Does Pre-login Search Matter? Evidence from a Mobile Commerce Platform

Qiaodan (Jordan) He
Lehigh University
qih320@lehigh.edu

Xingyue (Luna) Zhang
University of Washington Tacoma
xyzhang5@uw.edu

Dawei (David) Zhang
Lehigh University
daz215@lehigh.edu

Yuliang(Oliver) Yao
Lehigh University
yuy3@lehigh.edu

Abstract

An increasing number of consumers enjoy shopping through mobile devices. When consumers use a mobile app, they can choose whether to log in with their accounts. We argue that pre-login search plays a critical role in affecting consumers' purchase decisions, although it has largely been overlooked in the literature. Using clickstream data, we adopt different econometric models to examine whether and how pre-login search affects the likelihood of purchase. Our results show that pre-login search behaviors are as important as post-login search to consumers' purchase decisions. We also demonstrate that consumers' purchase propensity increases at a diminishing rate with an increasing search effort during both pre- and post-login periods. Based on recommender systems (RSs) and paradox of choice theory, our results contribute to the burgeoning literature on consumer behavior in mobile commerce and provide novel insights to the strategic usage of RSs. Finally, we discuss theoretical and managerial implications.

Keywords: mobile commerce, pre-login search, post-login search, recommender systems, choice overload.

1. Introduction

The growth of mobile commerce has been increasing sharply over the past few years. Global retail mobile commerce sales hit $359.32 billion in 2021, an increase of 15.2% from 2020. By 2025, retail mobile commerce sales are projected to be more than double to reach $728.28 billion and account for 44.2% of retail ecommerce sales in the U.S. (Insider Intelligence 2022).[1] It is not difficult to imagine that mobile commerce is going to be the next true frontier for shopping because mobile commerce continues to become more popular.

Since mobile commerce has become more prevalent in our daily lives, consumer behaviors in mobile commerce attract much attention from researchers. Research on mobile commerce involves many fields, including mobile app design (e.g., Ahmad and Ibrahim 2017), technology adoption (e.g., Chhonker et al. 2017), consumer trust in mobile commerce (e.g., Lin et al. 2014), and mobile app adoptions and purchase behaviors (e.g., Ghose and Han 2014, Han et al. 2016, Sun et al. 2019). All these studies focus on users' activities in mobile commerce. However, few studies have examined the relationship between consumers' pre-login search behaviors and their purchases in the context of mobile commerce. Consumers' purchase probability is determined by both pre- and post-login search. Ignoring pre-login search behaviors may overestimate the effect of post-login search on the consumers' purchase probability.

To narrow the literature gaps, we focus on comparing search behaviors before and after logging in. Specifically, we examine search behaviors of consumers who have accounts but search without logging into their accounts first as user-visitors (pre-login search), and then log into their accounts to search as users (post-login search). Following previous studies (Bucklin and Sismeiro 2003; Johnson et al. 2004), we develop three measures to capture consumers' search efforts: (1) search time, (2) search depth, and (3) search width. All three measures capture consumers search efforts independently and have a positive relationship with consumers search efforts. For example, the more search time consumers spend on viewing products, the more search efforts consumers put into.

To the best of our knowledge, no study has examined the relationship between consumers' pre-login search behavior and their purchase outcomes on a mobile app. We aim at answering an important research question: How do pre-login and post-login search efforts affect the likelihood of consumers' purchase outcomes differently? We collected consumer' click-stream data from a leading mobile commerce retailer and develop logit econometric

[1] https://www.insiderintelligence.com/

URI: https://hdl.handle.net/10125/103333
978-0-9981331-6-4

models to examine the effect of pre- and post-login search on consumers' purchase probability. We find that both pre-login and post-login search efforts significantly affect consumers' purchase probabilities. Strikingly, the differences between the effects of pre- and post- login search efforts are not statistically significant. This result suggests that pre-login search behaviors, which are often times neglected, are as important as post-login search behaviors. Furthermore, we test the second-order effect of search efforts on purchase probability and find that in both pre-login and post-login periods, with search efforts increasing, the marginal effect on consumers' purchase probability decreases.

Our paper makes several contributions to the Information System literature. First, we contribute to the burgeoning literature on online search by highlighting the effect of pre-login (can be regarded as nongoal-directed) search on consumers' purchase probability. Although there have been studies investigating consumers' online search behaviors due to the growth of e-commerce (e.g., Zhang et al. 2006; Bhatnagar and Ghose 2004; Bucklin et al. 2002), the effect of pre-login search has been largely neglected. The difference between search/deliberation visits and directed-buying visits is the timing of the purchase (Moe 2003). Not all search visits are goal-directed with a planned purchase in mind. It is often the case that some searches are purely based on consumers' general interests or inadvertently browsing. Our study is among the first to empirically study the relationship between user-visitors' search behavior and their purchase decisions.

Second, we contribute to the literature of consumers behavior in mobile commerce. Previous studies investigate mobile channel adopters search change (Park et al. 2020), the spillover effect of editor recommendations (Liang et al. 2019), and the real-time uses such as choice and duration across mobile apps (Wu et al. 2022). However, the understanding of consumers' pre-login behavior in mobile commerce is still lacking.

Third, we test the theory of choice overload in the context of mobile commerce, whereby we theorize how choice overload influences consumers purchase decision, and further provide insights to the strategic usage of RSs for a company. More broadly, our work extends the literature on the efficacy of choice overload and RSs in the mobile commerce contexts.

The rest of the paper is organized as follows. In Section 2, we develop our hypotheses based on recommender systems theory and paradox of choice. Next, we describe our clickstream data and empirical context in Section 3. In Section 4, we detail our empirical analysis by constructing several econometric models and discuss our empirical results. Finally, we make a conclusion with suggestions for future research.

2. Theory and Hypotheses

Consumer search behavior can be divided into two stages for consumers: pre-login search as user-visitors and post-login search as users. We posit that both stages play an important role in consumers' purchase decisions (H1). Specifically, pre- and post-login search behavior can be different as a result of how a RS works on an e-commerce platform. Figure 1 depicts the difference of a RS between these two stages. We develop two competing hypotheses to examine the importance of pre-login search and whether there is a statistically significant difference between pre- and post-login search efforts on consumers' purchase probability (H2a and H2b). Supported by the choice overload theory, we argue that there is a diminishing effect of both pre- and post-login search efforts on their purchase probability (H3a and H3b).

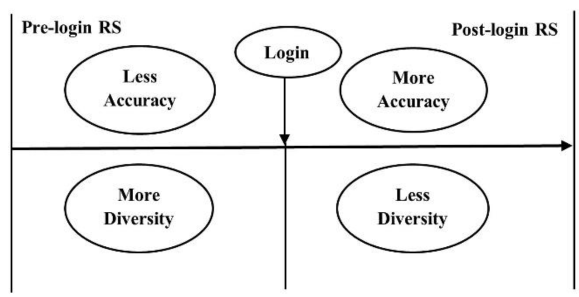


Figure 1. Illustration of the Difference between Pre-login and Post-login Periods.

2.1. Recommender Systems in Mobile Commerce

The RS is designed to assist consumers in product search and selection (Li and Karahanna 2015) and works for consumers in both pre- and post-login periods. RSs in online retail contexts benefit both consumers and companies. A mobile app server can record consumers' browsing and search behaviors as well as purchase histories, and present recommended products to specific consumers based on such information. Lee et al. (2020) examine differences of RSs between personal computer (PC)-based settings and mobile devices. They find that the use of RSs enhanced customer-level outcomes, such as number of views and sales of recommended products, clickthrough rate, and conversion. More importantly, marginal impacts of the RS are significantly higher for mobile users. In other words, RSs effectively reduce consumer search costs on mobile devices. Brynjolfsson et al. (2011) also argue that a retail website provided consumers with information technology (IT)- enabled search, discovery tools, and RSs, could lower consumers' costs of acquiring product information.

In addition, researchers have studied the combined effect of an RS' accuracy and diversity (e.g. Panniello et al. 2016). There have been many studies investigating the performance of RSs by adopting concepts such as accuracy, diversity, novelty and serendipity (e.g. Bobadilla et al. 2011; Patra et al. 2015; Adomavicius and Kwon 2014; Kaminskas and Bridge 2016; Kunaver and Požrl 2017; Silveira et al. 2019). Following previous studies and given our research questions, we define RSs' accuracy as the extent to which a recommended product fits a consumer's preference. The better a product is matched to a consumer's preference, the more accurate the RS is. RSs' diversity measures how dissimilar recommended items are for a customer and is defined as the number of unique products recommended to consumers regardless of popularity (e.g., Kunaver and Požrl 2017).

Many existing papers on context-aware RSs focus on accuracy metrics when measuring the performance of RSs and find that knowledge of certain relevant types of contextual information leads to a more accurate estimation of unknown ratings (Adomavicius and Tuzhilin 2011; Panniello et al. 2016). The concept of diversity of RSs has been discussed in mobile commerce literature because it is important to recommend a wide range of products that consumers have not seen before (Vargas and Castells 2011; Zhang and Hurley 2008). Both the accuracy of RSs and the number of recommended items have significant impacts on users' satisfaction (Liang et al. 2006) and consumers' purchase decisions.

Different from previous studies (e.g., Bradley and Smyth 2001; McSherry 2002; Adomavicius and Kwon 2012) (i.e., illustrating accuracy and diversity tradeoff in terms of the popularity of products), the degree of accuracy and diversity of a RS in our context depends on how much consumers' information a RS could access. Higher accuracy may be achieved by having access to more users' information, which can result in a reduction in diversity. Conversely, higher diversity can be obtained by recommending consumers products across categories as different as possible due to intentionally less-accurate RSs or a lack of access to user-visitors information. Not only are such systems focusing on providing wider range of products beneficial to users because they have more opportunities to get recommended such unintended items, but also the benefit of a RS being more diverse could be apparent to some business models as well (e.g., Brynjolfsson et al. 2011; Fleder and Hosanagar 2009). For instance, it would be profitable to Netflix if RSs encourage users to rent "long -tail" type of movies (i.e., obscure items located in the tail of the sales distribution) (Adomavicius and Kwon 2012). Nevertheless, incorporating diversity-enhancing techniques into RSs may be a pitfall causing issues with less accurate RSs (McGinty and Smyth 2003). Overall, a trade-off between RSs accuracy and diversity exists in a collaborative filtering.

2.2. Paradox of Choice in Mobile Commerce

From a psychology perspective, over-choice or choice overload is a cognitive impairment in which people have a difficult time deciding when faced with too many options (Schwartz 2004). This concept was first introduced by Alvin Toffler in his 1970 book. The popular theory - the Psychology Paradox of Choice was initially proposed by the American psychologist, Barry Schwartz in his book The Paradox of Choice, published in 2004. Prior studies on paradox of choice exist mostly in psychology fields and few papers investigate this theory in the mobile commerce context. Kinjo and Ebina (2015) study a specific example of this paradox with respect to consumer nonpurchase behavior and derive an optimal strategy for a firm selling goods or services for consumer purchase. Schwartz and Ward (2006) propose some facets, such as choice and happiness, freedom or commitment, and missed opportunities, to explain why people suffer from the choice overload. In sum, consumers facing too many choices may feel anxious, stressful, and unhappy because it takes more cognitive efforts for them to make the "best decision". There is evidence from Schwartz (2004) showing that choice overload is detrimental to people's emotional and mental well-being.

2.3. Hypothesis Development

Consumers receive benefits from RSs in three mechanisms – reduced search costs, the recall and retrieval process (i.e., recall a consumer's memory by presenting a focal product), and a signaling effect (Lee et al. 2020). With lower search costs of acquiring product information (Brynjolfsson et al. 2011), consumers are more likely to make a purchase using a mobile device. Li et al. (2021) conduct a field experiment and find that the presence of RSs increases consumers purchase probability. Regardless of the status (i.e., pre-login or post-login), the RS works all the time. Since an RS can access consumers' data such as browsing and purchase histories, it is expected to recommend products to fit consumer preference better (Adomavicius and Tuzhilin 2005; Xiao and Benbasat 2007; Li et al. 2021), leading to a higher likelihood of purchase.

Although the search cost for users may be lower than that for user-visitors due to higher accuracy of RSs in the post-login period, pre-login search behaviors cannot be simply ignored. Beyond reduced search costs, RSs are able to trigger the recall and retrieval process which has been demonstrated to influence purchase decisions positively (Lee et al. 2020; Bettman 1979; Lynch and Srull 1982). For example, based on what consumers viewed last time or purchase behavior, RSs can recommend a focal product and present related products (Linden et al. 2003). Even though RSs cannot access user-visitors' purchase histories and demographic information, RSs can recommend consumers with products based on their current search behaviors. Therefore, pre-login search effort can also increase consumers' purchase probability. Moreover, many consumers may be advertently or inadvertently concerned with mobile platforms collecting their personal information once logged in, and thus choose to do most of their searches without logging into their account (i.e., pre-login search).

Since RSs provide salient landing pages to a list of recommended products, these serve as decentralized signals of quality and potential demand, thereby enhancing the appeal of such products (Pathak et al. 2010; Lee et al. 2020). Such exposure may help consumers to "discover" more desirable items. Based on aforementioned arguments, we propose the following hypothesis.

Hypothesis 1: *Consumers' purchase probability is not only positively associated with their post-login search efforts as users but also positively associated with their pre-login search efforts as user-visitors*.

Pre- and post-login search behavior can be different depending on the accuracy and diversity of the RS. A RS can access users' purchase histories and demographic data, and thus can have more accurate recommendations for the users during the post-login period than for user-visitors in the pre-login period (Resnick and Varian 1997; Lee et al. 2020). As a result, consumers' purchase decisions are more correlated with their post-login search than their pre-login search. Using collaborative filtering algorithms, RSs can also present consumers who have logged in with a list of recommended products which have been viewed, searched or purchased by "similar" consumers in terms of gender, age, and other demographic characteristics. Such recommendations can be viewed as an indirect signal of quality (Linden et al. 2003), and can make the post-login search more influential on consumers' purchase decisions.

On the other hand, although RSs in the pre-login period may be less accurate and more diverse, pre-login search effect could have a larger effect on consumer purchase than post-login search effort. First, less accurate RSs lead to a higher diversity of recommended products during the pre-login period. Because not all search visits are goal-directed with a planned purchase in mind, higher diversity of the recommended products means a higher probability that consumers find some previously-unthought-of products they like, which may not be recommended during the post-login period. Second, since the RS is more precise after consumers log in, it may recommend more products similar to the consumers' focal products that they intend to buy. To select the "best" item among a larger number of similar alternatives, the comparison process could be more exhaustive. Consequently, users may need to spend more post-login search efforts in inspecting product attributes and comparing across the alternatives before making a purchase decision (Zhang et al. 2011). During this process, users are more likely to find drawbacks of the product they searched, which may hurt their desire to purchase this product. Thus, more post-login search efforts can have a negative effect on consumers purchases. In addition, users may get overwhelmed as a result of many similar products in the post-login period, leading to a lower likelihood of purchase for consumers. In sum, post-login search behaviors as users may have a smaller impact on their purchases than pre-login search behaviors as user-visitors.

Overall, whether consumers' pre-login search efforts have a greater or smaller impact on their purchase probability than post-login search efforts remains an empirical question. Thus, we propose the two facets about consumers' purchase probability as follows.

Hypothesis 2a: *Consumers' purchase probability is more associated with their post-login search efforts as users than their pre-login search efforts as user-visitors*.

Hypothesis 2b: *Consumers' purchase probability is more associated with their pre-login search efforts as user-visitors than their post-login search efforts as users*.

Search efforts increase the likelihood of consumers purchase directly and indirectly. With more search efforts, consumers are more likely to purchase because the probability of finding their targets is higher. Prior literature in psychology suggests affective connection between consumers and products built up by search efforts would increase the consumers' commitments to the products, inducing them to purchase (Traylor 1984; Beatty et al. 1988). Additionally, the consideration sets of consumers will become larger with more search efforts. Indeed, the results from Li et al. (2021) show that a larger consideration set could induce consumers to buy.

On the other hand, over-search in either pre- or post-login period may cause consumers to be overloaded with choices and impair their desire to purchase. Consumers seek to maximize their utilities when making decisions (Schwartz 2004). However, too many options may reduce their utilities. Iyengar and Lepper (2000) implemented three experimental studies in both field and laboratory settings, where they arranged an array of 6 choices, 24 varieties or 30 choices at a supermarket, and observed customer behavior. These experiments show that consumers are more likely to purchase when offered a limited array of 6 choices, rather than a more extensive array of 24 or 30 choices (Iyengar and Lepper 2000), indicating that having more choices is not necessarily more intrinsically motivating than fewer choices. Therefore, as the number of choices increase, the marginal utility of consumers' gain will decrease. We propose the following hypotheses on the second-order effects of search effort.

Hypothesis 3a: *Consumers' purchase probability is increasing with pre-login search efforts as user-visitors, but in a diminishing increasing rate*.

Hypothesis 3b: *Consumers' purchase probability is increasing with post-login search efforts as users, but in a diminishing increasing rate*.

Figure 2 illustrates our research framework and summarizes all hypotheses.

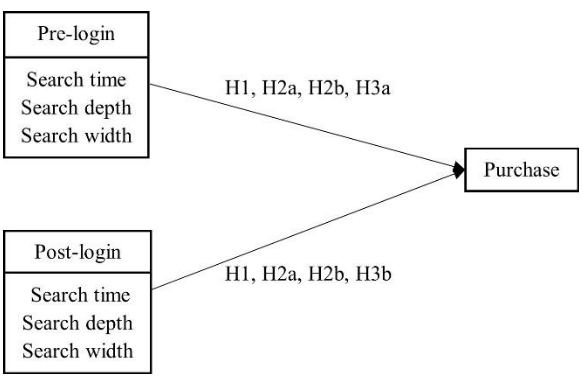


Figure 2. Research Model.

3. Empirical Context and Data

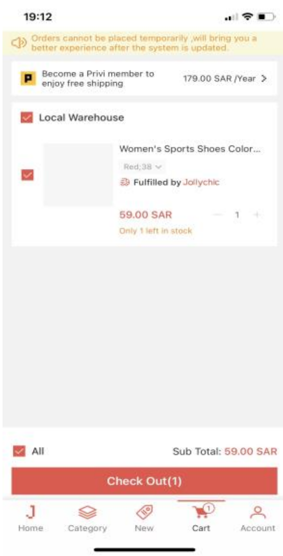
We collected data from a leading mobile commerce retailer that sells consumer products in 98 countries with a focus on the Middle Eastern market. The retailer provides over 1500 new items added every day and has 16,950 monthly mobile app downloads according to Apptopia statistics. [2] We collected mobile clickstream data for all consumers' activities from September 16, 2017 to September 18, 2017. Since there were no special promotions or significant changes on the mobile app affecting consumers' search and shopping behavior during these three days, we can eliminate selection bias given our research questions.

The mobile clickstream data record the consumers' dynamic browsing and purchasing histories, including consumers' activities on the app and the product information consumers viewed (e.g., login behavior, viewed product detail pages and product categories, time stamps for each activity, whether they searched for a product, whether they added a product to their shopping carts, what they purchased, how much they spent, etc.). We also obtained data on the consumers, such as consumers' IP addresses, whether they are new app users during a specific period, and what app and mobile devices they use (i.e., app version, operating system version).

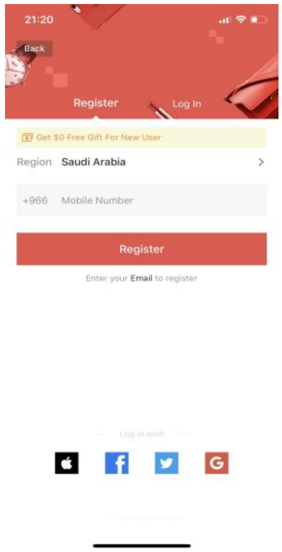
The retailer has a free mobile app to facilitate consumer shopping on mobile devices. Figure 3 shows some screenshots of shopping pages in the app. When a consumer opens the mobile app, the first page is homepage where she can start her shopping tour. The consumer can either search for a specific product by keywords or browse by categories. Once the consumer finds any product she is interested in, she can click on the product to access the product detail page. After searching for a while, the consumer may choose to check out with the shopping cart or leave. It is important to note that consumers cannot make a purchase if they do not log in with their accounts. We focus on consumers who did not log in in the beginning but logged in during their shopping process. After logging in, a consumer's status switches from user-visitor to user, although she may or may not make a purchase. Besides, we assume if a consumer is not active on the app for more than 1 day, the consumer's shopping session is finished. Since we have 3-day data, we use the data for consumers who are active only on the second day because we cannot track their shopping behaviors prior to the first day or after the

[2] Apptopia's data intelligence platform enables brands to analyze critical competitive signals and gain insights across mobile apps and connected devices.

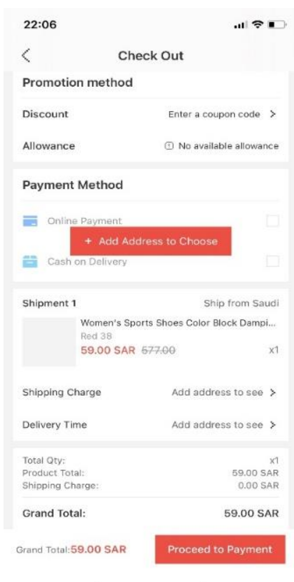
third day. Our final dataset includes 3,481,329 observations on consumers' activities from 23,278 consumers.



Checkout Page Before Login



Login/Register Page



Checkout Page After login

Figure 3. Screenshots of The Mobile App.

4. Empirical Analysis

4.1. Econometrics Models

We define the outcome variable *PURCHASE* as whether consumers have successfully checked out. Some check out attempts are not successful due to payment failure or unstable internet connections. PURCHASE is an indicator variable that equals to 1 if a consumer has successfully checked out, 0 otherwise.

Our goal is to examine the different effects of pre- and post-login search behaviors on consumer purchases. We create our variables of interest measuring consumer search behaviors as follows. First, we consider the time spent on product detail pages to measure consumer search efforts. Following previous studies (Bucklin and Sismeiro 2009; Moe 2003; Moe and Fader 2004), we exclude the time spent on administrative pages, such as account setting pages.[3] The more search time consumers spend on viewing products, the more search efforts consumers make. We calculate the time a consumer spent on viewing product detail pages before she logs in (*SEARCH_TIME_VISIT*) and that after she logs in (*SEARCH_TIME_USER*). Second, following Zhang et al. (2022), we use the numbers of clicks on goods detail page views per consumer to measure consumer search intensity for both pre- and post-login periods (*SEARCH_DEPTH_VISIT* and *SEARCH_DEPTH_USER*), respectively. Third, we measure consumer search diversity by calculating the number of unique goods views per consumer in two periods (*SEARCH_WIDTH_VISIT* and *SEARCH_WIDTH_USER*), respectively.

We control for product characteristics by including the average price of products (*AVGPRICE*), the number of brands customers have viewed (*BRAND*), the number of category level pages (*CATEGORY_VIEW*), the average number of unique products viewed per category (*PRODCAT*). All above control variables are calculated for pre- and post-login periods, separately. We also control consumer heterogeneity by including *NEWUSER*, an indicator variable that equals 1 if a customer registers a new account, and 0 otherwise. We test the VIF and there is no multicollinearity issue among variables. Table 1 presents the summary of variables.

[3] Some consumers have spent too long, for example, more than 1 hour, on one product detail page. It might be that these consumers have started to do something else or simply finished their shopping while keeping their shopping app on. The average time spent per product for user-visitors and users is 2 minutes and 12 minutes, respectively, thus we assume the maximum time a consumer can spend on a product page is 10 minutes.

Table 1. Summary of Variables (N = 23,278)

Variables	Definitions	Mean	SD	Min	Max
PURCHASE	=1 when consumer successfully checked out, and 0 otherwise	0.142	0.348	0	1
SEARCH_TIME_VISIT	Total time spent on product detail pages before login	176.895	799.753	1	26778
SEARCH_TIME_USER	Total time spent on product detail pages after login	742.947	1456.391	1	25717
SEARCH_DEPTH_VISIT	Number of clicks on goods detail page views before login	16.408	78.339	0	2465
SEARCH_DEPTH_USER	Number of clicks on goods detail page views after login	59.792	137.645	0	3522
SEARCH_WIDTH_VISIT	Number of unique goods views before login	5.526	20.904	1	634
SEARCH_WIDTH_USER	Number of unique goods views after login	19.115	37.354	1	714
AVGPRICE_VISIT	Average price of products on the detail pages a consumer has viewed before login	3.320	10.332	0	443
AVGPRICE_USER	Average price of products on the detail pages a consumer has viewed after login	17.635	23.563	0	1468
PRODCAT_VISIT	Average number of unique products viewed per category before login	1.222	5.008	0	139
PRODCAT_USER	Average number of unique products viewed per category after login	4.438	7.344	0	168
BRAND_VISIT	Number of brands a user-visitor has viewed before login	0.242	0.749	0	13
BRAND_USER	Number of brands a user-visitor has viewed after login	1.131	1.171	0	28
CATEGORY_VIEW_VISIT	Number of pages viewed that are category-level pages before login	2.348	16.277	0	708
CATEGORY_VIEW_USER	Number of pages viewed that are category-level pages after login	10.175	32.026	0	894
NEWUSER	=1 if a user-visitor registers a new account, and 0 otherwise	0.667	0.471	0	1

We construct Model 1 as shown in equation (1) to test our H1, H2a, and H2b:

$$\begin{aligned} PURCHASE_i = {} & \beta_0 + \beta_1 X\_VISIT_i + \beta_2 X\_USER_i \\ & + \gamma_1 AVGPRICE\_VISIT_i \\ & + \gamma_2 AVGPRICE\_USER_i \\ & + \gamma_3 PRODCAT\_VISIT_i \\ & + \gamma_4 PRODCAT\_USER_i \\ & + \gamma_5 BRAND\_VISIT_i \\ & + \gamma_6 BRAND\_USER_i \\ & + \gamma_7 CATEGORY\_VIEW\_VISIT_i \\ & + \gamma_8 CATEGORY\_VIEW\_USER_i \\ & + \gamma_9 NEWUSER_i + \varepsilon_i, \qquad (1) \end{aligned}$$

where i denotes the consumer who have both pre- and post-login search behaviors, and ε_i is the error term. We have three variations of Model 1 with the combination of *X_VISIT* and *X_USER* being (*SEARCH_TIME_VISIT*, *SEARCH_TIME_USER*), (*SEARCH_DEPTH_VISIT*, *SEARCH_DEPTH_USER*), and (*SEARCH_WIDTH_VISIT*, *SEARCH_WIDTH_USER*), respectively.

To test our H3a and H3b, we add the squared term for each independent variable and rewrite the equation (1) as equation (2) (Model 2):

$$\begin{aligned} PURCHASE_i = {} & \beta_0 + \beta_1 X\_VISIT_i + \beta_2 X\_USER_i \\ & + \beta_3 X\_VISIT_i^2 \\ & + \beta_4 X\_USER_i^2 \\ & + \gamma_1 AVGPRICE\_VISIT_i \\ & + \gamma_2 AVGPRICE\_USER_i \\ & + \gamma_3 PRODCAT\_VISIT_i \\ & + \gamma_4 PRODCAT\_USER_i \\ & + \gamma_5 BRAND\_VISIT_i \\ & + \gamma_6 BRAND\_USER_i \\ & + \gamma_7 CATEGORY\_VIEW\_VISIT_i \\ & + \gamma_8 CATEGORY\_VIEW\_USER_i \\ & + \gamma_9 NEWUSER_i + \xi_i. \qquad (2) \end{aligned}$$

Similarly, we have three variations of Model 2 with the combination of *X_VISIT*, *X_USER*, *X_VISIT*^{2}, and *X_USER*^{2} being (*SEARCH_TIME_VISIT*, *SEARCH_TIME_USER*, *SEARCH_TIME_VISIT*^{2}, *SEARCH_TIME_USER*^{2}), (*SEARCH_DEPTH_VISIT*, *SEARCH_DEPTH_USER*, *SEARCH_DEPTH_VISIT*^{2}, *SEARCH_DEPTH_USER*^{2}), and (*SEARCH_WIDTH_VISIT*, *SEARCH_WIDTH_USER*, *SEARCH_WIDTH_VISIT*^{2}, *SEARCH_WIDTH_USER*^{2}), respectively.

4.2. Empirical Results

We estimate Model 1 and Model 2 using logit model. Table 2 presents estimation results for Model 1. Results show that both pre- and post-login search efforts (including three measures) have statistically positive and significant impacts on consumers purchase. In column 1, the coefficient of *SEARCH_TIME _VISIT* is positive and significant ($\beta_1 = 0.00016, p < 0.001$), suggesting that if consumers increase their pre-login search time by 1,000 seconds (i.e., about 17 minutes), the probability of their purchase will increase by 16%. The coefficient of *SEARCH_TIME _USER* is also positive and significant ($\beta_2 = 0.00021, p < 0.001$), indicating that the likelihood of consumers' purchase will increase by 21% if they increase their post-login search time by 1,000 seconds. Similarly in column 2, the coefficients of *SEARCH_DEPTH _VISIT* and *SEARCH_DEPTH _USER* are both positive and significant ($\beta_1 = 0.0023, p < 0.001; \beta_2 = 0.00331, p < 0.001$). For *SEARCH_WIDTH* variable in column 3, we have similar results ($\beta_1 = 0.0038, p < 0.05; \beta_2 = 0.00456, p < 0.001$). These results are consistent with our expectation that not only post-login search, but also pre-login search activities have a positive and significant influence on consumers' purchase probability. Hence, our H1 is supported.

Table 2. Estimation Results for Model 1

	Model TIME	Model DEPTH	Model WIDTH
SEARCH TIME VISIT	0.00016***		
	(0.000)		
SEARCH TIME USER	0.00021***		
	(0.000)		
SEARCH DEPTH VISIT		0.00230***	
		(0.000)	
SEARCH DEPTH USER		0.00331***	
		(0.000)	
SEARCH WIDTH VISIT			0.00380*
			(0.002)
SEARCH WIDTH USER			0.00456***
			(0.001)
AVGPRICE VISIT	0.03178***	0.03352***	0.03170***
	(0.002)	(0.002)	(0.002)
AVGPRICE USER	0.00188*	0.00294***	0.00164*
	(0.001)	(0.001)	(0.001)
PRODCAT VISIT	0.02054***	0.01584**	0.02370***
	(0.005)	(0.005)	(0.005)
PRODCAT USER	0.01071***	0.001	0.01644***
	(0.003)	(0.003)	(0.003)
BRAND VISIT	0.26157***	0.21769***	0.27614***
	(0.039)	(0.041)	(0.043)
BRAND USER	0.06023**	0.005	0.10528***
	(0.020)	(0.021)	(0.022)
CATEGORY VIEW VISIT	0.00152	0.001	0.002
	(0.001)	(0.001)	(0.001)
CATEGORY VIEW USER	0.00244***	0.00149*	0.00233***
	(0.001)	(0.001)	(0.001)
NEWUSER	-0.00429	0.019	0.024
	(0.044)	(0.045)	(0.044)
INTERCEPT	-2.53473***	-2.49700***	-2.54656***
	(0.045)	(0.045)	(0.045)
Model Statistics			
N	23278	23278	23278
Log Likelihood	-8320.938	-8239.407	-8400.548

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

We use *t*-tests to test whether there is a significant difference between the effects of pre- and post-login search efforts. The test results (p = 0.377 for *SEARCH_TIME*; p = 0.1529 for *SEARCH_WIDTH*; p = 0.835 for *SEARCH_WIDTH*) cannot reject the null hypothesis $\beta_1 - \beta_2 = 0$, suggesting that the effects of pre-login search behaviors are at least as important as post-login search behaviors on consumers purchase. Neither Hypothesis 2a nor Hypothesis 2b is supported.

Table 3 shows the estimation results of Model 2. All three dimensions of search efforts (*SEARCH_TIME*, *SEARCH_DEPTH* and *SEARCH_WIDTH*) in both pre- and post-login periods have a statistically positive and significant impact on their purchase, which are consistent with our results from Model 1. The coefficients of all squared terms are negative and significant. For example, the coefficient of search diversity before login (*SEARCH_WDITH_VISIT*) is 0.01889, while the coefficient of the squared term (*SEARCH_WDITH_VISIT*^{2}) is -0.00006, indicating that the effect of pre-login search diversity on consumers purchase probability is increasing but at a diminishing rate. For post-login search behavior, we have similar results: the coefficients of *SEARCH_WDITH_USER* and *SEARCH_WDITH_USER*^{2} are $\beta_2 = 0.01975$ ($p < 0.001$) and $\beta_4 = -0.00006$ ($p < 0.001$), respectively. Hence, Hypotheses 3a and 3b are supported.

Table 3. Estimation Results for Model 2

	Model TIME	Model DEPTH	Model WIDTH
SEARCH TIME VISIT	0.00044***		
	(0.000)		
SEARCH TIME USER	0.00055***		
	(0.000)		
SEARCH TIME VISIT^2	-0.00000***		
	(0.000)		
SEARCH TIME USER^2	-0.00000***		
	(0.000)		
SEARCH DEPTH VISIT		0.00599***	
		(0.001)	
SEARCH DEPTH USER		0.00880***	
		(0.000)	
SEARCH DEPTH VISIT^2		-0.00000***	
		(0.000)	
SEARCH DEPTH USER^2		-0.00001***	
		(0.000)	
SEARCH WIDTH VISIT			0.01889***
			(0.002)
SEARCH WIDTH USER			0.01975***
			(0.001)
SEARCH WIDTH VISIT^2			-0.00006***
			(0.000)
SEARCH WIDTH USER^2			-0.00006***
			(0.000)
AVGPRICE VISIT	0.03018***	0.03250***	0.03040***
	(0.002)	(0.002)	(0.002)
AVGPRICE USER	0.00154	0.00309***	0.00149
	(0.001)	(0.001)	(0.001)
PRODCAT VISIT	0.01315**	0.00699	0.01236*
	(0.005)	(0.005)	(0.005)
PRODCAT USER	0.00401	-0.01214***	0.00592*
	(0.003)	(0.003)	(0.003)
BRAND VISIT	0.22322***	0.15060***	0.19158***
	(0.039)	(0.042)	(0.044)
BRAND USER	0.02315	-0.08386***	0.05408*
	(0.020)	(0.022)	(0.022)
CATEGORY VIEW VISIT	0.00077	0.000	0.0009
	(0.001)	(0.001)	(0.001)
CATEGORY VIEW USER	0.00230***	0.00145*	0.00273***
	(0.001)	(0.001)	(0.001)
NEWUSER	0.005	0.04965	0.05755
	(0.045)	(0.045)	(0.045)
INTERCEPT	-2.65809***	-2.63090***	-2.70856***
	(0.047)	(0.047)	(0.047)
Model Statistics			
N	23278	23278	23278
Log Likelihood	-8181.488	-7969.801	-8235.060

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

5. Conclusion and Future Research

In this paper, we examine the effect of consumer pre-login and post-login search behavior on consumer purchase. Our results extend the burgeoning literature on consumer behavior in mobile commerce and offer novel insights in this field. Previous studies neglect the effect of pre-login search on consumers' purchase probability. We empirically show that pre-login search behavior is as important as post-login search behavior. Furthermore, our results show that choice overload can pose a negative effect (i.e., cause anxiety, disappointment, and demotivation) on consumer purchase decisions so that they may give up their purchase. Less accurate RSs will not present consumers with too many similar and comparable products, which to some extent, alleviates consumers anxieties for making a choice. That is why the influence of users' search efforts on purchase is not statistically greater than that of user-visitors' search.

Our findings have important managerial implications. The importance of pre-login search behavior can make a difference in RS design. For example, should the algorithm of RSs be as accurate as possible? How does a manager balance the accuracy and diversity of RSs? A tradeoff between accuracy and diversity is a heated debate in the RS field (e.g. Liang et al. 2006; Adomavicius and Kwon 2012). Chasing more accuracy in the RSs may not always be optimal for companies. It not only requires a large investment, but also incurs a cost which could narrow down the diversity of products recommended for consumers. Therefore, companies should carefully balance between the accuracy and the diversity of RSs. Besides, too many recommended similar items may cause stress for consumers when making a choice. Managers should take this potential psychological effect into account and design an adaptive algorithm of RSs.

There are several limitations for this paper and a few future research directions we can explore. First, we only have three-day data. We do not observe long-term effect of pre- and post-login consumer search behavior on purchase. Future research can obtain panel data to examine the long-term effects. Second, although we study the relationship between user-visitors' search behavior and their purchase probability based on the difference of the recommender system before and after login, we could not empirically examine the performance of RSs directly due to our data limitation. Empirical analysis on such difference of RSs performance would be an interesting avenue for future research.

References

Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In Recommender systems handbook (pp. 217-253). Springer, Boston, MA.

Adomavicius, G., & Kwon, Y. (2012). Improving aggregate recommendation diversity using ranking-based techniques. IEEE Transactions on Knowledge and Data Engineering, 24(5), 896-911.

Adomavicius, G., & Kwon, Y. (2014). Optimization-based approaches for maximizing aggregate recommendation diversity. INFORMS Journal on Computing, 26(2), 351-369.

Ahmad, Z., & Ibrahim, R. (2017). Mobile commerce (M-commerce) interface design: A review of literature. J. Comput. Eng, 19, 66-70.

Beatty, S. E., Homer, P., & Kahle, L. R. (1988). The involvement—commitment model: Theory and implications. Journal of Business research, 16(2), 149-167.

Bettman, J. R. (1979). Information processing theory of consumer choice. Addison-Wesley Pub. Co..

Bhatnagar, A., & Ghose, S. (2004). An analysis of frequency and duration of search on the Internet. The Journal of Business, 77(2), 311-330.

Bobadilla, J., Ortega, F., Hernando, A., & Alcalá, J. (2011). Improving collaborative filtering recommender system results and performance using genetic algorithms. Knowledge-based systems, 24(8), 1310-1316.

Bradley, K., & Smyth, B. (2001). Improving recommendation diversity. Proceedings of the twelfth Irish conference on artificial intelligence and cognitive science, Maynooth, Ireland Vol. 85, pp. 141-152.

Brynjolfsson E, Hu Y, Simester D (2011) Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. Management science, 57(8):1373–1386.

Bucklin, R.E., Lattin, J.M., Ansari, A., Bell, D., Coupey, E., Gupta, S., Little, J.D.C., Mela, C., Montgomery, A., and Steckel, J. (2002) Choice and the Internet: From clickstream to research stream. Marketing Letters, 13, 3, 245–258.

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

Bucklin, R. E., & Sismeiro, C. (2009). Click here for Internet insight: Advances in clickstream data analysis in marketing. Journal of Interactive marketing, 23(1), 35-48.

Chhonker, M. S., Verma, D., & Kar, A. K. (2017). Review of technology adoption frameworks in mobile commerce. Procedia computer science, 122, 888-895.

Fleder, D., & Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. Management science, 55(5), 697-712.

Ghose, A., & Han, S. P. (2014). Estimating demand for mobile applications in the new economy. Management Science, 60(6), 1470-1488.

Han, S. P., Park, S., & Oh, W. (2016). Mobile App Analytics. MIS quarterly, 40(4), 983-1008.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? Journal of personality and social psychology, 79(6), 995.

Johnson, E. J., Moe, W. W., Fader, P. S., Bellman, S., & Lohse, G. L. (2004). On the depth and dynamics of online search behavior. Management science, 50(3), 299-308.

Kaminskas, M., & Bridge, D. (2016). Diversity, serendipity, novelty, and coverage: a survey and empirical analysis of beyond-accuracy objectives in recommender systems. ACM Transactions on Interactive Intelligent Systems (TiiS), 7(1), 1-42.

Kinjo, K., & Ebina, T. (2015). Paradox of choice and consumer nonpurchase behavior. Ai & Society, 30(2), 291-297.

Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems–A survey. Knowledge-based systems, 123, 154-162.

Lee, D., Gopal, A., & Park, S. H. (2020). Different but equal? A field experiment on the impact of recommendation systems on mobile and personal computer channels in retail. Information Systems Research, 31(3), 892-912.

Li, S. S., & Karahanna, E. (2015). Online recommendation systems in a B2C E-commerce context: a review and future directions. Journal of the Association for Information Systems, 16(2), 2.

Li, X., Grahl, J., & Hinz, O. (2021). How Do Recommender Systems Lead to Consumer Purchases? A Causal Mediation Analysis of a Field Experiment. Information Systems Research.

Liang, C., Shi, Z., & Raghu, T. S. (2019). The spillover of spotlight: Platform recommendation in the mobile app market. Information Systems Research, 30(4), 1296-1318.

Lin, J., Wang, B., Wang, N., & Lu, Y. (2014). Understanding the evolution of consumer trust in mobile commerce: a longitudinal study. Information Technology and Management, 15(1), 37-49.

Linden, G., Smith, B., & York, J. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. IEEE Internet computing, 7(1), 76-80.

Lynch Jr, J. G., & Srull, T. K. (1982). Memory and attentional factors in consumer choice: Concepts and research methods. Journal of consumer research, 9(1), 18-37.

McGinty, L., & Smyth, B. (2003, June). On the role of diversity in conversational recommender systems. International Conference on Case-Based Reasoning (pp. 276-290). Springer, Berlin, Heidelberg.

McSherry, D. (2002, September). Diversity-conscious retrieval. European Conference on Case-Based Reasoning (pp. 219-233). Springer, Berlin, Heidelberg.

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

Moe, W. W., & Fader, P. S. (2004). Dynamic conversion behavior at e-commerce sites. Management Science, 50(3), 326-335.

Panniello, U., Gorgoglione, M., & Tuzhilin, A. (2016). Research note—In CARSs we trust: How context-aware recommendations affect customers' trust and other business performance measures of recommender systems. Information Systems Research, 27(1), 182-196.

Park, Y., Bang, Y., & Ahn, J. H. (2020). How Does the Mobile Channel Reshape the Sales Distribution in E-Commerce?. Information Systems Research, 31(4), 1164-1182.

Pathak, B., Garfinkel, R., Gopal, R. D., Venkatesan, R., & Yin, F. (2010). Empirical analysis of the impact of recommender systems on sales. J. of Management Information Systems, 27(2), 159-188.

Patra, B. K., Launonen, R., Ollikainen, V., & Nandi, S. (2015). A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. Knowledge-Based Systems, 82, 163-177.

Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56-58.

Schwartz, B. (2004). The paradox of choice: Why more is less. New York: Ecco.

Silveira, T., Zhang, M., Lin, X., Liu, Y., & Ma, S. (2019). How good your recommender system is? A survey on evaluations in recommendation. International Journal of Machine Learning and Cybernetics, 10(5), 813-831.

Sun, T.,Shi, L.,Viswanathan,S., & Zheleva,E.(2019).Motivating effective mobile app adoptions: Evidence from a large-scale randomized field experiment. Information Systems Research, 30(2), 523-539.

Traylor, M. B. (1984). Ego involvement and brand commitment: not necessarily the same. Journal of consumer marketing.

Vargas, S., & Castells, P. (2011). Rank and relevance in novelty and diversity metrics for recommender systems. Proceedings of the fifth ACM conference on Recommender systems (pp. 109-116).

Wu, S., Tan, Y., Chen, Y., & Liang, Y. (2022). How Is Mobile User Behavior Different? —A Hidden Markov Model of Cross-Mobile Application Usage Dynamics. Information Systems Research.

Xiao, B., & Benbasat, I. (2007). E-commerce product recommendation agents: Use, characteristics, and impact. MIS quarterly, 137-209.

Zhang, "" Jennifer"" Jie, Xiao Fang, Olivia R Liu Sheng. 2006. Online consumer search depth: Theories and new findings. Journal of Management Information Systems, 23(3) 71-95

Zhang, M., & Hurley, N. (2008). Avoiding monotony: improving the diversity of recommendation lists. Proceedings of the 2008 ACM conference on Recommender systems (pp. 123-130).

Zhang, T.X., Agarwal, R., & Lucas Jr, H. C. (2011). The value of IT-enabled retailer learning: personalized product recommendations and customer store loyalty in electronic markets. Mis Quarterly, 859-881.

Zhang, X.Y., Cui, R.M., & Yao, Y.L. (2022). The Version Effect of Apps and Operating Systems in Mobile Commerce. (forthcoming)