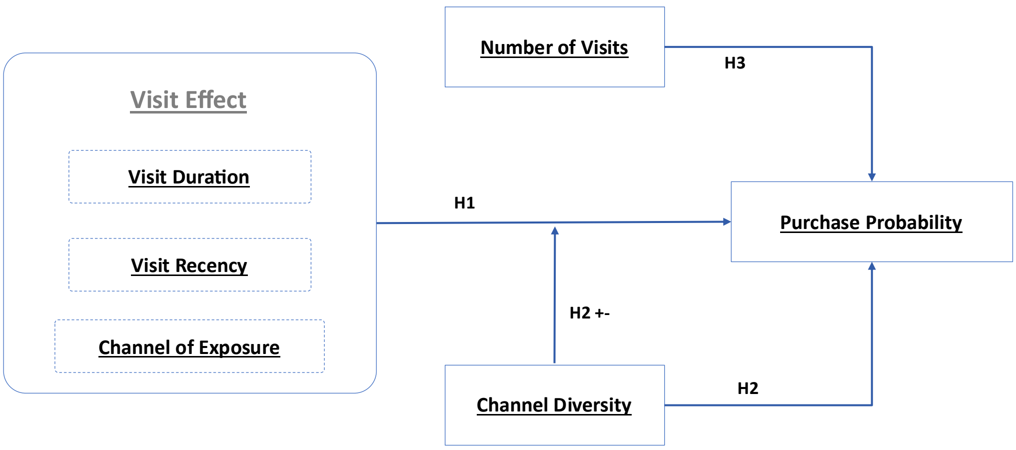

Figure 3-1 Theoretical Framework

Chapter 4 Research Design

4.1 Background

4.1.1 The Online Eyewear Market

The global eyewear market, made up of three main categories of eyeglasses (spectacles), sunglasses and contact lens, is valued at about USD 105.56 billion in 2020, of which 79.1% is from eye-glasses (spectacles) and about 15% from sunglasses. This market is estimated to grow from USD 152.95 billion in 2022 to USD 246.47 billion by 2030 according to Fortune Business Insight[2].

Prescription eyeglasses can be used to correct people's vision problems, such as near-sightedness, farsightedness and astigmatism, to improve the quality of life. In 2019, more than seven million pairs of prescription eyeglasses were sold online in the US. According to a report by the World Health Organization (WHO), more than 2.2 billion people worldwide have either near or distant vision impairments. The eyeglass segment can be further divided into frames and lens. Before buying a pair of prescription eyeglasses or sunglasses, the user needs provide the prescription including information about degree of near- or farsightness (denoted as S or Spherical), the degree of astigmatism (denoted as C or Cylinder), the orientation of the astigmatism that the user might have

[2] Refer to https://www.fortunebusinessinsights.com/industry-reports/eyewear-market-101749 for more up-to-date information about global eyewear market.

(denoted as Axis), and the amount of prismatic power measured in prism diopters (denoted as Prism).

Sunglasses can protect people against ultraviolet (UV) rays, glare, and debris during outdoor activities. The sunglass segment can be divided into prescription and plano sunglasses. With increasing awareness of the harmful effects of UV rays on the eyes and the importance of preventive eye care, along with the growing acceptance of sunglasses as a part of modern lifestyle accessories, the potential requirement for sunglasses is also huge and increasing. For example, about 80% of Americans wear sunglasses in the summer.

The global eyewear market continues to grow, considering the high population of individuals suffering from ocular disorders or diseases. Furthermore, due to the implementation of remote-work during the pandemic, people have to spend more time with electronic devices such as laptops. The popularity of social networks also increases the screen time on mobile devices with screens smaller than laptops. The longer screen time may cause eye strain, dry eyes, and related issues, resulting in more ocular problems and demands for antifatigue and vision-correction glasses. In addition, the aging problem also adds more demands to eyewear to address age-related eye problems such as cataracts, glaucoma, and macular degeneration. All of such factors are very likely to propel the entire eyewear market both online and offline. The latest research shows that the global eyewear market is expected to grow at a compound annual growth rate) of 4.37% in the period 2023-2027 period (Statista, 2023a). North America is the biggest market for eyewear in the word, with the 2020 market size was USD 30.87 billion. Figure 4-1 shows the actual and forecast size of the eyewear market size in North America from 2017 to 2018 according to Fortune Business Insight[3].

Figure 4-1 North America Eyewear Market Size

The main distribution channels for eyewear are brick-and-mortar retail stores, online websites, and clinics or supermarkets. The advantage with offline physical stores is that it is convenient for users to drop by and the in-store professionals can guide customers through the purchase process and provide help instantly. In addition, many of the elderly still prefer offline face-to-face communication. On the other hand, with the popularity of social media and e-

[3] [3] Refer to https://www.fortunebusinessinsights.com/industry-reports/eyewear-market-101749 for more up-to-date information about global eyewear market and year-over-year growth and breakdown by product type etc.

commerce, it is not surprising that more and more people will purchase eyeglasses or get related information from online stores. Online eyewear websites provide a wide range of products at competitive prices and often with virtual try-on tools. Online eyewear retailers also make use of social media and online advertising to attract users to their websites. In 2020, more than 44% of adults browsed eyewear websites to assist in their purchase of prescription eyeglasses, and 14.1% of eyeglass buyers made their purchase directly from online stores directly.

4.1.2 The User Path on Eyewear Website

People visit eyewear websites for many different purposes, to examine the brand and find suitable products according to their requirements, to check customer reviews across websites, to compare prices and features, or to make online purchases directly. according to a survey conducted by The Vision Council in 2020, 80% of recent buyers reported that they will use the internet to some extent for information retrieval before placing an order[4] (Vision Monday, 2021).

The user can register and log in to the website, or remains anonymous for browsing. According to business knowledge, a typical user path on an eyewear website can be illustrated in Figure 4-2.

[4] Refer to https://thevisioncouncil.org/blog/vision-council-releases-results-2020-internet-influence-report for more information about the use of Internet among consumers who recently purchased eyewear products.

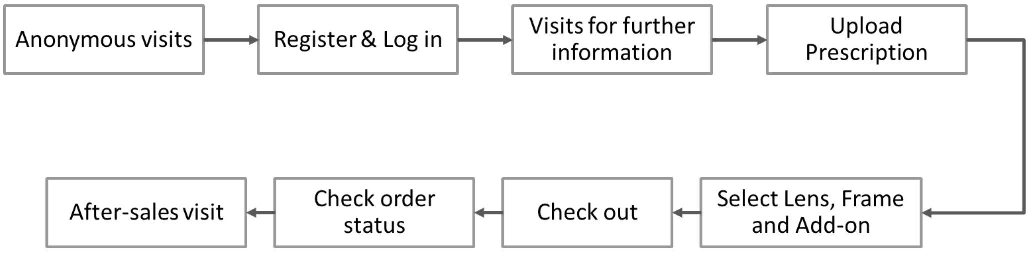

Figure 4-2 User Path on Eyewear Website

4.1.3 Company X

Company X is an online eyewear retailer that sells eyeglasses and sunglasses on its own website. It mainly targets users in Europe and North America and conducts on-line marketing activities through various channels including search engine, social media, display ads, retargeting and email, etc. To be specific, the marketing department launches several campaigns over online channels of Google Ads, Bing Ads, Google Display Network, Facebook, Instagram, Criteo, Steelhouse, Semrush, and Email etc. Potential users are directed from various channels to the website where they can browse products and make purchases directly. Table 4-1 demonstrates the percentage of marketing expenditures by channel of exposure of this company, with paid search and paid social being the top 2 focused channels.

Table 4-1 Markeing Spending across Channels

Channel of Exposure	% of Total Marketing Spending
Paid Search	49.26%
Paid Social	21.05%
Retargeting	14.06%
Affiliates	5.95%
Email, SMS, DM	5.85%
Organic Social	2.66%
Content Creation	0.65%
Organic Search	0.53%

4.2 Sample and Variables

In this research, the data is collected at user session level with detailed information about each visit to the website including visit time (at what time the visit happens), channel of exposure, geographic location, device, page views, time spent on browsing the website (visit duration or browsing time), and whether there is a purchase in this visit or not. The structure of the raw data is described in Table 4-2 below.

Table 4-2 Raw Data Fields at User Session Level

Level	Variable	Description
Visit	email_visitorId	The unique identifier of a user
Visit	visit_start_time	The start time of a visit
Visit	time_on_site	The duration (exposed time) of a visit in seconds
Visit	medium	The type of channel such as cpc, psoc and affiliate
Visit	source	The source of the channel such as google, bing or facebook
Visit	page_views	Total number of web pages browsed in a visit
Visit	is_newvisit	Whether the user has visited the website before
Visit	is_transaction	Whether there's a purchase made in this visit
Visit	deviceCategory	Category (Laptop or mobile) of the device used
Visit	operatingSystem	The device's operating system
Visit	mobileDeviceBrand	The brand of the device if it's mobile

Based on such raw data, some additional information can be further calculated and aggregated at user or channel level to conduct our research. The variables used in our research is illustrated below:

- Probability converted: Whether the user is finally converted (made a purchase) or not. The value is 0 or 1 at user level.
- Cummulative browsing time (s): The total browsing time (time_on_site in raw data) of all visits of a user, counted in seconds. In our research we first calculate this value at user level for all channels, then split this into multiple fields per channel, and discount it by recency with 1% hourly decay.
- Number of visits: How many visits a user has made along his visit journey.
- Number of unique channels: How many unique channels a user has gone through in his visit journey. It's the quantitative measurement of channel

diversity.

- Time per visit (s): Browsing time per each visit (in seconds).
- Recency relative to last visit (m): The time difference between a user's last visit and the vist before the last one (calculated in minute).
- Revenue if converted: The transaction revenue if the user is converted. This variable is at user level and only applicable to converted users.

Table 4-3 below shows the statistical description of these independent variables that will be used in the regression models in this research.

Table 4-3 Statistical Description of Independent Variables

Variable	N	Mean	Std. dev.	Min	Max
Probability converted	1,097,876	0.2018488	0.4013802		
Cumulative browsing time (s)	1,097,876	1102.029	3271.191	0	362676
Number of visits	1,097,876	1.888497	3.117575	1	460
Number of unique channels	1,097,876	1.211518	0.5170277	1	8
Time per visit (s)	2,000,000	569.4772	968.6214	0	36061
Recency relative to last visit (m)	887,653	11575.4	17196.58	1	87838
Revenue if converted	221,605	86.93971	65.78208	14.236	4367.819

4.3 Data Processing

4.3.1 Data Cleansing and Enhancement

In our research, the raw data is collected over a continuous 3-month period with over 3 million visits from the website of the online eyewear retailer. Technically, information about each user's each visit is captured by Google Analytics, and the raw data is stored in the underlying Google Big Query. For various reasons, the raw data contains records with “dirty” data that need be cleaned. The data quality issues and processing methods are explained in this section.

First, we exclude vists with 0 pages viewed. Such visits are possibly clicks by mistake without any purchase or browse intention, and the users close the landing web page immediately even before it’s fully loaded. Keeping such visits may add noise to our research so they are deleted. Records with missing or non-standard visitor identifier are also removed. The size of the data set drops from over 3,000,000 to around 2,300,000 after this step of cleansing.

Second, we map the channels of exposure (medium and source) in the raw data with categorized channels. There’re 10+ mediums (such as psoc, retargeting, referral, cse etc.) and approximately 100 sources (facebook, google, ask, pinterest, webpush, semrush, duckduckgo etc.) in the raw data and are not fit for regression by default. Based on business practice and nature of channels, we map them into standard channel categories as explained in Section 4.3.2. Some abnormal mediums (such as test, text, link etc.) are kind of noise and visits from such mediums are excluded. In addition, our research focuses on visit-to-purchase conversion without considering repurchase or customer loyalty. To avoid re-purchase data interfering with this topic, visits of each user after his purchase are excluded. Such data is still kept in the original data set and can be used in further research. After this step of data harmonization and cleansing, the size of the data set drops to approximately 2,000,000.

Third, we enrich the data set by adding more fields at user or visit level as required by our model (described in Chapter 3 and Section 4.2). At user level, number of unique channels is calculated to measure his channel diversity. The cummulative visit time is also calculated for each user by adding the browsing time of each of his visits together. In addition, this cummulative visit time of each channel is also calculated.

Recency plays an important role in measuring the visit's effect on conversion, as people tend to give more importance to information acquired or event happened more recently than that in the past, and the acquired information may decay over time. So we first calculate the difference between the time each visit happened and that of the last visit as,

recency_in_hour = datediff(hour, happen_time_of_each_visit, happen_time_of_last_visit_per_user).

Then we calculate the cummulative discounted time of all visits at user level. In our research we use 1% hourly discount rate and a continuous discounting model to calculate the discounted time as,

$$Time\_discounted = time\_of\_visit * e^{-0.01 * recency\_in\_hour} \quad (1)$$

Aside from conversion, this research also investigates the impact on revenue because of its importance to retailers - increasing revenue is the ultimate goal of marketing and conversion. In this step, we get the revenue for each user from his last visit which contains a transaction (purchase) and has a non-zero revenue.

Figure 4-3 below illustrates the data cleansing and processing steps:

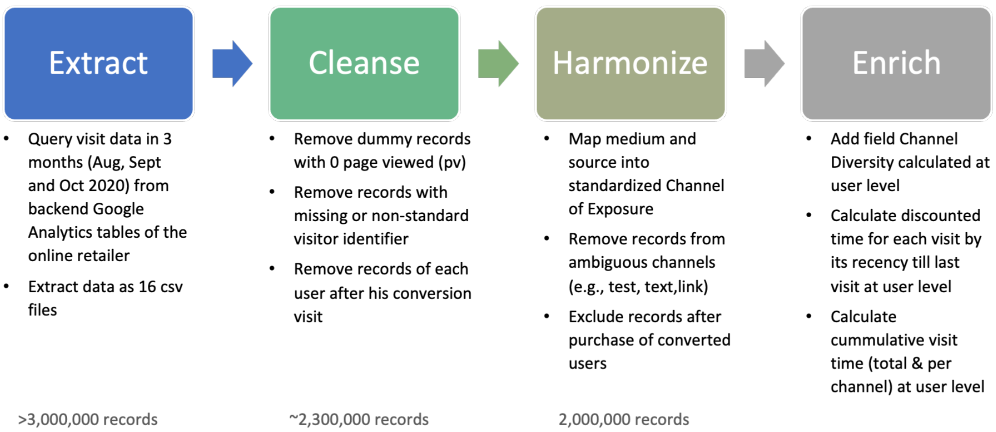

Figure 4-3 Data Cleansing and Processing Steps

4.3.2 Channel Categorization

In the raw data, the channels of exposure are very diverse. First, there're over 10 channels (mediums). Second, there're about 100 sources about the detailed vendor bringing the traffic (for example, there are multiple vendors such as steelhouse and criteo providing services of retargeting). They need be simplified and categorized to be able to be used in our research model.

Several ways have been tested to categorize them. For example, they can be categorized as firm-initiated (where the online retailer place paid ads or sponsored content, such as paid search, paid social or affiliate) or user-initiated (such as organic search, organic social, referral etc.), or as paid (the retailer pays for each visit or conversion) or organic (not paid). We find that such categorization methods are not practical in business – almost all marketing channels are planned or operated by firm and with budget allocated. Based on business practice and the nature of each channel, in this research we categorize all channels in the raw data into 8 groups which are illustrated in Table 4-4 below.

Table 4-4 Categorization of Channels in Raw Data

Code	Channel Category	Description	Paid or not	Initiated By
1	Paid Search	Pay search engines to place ads on search result pages to drive traffic. Channels of CPC, PCC or CSE (comparison shopping) from search engines like Google or Bing are included in this category.	Paid	Firm
2	Retargeting	Show targeted ads to potential users who have previously interacted with the brand. Retargeting platforms like Steelhouse are included together with Display network.	Paid	Firm
3	Affiliates	Pay affiliates on each visit or purchase made by clicking ads on affiliate network. Affiliate networks like Impact and CJ are included together with referral.	Paid	Firm
4	Paid Social	Pay social media platforms (Facebook, Twitter, LinkedIn etc.) to display sponsored posts or placements to targeted users to reach new customers or drive traffic.	Paid	Firm
5	Organic Social	Engage with audience and promote content on social media platforms without any paid promotion. Referral is also included in this category.	Not Paid	User
6	Email, SMS, DM	Send email or sms to their potential or existing customers to promote visit or purchases. Systems like listrak and ltdm are included in this category.	Not Paid	Firm
7	Organic Search	Obtain traffic naturally through search results determined by search engines without paying for ads placement.SEO is often used to optimize the website design for a higher rank to be displayed in the top after a search.	Not Paid	User
8	Direct("None")	No medium or source is captured and categorized as direct visit to the website.	Not Paid	User

Chapter 5 Results

5.1 Descriptive Statistics

After cleansing, the data set contains approximately 2,000,000 records of visit from ~1,000,000 users, of which about 220,000 are converted. Based on the different number of visits, number of channels gone through and how much time each user spends on the website in each visit and in all visits, the user's intention and purchase probability can be measured.

First, we check the number of visits from each channel of exposure. The result is illustrated in Figure 5-1. It's observed that paid search, on which the most of the company's marketing expense is spent, brings the most traffic to the website. Overall, search outperforms social media to a great extent, while precise message push (Email and SMS) also brings a lot of traffic. This reflects the normal behavior in eyewear industry that most users tend to retrieve information from search engines.

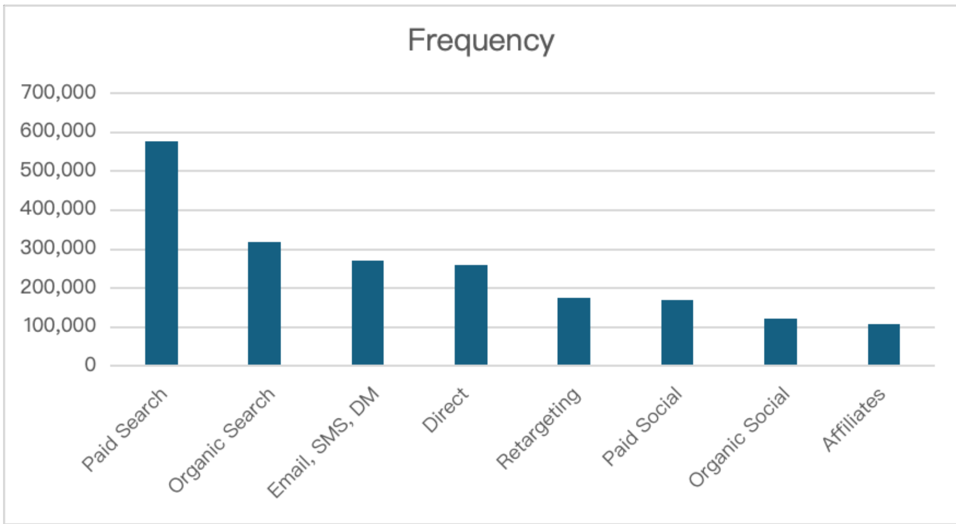

Figure 5-1 Distribution of Visits by Channel of Exposure

Second, we proceed to check the distribution pattern channel diversity at user level, or normally how many unique channels a user goes through. Table 5-1 below shows the statistics of number of unique channels. Most users go through only 1 to 3 channel categories in all their visits (remember that in our research the raw channels are categorized. For example, if a user clicks 2 paid ads on 2 social media platforms, his channel diversity is 1 as both are paid social). This can be explained that people can resort to different sources to retrieve information, while their attention is limited, so channel diversity won't be very high for most users.

Table 5-1 Distribution of User by Channel Diversity

Number of unique channels	Frequency	Percent	Cummulative percent
1	1,481,027	74.05	74.05
2	384,075	19.2	93.26
3	98,818	4.94	98.2
4	27,094	1.35	99.55
5	7,310	0.37	99.92
6	1,489	0.07	99.99
7	182	0.01	100
8	5	0	100

To examine whether there's any difference between converted and non-converted users, we furthur compare the distribution of channel diversity between converted and not-converted users. The results are illustrated in Figure 5-2 below. Our finding reveals that the distribution pattern still holds, while converted users tend to have a higher channel diversity.

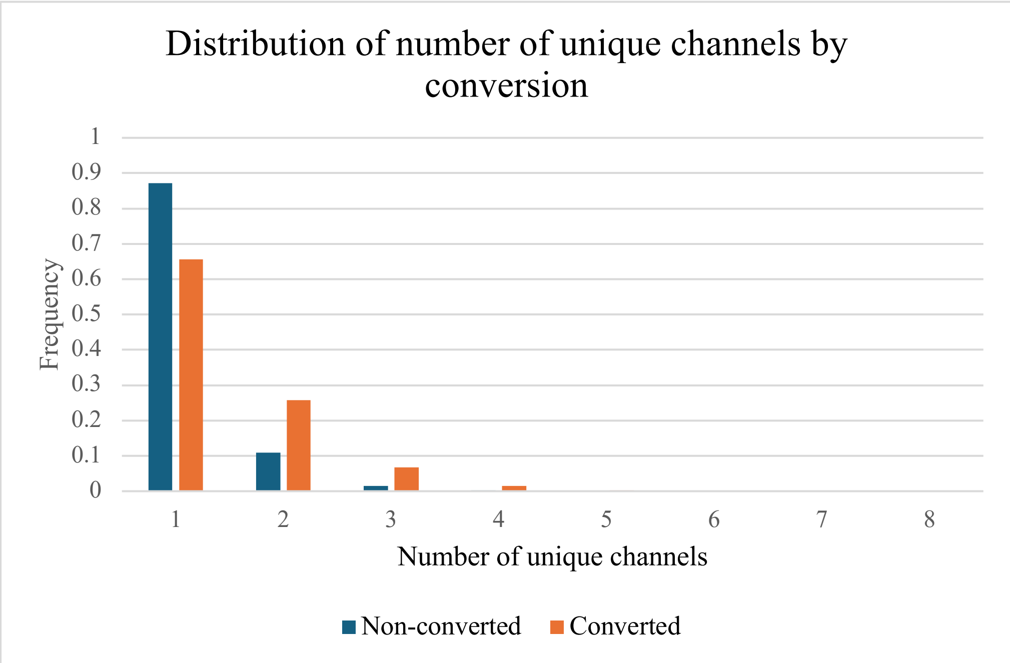

Figure 5-2 Distribution of Channel Diversity by Conversion

Figure 5-3 below shows the number of visits at user level. A power-law distribution pattern is observed that most users pay only 1 to 3 visits to the website before either purchasing or leaving. We conduct a comparison of such data between converted and non-converted users and find that converted users pay more visits than unconverted users while both show a power-law distribution pattern. Though a user may pay more visits in a longer time window than the 3-month period in our data set, the finding implies that the user's first several visits are key to retain or convert them and that online retailers need pay attention to them.

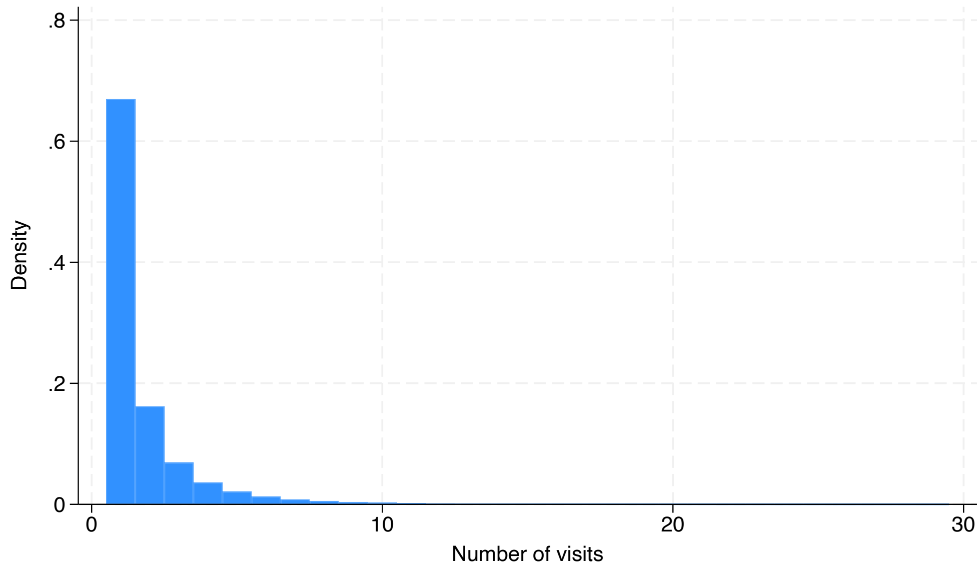

Figure 5-3 Distribution of Number of visits

In Chapter 3 we hypothesize that visit time, which reflects a user's attention, is an important factor the the visit's effect on conversion. Figure 5-4 shows the distribution of browsing time (how much time the user spends on the website in each visit) in seconds, which also shows a power-law distribution pattern. It's natural that most visits is less than 200 seconds (~3 minutes), and few visits will be longer than 5 minutes.

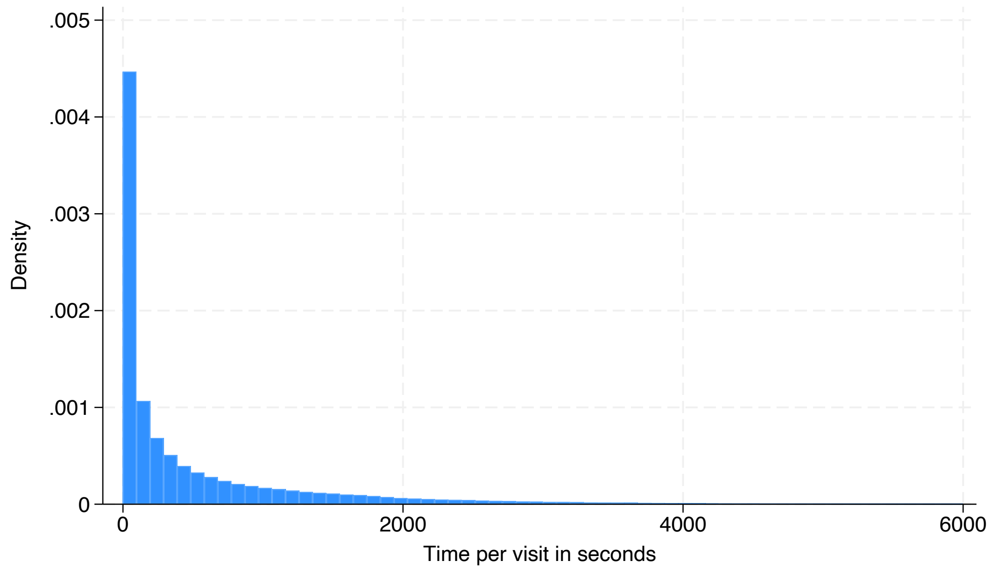

Figure 5-4 Distribution of Borwsing Time per Visit

To examine the difference in browsing time in each visit among channels, we calculate the average browsing time of all visits by channel of exposure. The result is illustrated in Figure 5-5 below. It's found that the average browsing time of visits from search engine (either paid or organic) is higher than most visits, showing that users who actively search for eyeglass related terms have a higher interest or intention, while some users from social media channels are simply triggered by posts shared by their friends and have no further interest after clicking the link.

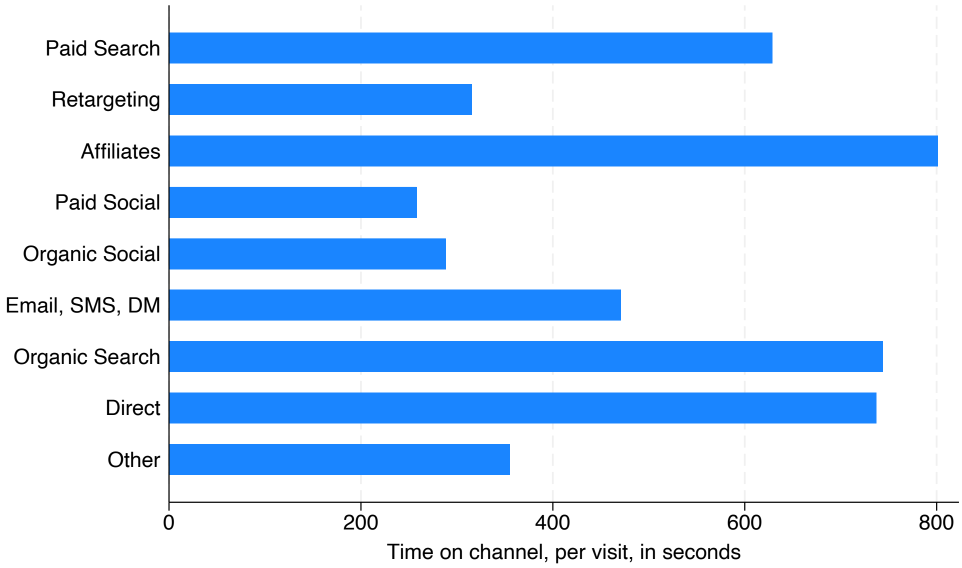

Figure 5-5 Average Time per Visit on Channel

The distribution of browsing visit time from all channels has been illustrated previously. Now we further extend the statistics by checking the value per channel and find that the distribution of browsing time of each visit follows a power-law distribution pattern for all channels despite the channel speciality, as summarized in Figure 5-6 below.

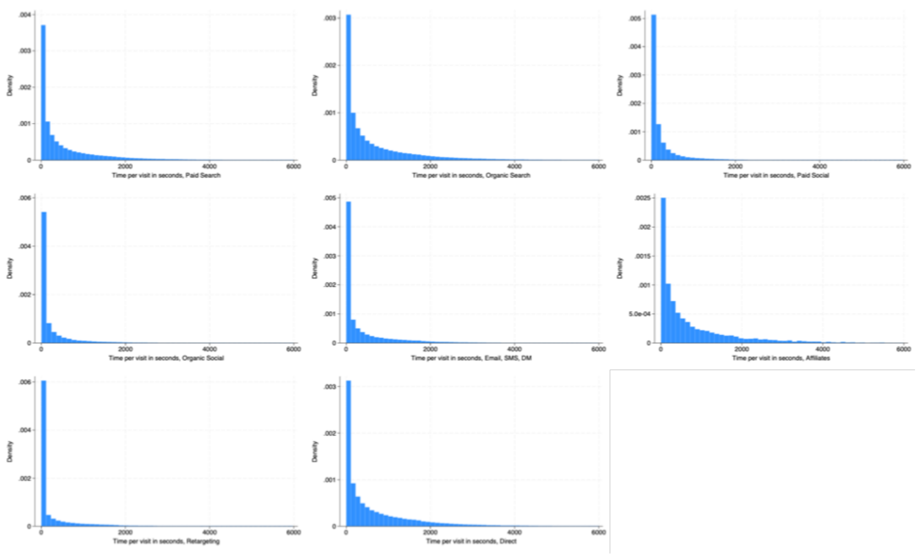

Figure 5-6 Distribution of Browsing Time per Visit for Each Channel

In Chapter 3 we hypothesize that the effects of previous visits along the user's visit journey can be accumulated on the user's final conversion. So, aside from the browsing time of each visit, we assume that the accumulated visit time at user level has a more direct impact on conversion. Table 5-2 below depicts the cummulative browsing time of all visits of each user by channel. Similar to that of browsing time per visit, visits from search engine also exhibits a higher total browsing time compared to those from other channels like social media platforms.

Table 5-2 Cummulative Time at User level per Channel

Cummulative per user last visit	**N**	**Mean**	**Std. dev.**
Paid Search	1,097,876	347.7971	1294.992
Retargeting	1,097,876	53.36138	569.6399
Affiliates	1,097,876	83.55583	2000.491
Paid Social	1,097,876	41.11382	303.8062
Organic Social	1,097,876	33.12629	465.3232
Email, SMS, DM	1,097,876	123.9929	1085.75
Organic Search	1,097,876	227.2996	1010.605
Direct	1,097,876	191.5124	1038.122
Other	1,097,876	0.269949	29.65231

Considering that converted and non-converted users may behave differently on the website, the cummulative time spent on visits from each channel should differ between converted and non-converted users. To check this difference, we further split the cummulative time of each user by conversion or not for each channel. The result is illustrated in Figure 5-7. We find that converted users show a higher average cummulative time in all channels. This aligns with the principles of attention theory that more time spent means more attention allocated which potentially increases the user's probability of purchase.

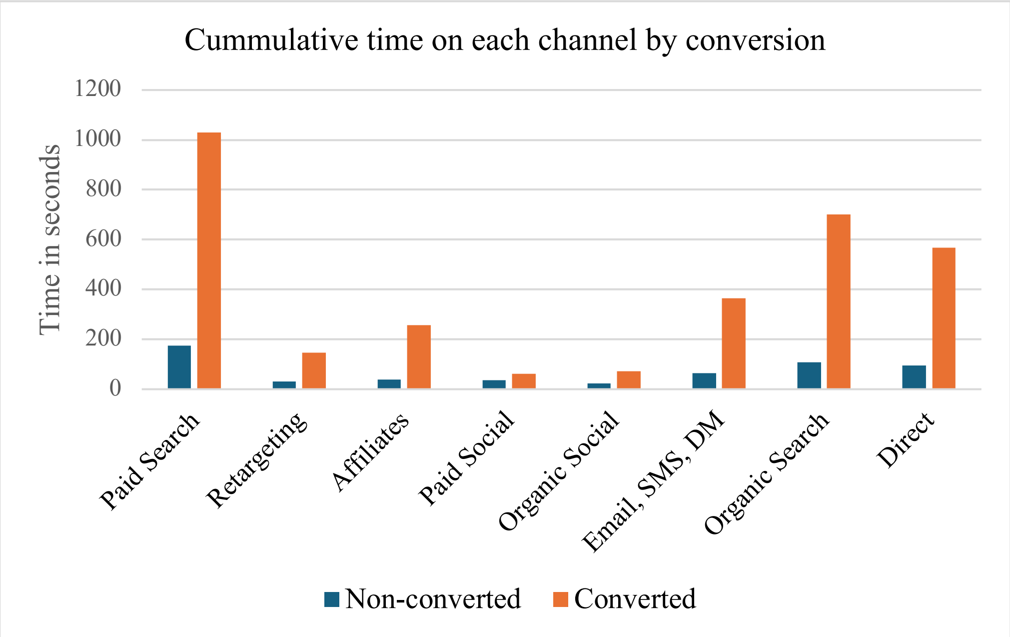

Figure 5-7 Cummulative Time per Channel by Conversion

Similar to what's demonstrated previously (in Figure 5-6), we examine the distribution of cummulative time at user and channel level to check the differce among channels of exposure. We find that same as browsing time of each individual visit, the cumulated time of each user also exhibits a power-law distribution pattern in all channels. Figure 5-8 below shows a summary, while figures of each channel is listed in appendix.

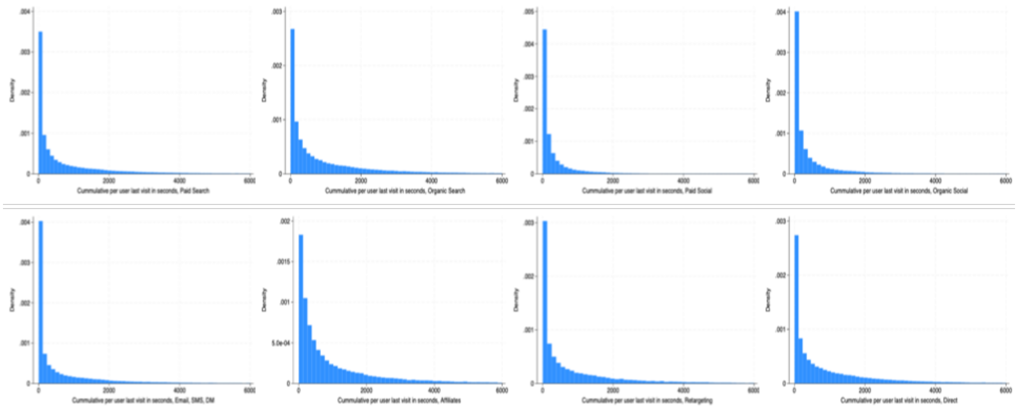

Figure 5-8 Cummulative Visit Time at User Level of Each Channel

According to recency effect, people tend to give more importance to events or information happened more recently than those happened in the past. In addition, the user's acquired knowledge of a product may decay over time. So, instead of simply adding the browsing time of each visit, it should make more sense to calculate the cummulative time of each user by discounting the browsing time of each visit by its recency compared to the user's last visit in the time period in our data set. The discounted time of each visit is used when calculating the cumulated time at user level. Here we examine the distribution pattern of recency (in hour) of each visit compared to the last one of each user. The results are depicted in Figure 5-9 which demonstrates a power-law distribution, indicating that most visits happen in the near-past of the last one. On the one hand, most users pay 1 or more visits within 3-5 days before either leaving or purchasing. On the other hand, users with visits in the long tail may be interested in or loyal to the brand, even though they don't purchase anything in this period.

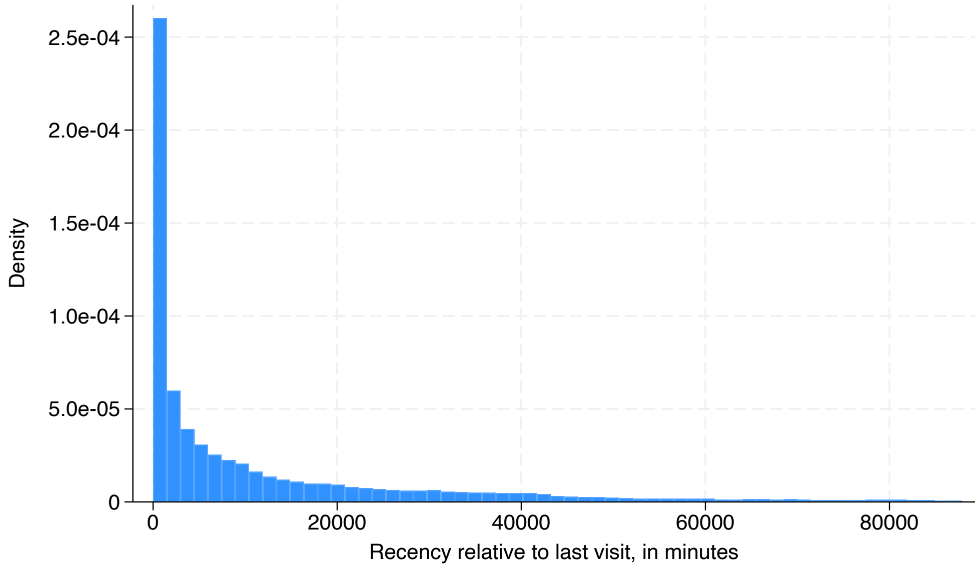

Figure 5-9 Distribution of Recency of Each Visit

Most research on conversion focuses on whether the user has made a purchase or not. However, in business practice, retailers want to understand each channel's contribution to revenue gained as the ultimate purpose of marketing is to drive revenue growth. Our research contribute to this area by checking the impact on revenues of converted users. The distribution of revenue for all converted users (in total 221,605) is depicted in Figure 5-10 which demonstrates a left-skewed normal distribution. Most order's revenue is between $20 and $100 which is consistent with the majority product (lens + frame) in the website. In addition, there're few products with a price less than $20 in the website, making the normal distribution left-skewed.

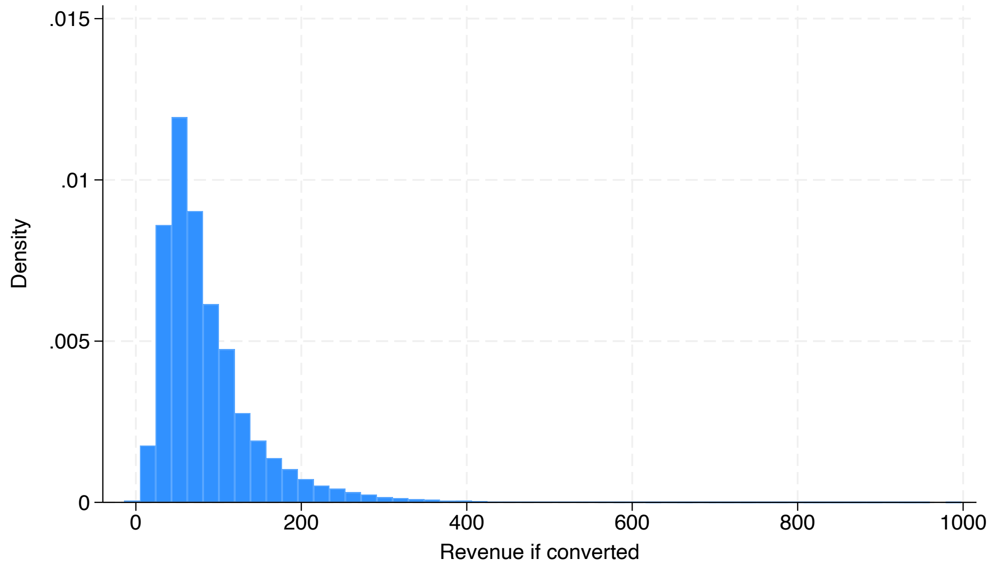

Figure 5-10 Distribution of Revenue at User Level

5.2 Regression Results

In this research we examine the effects of independently variables on user conversion and then on revenue as robustness check. The models are demonstrated in equations (1) and (2) below. The detailed explanation of each independent and dependent variables are listed in Table 5-3.

Equation 1 Equation for Conversion Probability

$$Probability(y=1) = \sum_{k=1}^{n} W_k \cdot X_k + \beta_0 Z + \sum_{k=1}^{n} \beta_k \cdot Z \cdot X_k + \beta_1 \mathrm{N} + \varepsilon$$

Equation 2 Equation for Revenue

$$Revenue = \sum_{k=1}^{n} W_k \cdot X_k + \beta_0 Z + \sum_{k=1}^{n} \beta_k \cdot Z \cdot X_k + \beta_1 \mathrm{N} + \varepsilon \qquad (2)$$

Table 5-3 List of Independent and Dependent Variables

Variable	Description
y	Dependent variable: Whether the user has converted or the conversion probability
Revenue	Dependent variable: revenue of converetd user
n	Total number of possible channels in the entire data set
X_k	Total browsing time from channel k for each user
W_k	Weight of channel k on conversion
Z	Channel diversity (number of unique channels for each user)
β_0	Weight of channel diversity on conversion
β_k	Weight of channel diversity on channel k's effect on conversion
N	Number of visits per user, indicating the extent of user's hesitation and price sensitiveness
β_1	Weight of Number of visits

We have 4 model specifications with progressive inclusion of independent variables as explained below:

- Model 1 checks the impact of number of visits (N) and the total visit time of all visits from all channels at user level.
- Model 2 checks the impact of channel diversity (Z) on conversion.
- Model 3 examines the weight of each channel (W_k) on the effect of each visit by its browsing time (X_k) with the total browsing time used in model 1 excluded.
- Model 4 furthur checks the impact of channel diversity on each channel (β_k) by interacting it with cummulative time per channel.

Each model specification is tested against two dependent variable separately: conversion and revenue, as illustrated in equations (1) and (2). The browsing time is not discounted in this section and will be discounted by recency with 1% hourly decay in Section 5.3 as additional analysis.

Table 5-4 below demonstrates the impact of number of visits, total browsing time and channel diversity at user level based on results of all models. The result keeps stable across all models. Number of visits exhibits a significantly negative impact on both conversion and revenue which supports Hypothesis 3 that the more times a user visits the website, the less likely he will make a purchase. This can be attributed to the fact that users with more visits tend to become more and more hesitated and also price-sensitive. As a resut, most of such users either abandon the purchase intention or switch to some cheaper alternative. On the other hand, the cummulative browsing time along the user's visit history is significantly positive to both conversion and revenue. This aligns with the principles of attention theory, that when users spend more time on a website, it indicates that they are allocating more attention to the brand or products of the online retailer. In turn, this increased attention can potentially lead to a higher probability of purchase conversion. In addition, more time spent on the website may help the user buy more products – for example, at first the user may just want to buy a pair of sunglasses which consist of lense and frames. But during navigating around the website, he finds some interesting accessories (such as case, lens cleaning kit) and is likely to add them to cart to buy together. As a result, more time spent on the website indicates higher order revenue. Note that this gross total of browsing time doesn't exist in models 3 and 4, where the cummulative time per channel is used instead (one variable for each channel). This supports Hypothesis 1.1 that the cummulative visit time has a positive impact on user conversion and the impact is significant. Putting the impact of these 2 independent variables (number of visits and cummulative browsing time) together, we find that the best users to the online retailer are those with few visits (not hesitating nor price-sensitive) but long time spent on the website (much attention allocated).

Table 5-4 Impact of Cummulative Visit Time, Number of Visits and Channel Diverisity

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Dependent variable:	Converted				Revenue			
Number of visits	-0.030***	-0.047***	-0.035***	-0.030***	-2.147***	-2.398***	-2.296***	-1.930***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.058)	(0.066)	(0.067)	(0.068)
Cummulative browsing time	0.062***	0.066***			3.720***	3.785***		
	(0.000)	(0.000)			(0.049)	(0.050)		
Number of unique channels		0.184***	0.147***	0.214***		1.837***	1.456***	4.485***
		(0.001)	(0.001)	(0.001)		(0.226)	(0.229)	(0.263)
Constant	0.190***	-0.006***	0.003***	-0.098***	81.290***	79.150***	78.747***	71.631***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.183)	(0.320)	(0.329)	(0.415)
Observations	1,097,876	1,097,876	1,097,876	1,097,876	221,605	221,605	221,605	221,605
R-squared	0.125	0.170	0.205	0.255	0.029	0.029	0.031	0.038

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Independent variables not shown: Effectiveness of individual channels (models 3 and 4), interaction of number of unique channels and individual channels (model 4)

Models 2, 3 and 4 also reveal that number of unique channels for each user, or channel diversity, has a significantly positive impact on both user conversion and revenue. This supports Hypothesis 2 that channel diversity has a significantly positive impact on conversion. Such effect can be explained as when users are exposed to a variant of channels, they retrieve more information about the product and the brand which enhances their confidence in making the purchase decisions. On the other hand, retrieving information from various channels may lead to the problem of information overload or cognitive fatigue as mentioned in Section 2.3. When users are exposed to different channels, they may encounter much information about the product itself, reviews, recommendations or promotional message which are sometimes beneficial while some other times conflicting or inconsistent which could harm their purchase probability. In our research, the problem of information overload is not found in models 2, 3 and 4. This is consistent with the finding of an extant research investigating the impact of added northern ads on the click-through rate and overall clicks of existing ads (Reiley et al., 2012), that increasing the number of north ads displayed actually benefit existing ads by increasing the CTR and boosting overall clicks despite the potential information overload problem of more north ads.

In models 3 and 4, the cummulative browsing time of all visits per user is split into that per channel to check the different effectiveness on conversion and revenue of each channel of exposure. For each user, it's cummulative browsing time of all visits of each channel is calculated. If the user doesn't have any visit exposed from some channels, the corresponding values will be 0. As demonstrated in Table 5-5 below and same as total cummulative time investigated above, the cummulative time spent on the website of each channel is significantly positive on either conversion or revenue. This supports Hypothesis 1.1 that browsing time of a visit, despite its channel of exposure, has a positive effect on conversion. It can be attributed to the fact that despite channel characteristics, time spent from each channel on the website reflects user's attention allocated, and consequently has positive influence on his interest in the brand or product and his purchase probability.

Table 5-5 Regression Result for Channel of Exposure

	(3)	(4)	(3)	(4)
Dependent variable:	Converted		Revenue	
Cummulative time on channel:				
Paid Search	0.091***	0.174***	4.661***	6.127***
	(0.000)	(0.001)	(0.074)	(0.143)
Retargeting	0.047***	0.052***	4.169***	4.343***
	(0.001)	(0.001)	(0.156)	(0.407)
Affiliates	0.034***	0.030***	3.888***	9.101***
	(0.000)	(0.000)	(0.099)	(0.258)
Paid Social	0.076***	0.112***	2.468***	1.767***
	(0.001)	(0.002)	(0.300)	(0.594)
Organic Social	0.049***	0.095***	4.957***	7.524***
	(0.001)	(0.002)	(0.244)	(0.547)
Email, SMS, DM	0.042***	0.130***	2.794***	7.100***
	(0.000)	(0.001)	(0.079)	(0.199)
Organic Search	0.097***	0.151***	3.934***	5.114***
	(0.000)	(0.001)	(0.084)	(0.163)
Direct	0.085***	0.170***	3.647***	6.841***
	(0.000)	(0.001)	(0.092)	(0.191)
Other	0.062***	0.124***	7.722***	6.182
	(0.012)	(0.029)	(2.445)	(6.559)
Constant	0.003***	-0.098***	78.747***	71.631***
	(0.001)	(0.001)	(0.329)	(0.415)
Observations	1,097,876	1,097,876	221,605	221,605
R-squared	0.205	0.255	0.031	0.038

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Independent variables not shown: Number of visits and number of unique channels (models 1 an 2), interaction of number of unique channels and individual channels (model 4)

The difference lies in the effectiveness (influencing power) of each channel, as revealed by model 3 and illustrated in Figure 5-11 below. This supports

Hypothesis 1.3 that channel of exposure reflects the user's intention and decision-making pattern and has a significant influence on each visit's effect on conversion. Figure 5-11 carries many interesting findings. First, we observe that search engines outperforms social media platforms in terms of their effectiveness on purchase conversions. That's, an equivalent increase in browsing time from visits originating from search engines yields a higher increase in the user's purchase probability compared to that from social media platform. This is consistent with our common knowledge that search engines are better at harvesting demands of high-intent users (they are searching for related products), while social media platforms are better at generating demands. What's more, organic search exhibits a high weight on conversion in our research, partially because the company has spent much effort on search-engine optimization (SEO) with a large IT team. As a result, products from our target company's website are likely to be in the top of the result page when users conduct a search. Second, when assessing the impact on revenue, we find that organic social shows the highest influencing power on converted user's order amount, even a little higher than paid search which normally generates the most revenue. This effect can be attributed to the inherent strength of social media platform that it's a good influential place where consumers can showcase luxury or expensive products, which influences their friends' or followers' purchasing decisions on what to buy. This also implies business practitioners that organic social marketing, which is about maintaining brand awareness and fostering client relationship in the long run instead of boosting traffic or revenue quickly, is worthy of more investment. Third, paid search exhibits a stable strong impact on both conversion and revenue indicating it a good channel for online retailers to invest on. In our target company, paid search is the channel where the most marketing expense is allocated on and the most revenue is generated. Last but not least, paid social channels are helpful in converting users, but such users are very likely to be very price-sensitive with low revenue contributed. Business practitioners can refer to our findings in their marketing stategy considering the difference in each channel's effectiveness on both conversion and revenue.

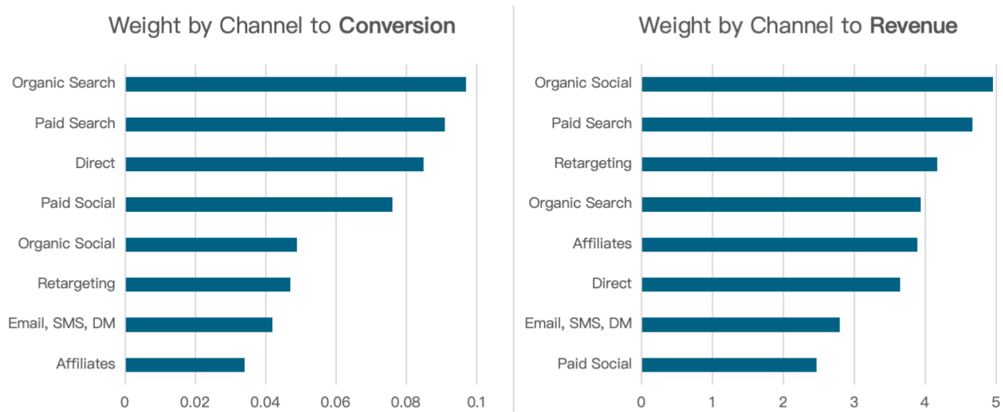

Figure 5-11 Comparison of Channel Effectiveness

In models 2 and 3, we has investigated the impact of channel diversity and each individual channel separately. In model 4, we extend our analysis to examine the impact of channel diversity on the effectiveness of each channel in driving conversions, as stated in Hypothesis 2. To do this, we interact channel diversity (measured as number of unique channels) with time on each channel. The results are illustrated in Table 5-6, revealing that though channel diversity itself is positive to conversion, it shows a negative interaction effect through most channels. At first we hypothesize that channel diversity may strength the effectiveness of some channels on conversion while weaken some others, but the model result shows that channel diversity weakens almost all individual channels. This supports Hypothesis 2 that channel diversity weakens the influenccing power of each individual channel. Business practitioners need be careful when launching marketing activities or engaging the same users over multiple channels. The effect could potentially be explained by the diluting effect on a single channel when the user acquire information in multiple ways. In other words, when users are exposed to a larger variety of channels, their attention may become divided among them. As a result, the impact of each individual channel on the users' decision-making and purchse conversion gets split and diluted. With limited attention allocated in each single channel, users may not give equal consideration as in a single channel context, reducing the effectiveness of any single channel in driving conversions. Another potential explanation is that in some cases, different channels may provide users with similar or redundant information about product or promotion. Consequently, the incremental impact of each additional channel decreases. As a result, though additional channel may increase the overall conversion probability, the power of each individual channel gets decreased.

Table 5-6 Regression Results of Model 4

	(4)	(4)
Dependent variable:	Converted	Revenue
Number of unique channels interact with time on channel:		
Paid Search	-0.047***	-0.724***
	(0.000)	(0.072)
Retargeting	0.002***	-0.018
	(0.001)	(0.151)
Affiliates	0.001***	-2.142***
	(0.000)	(0.098)
Paid Social	-0.018***	0.503**
	(0.001)	(0.256)
Organic Social	-0.017***	-1.037***
	(0.001)	(0.219)
Email, SMS, DM	-0.029***	-1.536***
	(0.000)	(0.072)
Organic Search	-0.029***	-0.531***
	(0.000)	(0.084)
Direct	-0.054***	-1.786***
	(0.000)	(0.109)
Other	-0.017*	0.906
	(0.010)	(2.267)
Constant	-0.098***	71.631***
	(0.001)	(0.415)
Observations	1,097,876	221,605
R-squared	0.255	0.038

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Independent variables not shown: Number of visits and number of unique channels (models 1 and 2), effectiveness of individual channels (model 3)

5.3 Additional Analysis

In Chapter 3 we hypothesize that recency is an important factor in measuring the effect of visits on conversion, assuming that more recent visits have a stronger effect than those happened earlier. In this section we extend our research by applying a continuous discounting method to browsing time with 1% hourly discount rate as additional analysis. The discounted browsing time of each visit is then calculated as,

$$Time\_discounted = time\_of\_visit * \ e^{-0.01 * recency\_in\_hour}$$

where *time_of_visit* indicates the time spent on the website of each visit (browsing time), *recency_in_hour* indicates the difference between each visit and the last one of each user (calculated in hour), and *time_discounted* is the discounted visit time which will be used in calculating the cummulative browsing time at user or channel level.

With this discounted browsing time, we re-run the 4 models as mentioned in Section 5.2 in the 2 equations. The results are explained below.

Same as the method described in Section 5.2, we first check the impact of number of visits, total discounted browsing time and channel diversity at user level. The result is depicted in Table 5-7 below. Consistent with section 5.2 where recency is not considered, number of visits is still significantly negative on conversion which supports Hypothesis 3. Cummulative browsing time with discount still reveals a significantly positive effect on conversion. What's more, when recency is considered with continuous time discounting, the impact becomes even stronger. This again supports Hypothesis 1.1 and 1.2 that both visit time and recency affects user's conversion probability and the effects are cummulative. However, channel diversity exhibits a negative impact on revenue in models 2 and 3 in robustness check. This can be attributed to the fact that when users have been exposed to a large variety of channels, they may retrieve much information about the price, coupon, promotion as well as reviews. As a result, the user may then choose to use the most cost-effective promotion if he decides to buy. As a result, though the conversion probability is increased with additional channels, the order amount gets decreased. This also implies business practitioners that when engaging users with multiple channels, they need balance the goals of converting them or achieving a higher revenue.

Table 5-7 Impact of Number of Visits, Cummulative Time and Channel Diversity with Time Discounted

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Dependent variable:	Converted				Revenue			
Number of visits	-0.004***	-0.009***	-0.009***	-0.008***	-0.470***	-0.278***	-0.325***	-0.209***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.043)	(0.049)	(0.049)	(0.050)
Cummulative time (discounted)	0.129***	0.125***			5.586***	5.611***		
	(0.000)	(0.000)			(0.062)	(0.062)		
Number of unique channels		0.078***	0.079***	0.167***		-1.883***	-2.406***	1.831***
		(0.001)	(0.001)	(0.001)		(0.222)	(0.227)	(0.292)
Constant	0.097***	0.015***	0.013***	-0.104***	73.462***	75.567***	76.516***	69.322***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.212)	(0.326)	(0.336)	(0.460)
Observations	1,097,876	1,097,876	1,097,876	1,097,876	221,605	221,605	221,605	221,605
R-squared	0.284	0.292	0.296	0.336	0.039	0.039	0.041	0.044

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Independent variables not shown: Effectiveness of individual channels (models 3 and 4), interaction of number of unique channels and individual channels (model 4)

Model 3 examines the weight of each channel by replacing the total cumulated discounted visit time by that per channel and the results are depicted in Table 5-8 below. We find that consistent with the results as demonstrated in Section 5.2, the cummulative visit time on all channels exhibits a significantly positive impact on both conversion and revenue. In fact, after discounting visit time by its recency relative to the user's last visit, the weight of each channel becomes stronger than that in Section 5.2 (Table 5-5) in model 3. The difference lies in the different weight of each channel on their contribution or influencing power to conversion and to revenue as demonstrated in Figure 5-12 below.

Table 5-8 Weight of Channels with Time Discounted by Recency

	(3)	(4)	(3)	(4)
Dependent variable:	Converted		Revenue	
Cummulative time (discounted) on channel:				
Paid Search	0.137***	0.223***	5.815***	7.664***
	(0.000)	(0.001)	(0.088)	(0.188)
Retargeting	0.090***	0.128***	5.999***	7.749***
	(0.001)	(0.002)	(0.212)	(0.566)
Affiliates	0.144***	0.217***	7.528***	12.850***
	(0.001)	(0.002)	(0.166)	(0.355)
Paid Social	0.110***	0.172***	2.768***	2.732***
	(0.001)	(0.003)	(0.372)	(0.746)
Organic Social	0.099***	0.162***	6.483***	10.853***
	(0.001)	(0.003)	(0.297)	(0.675)
Email, SMS, DM	0.115***	0.188***	6.361***	8.208***
	(0.001)	(0.001)	(0.138)	(0.295)
Organic Search	0.135***	0.213***	4.909***	6.816***
	(0.000)	(0.001)	(0.100)	(0.212)
Direct	0.110***	0.261***	4.787***	8.526***
	(0.000)	(0.001)	(0.109)	(0.252)
Other	0.062***	0.124***	7.390**	7.669
	(0.012)	(0.029)	(2.883)	(8.440)
Constant	0.013***	-0.104***	76.516***	69.322***
	(0.001)	(0.001)	(0.336)	(0.460)
Observations	1,097,876	1,097,876	221,605	221,605
R-squared	0.296	0.336	0.041	0.044

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Independent variables not shown: Number of visits and number of unique channels (models 1 and 2), interaction of number of unique channels and individual channels (model 4)

Figure 5-12 compares the weight of each channel with cummulative visit time discounted by recency. The channel of affiliate exhibits the highest weight on both conversion and revenue which is quite different from the observations in Section 5.2. One potential explanation is that visits from affiliate channels for converted users are in general more adjacent to the last one, making every time increase of the same amount result in higher increase in both conversion probability and revenue gained from converted users. The relative weight of other channels remains consistent with that observed in Section 5.2, that search engines outperform social media platforms in conversion while organic social media demonstrates outstanding effectiveness in revenue generation. This implication encourages online retailers to invest more effort on organic social marketing to build long-term brand awareness, impress more potential customers and foster long-term client relationship aside from boosting traffic and revenue in a short period from channels like paid search or paid social. Of course, both paid and organic search cannot be overlooked as they are very efficient in converting users into buyers and bringing a lot of traffic.

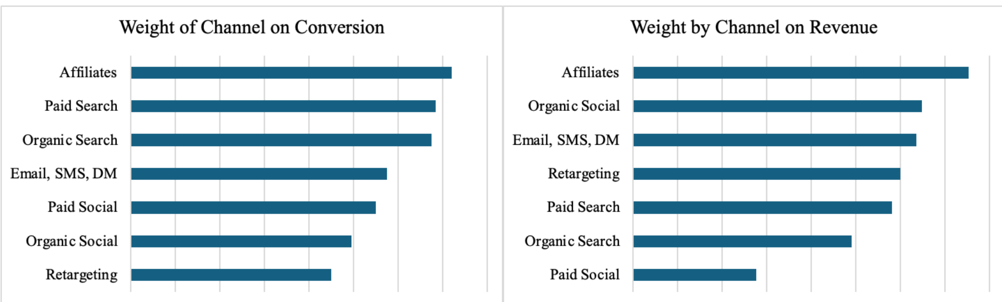

Figure 5-12 Comparison of Weight by Channel with Discount by Recency

Same as in Section 5.2, we further examine the impact of channel diversity on each individual channel of exposure by interacting the cummulative discoutned time per channel with number of unique channels in model 4. The result is depicted in Table 5-9 which is overall consistent with that of model 4 in Section 5.2 that channel diversity weakens the effectiveness of each individual channel on either conversion or revenue. This could also be explained by the diluting effect of a single channel with additional channels. When the user is exposed to multiple channels, he retrieves information in a variety of forms. On the one hand, his attention paid on one channel may be decreased as it becomes divided among multiple channels. Consequently, with limited attention allocated in each single channel, the user may not give equal consideration as in a single channel context, reducing the effectiveness of any single channel in driving conversions. On the other hand, the user may get similar or conflicting, redundant or less-persuasive information about product, promotion, coupon or user review from different channels. This may introduce confusion, hesitation or uncertainty to the user's purchase decision. As a result, the impact of each individual channel on the users' decision-making and purchse conversion gets split and diluted, and the incremental impact of each additional channel decreases. In a word, though additional channel may increase the overall conversion probability, the interacted effectiveness of each individual channel gets decreased .

Table 5-9 Regression Result of Model 4 with Time Discounted by Recency

	(4)	(4)
Dependent variable:	Converted	Revenue
Number of unique channels interact with time (discounted) on channel:		
Paid Search	-0.055***	-1.031***
	(0.000)	(0.108)
Retargeting	-0.014***	-0.761***
	(0.001)	(0.227)
Affiliates	-0.035***	-2.739***
	(0.001)	(0.167)
Paid Social	-0.035***	0.165
	(0.001)	(0.356)
Organic Social	-0.029***	-2.038***
	(0.001)	(0.293)
Email, SMS, DM	-0.037***	-0.952***
	(0.001)	(0.144)
Organic Search	-0.047***	-1.035***
	(0.000)	(0.118)
Direct	-0.099***	-2.425***
	(0.001)	(0.159)
Other	-0.036***	0.035
	(0.012)	(3.151)
Constant	-0.104***	69.322***
	(0.001)	(0.460)
Observations	1,097,876	221,605
R-squared	0.336	0.044

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Independent variables not shown: Number of visits and number of unique channels (models 1 and 2), effectiveness of individual channels (model 3)

In model 2 we find that channel diversity has a positive direct impact on conversion. However, the results of model 4 reveal that channel diversity poses a significantly negative effect on each channel's effectiveness. Then we wonder if there's a "perfect" channel diversity that yields the best overall conversion probability. That's, is there a reverse U-shape in the relationship between channel diversity and conversion? To address this question, we add a quadratic term (Z^2, where Z stands for channel diversity) into the regression model. What's more, in model 4 channel diversity is intereacted with browsing time of each channel and the results report a negative effect, which inspires us to examine channel diversity's moderating effect on the gross total browsing time regardless of channel. To address this problem, we interact channel diversity with total browsing time and add it into the model. We re-run the models 1, 2 and 3 with these 2 new independent variables included and the regression results are illustrated in Table 5-10 below. The findings are explained after the table.

Table 5-10 Regression Results with Channel Diversity Squared and Its Interactions with Total Time

	(1)	(2)	(3)
Dependent variable:		Converted	
Number of visits	-0.030***	-0.046***	-0.034***
	(0.000)	(0.000)	(0.000)
Cummulative browsing time	0.062***	0.096***	
	(0.000)	(0.000)	
Time * number of channels		-0.016***	
		(0.000)	
Number of unique channels		0.230***	0.315***
		(0.001)	(0.003)
Number of unique channels squared			-0.043***
			(0.001)
Cummulative time on channel:			
Paid Search			0.091***
			(0.000)
Retargeting			0.049***
			(0.001)
Affiliates			0.034***
			(0.000)
Paid Social			0.077***
			(0.001)
Organic Social			0.050***
			(0.001)
Email, SMS, DM			0.042***
			(0.000)
Organic Search			0.097***
			(0.000)
Direct			0.085***
			(0.000)
Other			0.074***
			(0.011)
Constant	0.190***	-0.067***	-0.127***
	(0.000)	(0.001)	(0.002)
Observations	1,097,876	1,097,876	1,097,876
R-squared	0.125	0.189	0.209

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The impact of Time * number of channels, or the interaction between total browsing time and channel diversity, is reported to be negative in Table 5-10. This indicates that, consistent with the effects when interacting channel diversity with browsing time per channel, the gross mediation effect of channel diversity on the total browsing time (regardless of the specific channel) is also significantly negative. These findings highlight the importance of considering the interaction between browsing time and channel diversity when examining the visits' effects on conversion.

The other newly-included quadratic term Z^2 (Channel diversity squared, measured as Number of unique channels squared), shows a significantly negative effect on conversion. Remember that previous model results have shown that channel diversity itself has a positive effect on conversion. These results suggest a curvilinear relationship between channel diversity and conversion. Specifically, as channel diversity increases, the user's conversion probability tend to increase initially. However, beyond a certain threshold, the increase of channel diversity shows a diminishing or negative effect on conversion. This indicates that there is an optimal level of channel diversity that maximizes the conversion probability. Although the absolute value of the optimal channel diversity is not determined in our current model, this presents an opportunity for future research to delve deeper into understanding and quantifying the precise level of channel diversity that leads to optimal conversion rates.

5.4 Conclusion

This paper investigates the cummulative visit effect on purchase conversion as well as on revenue. Factors like channel of exposure, time spent on website (browsing time) of each visit and cumulated at channel or at user level, recency of each visit compared to the last one of each user and channel diversity are considered. Their impacts are evaluated by 4 model specifications with progressive inclusion of independent variables. All of the 4 model specifications are firstly run against dependent variable of conversion as the main model, and then on revenue as robustness check.

As illustrated in sections 5.2 and 5.3, the regression results indicate that the cummulative visit time has a positive impact on purchase conversion which supports Hypothesis 1.1. In addition, recency, which is measured as the time difference of each visit compared to the last one at user level, plays an important role in measuring the visit's cumulative effect on final conversion which supports Hypothesis 1.2. We recommend discounting time of each visit by its recency in future research on conversion when calculating cummulative visit effect for either all or individual channel. The impact of channels of exposure is also examined and the results imply that different channels show different influencing power on visits' effect on conversion which supports Hypothesis 1.3. For example, search engines outperform other channels in conversion, while organic social shows its high effectiveness in increasing revenue in both cases when vist time is discounted or not. This encourages business practitioners to invest more on social media platforms in building long-term brand awareness and client relationship.

The results also support Hypothesis 2 that channel diversity itself has a positive effect on conversion, indicating that retrieving information from multiple sources may increase the user's conversion probability. However, when interacting channel diversity with browsing time of each channel of exposure, negative impacts are demonstrated on most channels on their effect on both conversion and revenue. This could potentially be explained by the diluting effect of each single channel by additional channels when the user is acquiring information from multiple channels. That's, though additional channel may increase the overall conversion probability, the interacted effectiveness of each individual channel gets decreased. What's more, by interacting channel diversity with total browsing time across all channels, the results show that the gross mediation effect of channel diversity on the total browsing time is still negative, consistent with that on browsing time of each channel.

The positive impact of channel diversity itself on conversion, together with the negative impact of channel diversity squared, demonstrates a curvilinear relationship between channel diversity and conversion. In other words, increasing channel diversity may increase conversion probability at first, but beyond a certain threshold, the increase of channel diversity will have a diminishing or negative effect on conversion.

Chapter 6 Discussion

6.1 Theoretical Contribution

This study has made several theoretical contributions to research on user conversion and channel effectiveness in the area of e-commerce and online marketing.

First, our research is among the first to introduce the concept of channel diversity, quantify it, and to investigate its direct impact on conversion and its indirect effect on each channel's effectiveness. Based on the comprehensive analysis, this research reports a curvilinear relationship between channel diversity and user conversion.

Second, it contributes to conversion and attribution research by proposing a new regression model, highlighting the importance of browsing time and recency with a mechanism to measure and compare the difference of each channel's effectiveness on conversion and revenue in a multi-channel environment. Each channel has its pros and cons with different effectiveness in capturing user attention, driving conversion or increasing revenue.

Third, it develops a method to discount visit effect by recency when measuring the cummulative effects of prior visits, offering a recency-weighted approach for multi-channel attribution. People tend to give more importance to events happened more recently than those happened in the past. In addition, the user's acquired knowledge of a product may decay over time. However, there has been no standard way to measure the decay over time. Our research contributes to this topic by providing a continuous discounting model with 1% hourly discount rate. Future research can refer to this method to check effects of prior visits.

Fourth, it provides empirical support in favor of the information overload theory in consumer decisions that channel diversity may diminish or dilute the impact of each individual channel. Users need retrieve information to make informed purchase decisions. When they are exposed to information from multiple sources, the additional information can be either supportive, informative or redundant and inconsistent. Our research also tests the implications of attention theory in online retail context that more time spent indicates greater attention allocated, translating into higher conversion probability.

Last not but least, our research examines the impact on revenue while most extant research focus on visit-to-purchase conversion. Retailers eventually pursue high revenue, so we call for more research on factors influencing revenue generated by converted users aside from simply conversion.

6.2 Managerial Implication

With the popularity and growth of online commerce and communication, more and more commercial activities happen online. Consumers retrieve plenty of information from various online sources within limited attention before making a purchase decision. To attract and convert them, online retailers spend a notable portion of their revenue on digital online marketing including paid ads on search engines, sponsored posts or videos on social media platforms, tailored content sent via email or sms, and promotions on a huge variety of web pages. However, retailers often lack a data-driven way to determine the effectivess of each channel, known as ROAS (return of advertising spend). In addition, in face of so many online marketing channels, retailers often face the challenging of whether it's worthy of investing on more channels, and how to better allocate the marketing budget based on each channel's effectiveness. Our research provides several useful managerial implications to online retailers.

First, this research provides online retailers with an actionable model to measure and compare the difference of each channel's effectiveness on conversion and revenue. Online retailers can run this model with their own data, and the result can be used in channel attribution so that that can allocate marketing budget across a variety of channels in a more informed and data-driven way to align budget with channel effectiveness. Online advertising agencies (such as paid search vendors, search engine optimization service providers, social media advertising operators) can also use the estimated effectiveness on conversion and revenue to price its ads on different channels for its customers.

Second, this research provides evidence on the effectiveness of some typical channels. Consistent with business common sense, search engine (either paid or organic) is the most effective in bringing traffic and converting visitors to buyers. That's, online retailers can resort to search engine advertising when they are launching a new product or pursing a traffic growth. On the other hand, organic social shows its unneglectable effectiveness in revenue generation. Online retailers can follow this implication to build brand awareness and foster high-value client relationship on social media platforms for revenue increase in the long run. Wht's more, in this digital era with various marketing channels, investing in more channels may not always be cost-effective. Instead, online retailers should be aware of the curvilinear relationship between channel diversity and conversion to find the optimal level of channel diversity.

Third, the research also reveals that most users won't pay more than 3 visits (in less than 7 days) before either leaving or purchasing, so it's vital to engage them in the first few visits. Online retailers can engage them with promotions like coupons in this golden period for a higher conversion, or wait for a longer time before engaging them again after this period.

Fourth, our research demonstrates the curvilinear relationship between channel diversity and conversion. As channel diversity increases, the user's conversion probability tend to increase at first. However, beyond a certain threshold, the increase of channel diversity shows a diminishing or negative effect on conversion. As a result, it may not be that cost-effective to for online retailers to invest in as more marketing channels as possible. Marketers should aim for the optimal level of channel diversity that maximizes the conversion probability.

6.3 Limitations and Future Research

Although this research has made several unique contributions to both theoretical research and managerial practice, there are some limitations which can be directions for future research. We hope that future research will address these limitations and make improvements in the following areas.

First, there's no product-level information in our data set. In general we can check if this user has converted or not and the revenue when converted, but the information about what products he has browsed and what are finally purchased is unknown. We hope such information, including product identifier, category, price and purchased volume etc., can be augumented in future research so that the impact at product level can be investigated.

Second, visits of a user after he has converted are excluded in this research. The reason is that this research focuses on conversion and revenue on conversion, while visits after conversion are about re-purchase or customer loyalty so they are not considered. Future study can extend our research by including visits after purchase to test and refine our model in the context of re-purchase and customer life-cycle value.

Third, the data set used in this research is sampled at a 3-month period in 2020, while it may not be long enough to capture all visits along the user's visit-to-purchase journey. Future research can base on data set in a longer period to capture more visits till purchase as well as re-purchase to test and refine our model in a larger data set.

Fourth, the research is conducted in the context of eyeglasses or sunglasses. People usually buy a pair of eyeglasses or sunglasses once or twice per year, and the average order value is around $80. The result may not apply to other industries where consumer behavior pattern, purchase frequency and product price may be different. We encourage researchers test and refine our model in other industries.

Last but not least, this research examines the conversion probability of users without any prior purchase. That's why visits happened after a user has made a purchase are excluded. However, a limitation arises from the inability to identify users who has made purchases before the dataset period of August 2020 to October 2020. For example, if a user has made a purchase in 2019, his visits in our dataset will not be excluded since his prior purchase record is unavailable. Future research can mitigate this issue by retrieving a comprehensive list of users who has made purchases over an extended period and then remove records of such users in sample dataset to avoid the noise of visits from re-purchasing.

Another important noteworthy point is about the optimal level of channel diversity. This research has examined the impact of several independent variables in this point, including channel diverisity itself, its quadratic term (number of unique channels squared), its interaction with total browsing time across all channels, and its interaction with browsing time per channel of exposure. The results has demonstrated a curvilinear relationship between channel diversity and conversion probability, suggesting that investing in more channels (an omni-channel approach) may not be that cost-effective in this digital marketing context. Instead, business practitioners should consider the balance between benefits of engaging users through additional channels and the diluting effects on each individual channel as well as the dimishing or negative effects on conversion. However, the absolute value of the optimal channel diversity is not determined in this research. We hope future research could furthur investigate this area to understand and quantify the precise level of channel diversity that leads to optimal conversion output. Research in different industries or product categories may report different level of optimal channel diversity, which will be practical and beneficial to business.

References

Armstrong, M., & Chen, Y. (2009). Inattentive consumers and product quality. Journal of the European Economic Association, 7(2-3), 411-422.

Arsath, M. A. (2018). Social Media Marketing: Advantages and Disadvantages. Shanlax International Journal of Management, 6(1), 152-158.

Bawden, D. & Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. Journal of Information Science, 35(2), 180-191.

Bawden, D., & Robinson, L. (2020). Information overload: An overview

Bucklin, R. E., & Sismeiro, C. (2003). A model of web site browsing behavior estimated on clickstream data. Journal of Marketing Research, 40(3), 249-267.

Chen, H. S., Tsai, B. K., & Hsieh, C. M. (2017). Determinants of consumers' purchasing intentions for the hydrogen-electric motorcycle. Sustainability, 9(8), 1447.

Dai, J., & Sheng, G. (2022). Advertising strategies and sustainable development: The effects of green advertising appeals and subjective busyness on green purchase intention. Business Strategy and the Environment, 31(7), 3421-3436.

Falkinger, J. (2007). Attention economies. Journal of Economic Theory, 133(1), 266-294.

Hauser, J. R., Urban, G. L., Liberali, G., & Braun, M. (2009). Website morphing. Marketing Science, 28(2), 202-223.

Ismail, A. R. (2017). The influence of perceived social media marketing activities on brand loyalty: The mediation effect of brand and value consciousness. Asia pacific journal of marketing and logistics, 29(1), 129-144.

Jacoby, J., Speller, D. E., & Kohn, C. A. (1974). Brand choice behavior as a function of information load. Journal of Marketing Research, 11(1), 63-69.

Jansen, B. J., & Schuster, S. (2011). Bidding on the buying funnel for sponsored search and keyword advertising. Journal of Electronic Commerce Research, 12(1), 1.

Ji, W., Wang, X., & Zhang, D. (2016). A probabilistic multi-touch attribution model for online advertising. In Proceedings of the 25th acm international on conference on information and knowledge management (pp. 1373-1382)

Kannan, P., Reinartz, W., & Verhoef, P. (2016). The path to purchase and attribution modeling: Introduction to special section. International Journal of Research in Marketing, 33(3), 449–456.

Kobayashi, M., & Takeda, K. (2000). Information retrieval on the web. ACM computing surveys (CSUR), 32(2), 144-173.

Kuan, H. H., Bock, G. W., & Vathanophas, V. (2008). Comparing the effects of website quality on customer initial purchase and continued purchase at e-commerce websites. Behaviour & Information Technology, 27(1), 3-16.

Kuhlthau, C. C. (2005). Information search process. Hong Kong, China, 7(2005), 226.

Kuhlthau, C. C., Heinström, J., & Todd, R. J. (2008). The ‘information search process’ revisited: Is the model still useful. Information research, 13(4), 13-4.

Li, H., & Kannan, P. K. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. Journal of marketing research, 51(1), 40-56.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of marketing, 80(6), 69-96.

McKinsey & Company. (2009). The consumer decision journey. https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/the-consumer-decision-journey

Moe, W. W. (2003). Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. Journal of Consumer Psychology, 13(1-2), 29-39.

Moe, & Fader, P. S. (2004a). Capturing evolving visit behavior in clickstream data. Journal of Interactive Marketing, 18(1), 5–19.

Moe, & Fader, P. S. (2004b). Dynamic Conversion Behavior at E-Commerce Sites. Management Science, 50(3), 326–335.

Molla, A., & Heeks, R. (2007). Exploring e-commerce benefits for businesses in a developing country. The Information Society, 23(2), 95-108.

Montgomery, Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling Online Browsing and Path Analysis Using Clickstream Data. Marketing Science (Providence, R.I.), 23(4), 579–595.

Nottorf, F. (2014). Multi-channel Attribution Modeling on User Journeys. Communications in Computer and Information Science, 456, 107–125.

Park, & Park, Y.-H. (2016). Investigating Purchase Conversion by Uncovering Online Visit Patterns. Marketing Science (Providence, R.I.), 35(6), 894–914.

Pieters, R., Warlop, L., & Wedel, M. (2002). Breaking through the clutter: Benefits of advertisement originality and familiarity for brand attention and memory. Management science, 48(6), 765-781.

Reiley, D., Li, S., & Lewis, R. A. (2012). Northern Exposure: A Field Experiment Measuring Externalities between Search Advertisements. Social Science Research Network. https://doi.org/10.2139/ssrn.2190189

Shao, X., & Li, L. (2011). Data-driven multi-touch attribution models. Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 258–264.

Statista. (2022, August). U.S. online shopper conversion rate. https://www.statista.com/statistics/439558/us-online-shopper-conversion-rate/

Statista. (2023, February). Online advertising spending worldwide. Statista. https://www.statista.com/statistics/237974/online-advertising-spending-worldwide/

Statista. (2022, June). Number of social media users worldwide from 2017 to 2027. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Statista. (2023a). Eyewear – Worldwide. https://www.statista.com/outlook/cmo/eyewear/worldwide

Statista. (2023, August). Online shopping - statistics & facts. Statista. https://www.statista.com/topics/871/online-shopping/

Sun, Q., & Spears, N. (2012). Frustration and Consumer Evaluation of Search Advertising and Search Engine Effectiveness: The Case of Hedonic versus Utilitarian Product. Journal of Electronic Commerce Research, 13(2), 122-134.

Skiera, B., Eckert, J., & Hinz, O. (2010). An analysis of the importance of the long tail in search engine marketing. Electronic Commerce Research and Applications, 9(6), 488-494.

Wan, Yun, Satya Menon, and Arkalgud Ramaprasad. 2009. "The Paradoxical Nature of Electronic Decision Aids on Comparison-Shopping: The Experiments and Analysis" Journal of Theoretical and Applied Electronic Commerce Research 4, no. 3: 80-96.

Zhao, X., & Stecke, K. E. (2010). Pre-orders for new to-be-released products considering consumer loss aversion. Production and Operations Management, 19(2), 198-215.