

# WeStore or AppStore: How Customers Shop Differently in Mobile Apps vs. Social Commerce

Kejia Hu

Owen Graduate School of Management, Vanderbilt University, Kejia.Hu@owen.vanderbilt.edu

Nil Karacaoglu

Fisher College of Business, The Ohio State University, karacaoglu.2@osu.edu

In the dynamic e-commerce environment, social commerce has emerged as a revolutionary force, transforming how consumers interact and transact online. This paper investigates the differences in customers' search and purchase patterns between a prominent online retailer's burgeoning social commerce channel, the WeChat mini-program, and its native mobile app. We analyze the customers' entire journey through a sequential search model that encapsulates decisions from channel selection to product search, search termination, and the final purchase. This study contributes to the search model literature by being the first to estimate both fixed and marginal search costs in a sequential search model in an omnichannel retail environment. We calculate fixed search costs, marginal search costs, and preferences for each channel, revealing differences in customers' behaviors across channels. Our analysis shows that customers' fixed search costs are higher, but marginal costs are lower on WeChat channel compared to the App channel. Also, customer characteristics like historical spending levels and search timing influence their search costs. From these insights, we suggest strategies tailored to each channel capitalizing on the differences in customers' search costs. The first strategy encourages search initiation by lowering fixed search costs through peer-to-peer link sharing in the WeChat channel. The second strategy aims to minimize marginal search costs using search-triggering coupons in the App channel. Implementing these strategies significantly boosts conversion rates and profits for the online retailer. This research is one of the first to explore the differences between traditional retail channels and emerging social commerce channels.

*Key words:* Social Commerce, Omnichannel Retailing, Service Operations, Sequential Search Model, Propensity Score Matching.

---

## 1. Introduction

The social commerce industry, valued at a staggering \$474.8 billion in 2020, continues to reshape the retail landscape and to transform the way customers discover and purchase products without ever leaving the social media platforms (GrandviewResearch 2021). This emerging trend has led major social media platforms to adapt and integrate social commerce functionalities to their platforms since 2018. For example, Instagram and Pinterest now enable customers to directly purchase products shown in posts, while Facebook and Twitter have added "Buy" buttons to their user interfaces (Tech 2019). Messaging-based social media platforms such as WhatsApp and WeChat have also become essential players in the social commerce arena. Facebook sees significant business potential in WhatsApp, as retailers can create storefronts using WhatsApp Business and interact with customers through messaging functionalities (Reuters 2019).

WeChat, a widely-used social messaging app with over one billion monthly users, introduced mini-programs in 2017, which have since become a significant and growing social commerce channel (Kharpal 2019). Operating within the WeChat ecosystem, these sub-applications have attracted more than 300 million daily users and facilitated transactions worth over \$250 billion in 2020 (Liao 2021). Social commerce mini-programs rank among the most popular applications in the WeChat ecosystem (Chu 2019). WeChat mini-programs enable users to access e-commerce sites without installing new apps or leaving WeChat. Within the social network app, users can also effortlessly share products with peers or chat groups via WeChat messages or the embedded “share” function in the mini-programs. Retailers engage with users on WeChat mini-programs differently than on native apps. While they cannot send notifications directly to customers, they can leverage social networks and peer-to-peer messaging to reach customers through content creators, influencers, and friends on WeChat. Although accessing WeChat mini-programs requires more effort than native apps (see E-companion Appendix 1), the mini-programs demand less processing power on the phone and thus offer faster loading speeds for shopping pages. Table 1 summarizes the differences between native apps and WeChat mini-programs (Chu 2019, QPSoftware 2021). Despite the differences in features, there are many similarities across these channels. For instance, numerous retailers—including the one in our study—offer the exact same assortment and prices in their WeChat mini-program channels and their native apps.

As social commerce channels, such as WeChat mini-programs, continue to flourish, they offer retailers innovative opportunities to tap into previously unexplored markets. However, because of the ease with which customers can access competing mini-programs within the WeChat ecosystem, competition among retailers is intense, making it crucial to understand the intricacies of customer behavior in this channel. Native mobile apps, which constitute the largest channel for the retailer in our dataset, also maintain their vital role in mobile e-commerce. As retail operations increasingly transition to mobile platforms, the resulting abundance of customer data presents valuable opportunities for retailers, as highlighted by Kumar and Venkatesan (2021). Consequently, it is vital for retailers to understand the nuances of customer behavior in each channel in order to refine their marketing and operational strategies. Our study investigates the differences in customers’ search and purchase patterns between a major online retailer’s mobile app and its WeChat social commerce channel. By examining these variations, we aim to provide crucial insights that enable retailers to tailor their strategies for native mobile apps and social commerce channels and achieve sustained success in the ever-evolving e-commerce landscape.

To obtain a comprehensive understanding, we not only examine customers’ final purchases, but also focus on their overall shopping journeys utilizing detailed customer-level clickstream data. This dataset delineates the channels customers visit, the products they view, the sequence in which they

view them, and their eventual purchases. Our counterfactual analyses examine channel-specific operational strategies that retailers can employ to leverage the differences in customer behavior across channels. Herein, we use the App channel to refer to the retailer's native app.

To characterize customers' search processes and evaluate potential operational and marketing strategies for increasing sales, we construct a structural model grounded in search theory literature. Specifically, we examine whether and how differences in customer behavior across channels are affected by search cost differences in these channels. By estimating a sequential search model, we investigate the differences in customers' product preferences, fixed search costs (incurred when initiating a search), and marginal search costs (incurred at each additional search step) between the WeChat and App channels. Our findings indicate that WeChat customers experience higher fixed search costs but lower marginal search costs compared to those in the App channel, leading to more extensive browsing and more frequent conversions. Our structural model includes details on customer characteristics and browsing times, demonstrating that search costs exhibit heterogeneity across users. Customers with higher historical spending levels incur higher fixed and marginal search costs. In Section 5, we discuss the factors contributing to the observed search cost differences in these two channels in more detail.

To corroborate the validity of our model, we assess within-sample goodness of fit and compute out-of-sample predictions, both of which substantiate our model's proficiency in estimating and forecasting customer behavior. As in many omnichannel settings, customers in our setting can self-select which channel they would like to use. We mitigate this issue by utilizing the propensity score matching technique to create customers and sessions that are similar across their observed characteristics across the WeChat and App channels. Additionally, using statistical tests, we confirm the sample similarity across channels after matching. We also perform an array of robustness analyses to eliminate potential alternative explanations.

Finally, our counterfactual analyses explore the practical implications of our findings. Based on the observed differences in customer search costs across the WeChat and App channels (higher fixed search costs in WeChat and higher marginal search costs in the App channel), we propose two channel-specific marketing coupon strategies. The first strategy promotes peer-to-peer link sharing by offering friend referral discounts to reduce customers' fixed search costs in the WeChat channel. The second strategy utilizes search-triggering coupons to encourage customers to continue their searches in the App. Both strategies can bolster conversion rates and—most importantly—significantly augment retailer profits if the discounts are optimally designed. According to our estimates, the retailer can generate approximately 0.6 Chinese Yuan (CNY) in additional profits per WeChat customer by implementing a 2 CNY friend referral coupon. Similarly, the retailer can yield 0.7 CNY in additional profits per App customer by offering a 1-2 CNY search-triggering

coupon sent after their second search step in a session. Both strategies are easily implemented and present no ethical concerns. Additionally, given the rarity of multi-homing in our setting, these interventions are unlikely to incentivize strategic customer behavior. Considering the retailer's substantial customer base on both WeChat and App channels, these strategies hold significant potential for increasing profits.

Our research makes significant contributions to both the academic literature and practical applications in the fields of omnichannel retail operations, customer choice models, and promotion design. Furthermore, it connects two academic fields—omnichannel retail operations and search models in marketing and economics—by introducing structural search models to operations management (OM) literature. Below, we elaborate on the key contributions of our paper to the academic literature.

First, our paper makes a significant contribution to the field of omnichannel retailing in OM by investigating the differences in customer search costs between two burgeoning digital channels: social commerce and native apps. While previous research has primarily focused on the differences between online and offline channels (Gallino and Moreno 2014) or mobile and desktop channels (Gallino et al. 2022b), we focus on WeChat and App channels, both accessible via smartphones. By examining separate channels accessed through the same device, we uncover novel insights about the differences in customer search behavior and cost across these channels. Social commerce channels, exemplified by WeChat mini-programs, present distinct challenges for retailers when initiating dialogues, while concurrently offering opportunities to boost profitability by capitalizing on the influence of social networks on customers' preferences. To the best of our knowledge, our study serves as the first to explore the differences between the burgeoning social commerce channel and the established app channel within the omnichannel retailing literature.

Second, compared to previous studies on customer choice modeling in OM and marketing, our research leverages an innovative sequential search model that thoroughly captures customers' shopping journeys in an e-commerce environment. Traditional literature on customer choice in OM has primarily relied on transaction data and employed discrete choice models for analyzing customer preferences, which often leads to biased estimates due to neglecting endogenous formation of choice sets during the search process (Koulayev 2014). In contrast, our structural search model accounts for choice set formation, providing more accurate estimates of search cost parameters (Honka 2014). Furthermore, we contribute to the literature on search models in marketing and economics as the first to estimate both fixed and marginal search costs in a sequential search model. This advancement is made possible by leveraging data on customers' search paths in an omnichannel setting. By interconnecting every customer decision—from channel choice to which products to search, search termination, and final purchase—we are able to quantify customers' search costs and preferences

**Table 1 Comparison of Native Mobile Apps and WeChat Mini-Programs**

Comparison Factors	Mobile Apps	WeChat Mini Programs
<b>Installation</b>	Downloaded and installed separately from app stores	Do not require separate installation; accessed within WeChat
<b>Access</b>	Quick tap on the app icon	Search and multiple clicks required
<b>User Attraction</b>	Typically attract users through push notifications	Typically attract users through social network sharing
<b>User Experience</b>	Complex design and frequent back-end operations that may lead to slow-downs, overheating, and battery drain	Lightweight and streamlined design that provides a fast loading speed and an efficient experience

in each channel, providing a comprehensive understanding of customer behavior in the channels under investigation.

Lastly, our research contributes to the promotion design literature in OM by suggesting channel-specific promotion strategies that capitalize on customers' search behavior and search cost differences across channels in an omnichannel environment. Previous studies on pricing and promotion strategies have primarily focused on a single channel of a retailer and employed transaction and demand patterns to develop promotion strategies (see Mišić and Perakis 2020). In contrast, by incorporating customers' search costs into the promotion design process, we are able to create tailored promotions for each channel that align with customers' search behavior. Our findings carry significant implications for practitioners aiming to optimize operations and marketing strategies across various channels. Specifically, we recommend two easily implementable channel-specific promotion strategies: friend-referral coupons for WeChat and search-triggering coupons for the App, which assist retailers in enhancing profitability. Equipped with this understanding, retailers can devise channel-specific marketing strategies that effectively boost their profitability.

## 2. Literature

This paper is closely related to the literature on omnichannel retail in operations management and the literature on customer search and choice.

### 2.1. Omnichannel Retail

As the retail and online service landscapes evolve, understanding the differences in customer behavior across various channels has become crucial. Recent studies have primarily explored online channels and examined how customer behavior differs between channels on different mediums, such as mobile (smartphones/tablets) and desktop channels. These papers suggest that factors such as screen size, internet speed, and device portability drive the differences in customer behavior across mobile and desktop channels (Gallino et al. 2022b).

Our research contributes to this growing body of literature on omnichannel retail (Ertekin and Agrawal 2021), mobile commerce (Zhang et al. 2023), and platforms (Rong et al. 2022, Gallino et al. 2022a). It focuses on two channels that coexist on the same medium (i.e., the mobile phone) and explores how social aspects and retailer-consumer interaction differences affect shopping behavior across these two channels. As the WeChat mini-programs and the retailers' native mobile Apps are both accessed via smartphones, the previously highlighted medium-specific characteristics that explain customer behavior differences across channels are kept fixed in our setting. This enables us to cleanly study customer behavior and search pattern differences from the perspective of the interactions between the retailer and the customers and between customers. Retailers can send promotional messages and notifications to lure shoppers to a traditional mobile app, but they cannot initiate conversations in WeChat unless the customer first opens the mini-program. However, there are other social nudging avenues in WeChat, as explored in our counterfactual analyses. For instance, a shopping session might be initiated by a message from a family member or friend, or by promotional marketing messages sent by social media influencers, who command significant WeChat-specific follower bases. The social aspect of the WeChat mini-programs is unique to emerging social commerce channels such as WeChat, WhatsApp, Facebook Marketplace, and Instagram. However, many other companies, such as the online grocer Weee, attempt to use the power of social commerce—namely, social marketing through peer-to-peer sharing—to increase their revenues.<sup>1</sup>

Our research also differs from previous work due to the precise information and richness of our dataset. First, we leverage clickstream data from both customers who made purchases and customers who did not purchase, which provided a comprehensive, step-by-step understanding of customers' shopping process. By contrast, previous literature has only explored shopping outcomes such as total sales, due to a lack of data on customers who did not finalize their purchases. Second, previous literature on customers' choice sets assumes that every available item in a store belongs to customers' choice sets (Boada-Collado and Martínez-de Albéniz 2020). By observing the exact items that customers search for throughout their journeys, we can relax this assumption and effectively identify the sets of products that the customers considered to be alternatives. Both aspects provided our research with a precise characterization of the customer behavior in our studied channels. In this way, we contribute to the rising literature on data-driven decision-making in operations management (Kong et al. 2022, Hu et al. 2019).

## 2.2. Customer Search and Choice

Our research contributes to the literature on customer choice and search patterns. Previous studies explored customer choice using static choice models and customer search strategies using simultaneous and sequential search models.

<sup>1</sup> <https://www.wsj.com/articles/online-grocer-weee-hires-first-cfo-as-it-looks-to-grow-its-business-11613048400>

The operations management literature has examined how retailers can leverage data on aggregate demand, staffing levels (Lee et al. 2021), and social media (Cui et al. 2018) to optimize their inventory and pricing strategies. Most of these studies analyze customers' final purchases using static choice models. In contrast, we leverage customers' search paths and consideration sets to advance our understanding of customer preferences, price elasticities, and choice processes.

The traditional choice models assume that customers are aware of all products in the marketplace prior to purchase. This assumption might be unwarranted in online marketplaces, where hundreds of items are sold by various sellers in a product category. Previous studies show that not considering the search process can bias customer preference estimates when customers' choice sets are endogenously formed. The magnitude of this bias can be as large as 30%, and it is not possible to predict the direction of this bias (Koulayev 2014). Hence, it is important to model and incorporate customers' search processes into the model to recover unbiased estimates of customer preferences (De los Santos et al. 2012, Honka 2014, Kim et al. 2010). Our paper contributes to the literature on customer choice in operations management by integrating customers' search processes into the decision-making framework.

Although examining customers' search process—rather than focusing solely on their final purchases—enables improvements in estimation accuracy, it is essential to correctly identify customers' search strategies to pinpoint their preference parameters (Honka and Chintagunta 2017). Previous studies have predominantly focused on two search strategies: simultaneous search and sequential search. In a simultaneous search, customers predetermine their search set and continue to search until they exhaust all items in the set. In contrast, in a sequential search, customers' search sets endogenously evolve during the search process.

In most studies, researchers have access to limited data on customer searches. For instance, they might have either browsing data or aggregate data on view count and purchase popularity ranking (Kim et al. 2010), but not purchase data (Koulayev 2014). Similarly, they might have access to the search set but not to the search order and revisit patterns (Honka and Chintagunta 2017). Hence, these studies cannot cleanly identify the underlying search strategy. By contrast, we observe which products customers searched, the order of these searches, and the revisit and purchase patterns for many customers. This rich dataset allows us to estimate the parameters governing customers' search strategies, rather than assuming them. We also recover customers' preference parameters.

Previous studies in the customer search literature have analyzed customer searches in offline channels, such as grocery stores (Seiler and Pinna 2017), and online channels, such as hotels (Koulayev 2014), electronic devices (Bronnenberg et al. 2016), and books (De los Santos et al. 2012). In most of these studies, customers search over one channel, and the fixed costs of starting a search are assumed to be negligible. Zhang et al. (2019) constitute an exception, having analyzed

customers' search patterns over both mobile and desktop channels. However, they do not have data on customers' search orders or revisit patterns; thus, they assume that customers search simultaneously. In contrast, we analyze customer search behavior across the different channels of the same retailer using sequential search model. As customers increasingly interact with retailers over multiple channels, it is crucial to understand how customer searches differ across these channels in order to design channel-specific promotions and customer-targeting strategies.

### 3. Data Description and Exploratory Analyses

In this section, we provide a comprehensive overview of our dataset and discuss the details of the propensity score matching algorithm employed to create comparable customers and sessions across WeChat and App channels. Additionally, we present analyses that shed light on whether customers engage in simultaneous or sequential search. Finally, before implementing the structural model, we use a model-free approach to explore how customer search costs differ for native mobile Apps and WeChat mini-programs.

#### 3.1. Data Description

We obtained the raw data from the 2020 MSOM Data-Driven Research Challenge for one of the biggest online retailers in China, JD. As noted by Shen et al. (2020), "JD.com is China's largest retailer with a net revenue of US\$67.2 billion in 2018 and over 320 million annual active customers." This dataset provides a detailed picture of customers' entire shopping journeys for one specific product category in March 2018.

We combined the following relevant datasets to study our research question: clicks, orders, users, and SKUs. The clicks dataset includes detailed clickstream data. In this dataset, user ID, the product the customer clicked, the timestamp of the click, and the channel the customer is in. In the orders dataset, we can observe detailed order information on customer ID, the purchased product, the timestamp of the purchase, and price details. Using these datasets, we were able to construct entire search and purchase paths for both customers who made purchases and customers who did not.

To create our final dataset, we followed a detailed data cleaning and restructuring process. First, we removed the clicks of unidentified users, the duplicated order records, and consecutive clicks by the same user on the same product (as these were either simply refreshing webpages or clicks on different sections of a product page). We provide a detailed explanation of the dataset in E-companion Appendix 2, and data cleaning steps and data restructuring process in E-companion Appendix 3.

The restructured dataset is organized at the user and search session level. We also created additional variables. Detailed definitions of these key variables and terms are presented in Table



**Table 2** Term Dictionary

Term	Definition
<b>Search Session</b>	A customer's continuous actions (i.e., clicks and order) on a channel where the time gaps between successive actions are no more than 60 minutes
<b>Search Step</b>	A visit to a product's information page during a customer's search session
<b>Session Length</b>	Number of search steps within a session
<b>Action Time</b>	A session's start time and date
<b>Fixed Search Cost</b>	A one-time upfront cost incurred in the first search step to initiate a search
<b>Marginal Search Cost</b>	A cost incurred at each additional search step after the first search step

2. A search session corresponds to a user's continuous search activity in a channel where the time between successive actions (i.e., clicks and order) is not more than 60 minutes. Using the timestamps of the clicks and orders, we can trace the entire shopping journey of each customer; namely, we can identify the channel the customer is in, the time of the search session and clicks, the products the customer browsed (i.e., clicked), the sequence of browsing, whether the customer purchase an item, and if they do, which one. We also enrich this user search session data with user and SKU characteristics. The user data denotes customers' historical spending levels in JD (*user-level*). The SKU dataset includes information on the values of two quality attributes (*Attribute 1* and *Attribute 2*), for which higher values indicate better quality. We can identify the price of SKUs from the order dataset. A sample user-search session is presented below in Table 3.

Among four shopping channels of JD (App, WeChat, Desktop Browser, and Mobile Browser), the App and WeChat channels are the most important ones, accounting for 81.2% and 12.4% of search sessions, respectively. Moreover, in our analyses we focus on sessions in which customers exclusively use one channel throughout the session to obtain a clean identification for channel-specific search costs and benefits. Focusing on sessions conducted in one channel is innocuous as only 0.5% of customers switch channels within a search session. In addition, only 0.25% of customers purchase more than one unique product from this product category during a session, and thus, we assume unit demand and focus on sessions in which customers purchased no more than one unique product from this product category during the session. We also removed clicks after a purchase, as these clicks might not correspond to a new search activity. As we use customers' historical spending levels in JD (*user-level*), the average price of browsed items within a session, the average Attribute 1 of browsed items within a session, the average Attribute 2 of browsed items within a session as matching variables in our propensity score matching procedure, we removed sessions for which any of this information is not available. Our final dataset is composed of 3,354,056 clicks, 871,763 search sessions, and 247,842 orders pertaining to 328,618 customers.

We observe that customers conduct deeper searches in the WeChat when we compare the number of clicks per session and number of unique products browsed per session. In this channel, customers click more (4.35 on WeChat versus 3.46 in the App, on average  $t\text{-stat}=41.81$ ,  $p\text{-val}<0.0001$ ) and

**Table 3 Illustration of One Session**

User_ID	Session_ID	SKU_ID	User_Level	Channel	Action	Time	Price*	Attribute 1	Attribute 2
000c6e472d	472d2	f06e13a877	1	WeChat	click	2018-03-09 19:45:12	15.9	2	40
000c6e472d	472d2	c4ff8911d9	1	WeChat	click	2018-03-09 19:45:51	29.9	2	60
000c6e472d	472d2	b616c3b114	1	WeChat	click	2018-03-09 19:46:06	32	2	60
000c6e472d	472d2	b616c3b114	1	WeChat	order	2018-03-09 19:47:04	32	2	60

\*Note: Price is calculated as (Original Price-Direct Discount)

explore more products during a search session (3.11 on WeChat versus 2.64 in the App, on average  $t\text{-stat}=37.64$ ,  $p\text{-val}<0.0001$ ). Customers are also more likely to convert on WeChat than the App, 41% versus 26.8% ( $t\text{-stat}=85.93$ ,  $p\text{-val}<0.0001$ ). We present summary statistics for the relevant metrics before and after propensity score matching in Table 4.

The *user-level* variable captures customers' historical spending levels in JD with higher user-level values indicating greater historical spending. The timing of the sessions is captured by the *action time* variable, which denotes the time and date of a search session. We incorporate the *user-level* and *action time* variables into our matching procedure. Our structural model also incorporates these variables and explores how customers' search costs vary according to these variables.

### 3.2. Propensity Score Matching

One challenge in our setting is that customers can self-select whether they want to use the WeChat or the App channel. An ideal experiment would randomize customers into either only the WeChat channel or only the App channel. Unfortunately, restricting customers' access to a particular channel is infeasible, if not impossible. Customer self-selection of a channel is a common feature in studies that examine retailers' different channels. Many papers use matching techniques to address and alleviate the endogeneity of channel choice (Manchanda et al. 2015, Xu et al. 2017). In line with the previous literature, we utilize propensity score matching (Rosenbaum and Rubin 1983) to create comparable customer groups and sessions that are similarly distributed on the observed metrics in the WeChat and App channels.

The application and goal of the propensity score matching in our paper slightly differ from settings in which the propensity score matching is used in the context of a difference-in-differences (DiD) analysis to examine the effect of a treatment on various metrics of interest. In such settings, treatment is often endogenous, and treatment assignment might depend on the characteristics of the treated and control units prior to the treatment. Hence, when applying propensity score matching in these settings, researchers aim to create treatment and control groups that are similarly distributed across observable metrics before treatment to be able to credibly attribute post-treatment differences to the effect of the treatment itself. Instead, in our setting, we explore how customers' search patterns and search costs vary across the WeChat and App channels by modeling customer behavior and estimating the parameters using structural estimation. Hence, our goal with the application of propensity score matching is to create comparable customers and sessions across the

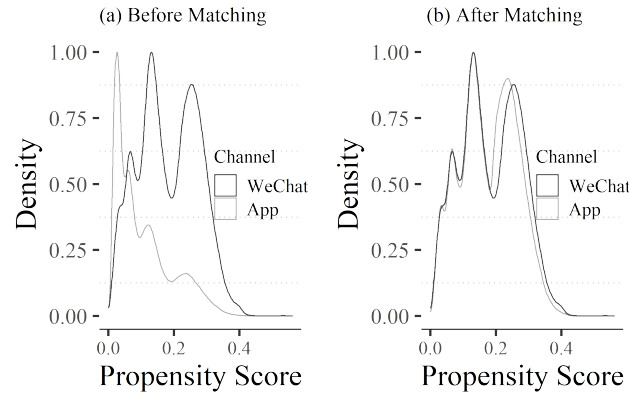
WeChat and App channels. Using propensity score matching allows us to ensure that customers and sessions in the WeChat and App channels are similarly distributed in observable metrics so that the differences in behavior can be more credibly attributed to channel effects. Similar applications of matching where matching is used to create sample similarity between two groups outside of natural experiment settings can also be found in other papers such as Brynjolfsson et al. (2011), Donovan (2021).

Our structural model leverages a customer's search steps within a session for identification. Hence, in the matching procedure, we matched sessions in the WeChat channel to sessions with similar characteristics in the App channel. The idea behind this matching procedure is that each session in the WeChat channel is matched with sessions with similar characteristics in the App channel allowing us a cleaner comparison of search cost differences across the two channels.

In line with Rosenbaum and Rubin (1983), the sessions in the WeChat channel are matched with sessions in the App channel that are closest to them in terms of propensity scores. To calculate the propensity scores for each session, we used the following covariates: session action date (i.e., the day of the month of a session's start time), session action time (i.e., the hour of the day of a session's start time), the average price of the browsed items within a session, the average Attribute 1 of the browsed items within a session, the average Attribute 2 of the browsed items within a session, and the user-level of the customer conducting the session.

The rationale for using these variables is as follows. The customer's user-level or the time of his session might affect channel choice, and thus we control for these variables in matching. Similarly, controlling for the prices and attributes of the browsed items within a session—as well as the time of the session—allows us to match sessions in which customers browsed items with similar price levels and characteristics across the two channels during similar time periods. As many more sessions are conducted in the App channel, we matched each WeChat session to the nearest three neighboring App sessions without replacement to generate a more representative sample of our original dataset while generating comparable session groups across the two channels. This matching method also allowed us to keep all the sessions on the WeChat channel. We also conducted extensive analyses to ensure that our results are consistent under alternative sets of covariates and more stringent matching methods. Our findings are robust to different matching techniques.

Table 4 shows the mean values of covariates before and after matching in the WeChat and App channels. As can be observed, matching significantly improved the sample similarity for the observed metrics between the two channels. To further evaluate the matching performance, we compared the propensity score distribution between WeChat and App channels before and after matching. As shown in Figure 1, the App sessions' propensity score distribution is significantly closer to that of the WeChat sessions after matching. Following Haviland et al. (2007), we also

**Figure 1** Propensity Score Distributions Before and After Matching**Table 4** Mean Statistics Before and After the Matching

Variable	App Pre-Matching	WeChat Pre-Matching	App Post-Matching	WeChat Post-Matching
Action Time Hour	14.17	14.53	14.47	14.53
Action Time Date	15.53	16.51	16.29	16.51
User-Level	2.50	1.66	1.66	1.66
Price	115.46	90.13	92.44	90.13
Attribute 1	2.98	2.84	2.88	2.84
Attribute 2	83.09	78.66	79.61	78.66

**Table 5** Standard Bias Before and After the Matching

Variable	Std. Bias Pre-Matching	Std. Bias Post-Matching	Percentage Improvement
Action Time Hour	0.07	0.01	83.21
Action Time Date	0.11	0.02	77.47
User-Level	0.87	0.01	99.28
Price	0.39	0.04	89.93
Attribute 1	0.24	0.07	72.54
Attribute 2	0.24	0.05	78.72

evaluated standardized bias before and after matching.<sup>2</sup> As presented in Table 5, standardized bias between the sessions in WeChat and App channels significantly decreases after matching, with a 72.54% to 99.28% reduction in standardized bias across all covariates.

In summary, the improvements in covariate balance, propensity score distributions, and reduction in standardized bias indicate that the matching procedure creates comparable sessions and user groups across the WeChat and App channels. We used the matched dataset for all analyses presented in the remainder of the paper.

### 3.3. Examination of Search Types: Sequential or Simultaneous

In the search literature, it is assumed that customers might follow two distinct search strategies: simultaneous or sequential (Honka and Chintagunta 2017). In a simultaneous search, customers determine the set of products they will search before initiating the search session and make purchase decisions after viewing all products in their pre-determined search set. As customers make purchase

<sup>2</sup> The details of standardized bias calculation are provided in E-companion Appendix 4.

decisions after exhausting their search set, the search order does not matter in a simultaneous search. One major issue with this search strategy is that it assumes customers continue to search even if the outcomes of the previous search steps are satisfactory (Honka et al. 2019). By contrast, in a sequential search, the set of products customers search depends on the outcomes of previous search steps. After each search step, customers decide whether to continue or to stop by comparing the expected benefits of an additional search step against the cost of acquiring information. Hence, the search sequence is critical for determining the next search steps or purchases.

Understanding whether customers use a simultaneous or sequential search strategy in our setting is essential for ensuring that our structural model is built on correct assumptions. We use an analysis similar to Bronnenberg et al. (2016) to identify customers' search strategies.

In this analysis, we examine whether customers' search patterns exhibit dependence on the previous search steps. In a simultaneous search, the search set is predetermined and the search order within the session is expected to be random. Hence, the customer's moves among products within that session should not depend on the search path: in other words, the product attributes of past searches should not affect the product attributes of the current search. On the other hand, under a sequential search, we expect attributes to have carryover effects.

To test the dependence of search patterns on the previous search steps within a session, we estimate the following dynamic panel regression for Price, Attribute 1, and Attribute 2:  $x_{sjd} = \alpha_{sj} + \beta_j x_{sj(d-1)} + \epsilon_{sjd}$  where  $\alpha_{sj}$  captures session fixed effects for attribute  $j$ ,  $x_{sjd}$  denotes the level of attribute  $j$  in the search step  $d$  within the session  $s$ , and  $\epsilon_{sjd}$  is the error term. For instance, if attribute  $j$  is price,  $x_{s,price,3}$  denotes the price level in search step 3 in session  $s$ . Due to the correlation between the fixed effects and the parameters of interest ( $\beta_j$ ), we use the Arellano-Bond GMM estimator with finite sample correction (Windmeijer 2005). We control for session-specific and time-invariant fixed effects. In this analysis, we consider search steps up to the eleventh search step and remove outlier search steps. Our results are robust to including all search steps.

We run this analysis for each attribute (Price, Attribute 1, and Attribute 2) separately. The results show a strong dependence for all attribute levels (p-vals  $\leq 0.01$ ) within a search session (see Table 6). This result suggests that customers search sequentially. We confirm the robustness of our findings by conducting a series of checks.<sup>3</sup> Based on the results of this test, we conclude that customers in our dataset search sequentially, and we construct our structural model accordingly.

<sup>3</sup> We also analyzed the App and WeChat channels separately, as well as analyzing only sessions that ended with a purchase. All the results are robust.

**Table 6** Test – Whether Future Search Steps Depend on Previous Steps

	Price	Attribute 1	Attribute 2
First Lag	-0.054*** (0.00356)	-0.184*** (0.00359)	-0.154*** (0.00358)
$\chi^2$	225.46	2633.35	1863.20

Standard errors are in parentheses. P-value: \*p<0.1;

\*\*p<0.05; \*\*\*p<0.01.

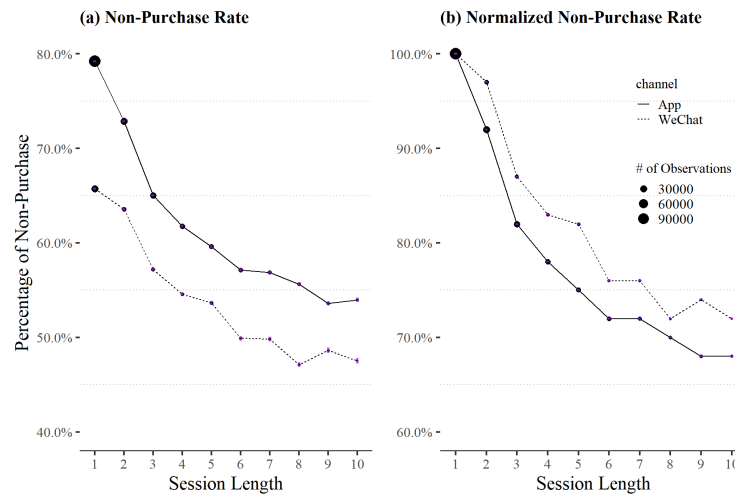
Remark: Both price and attributes show significant dependence along the search session.

### 3.4. Model-Free Evidence for Fixed and Marginal Search Costs Across the Two Channels

In this section, we provide exploratory evidence of the search cost differences between the App and WeChat channels. We want to note that these analyses provide suggestive evidence and we quantify the exact magnitude of these search costs using our structural model in Section 4. Previously, Zhang et al. (2019) examined channel that existed in different mediums, namely the mobile and desktop channels. They found that fixed search costs are lower, but marginal search costs are higher for mobile devices because mobile devices are portable. It is not a priori clear whether and how the fixed and marginal search costs vary between the App and WeChat channels that coexist on a smartphone.

During the search process, the customers might be influenced by their fixed search cost, marginal search cost, overall expected valuation, and the value of the outside option when they make their search decisions. Customers incur a fixed search cost when they initiate a search and only incur marginal costs if they continue to search for more products within a session. Hence, to explore fixed search cost differences across the App and WeChat channels, we restrict our attention to customers who browsed only one product. The behavior of these customers is mostly influenced by the upfront fixed cost of initiating the search but not by the marginal costs incurred from additional search steps beyond the first. Moreover, keeping a customer's expected outside utility and overall expected valuation constant, he might be more likely to initiate a search when his fixed search cost is lower. Thus, when a customer's fixed search cost is lower, he might be more likely to start the search and then leave it without buying anything. Customers with a low fixed search cost are more likely to initiate casual browsing (i.e., browsing without an intent to purchase). In our data, we find that customers engage in such casual browsing more often in App than WeChat. As shown in Figure 2(a), non-purchase rates among customers who browsed only one product in a search session are much higher in the App channel. The higher proportion of casual browsers in the App channel suggests that customers have lower fixed costs to initiate a search in the App channel compared to the WeChat channel. In addition, the percentage of sessions where customers browse only one item might be higher in a channel if customers' fixed search costs are lower in that

**Figure 2 Non-Purchase Rate**



channel. In line with the analysis presented in Figure 2(a), we also observe that the percentage of sessions where users only browsed one item is higher in the App channel (49%) compared to the WeChat (45.1%) channel. These two model-free analyses suggest that customers' fixed search costs are higher in the WeChat channel compared to the App channel.

Second, to test the marginal cost differences between the App and WeChat channels, we calculate non-purchase probabilities in each channel for each session length (measured by the number of search steps in a session). For each channel, we normalize the non-purchase probability of the  $k$ -step session length with respect to its non-purchase probability at the 1-step session length (i.e., the non-purchase probability of customers who conducted 1-step sessions.). Thus, both channels' normalized non-purchase probabilities start at 100%, as shown in Figure 2(b). By normalizing the 1-step session length of these two channels at 100%, we mute the influence of the fixed costs. This allows us to solely compare the two channels' marginal costs. Customers with high marginal search costs tend to end their searches earlier. They may either abandon the search altogether (in which case, the number of observations quickly shrinks over the search length) but they might also purchase at an earlier stage (in which case, the non-purchase rate would quickly decrease as the search length increases). Figure 2(b) shows that the App channel's observations shrink more notably (as indicated by the size of the points), and the App channel's normalized non-purchase rates decrease faster than that of the WeChat channel. Both pieces of evidence indicate that marginal search costs are higher in the App channel.

Our explanatory analyses suggest that fixed search costs are higher in the WeChat channel whereas marginal search costs are higher in the App channel and in the next section we build our structural model to quantify the search costs in both channels.

## 4. Structural Model

In this section, we model customers' search and purchase decisions for the two mobile channels under consideration. Based on the evidence provided in Section 3.3, we assume customers follow a sequential search model; namely, after each search step, a customer determines whether to continue searching by weighing up the expected marginal benefits and marginal costs of an additional search step. When an additional search step is deemed no longer beneficial, the customer either makes a purchase from the previously searched options or selects the outside option (i.e., no purchase from the e-commerce platform).

We begin by outlining the customers' multi-stage decision process, including the channel to use, the products to search, and the purchase decision (whether and what to purchase). Particularly, we explore how the fixed and marginal search-costs drive decisions at each step of the shopping journey. Finally, we explain the identification strategy that we devise to estimate the parameters of our sequential search model.

### 4.1. Customer Multi-Stage Decisions

**The Channel of Search** Before starting a search session, customers first decide which channel to use, App or WeChat, and then incur a fixed search cost when they begin to search. This fixed search cost is a one-time upfront cost and does not depend on the number of products that the customer browses during the search session. Based on transaction cost economics (Chintagunta et al. 2012), such fixed costs include the opportunity costs of committing the customer's attention to shopping on JD rather than engaging in another activity, the psychic costs in exerting effort and enduring the inconvenience of completing the necessary steps to access JD's e-commerce App or the WeChat mini-program.

These fixed search costs vary between individuals with different user-levels (i.e., customers' historical spending levels in JD), the different action times (i.e., session's start time and date), and different channels. We benchmark customer  $i$ 's fixed search cost for the App channel at 0 (i.e.  $FC_{it,App} = 0$ ) where  $t$  denotes the action time. We specify his fixed search cost for the WeChat channel as follows:

$$FC_{it,WeChat} = f + Z_{it}B_{itc} + e_{it}^f \quad (1)$$

where  $f$  is a constant term for all customers.  $Z_{it}$  denotes customer  $i$ 's characteristics, such as his user-level and his action time  $t$ , with the  $B_{itc}$  parameters capturing their influences on the fixed search cost.

In line with the transaction cost economics, we anticipate that the opportunity cost component of the fixed costs depends on the time of the shopping activity (i.e., action time).  $Z_{it}$  include a set



of shopping time variables, namely the time of day and the day of week, to capture these effects. Moreover, the fixed costs also capture the psychic cost of accessing the channel.

The error term  $e_{it}^f$  further captures the unobservables in the fixed costs. Following Chintagunta et al. (2012), we use a customer-time-specific error term that follows a standard normal distribution to reflect the unobservable fixed costs. We denote the set of parameters associated with fixed costs as  $\theta_{itc}^f = \{f, B_{itc}\}$ .

The customer selects the search channel  $c_i$  that offers him the maximum expected utility for the search session by solving  $\max_c E(U_{it}|c) - FC_{itc}$  where  $c \in \{\text{WeChat}, \text{App}\}$  and  $E(U_{it}|c)$  is the expected utility of a search session in channel  $c$  for customer  $i$  at time  $t$ .

**Products to Search and Purchase Decision** We assume that customers follow an optimal sequential search strategy throughout a search session. Specifically, they continue to search as long as the marginal benefit of an additional search step outweighs the marginal search cost. The marginal costs capture the search-associated costs in each additional search step including the efforts to scroll down to locate the next product to search, the effort to evaluate the newly searched product, the psychic cost (such as fatigue and frustration) associated with one more search, and the opportunity cost of time in searching one more product.

As with the fixed search costs, we model the variation in marginal search costs across individuals, action times, and channels. The marginal search cost of customer  $i$  at action time  $t$  on channel  $c$  is

$$MC_{itc} = \mu_c + Z_{it}\Lambda_{itc} + e_{itc}^m \quad (2)$$

where  $\mu_c$  is a constant term for all customers using the search channel  $c$ ,  $Z_{it}$  denotes customer characteristics, such as user-level, and action times and the  $\Lambda_{itc}$  parameter reflects their influences on the marginal search costs. Based on transaction cost economics, we anticipate that opportunity costs and efforts depend on the customers' characteristics such as their experiences of online shopping or historical average spending levels (i.e., user-level). Moreover, we anticipate that, even for the same individual, the opportunity cost of search could vary by the time of the day and the day of the week. Hence, we also include a set of time-related variables when characterizing the marginal costs. Like Chintagunta et al. (2012), we use a customer-time-and-channel-specific error term,  $e_{itc}^m$ , that follows a standard normal distribution to capture the unobserved part in the marginal costs. We denote the set of parameters associated with marginal costs as  $\theta_{itc}^m = \{\mu_{\text{WeChat}}, \mu_{\text{App}}, \Lambda_{itc}\}$ .

Customer  $i$ 's utility for product  $j$  on channel  $c$  is defined as follows:

$$u_{ic,j} = X_j\Theta_{ic} + \epsilon_{ic,j} \quad (3)$$

where  $X_j$  captures product  $j$ 's characteristics such as the product's price and quality level. The  $\Theta_{ic}$  parameter thus reflects customer  $i$ 's preferences on channel  $c$ , and  $\epsilon_{ic,j}$  refers to customers'

idiosyncratic preferences or special tastes for the product  $j$  on channel  $c$ .  $\epsilon_{ic,j}$  is assumed to be i.i.d. across customers, products, and channels.

In our setting, we assume that product characteristics ( $X_j$ ) are known by the customer prior to the search step, as the products are displayed directly in the search results page in a list or grid format. When the customer clicks into a product's detailed product page, the unknown  $\epsilon_{ic,j}$  and consequently the realized utility  $u_{ic,j}$  are revealed to customers. We denote the set of parameters associated with customers' preferences as  $\theta_{ic}^s = \{\Theta_{ic}\}$ .

We now use an example to illustrate how a customer chooses whether and which product to search based on the associated marginal benefit and marginal cost of taking an additional search step. Consider a customer who already conducted ten search steps, and the highest realized utility from his search set in these ten search steps is equal to  $u^*$ . To decide whether to take the eleventh search step, he evaluates his marginal search benefits for the unsearched products. For product  $l$  in his unsearched product set, the expected marginal benefit of the search step is defined as follows:

$$B_l(u^*) = \int_{u^*}^{\infty} (u_l - u^*) f_l(u_l) du_l \quad (4)$$

where  $f_l(u_l)$  is the probability density function of  $u_l$ , a random variable capturing the utility of the unsearched product  $l$ .

Suppose that a product in the unsearched set provides an expected marginal benefit that is higher than the marginal search cost. In that case, the customer continues his eleventh search step by examining the product with the highest expected marginal benefit. If the newly searched product offers a higher utility, the highest realized utility  $u^*$  is updated.

When no product in the unsearched product set provides a marginal benefit exceeding the marginal search cost, the customer stops the search process. At this point, the entire sequential search path,  $SS_{itc}$ , for customer  $i$ 's session with action time  $t$  on channel  $c$  is completed. Customers then make a purchase decision by comparing the product generating the maximum realized utility from the search set to the outside option (i.e., no purchase or purchase from other sources). Specifically,

$$m = \arg \max_{j \in S \cup S_0} u_{i,j} \quad (5)$$

where  $S$  is the search set covering all products searched in the session, and the utility associated with the outside option  $S_0$  is distributed as  $u_{i,0} \sim N(\mu_0, \sigma^2)$ .

In his seminal paper on customer search behavior, Weitzman (1979) shows that if the stochastic term of the utility is uncorrelated across products, the customers' search process can be viewed as a process based on products' reservation utilities. Product  $l$ 's reservation utility,  $r_l$ , indicates the hypothetical level of utility that makes the customer indifferent between searching or not

searching, i.e.,  $MC = B_l(r_l)$ . Weitzman (1979) proves that the optimal sequential search process is equivalent to selecting the product with the highest  $r_l$  from the unsearched set at each search step and terminating the search process if no product's reservation utility exceeds the highest realized utility from the searched set.

Similar to Weitzman (1979), in our setting we can equate the marginal benefit of a product with the marginal search cost to obtain the product's reservation utility. These reservation utilities characterize the customers' search and purchase decisions. For customer  $i$  searching in channel  $c$  at time  $t$ , the reservation utility for product  $j$  is  $r_{itc,j}$ . By equating the marginal search cost with the marginal benefit of this search, we obtain an equation that the reservation utility  $r_{itc,j}$  should satisfy:

$$MC_{itc} = \int_{r_{itc,j}}^{\infty} (u_{itc,j} - r_{itc,j}) f(u_{itc,j}) du_{itc,j} \quad (6)$$

After  $k$  search steps, customer  $i$ 's search set is  $S_{itc,k}$  for the current search session, starting at action time  $t$  on channel  $c$ . This customer will continue to search if the highest reservation utility of the unsearched products set exceeds the highest realized utility, i.e., if:

$$\max_{j' \notin S_{itc,k}} r_{itc,j'} \geq \max_{j \in S_{itc,k}} u_{itc,j}. \quad (7)$$

Then the customer selects the product  $j$  which has the highest reservation utility for further exploration in his search process, i.e.  $r_{itc,j} > r_{itc,j'}, \forall j' \notin S_{itc,k}$ , and the actual utility  $u_{itc,j}$  is revealed.

The search comes to an end when the reservation utilities of all unsearched products are smaller than the highest realized utility. Customer  $i$  searching in channel  $c$  at time  $t$  either purchases the product with the highest realized utility or chooses the outside option:

$$m = \arg \max_{j \in S \cup S_0} u_{itc,j} \quad (8)$$

Linking with the channel choice, the channel-specific expected utility for the corresponding search session can be written as follows:

$$E(U_{it}|c) = E\left(\max_{j \in S_{itc} \cup S_0} u_{itc,j} - K_{itc}(\theta_{itc}^m, \theta_{itc}^s) MC_{itc}\right) \quad (9)$$

where  $K_{itc}(\theta_{itc}^m, \theta_{itc}^s)$  is the number of sequential search steps (i.e., the session length).  $K_{itc}(\theta_{itc}^m, \theta_{itc}^s)$  is a random variable that depends on customer  $i$ 's marginal search cost parameters,  $\theta_{itc}^m$ , and customer utility preference parameters,  $\theta_{itc}^s$ , on channel  $c$  at time  $t$ . The parameters of the marginal search costs determine the marginal cost  $MC_{itc}$ , and the parameters in the utility preferences determine the reservation utilities. As customers incur the marginal search cost at each additional search step, we multiply session length,  $K_{itc}(\theta_{itc}^m, \theta_{itc}^s)$ , by marginal search cost,  $MC_{itc}$ ,

to determine the total marginal search cost for the session. The larger the marginal cost and the smaller the reservation utility, the shorter the session (i.e., a smaller  $K_{itc}$ ).

**Likelihood** We now define the likelihood of observing the customers' decisions. For customer  $i$  searching in channel  $c$  at time  $t$ , his channel choice  $C_{itc}$ , sequential search path  $SS_{itc}$ , and purchased product  $D_{itc}$  are random variables with the corresponding realized observations as  $c^*$ ,  $ss^*$ , and  $j^*$ , respectively. The corresponding likelihood of observing such a combination of channel choice, search sequence, and purchase decision is then as follows:

$$Pr(C_{itc} = c^*, SS_{itc} = ss^*, D_{itc} = j^* | \Phi_{itc}, X_{it}, Z_{it}) \quad (10)$$

where  $\Phi_{itc} = \{\theta_{itc}^f, \theta_{itc}^m, \theta_{itc}^s\}$  contains the parameters associated with fixed search costs, marginal search costs, and customer utility preferences.  $Z_{it}$  is a list of customer characteristics and action time attributes, and  $X_{it}$  denotes the characteristics of the products available to search.

To further express the likelihood function using reservation utilities and realized utilities, we use the following:

$$\begin{aligned} & Pr(C_{itc} = c^*, SS_{itc} = ss^*, D_{itc} = j^* | \Phi_{itc}, X_{it}, Z_{it}) \\ &= Pr(E(U_{it}|c^*) - FC_{itc^*} \geq E(U_{it}|c') - FC_{itc'}, \forall c' \in \{\text{WeChat, App}\}, \\ & \quad r_{itc, ss_{itc,k}^*} \geq \max_{j \in S_{itc,k}^*} u_{itc,j}, r_{itc, ss_{itc,k}^*} > r_{itc,j'}, \forall j' \notin S_{itc,k}^*, \forall k \in 1, \dots, K_{itc}, \\ & \quad u_{itc,j^*} > u_{itc,j}, \forall j \in S_{itc, K_{itc}} \cup S_0) \end{aligned} \quad (11)$$

Since we observe limited instances in which one customer conducts multiple search sessions or conducts search sessions in multiple channels, we simplify the overall likelihood function for  $N$  customers as follows:

$$L = \prod_{i=1}^N Pr(c_i^*, ss_i^*, j_i^* | \Phi_{itc}, X_{it}, Z_{it}) \quad (12)$$

#### 4.2. Model Estimation and Identification Strategies

Our model captures the intricacies of the customer's decision-making process during his shopping journey, but there are several estimation challenges. This section highlights these estimation challenges and explains how we overcome them to identify the parameters governing the structural search model.

**Estimation Challenges** First, in our setting, unlike in traditional discrete choice models, the probabilities of the observed choices have closed-form solutions, subject to appropriate distribution assumptions for the random variables. This leads to computational convenience in maximizing the likelihood function. On the other hand, for a sequential search algorithm, there is no closed-form solution for the search set probability or the conditional purchase probability.

Although we can theoretically calculate a customer's search probability by writing the reservation utilities and expected realized utilities for all products available for a customer to search, the computational complexity quickly increases with the number of search steps and the size of the searchable product set. For instance,  $P_{30}^{10} = 1.09E + 14$  potential search sequence paths exist for a customer who searches 10 products from a set of 30.

In our model, each customer's choice depends on his previous decisions. This interdependence of channel choice, search process, and purchase decision also increases the computational complexity. For example, customers' search process is governed by the marginal cost incurred on the chosen channel. Similarly, throughout the search process, each search step depends on the realized utilities of previously searched products. Finally, the purchase decision is conditioned by the searched set. Such dependency requires a characterization of the joint probability of all decisions involved in a search session.

To overcome these aforementioned challenges, we follow a likelihood-based estimation method (Chung et al. 2019). This method recursively makes random draws for each dimension that requires numerical integration to simulate the likelihood associated with the decisions in the sequential search process. Compared to traditional frequency simulators (Honka and Chintagunta 2017, Chen and Yao 2017), this method improves computational efficiency and estimation accuracy in the following aspects. (i) This method only draws random values from the set that satisfies the search set condition (as shown in Equation 11), and thus it requires a relatively small number of random values. (ii) It approximates the likelihood directly using the probabilities of the events of interest, rather than using some arbitrary distribution chosen by the modeler. Hence, potential biases introduced by the modeler are less of a concern in this model. (iii) This method's estimates are precise and accurately predict customer behavior (Chung et al. 2019).

We now use an example to explain the core of the method and discuss the algorithm in detail in E-companion Appendix 5. Take an example of a search session that ends with a purchase. As discussed in Section 4.1, the customer's decisions throughout the search session and purchase are entirely driven by reservation utilities and realized utilities. Therefore, we begin with the realized utility of the purchased product and then simulate other realized utilities and reservation utilities in the search process that satisfy the inequalities in the search and purchase decisions. The constraints bounded by the inequality conditions (as shown in Equation 11) and the observed events allow us to make random draws from limited areas of the multi-dimensional space. This significantly reduces computational complexity. Moreover, by deriving the likelihood from actual observed events, the method is more likely to find estimates that coincide with the true values of the parameters.

**Identification Strategies** The parameters to be estimated are  $\Phi_{itc} = \{\theta_{itc}^f, \theta_{itc}^m, \theta_{itc}^s\}$  which represent fixed search costs, marginal search costs, and customers' utility preference parameters,

respectively. The channel-specific customer utility preference parameters,  $\theta_{itc}^s$ , are identified by both the search path and the purchase choice. We observe the attributes of both the purchased product and the other products that the customer searches along the path. We then can compare the purchased product's attributes with the attributes of all other products in the search set, and this comparison contributes to the identification of customers' utility preference parameters. Similarly, across customers, we leverage the correlation between the products' search popularity and their attributes to pin down the mean values of the preference parameters. Leveraging the heterogeneous composition of search sets across individuals, we pin down the variance of the estimates across individuals.

The marginal search costs,  $\theta_{itc}^m$ , are identified using the average length of the search sequence in each channel and action time (its mean) and the variation in search length across individuals (its variance). Search length is proxied by the number of clicks in a sequential search path. The fixed search costs,  $\theta_{itc}^f$ , are identified using the proportion of customers in each channel after controlling for their search lengths and utilities drawn from the purchase option (including the no purchase option).

## 5. Results

We present the results of our structural model in this section. As discussed in Section 3.1, we removed all search sessions with multiple purchases and only considered sessions with at most one unique product purchase in our analyses. We also removed customers who do not have historical records for the user-level variable.<sup>4</sup>

We estimate the model in several computationally efficient ways, while still capturing as much heterogeneity as possible in the key variables. First, rather than using the thousands of SKUs observed in the dataset to construct the potential search set, we use the joint set of customers who have searched for one common product. For example, instead of putting all “hair shampoo” products and “hair conditioner” products in the potential search set, we only include “hair shampoo” products or hair shampoo alternatives when constructing the search set. As a result, we have multiple disjoint search sets ranging from 5 SKUs to 39 SKUs. Second, we estimate search costs (both fixed and marginal) and allow them to vary across channel choice, customer user-level, and action time. To account for the effect of action time, we use harmonic regression terms similar to Bharadwaj et al. (2022). Using harmonic regression terms for time effects (i.e., time-of-the-day and day-of-the-week effect) allows us to capture the periodicity of time in a granular yet parsimonious way. Specifically, we include harmonic regression terms into our model as follows: let  $t_i$  represent

<sup>4</sup> i.e., user-level variable takes value -1, 10, or NA.

the action time  $i$ 's second of the day, and the two periodic regressors included for the time-of-the-day effect are  $(\sin[\frac{2\pi t_i}{86400}], \cos[\frac{2\pi t_i}{86400}])$ .<sup>5</sup> Similarly, to capture the day-of-the-week effect, we include two regressors  $(\sin[\frac{2\pi t_{iw}}{604800}], \cos[\frac{2\pi t_{iw}}{604800}])$  where  $t_{iw}$  denotes the action time  $i$ 's seconds of the week.<sup>6</sup>

Third, the preferences for price and quality are estimated at the customer-channel level and are constant across action times. We convert the attributes to price premiums that customers are willing to pay. Thus, instead of searching sequentially in a three-dimensional space (price, Attribute 1, and Attribute 2), customers are assumed to search on a unified price-quality dimension. We estimate how customers trade off between price and quality using a linear regression for all purchases where price is the dependent variable and two attributes are the explanatory variables. The estimated coefficients for these two quality attributes reflect the amount of price premium that customers are willing to pay for a unit increase in the quality attributes. We calculate the price-equivalent variable by subtracting the premiums for the two quality attributes from the product's price (as defined in Section 3). This method is in line with quality adjusted price, and similar methods are used in recent papers such as De Corniere and Taylor (2019).

In an online retail setting, the prices can be adjusted quickly in response to the prevailing demand levels. This might lead prices to be endogenous. To address this, we included price copulas following Park and Gupta (2012). The copula method is an instrument-free and semi-parametric method, and has been widely used to address endogeneity concerns (Heitmann et al. 2020). As explained by Park and Gupta (2012), this method models the joint distribution of the endogenous variable and the error term, and thus accommodates correlation between them. Following Park and Gupta (2012), we compute the price copulas as  $\tilde{P}_{it} = \Phi^{-1}[H(P_{it})]$  where  $P_{it}$  is the price of the item  $i$  on date  $t$ ,  $\Phi^{-1}$  is the inverse of the normal cumulative distribution function, and  $H(P_{it})$  is the empirical cumulative distribution function of prices. To estimate the empirical cumulative distribution function of prices, similar to Park and Gupta (2012), we used an Epanechnikov kernel function with data-driven bandwidth  $b = \frac{0.9}{n^{1/5}} \times \min(s, \frac{IQR}{1.349})$  where  $n$  is the number of observations in the sample,  $s$  is the sample standard deviation and  $IQR$  is the interquartile range of the data. It is important that prices are not normally distributed for identification purposes. The Anderson-Darling normality test shows that prices are not distributed normally ( $p < 0.0001$ ). The above-mentioned techniques greatly improve computational efficiency without sacrificing our ability to quantify our parameters of interest.

The key results for fixed search costs are presented in Table 7. Overall, we see that the fixed search costs are higher in the WeChat channel compared to the App channel. This is reflected in

<sup>5</sup> A day is composed of 86400 seconds. Hence, 86400 seconds corresponds to  $2\pi$  radians for a day.

<sup>6</sup> A week is composed of 604800 seconds. Hence, 604800 seconds corresponds to  $2\pi$  radians for a week.

the significant positive estimate for **Base** on Table 7. Several factors might contribute to these results. First, it is easier to access the App channel than the WeChat channel. Customers can access retailers' mobile apps just by tapping the app, while, as shown in E-companion Appendix 1, accessing the WeChat channel involves more steps and is more cumbersome. These differences in the ease of access between the WeChat and App channels could be one of the drivers of the fixed search cost differences that we estimated. In addition, retailers can send notifications to customers in the App channel, but they cannot do so in the WeChat channel. This allows them to play a more active role in reaching out to customers in the App channel, thereby reducing their fixed search costs through the use of push notifications. In contrast, as retailers cannot actively initiate conversations with customers in the WeChat channel, they must instead rely on peer-to-peer communication and organic customer search initiation. Thus, retailers cannot take actions to reduce customers' fixed search costs in the WeChat channel. This can also explain why fixed search costs are higher in the WeChat channel.

**Table 7 Fixed Cost: WeChat is Higher than App**

	Estimate	Std. Error	P-value
Base (User-level 1)	2.063	(0.104)	***
User-level 2	2.622	(0.085)	***
User-level 3	3.319	(0.060)	***
User-level 4	4.015	(0.044)	***
Sin.day	2.549	(0.672)	***
Cos.day	-0.454	(0.546)	
Sin.week	-2.412	(0.529)	***
Cos.week	2.311	(0.533)	***

P-value: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

The results for marginal search costs are given in Table 8. Overall, customers have higher marginal search costs in the App channel compared to the WeChat channel. The estimates in the first row of Table 8 indicate that the difference is statistically significant for the base-level customers in these two channels. Several factors might contribute to this result. As discussed in Section 4, marginal search costs capture efforts exerted to locate a product, the psychic costs of searching for a product, and the opportunity cost of time spent on searching for products. In many instances, including our setting, a retailer's WeChat and App channel user interfaces are designed in a very similar way, with the WeChat mini-programs being smaller and lighter compared to Apps. Apps usually occupy more memory space in the phone and include additional access and third-party user tracking features such as accessing users' contact information and pictures. In some cases, apps can even access users' microphones and cameras. Hence, Apps might have slower loading speeds and higher data costs. The faster loading speeds and lower data usage of the WeChat mini-programs can reduce customers' marginal search costs, especially when internet coverage is poor.

In addition, we also observe that both the fixed search costs and marginal search costs are higher for customers at higher user-levels (i.e., customers with high historical spending in JD) in both



**Table 8 Marginal Cost: WeChat is Smaller than App**

	WeChat			App			One-sided K-S Test Hypo: Wechat MC < App MC
	Estimate	Std. Error	P-value	Estimate	Std. Error	P-value	
Base (User-level 1)	0.087	(0.017)	***	0.168	(0.025)	***	***
User-level 2	0.502	(0.169)	**	1.427	(0.094)	***	***
User-level 3	1.506	(0.114)	***	2.107	(0.088)	***	***
User-level 4	2.288	(0.083)	***	3.049	(0.057)	***	***
Sin.day	-0.083	(0.220)		0.673	(0.314)	**	
Cos.day	0.125	(0.179)		0.479	(0.239)	**	
Sin.week	-0.474	(0.175)	***	-0.271	(0.242)		
Cos.week	-0.339	(0.175)	*	0.298	(0.242)		

P-value: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

the WeChat and the App channels. It is not a priori clear whether customers with higher user-levels exhibit higher or lower search costs. On the one hand, these customers might have lower search costs as they might be more familiar with the platform. On the other hand, more affluent customers might value their time more highly and have higher search costs. Our findings suggest that the second force dominates in this setting because we estimate that affluent customers who have historically spent more money on JD have higher search costs. These customers might value their time more; thus, the extra time spent in clicking through products is considered costly for affluent customers. Moreover, the significant estimates of time effects at the daily and weekly levels confirm the influence of customers' shopping time on search costs. Finally, the channel-specific utility parameters in Table 9 also show that the base utility from a purchase increases with a customer's historical spending level (i.e., their user-level). To further strengthen the validity of our finding, we conducted three pair-wise comparisons using the one-sided Kolmogorov-Smirnov test. Specifically, we verified that the base utility of purchase for level-2 customers is significantly larger than that of level-1 customers, the base utility for level-3 customers is larger than that of level-2 customers, and the base utility for level-4 customers is larger than that of level-3 customers. These three pair-wise comparisons confirm a clear increasing trend in base utility as user-spending level increases.

**Table 9 Utility Preference Parameter**

	WeChat			App		
	Estimate	Std. Error	P-value	Estimate	Std. Error	P-value
User_Level 1 Base Utility	1.694	(0.010)	***	1.983	(0.007)	***
User_Level 2 Base Utility	1.854	(0.017)	***	2.062	(0.011)	***
User_Level 3 Base Utility	2.243	(0.027)	***	2.141	(0.018)	***
User_Level 4 Base Utility	2.783	(0.046)	***	2.141	(0.029)	
User_Level 1 Price-equivalent Sensitivity	-0.004	(0.0002)	***	-0.003	(0.0001)	***
User_Level 2 Price-equivalent Sensitivity	-0.004	(0.0003)		-0.003	(0.0002)	
User_Level 3 Price-equivalent Sensitivity	-0.004	(0.0004)		-0.002	(0.0003)	**
User_Level 4 Price-equivalent Sensitivity	-0.004	(0.001)		-0.0005	(0.0003)	***
Price-copula	55.926	(0.159)	***	60.330	(0.230)	***

Price-equivalent: a combination of price and quality attributes.

P-value: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

Finally, we conduct a series of analyses to assess the robustness of our findings to alternative explanations. First, we explore the assortment similarity across the two channels. Our analysis

shows that both the product assortments and popular products are very similar across the two channels. This suggests that product assortment differences are not a significant factor that impacts our findings. Similarly, we do not see any difference in prices across the two channels. Second, we restrict our analysis to sessions which only include the items clicked at least once in both channels. The results of this analysis are in line with the main findings. Moreover, like other papers in the search cost literature (De los Santos et al. 2012, Bronnenberg et al. 2016, Zhang et al. 2019), our analysis focuses on one product category. This allows us to better assess how customers trade off search costs and product characteristics. However, customers might browse products across different categories within a search session. To focus on customer groups in which cross-category browsing behavior is less likely, we sample sessions that last less than fifteen minutes. Concentrating on shorter sessions can help us focus on customers who are less likely to browse products across multiple categories during a session. The results of this analysis are in line with our key findings, which suggest that cross-category browsing is not a significant concern in our setting.

**Model Validation** We evaluate the performance of our structural model with respect to within-sample goodness of fit and out-of-sample predictions. First, we measure within-sample goodness of fit. Using the parameter estimates, we simulate customers' search sequences and purchase decisions in the estimation sample. We begin by calculating the probabilities for channel choice, search choice, and purchase choices for 1000 sampled customers. To calculate these probabilities, we compute the search costs and preference parameters based on our model estimates and the known customers' features such as user-level and action time for each individual customer. Then, conditioning on the product pool and a random draw of the idiosyncratic shock from our estimated distribution, we can calculate the probabilities of each customer's search choices and purchase choices along his search path for both the possibility of the customer using the WeChat channel or the App channel. Combining these two utilities (i.e., from the search and channel choices), we determine the customer's channel choice as the channel generating the higher utility. We repeated this simulation procedure 30 times for each customer to calculate the associated probabilities for this customer. The results for prediction accuracy for each user-level and channel are provided in E-companion Appendix Table 1. Reassuringly, the root-mean-squared errors (RMSE) of the predictions for the search sequences and purchase probabilities are below  $10^{-3}$ .

In addition to within-sample goodness of fit, we also conduct 10-fold cross-validation to evaluate the model's predictive power in new samples. To measure the out-of-sample fit, we repeated the following procedure ten times. At each iteration, we sampled 90% of the sessions to estimate the parameters of interest and used the estimates to simulate the behavior in the remaining 10% of the sessions. Finally, we compared the discrepancies between the customers' predicted and actual decisions. The results for the prediction accuracy for each user-level and each channel are provided

in E-companion Appendix Table 2. Reassuringly, the root-mean-squared errors (RMSE) of the predictions for the search sequences and purchase probabilities are below  $10^{-3}$ .

## 6. Counterfactual Analyses of Two Channel-Specific Operation Strategies

In this section, we leverage the findings on the differences between the customer search costs in the WeChat and App channels to propose strategies that JD can utilize to improve its multi-channel strategy. Our results show that customers have higher fixed costs in WeChat compared to App for initiating a search session but lower marginal search costs per additional search step. We explore the profitability of potential business practices that JD can implement in light of these differing fixed and marginal search costs across the two channels and we propose two strategies. The first strategy is to encourage customers to distribute shareable links in the WeChat channel to reduce the associated fixed search cost. The second strategy is to leverage search-triggering coupons in the App channel to mitigate the marginal search costs.

### 6.1. Sharable Links to Lower Fixed Search Costs in WeChat

The first strategy is to promote shareable links in the WeChat channel. A customer who receives a link shared by a friend can click this link and be taken directly to the product page in the WeChat mini-program without needing to complete the multi-step opening procedure shown in E-companion Appendix 1. In this way, the customer's fixed cost of initiating a search in the WeChat channel is reduced.

Many e-commerce platforms began exploring social commerce and encouraging customers to share product links with their friends. For example, Taobao (owned by Alibaba) motivates its customers to share product links with their friends by offering a commission fee on purchases that occur through these shared links. Moreover, the customers who make purchases using these shareable links also receive discounts at the check-out. Similarly, Weee and FreshGoGo, two rising Asian grocery e-commerce platforms in North America, offer credits to both link distributors and link users if these links generate a purchase. As WeChat is a social commerce channel, promoting link-sharing between WeChat friends can be an effective strategy for reducing fixed search costs and increasing conversion rates.

Our counterfactual analysis for WeChat is inspired by the business practices of Taobao, Weee, and FreshGoGo. In this analysis, we incentivize a customer to share links by offering him purchase credits for each time a purchase is made via his shared link. We also incentivize the link receiver to purchase by providing the equivalent amount in purchase credits to them in the event of a purchase. We conduct the following simulation to evaluate the effectiveness of the shareable link strategy and to determine the optimal purchase credit level. First, we generate 1000 customers in

four user-level and across shopping times following the distribution in our data. We then compute each customer's fixed search cost based on estimates from Table 7; his marginal cost based on estimates from Table 8; and his utility preference parameters based on estimates from Table 9, according to his user-level and action time. As sharing links has virtually no cost for either the platform or the customer who shares the link (the return is positive if the link leads to a purchase and zero if there is no purchase), we assume that all 1000 customers receive a shareable link. Hence, their fixed search cost becomes 0 as they can access the WeChat mini-program directly via the link, without completing the multi-step procedure needed for access. In addition, these 1000 customers would receive purchase credits if they choose to make a purchase. Thus, these purchase credits would lower their product purchasing costs and increase their probability of converting the search into a purchase. Ideally, we would measure how the probability of link sharing increases with the given incentives and purchase credits. However, we cannot estimate this relationship using our current data.

Using the behavior models introduced in Section 4, we simulate the customers' search paths, channel choices, and final purchases. We compute the conversion rates and the net returns (extra profit generated minus the additional costs) for each of the four user-level and purchase credit levels. Using these simulated customers, we evaluate the impact of varying levels of purchase credits associated with each sharable link, ranging from 0 to 5 CNY. The 0 CNY case constitutes the status quo (i.e., no incentives for link-sharing and customers face positive fixed search costs for accessing the WeChat mini-program). The purchase credits were capped at 5 CNY because of JD's reported average net profit margin of 10%.<sup>7</sup> For example, for a product that sells at 100 CNY (a typical price level in our data), JD's net profit margin is around 10 CNY. As the purchase credits are offered to both the link distributor and the link user, we cap the purchase credit at 5 CNY (where  $2 \times 5\text{CNY} = 10\text{CNY}$ ) to ensure that JD does not lose money by offering these purchase credits.

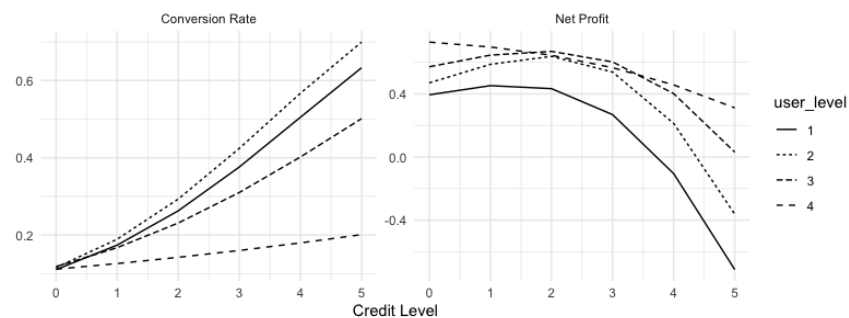
Based on the simulated behavior at each credit level and user-level, we compute the conversion rate as the ratio of customers who complete their search by making a purchase to all customers in this group. Figure 3 shows the changes in conversion rates in relation to changes in purchase credits. The monotone increasing trend in the left panel shows that, as expected, the likelihood of purchase increases as the value of the purchase credits increases. Moreover, customers with lower user-level are more likely to be motivated to purchase by the credits.

To assess how the referral discounts might benefit JD, we calculate the net return per customer at each purchase credit level and user-level. If a link user does not make a purchase, there would

<sup>7</sup> <https://xueqiu.com/4459946931/86998424>

be no return or cost incurred for JD. If a link receiver makes a purchase, JD receives the net profit (10% of the product price) minus the value of the two purchase credits. We compute the net return per customer by using the total net return for customers at a particular user-level divided by the number of customers at this user-level. As shown in Figure 3, customers are more likely to purchase when the purchase credits are higher. However, if the credit level is too high, this might hurt JD's profits. In the right panel of Figure 3, we see that the optimal credit level is around 2 CNY for customers of user-level 1-3, leading to an average additional profit for JD of 0.6 CNY. Offering a 5 CNY discount induces the most significant conversion jump but decreases the generated margin per customer. Considering the volume of customers using JD's WeChat channel, an additional profit of 0.6 CNY per customer would amount to millions of CNY.

**Figure 3** Improvement with Referral Discount – WeChat



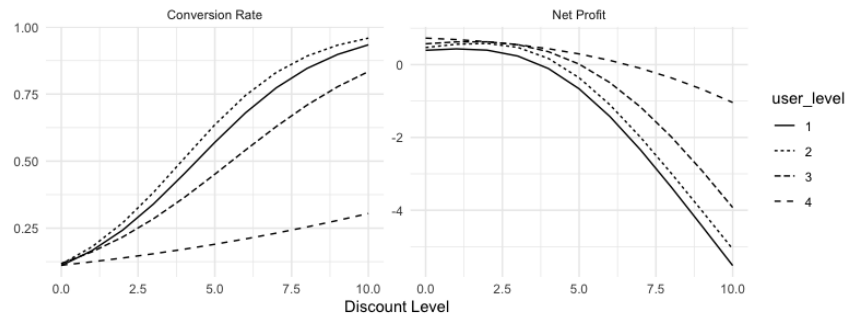
## 6.2. Using Search-Triggering Coupons to Lower Marginal Search Costs in the App

Our results indicate that customers have higher marginal costs in App compared to WeChat. These high marginal costs induce customers to stop their search sessions early in the App if the remaining products in the search set are not expected to yield sufficient marginal benefits. This early search process termination results in lower conversion rates. To prevent customers from exiting the search early, we propose a nudging coupon for App customers along their search paths. This coupon would be provided to the users after their second click. We choose the second click because this point corresponds with the highest drop-out rate. However, the timing of this nudging coupon can potentially be personalized, based on the customer's marginal costs.

As in the customer-generation steps described in Section 6.1, we simulate 1000 App customers, with characteristics representing the observed composition of the data. We then sample their fixed costs, marginal costs and preferences from the estimates' confidence intervals in Tables 7, 8, and 9. We evaluate the impact of various levels of search-nudging discounts ranging from 0 to 10 CNY. The coupon with 0 CNY represents the status quo and is used as a benchmark.

Figure 4 shows the changes in the conversion rate with such search-triggering discounts. As expected, customers' purchase likelihood increases with an increase in the search-triggering discounts. In the right panel of Figure 4 we see that JD's profit margin is non-monotonic. The optimal coupon level is around 1-2 CNY per customers. For each customer, we can generate an average of approximately 0.7 CNY, which is a sizable amount given the large volume of JD customers.

**Figure 4** Improvement with Search-Triggering Discount – App Customers



In this section, we proposed two channel-specific strategies in response to customers' differing search costs across the WeChat and App channels. The first strategy is to use friend-referral discounts to lower the fixed search costs in the WeChat channel. The second is to use search-triggering coupons to nudge App customers to continue their search. Both strategies can increase conversion rates, and most importantly, JD's profits if the discounts are correctly quantified.

## 7. Conclusion

As social commerce gains prominence, retailers are increasingly exploring new channels to engage customers. Our study is the first to analyze customer behavior differences across a retailer's social commerce and native app channels, uncovering notable distinctions in search costs. Our paper not only contributes to the growing literature on omnichannel retail by examining two innovative mobile channels but also adopts a unique approach by analyzing customers' entire shopping journeys in WeChat and App channels using a sequential search model. This enables us to better understand customer decision-making processes and develop tailored promotion strategies. Furthermore, our work is the first to estimate both fixed and marginal search costs in a sequential search model, adding valuable insights to the search model literature.

In summary, our findings reveal significant differences in the search costs associated with the two channels, with customers' fixed search costs being higher in the WeChat channel and marginal search costs being lower in WeChat. Leveraging these fixed and marginal search cost differences across channels, we design channel-specific customer targeting strategies for JD through counterfactual analyses. By offering 2 CNY friend referral coupons, JD can reduce the fixed cost of

initiating a WeChat search session and increase their profit margins by an average of 0.6 CNY per user per search session. Similarly, a 1-2 CNY coupon designed for the App channel that is sent to customers after their second click in a search session can increase profit margins by an average of 0.7 CNY per user per search session. These findings suggest that appropriately catering coupons of this nature could create a substantial revenue stream for the retailer, considering the volume of customers visiting this specific product category over a one-month period.

In conclusion, our paper offers valuable insights for retailers regarding channel-specific promotion design and beyond. While our study focuses on an e-commerce retailer that designed WeChat and App channels identically and offered the same assortment, we think our findings provide practical implications for channel design and actionable insights. For instance, retailers can allocate resources to innovations that decrease fixed search costs in WeChat and marginal search costs in App channels. Moreover, ignoring search costs in assortment planning may lead to significant revenue loss, as prior literature suggests (Wang and Sahin 2018). Future research could investigate how retailers can adjust assortments based on varying search costs across channels. Our study's results can also help develop channel-specific marketing campaigns. Campaigns facilitating decrease in fixed search cost in the WeChat channel could be effective, while those aimed at reducing marginal search costs may work better in the App channel. Overall, our findings hold significant implications for channel design and marketing strategies, potentially improving customer experience and driving sales.

## References

- Bharadwaj, Neeraj, Michel Ballings, Prasad A Naik, Miller Moore, Mustafa Murat Arat. 2022. A new livestream retail analytics framework to assess the sales impact of emotional displays. *Journal of Marketing*, 86 (1), 27-47.
- Boada-Collado, Pol, Victor Martínez-de Albéniz. 2020. Estimating and optimizing the impact of inventory on consumer choices in a fashion retail setting. *Manufacturing & Service Operations Management*, 22 (3), 582-597.
- Bronnenberg, Bart J, Jun B Kim, Carl F Mela. 2016. Zooming in on choice: How do consumers search for cameras online? *Marketing Science*, 35 (5), 693-712.
- Brynjolfsson, Erik, Yu Hu, Duncan Simester. 2011. Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, 57 (8), 1373-1386.
- Chen, Yuxin, Song Yao. 2017. Sequential search with refinement: Model and application with click-stream data. *Management Science*, 63 (12), 4345-4365.
- Chintagunta, Pradeep K, Junhong Chu, Javier Cebollada. 2012. Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Science*, 31 (1), 96-114.

- Chu, Franklin. 2019. Why china e-commerce is going crazy for wechat mini-programs. URL <https://www.digitalcommerce360.com/2019/04/16/why-china-ecommerce-is-going-crazy-for-wechat-mini%E2%80%91programs/>. (Accessed on 10/12/2021).
- Chung, Jae Hyen, Pradeep K Chintagunta, Sanjog Misra. 2019. Estimation of sequential search model. *Available at SSRN 3203973*, Working Paper.
- Cui, Ruomeng, Santiago Gallino, Antonio Moreno, Dennis J Zhang. 2018. The operational value of social media information. *Production and Operations Management*, 27 (10), 1749-1769.
- De Corniere, Alexandre, Greg Taylor. 2019. A model of biased intermediation. *The RAND Journal of Economics*, 50 (4), 854-882.
- De los Santos, Babur, Ali Hortaçsu, Matthijs R Wildenbeest. 2012. Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102 (6), 2955-80.
- Donovan, John. 2021. Financial reporting and entrepreneurial finance: Evidence from equity crowdfunding. *Management Science*, 67 (11), 7214-7237.
- Ertekin, Necati, Anupam Agrawal. 2021. How does a return period policy change affect multichannel retailer profitability? *Manufacturing & Service Operations Management*, 23 (1), 210-229.
- Gallino, Santiago, Nil Karacaoglu, Antonio Moreno. 2022a. Algorithmic assortment curation: An empirical study of buybox in online marketplaces. *Available at SSRN 4218952*, Working Paper.
- Gallino, Santiago, Nil Karacaoglu, Antonio Moreno. 2022b. Need for speed: The impact of in-process delays on customer behavior in online retail. *Operations Research*, Forthcoming.
- Gallino, Santiago, Antonio Moreno. 2014. Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information. *Management Science*, 60 (6), 1434-1451.
- GrandviewResearch. 2021. Social commerce market size, share trends analysis report by business model (b2c, b2b, c2c), by product type (apparel, personal beauty care, accessories, home products), by region, and segment forecasts, 2021 - 2028. URL <https://www.grandviewresearch.com/industry-analysis/social-commerce-market>. (Accessed on 04/20/2021).
- Haviland, Amelia, Daniel S Nagin, Paul R Rosenbaum. 2007. Combining propensity score matching and group-based trajectory analysis in an observational study. *Psychological Methods*, 12 (3), 247.
- Heitmann, Mark, Jan R Landwehr, Thomas F Schreiner, Harald J van Heerde. 2020. Leveraging brand equity for effective visual product design. *Journal of Marketing Research*, 57 (2), 257-277.
- Honka, Elisabeth. 2014. Quantifying search and switching costs in the us auto insurance industry. *The RAND Journal of Economics*, 45 (4), 847-884.
- Honka, Elisabeth, Pradeep Chintagunta. 2017. Simultaneous or sequential? search strategies in the us auto insurance industry. *Marketing Science*, 36 (1), 21-42.



- Honka, Elisabeth, Ali Hortaçsu, Matthijs Wildenbeest. 2019. Empirical search and consideration sets. *Handbook of the Economics of Marketing*, vol. 1. Elsevier, 193-257.
- Hu, Kejia, Jason Acimovic, Francisco Erize, Douglas J Thomas, Jan A Van Mieghem. 2019. Forecasting new product life cycle curves: Practical approach and empirical analysis: Finalist-2017 m&som practice-based research competition. *Manufacturing & Service Operations Management*, 21 (1), 66-85.
- Kharpal, Arjun. 2019. Everything you need to know about wechat — china's billion-user messaging app. URL <https://www.cnbc.com/2019/02/04/what-is-wechat-china-biggest-messaging-app.html>. (Accessed on 05/15/2021).
- Kim, Jun B, Paulo Albuquerque, Bart J Bronnenberg. 2010. Online demand under limited consumer search. *Marketing Science*, 29 (6), 1001-1023.
- Kong, Lu, Kejia Hu, Rohit Verma. 2022. Service chains' operational strategies: Standardization or customization? evidence from the nursing home industry. *Manufacturing & Service Operations Management*, 24 (6), 3099-3116.
- Koulayev, Sergei. 2014. Search for differentiated products: identification and estimation. *The RAND Journal of Economics*, 45 (3), 553-575.
- Kumar, V, Rajkumar Venkatesan. 2021. Transformation of metrics and analytics in retailing: The way forward. *Journal of Retailing*, 97 (4), 496-506.
- Lee, Hyun Seok, Saravanan Kesavan, Vinayak Deshpande. 2021. Managing the impact of fitting room traffic on retail sales: Using labor to reduce phantom stockouts. *Manufacturing & Service Operations Management*, 23 (6), 1580-1596.
- Liao, Rita. 2021. Wechat advances e-commerce goals with \$250b in transactions. URL <https://techcrunch.com/2021/01/19/wechat-2020>. (Accessed on 04/20/2021).
- Manchanda, Puneet, Grant Packard, Adithya Pattabhiramaiah. 2015. Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Science*, 34 (3), 367-387.
- Mišić, Velibor V, Georgia Perakis. 2020. Data analytics in operations management: A review. *Manufacturing & Service Operations Management*, 22 (1), 158-169.
- Park, Sungho, Sachin Gupta. 2012. Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, 31 (4), 567-586.
- QPSSoftware. 2021. Wechat mini program - all you need to know. URL <https://qpsoftware.net/blog/wechat-mini-program-all-you-need-know>. (Accessed on 10/12/2021).
- Reuters. 2019. Whatsapp adds shopping catalog feature, courting e-commerce — reuters. URL <https://www.reuters.com/article/us-facebook-whatsapp-shopping-idUSKBN1XI050>. (Accessed on 10/01/2020).

- Rong, Ke, Di Zhou, Xinwei Shi, Wei Huang. 2022. Social information disclosure of friends in common in an e-commerce platform ecosystem: An online experiment. *Production and Operations Management*, 31 (3), 984-1005.
- Rosenbaum, Paul R, Donald B Rubin. 1983. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45 (2), 212-218.
- Seiler, Stephan, Fabio Pinna. 2017. Estimating search benefits from path-tracking data: measurement and determinants. *Marketing Science*, 36 (4), 565-589.
- Shen, Max, Christopher S Tang, Di Wu, Rong Yuan, Wei Zhou. 2020. Jd. com: Transaction-level data for the 2020 msom data driven research challenge. *Manufacturing & Service Operations Management*, Forthcoming.
- Tech, Thunder. 2019. Does the future of retail start with social media? URL <https://www.thundertech.com/blog/February-2019/Does-the-Future-of-Retail-Start-with-Social-Media>. (Accessed on 10/01/2020).
- Wang, Ruxian, Ozge Sahin. 2018. The impact of consumer search cost on assortment planning and pricing. *Management Science*, 64 (8), 3649-3666.
- Weitzman, Martin L. 1979. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 641-654.
- Windmeijer, Frank. 2005. A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126 (1), 25-51.
- Xu, Kaiquan, Jason Chan, Anindya Ghose, Sang Pil Han. 2017. Battle of the channels: The impact of tablets on digital commerce. *Management Science*, 63 (5), 1469-1492.
- Zhang, Shuo, Zhenling Jiang, Hai Che. 2019. Online shopping with endogenous pc and mobile channel choice. Working Paper.
- Zhang, Xingyue, Ruomeng Cui, Oliver Yao. 2023. The version effect of apps and operating systems in mobile commerce. *Production and Operations Management*, 32 (2), 637-654.