

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/271373754>

Learning User Real-Time Intent for Optimal Dynamic Web Page Transformation

Article in *Information Systems Research* · January 2015

DOI: 10.1287/isre.2015.0568

CITATIONS

43

READS

1,728

3 authors:



Amy Wenxuan Ding

emlyon business school

33 PUBLICATIONS 137 CITATIONS

[SEE PROFILE](#)



Shibo Li

Indiana University Bloomington

33 PUBLICATIONS 1,326 CITATIONS

[SEE PROFILE](#)



Patrali Chatterjee

Montclair State University

35 PUBLICATIONS 2,797 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Humanity-enriched computing [View project](#)

Learning User Real-Time Intent for Optimal Dynamic Webpage Transformation¹

Amy Wenxuan Ding

Shibo Li

Patrali Chatterjee

(Accepted and forthcoming: *Information Systems Research*)

¹ The authors thank comScore Networks for its generous contribution of data, without which this research would not have been possible.

Learning User Real-Time Intent for Optimal Dynamic Webpage Transformation

Abstract

Many e-commerce websites struggle to turn visitors into real buyers. Understanding online users' real-time intent and dynamic shopping cart choices may have important implications in this realm. This study presents an individual-level, dynamic model with concurrent optimal page adaptation that learns users' real-time, unobserved intent from their online cart choices, then immediately performs optimal webpage adaptation to enhance the conversion of users into buyers. To suggest optimal strategies for concurrent page adaptation, the model analyzes each individual user's browsing behavior, tests the effectiveness of different marketing and web stimuli, as well as comparison shopping activities at other sites, and performs optimal webpage transformation. Data from an online retailer and a laboratory experiment reveal that concurrent learning of the user's unobserved purchase intent and real-time, intent-based optimal interventions greatly reduce shopping cart abandonment and increase purchase conversions. If the concurrent, intent-based optimal page transformation for the focal site starts after the first page view, shopping cart abandonment declines by 32.4%, and purchase conversion improves by 6.9%. The optimal timing for the site to intervene is after three page views, to achieve efficient learning of users' intent and early intervention simultaneously.

Keywords: Learning, shopping intent, optimization, concurrent page adaptation, website productivity, hierarchical Bayes models, hidden Markov models

1. Introduction

To increase their revenues and profits, many retail websites use advanced information technologies, such as collaborative filtering, intelligent recommendations, customization with machine learning, and so forth. Such technologies attempt to attract, retain, and convert users into buyers (Albert et al. 2004; Cho et al. 2002; Retail Week 2010). Yet adopting such conventional technologies apparently converts only about 3.96% of users; even the most successful site converts only 8% of online visitors into paying customers (eMarketer 2009; Forrester Research 2012). Thus, how to increase online conversion rates is a difficult task for academics and practitioners (Palmer 2002; Savitz 2011; Straub and Watson 2001). Industrial and academic studies suggest that website design can become a major driver of profits, and exposures to irrelevant marketing or web stimuli and lack of understanding of users' shopping intent and behavior can distract consumers from their shopping goals (eMarketer 2009; Rajamma et al. 2009). Jeff Bezos, CEO of Amazon.com, said that "if I have 3 million customers on the Web, I should have 3 million stores on the web". This suggests that how to effectively generate real-time dynamically customized webpages for each customer is of utmost interest to online managers.

Currently many websites that offer customized information and/or recommendations to users rely on users' historic information and preferences, and provide mass customization assuming that users' preferences and browsing behaviors are static. For example, Amazon.com exploits users' purchase histories to suggest similar products, as well as items selected by other users with similar buying behavior, with the implicit assumption that users with similar buying records like the same products and share similar product interests. However, users could change their intentions over time or in response to stimuli and information they encounter while browsing during a website visit. This dynamic property of intentions is often ignored by conventional recommendation or customization methods (e.g., collaborative filtering, data mining for content-based filtering). Moreover, conventional customization and recommendation models usually provide product recommendations on the side or bottom of a focal page without considering (1) performing dynamic content changes on the page, (2) whether product recommendations meet each user's unobserved purchase intent, and (3) whether such recommendations

are optimal in terms of increasing purchase conversion while reducing shopping cart abandonment. Although users' navigation goals, changing intentions, desire to comparison shop, and reactions to marketing and web stimuli may influence their decisions on shopping cart choices, information systems (IS) and computer science literature has not addressed these issues, particularly on the impact of individual level real-time dynamic and optimal webpage transformation (eMarketer 2009; Parboteeah et al. 2009). This study seeks to add to extant literature by investigating learning user real-time intent for dynamic and optimal page transformation to improve the website performances.

In this paper, we propose a new individual-level, dynamic learning model that first observes individual users' shopping cart choices and how they navigate during the course of their site visits to infer each user's unobserved real-time intent. Then the model automatically performs optimal intent-based webpage transformations for the next page before the user exits the site in order to increase purchase conversion while reducing cart abandonment rates. We show that having an individual level optimal dynamic webpage transformation during a user's visit can greatly improve the site performance.

Specifically, we note that users visiting a retailer's website have different goals or intentions (Gollwitzer 1999; Moe and Fader 2004). Some engage in pure browsing or compare competing products, then end the session without using the shopping cart at all. Others enter the website to look for a product and place items in shopping carts, but a subset of these users abandon those carts without purchase, and the remainder completes a purchase before exiting the site. The retailer lacks a priori knowledge of a user's real-time intent during the course of the visit to the website, because those intentions are not directly observable. We propose a learning model to infer users' unobserved, real-time intent from the series of activities that they perform online, to understand shopping cart choices, and then perform optimal webpage transformation for the next page, in line with the stimulus-organism-response (S-O-R) framework (Donovan and Rossiter 1982), as used in environmental psychology (Meharabian and Russell 1974). We assert that retail environmental stimuli affect consumers' emotional states, prompting approach and avoidance behaviors (e.g., Baker 1986; Bitner 1992; Spangenberg et al. 1996). Eroglu et al. (2001) shows that an online store's environmental stimuli S (on shoppers' computer screens) influence

affective and cognitive internal states O, which then alter various shopping outcomes R. In this paper, we apply a similar S-O-R paradigm to improve website performances. Specifically, we examine a user's shopping cart choices (R) and perform reverse reasoning to infer the user's shopping intent states (O) with a hidden Markov model, given that we cannot directly observe O but know R and can manage S. Researchers already have demonstrated that internal mind states determine behavior, such that O is an immediate antecedent of response or actions (e.g., Ajzen 1991; Davis et al. 1989; Mathieson 1991; Pavlou and Gefen 2004; Sheppard et al. 1988; Shimp and Kavas 1984). Thus, after identifying the user's intent state O, the proposed model implements optimal page transformation immediately by adjusting the marketing and web stimuli (S) on the next page to influence O to generate positive outcomes R.

As currently the contents on most websites are static and predesigned to appeal to a wide user base, we propose an individual-level model system with a theoretical foundation and mathematical descriptions to show that employing real-time, dynamic, optimal page transformation that reflects a user's intent leads to more positive outcomes, including higher purchase conversion rates and lower shopping cart abandonment rates. In particular, our model system automatically makes real-time page transformations which are dynamic and optimal such that it maximizes the user's probability of placing item(s) into his/her shopping cart if there is no item in the cart, and the probability of making a purchase for those with items in the cart. Our test of the model includes both an empirical investigation using real-world data and a small laboratory experiment. Purchase conversion and shopping cart abandonment rates offer the criteria for measuring the model's effectiveness. We follow environmental psychology (Mehrabian and Russell 1974) to identify stimuli that are likely to affect behavior and thus examine the effectiveness of various marketing and web stimuli, users' comparison shopping activities, and past purchase and browsing behaviors. The empirical and laboratory results affirm that our proposed model effectively differentiates each user according to her or his real-time intent. In doing so, it can reduce shopping cart abandonment by 32.4% and improve purchase conversion by 6.9%, if the retailer initiates optimal page adaptation immediately after the first page view. For the site we study, the optimal timing for an

intervention is after three page views, at which point, it can achieve efficient learning of user intents and early intervention simultaneously.

Our research thus contributes to IS literature and practices. Retailers need a means to recognize when shopping carts are abandoned, to be able to show users appropriate contents and offers automatically (Forrester Research 2012). Methodologically, our model can automatically analyze individual user's browsing behavior and cart choices, and perform concurrent learning and optimal page transformation, representing a new approach to generating intent-based site dynamism. Theoretically, we extend the S-O-R framework by using backward reasoning with dynamic learning. Managerially, site managers can use the proposed approach to improve website performances without the need of advanced tracking systems (e.g., video devices, eye tracking, fMRI) to infer users' mindset changes. Substantively, our results suggest that site managers should provide users with concurrent, intent-based dynamic contents before they exit the site. We show that simple observation data (i.e., navigation paths and shopping cart choices) can effectively support this effort. To the best of our knowledge, the proposed model system is the first in IS and Computer Science literature as well as the online retailing industry to realize a real-time learning and individual level concurrent optimal page transformation before the user exits the site.

In Section 2, we review related literature and present our theoretical model, with a high-level framework to provide an overview and clarify our contributions. We then present a detailed mathematical description of the proposed theoretical model, including system components for concurrently learning users' real-time intent with dynamic page adaptation. In Section 4, we present our empirical tests, report the results from simulations with a real-world data set and a lab experiment, and discuss some managerial implications. Finally, we conclude with some limitations and research directions in Section 5.

2 Theoretical Development

Our research setting includes two types of players: users and retail websites. Research from various disciplines offers insights into user online shopping behavior and website profitability, including marketing (focused on the user side), computer science (website side), and IS (user and website sides).

2.1 User Behavior

Research in IS and marketing on user online shopping comprises two broad and interdependent streams, focused on how users respond to stimuli encountered in the shopping environment or how users' shopping goals drive their search behavior. Website characteristics, such as product presentations, web page formatting, and usability, can influence users' perceptions, shopping decisions, and purchases (Cho et al. 2006; Currim et al. 2006; Everard and Galletta 2006; Jiang and Benbasat 2007; Koufaris 2002; Mandel and Johnson 2002; Mithas et al. 2007; Palmer 2002; Parboteeah et al. 2009; Song and Zahedi 2005; Tam and Ho 2005), as can the price, promotions, and product information provided online (Eroglu et al. 2001; Ratchford et al. 2003; Viswanathan et al. 2007; Zettelmeyer et al. 2006). Most studies indicate that users have certain shopping goals in mind that drive their navigation behavior during the site visit (Bucklin and Sismeiro 2003; Lee and Ariely 2006; Moe 2003; Moe and Fader 2004; Montgomery et al. 2004; Novak et al. 2003; Putsis and Srinivasan 1994; Sismeiro and Bucklin 2004). Thus users might conduct a goal-directed search, to browse, compare, or purchase, or they might just enjoy experiential browsing (Hoffman and Novak 1996; Nadkarni and Gupta 2007; Novak et al. 2003). Furthermore, prior studies consider the paths consumers take across websites (Johnson et al. 2004; Park and Fader 2004) or user behavior within a website; among this latter set, some examine search within a session (Moe 2003; Sismeiro and Bucklin 2004; Ramachandran et al. 2010), whereas others model sessions over time (Moe and Fader 2004). Most studies suggest using a combination of measures related to the breadth, depth, and intensity of search within or across sessions to differentiate the underlying intentions that drive user behavior (Hoffman and Novak 1996; Moe 2003; Wolfenbarger and Gilly 2001). For example, Montgomery et al. (2004) provides a method to demonstrate that a user's browsing path can be used to capture the user's intention and predict the next webpage category that the user might like to view. Following this direction, Ramachandran et al. (2010) investigate how the breadth, depth, and intensity of user navigation search about product category information within a session can be used to capture users' unobserved shopping goals.

However, most of these studies also tend to assume that a user's goal remains static throughout the shopping process, even though Mandel and Johnson (2002) show that users dynamically adapt their

behaviors to the page-by-page stimuli they see, without being consciously aware of this behavior. Page content, hyperlinks, and marketing stimuli can cause interruptions, diversions, or abandonment of original shopping goals. Because we consider online users' unobserved intentions as dynamic across an optimal number of states (to be determined empirically for each context), both within and across sessions and in response to marketing and web stimuli on each page encountered during a visit to a retailer's website, we propose a hidden Markov model. Meanwhile, we conduct a page-level analysis and examine each individual user's online shopping cart choices to reverse infer the user's unobserved purchase intent state, which was then used to recommend dynamic and optimal changes on the marketing and web stimuli on the next page to increase purchase conversion while reducing cart abandonment, which cannot be performed by the models proposed in both Montgomery et al. (2004) and Ramchandran et al. (2010).

Though Montgomery et al. (2004) also uses a hidden Markov model to capture the user's intention, our study differs substantively from it in terms of research focus, model development, data range and the level of data analysis. First, the two papers' research focuses are different. Montgomery et al. (2004) aims to show whether a user's browsing path can be used to capture the user's intention so as to predict the next webpage category that the user might like to view. Our research continues in this direction but focuses on exploring the dynamics of a user's real-time intention (as people's minds may change) based on the user's footprints (browsing path and shopping cart activities) and how such intention can be influenced by environmental stimuli such that concurrent optimal page transformation can be performed on the next page to increase the website performances.

Second, the models are different. Though both papers have deployed HMM models, Montgomery et al. (2004) assumes users' unobserved intention are stationary; thus their homogenous HMM model cannot capture the non-stationary dynamics of user task goals. In our paper, however, our model with a heterogeneous HMM not only captures the dynamics of the user's intent but also identifies

an optimal policy for realizing real-time webpage transformation. So our model includes both dynamic intention identification module and optimization module for concurrent page transformation.

Third, despite using the similar dataset, the data ranges used in these two papers are somewhat different. Montgomery et al. (2004) only uses users' browsing path and webpage content information. In our paper, however, as we capture the dynamics of user intent, we augment the dataset with additional detailed shopping cart activity information extracted from the URLs. Specifically, we use not only webpage content information (such as web and marketing stimuli) and the user's footprints on visiting the website such as browsing path, but also a complete set of the user's shopping cart activities during the course of the site visit.

Fourth, The level of data analysis is different. The dependent variable in Montgomery et al. (2004) is at the web page category level (i.e., information page, product page, etc.). They neither consider users' shopping cart activities nor the impact of marketing and web stimuli as covariates on the evolution of individual user unobserved intent, as examined in our paper. So the model proposed by Montgomery et al. (2004) cannot capture the dynamics of users' shopping cart choice decisions. In our paper, we analyze each individual user's shopping cart choices and investigate how environmental stimuli influence user unobserved intent. On the basis of this analysis, we build an optimization module to transform the next webpage concurrently. To further test our proposed model, we also conduct a lab experiment. Hence, our study has broader managerial implications.

Therefore, unlike all previous research, our research captures the real-time dynamics of a user's latent intention with optimal webpage transformation, and performs remedial actions to increase the site performance during the session while the user remains on the site, instead of trying to win the user back after abandonment.

2.2 Website Customization/Personalization

Many retail websites rely on customization, recommendation, or personalization technologies to increase their profitability (Ansari and Mela 2003). One method relies on users' explicit inputs, such as registration information, search terms, or user-specified preferences, to produce customized information and products (Bolin et al 2005; Kobsa et al 2001; Schafer et al. 2001). Another approach is the content-based filtering method, which uses machine learning techniques to construct a user's preferences by examining navigation across different websites that contain specific items and organizing them by similarity. This interface constructs customized information on the basis of the preferences exhibited by users' past purchase behaviors (Passani and Billsus 1997; Ricci 2002). A third category of collaborative filtering systems searches for commonalities among preferences expressed by different users and provides customized pages by analyzing similar preferences (Herlocker et al. 2004). Regardless of the approach adopted, performance evaluations usually focus on whether the system can find items or retrieve products that a user is likely to evaluate positively, rather than considering purchase conversion or shopping cart abandonment issues.

However, online consumers increasingly worry about revealing personal information when they log in, for fear it may be misused. Interfaces that anticipate users' needs on the basis of their profiles or registration information, then provide them with customized information, often trigger concerns about privacy. In addition, the content-based approach relies on a user's historical behaviors, leaving no room for variability or changes in the user's interests and goals across sessions. With collaborative filtering technology, the interface generates recommendations on the basis of common patterns across similar users, such that it fails to capture any individual user's unique preferences. Similarly, using historical records in collaborative filtering may not reflect users' current online visiting interests; nor can simply noting a user's preferences explain why and when he or she might abandon a cart. These conventional methods thus ignore the dynamics of users' unobserved shopping intentions, whereas we propose an individual-level model that relies on each user's initial navigation information in the current visit to infer unobserved intent and then generate a corresponding, intent-based, optimal page adaptation.

Because what a user seeks depends on who the user is, and site designers cannot predict with full accuracy users' shopping intent, the site must recognize real-time intentions quickly to generate corresponding optimal marketing and web stimuli on successive page views immediately, before the user exits, rather than aggregating experiences of many users over time to present a single version of the site to various users.

2.3 Proposed Theoretical Framework

Our theoretical framework is based on and extends the S-O-R model in environment psychology, which suggests that an environmental stimulus S influences cognitive internal states O, which then affect response behavior R (Mehrabian and Russell 1974). This paradigm is well-established, applied and validated in many studies in consumer psychology, marketing and IS that investigate both offline and online consumer behaviors (e.g., Adelaar et al. 2003; Jacoby, 2002; Jarboe and McDaniel 1987; Koufaris et al. 2001; Parboteeah et al 2009). In the offline setting, several studies suggest that retail environmental stimuli in physical store context impact consumers' emotional states, which then result in approach or avoidance behaviors toward the store (Bitner 1992; Donovan and Rossiter 1982; Mehrabian and Russell 1974). Recent research furthermore examines the influence of multiple retail atmospheric cues on consumer responses based on the S-O-R framework (Hulten, 2012; Mattila and Wirtz, 2001; Parsons, 2011; Spangenberg et al., 2005). In the online setting, Eroglu et al (2001) find that atmospheric cues of online stores, through the intervening effects of affective and cognitive states, influence the outcomes of online retail shopping. Mazaheri et al (2012) show that consumers' emotions influence their perceptions of site atmospheric cues, which, in turn, impact consumers' site attitudes and purchase intention. Parboteeah et al. (2009) demonstrate task and mood-related cues impact a consumer's impulse purchase online. According to S-O-R paradigm, S arouses the individual; O represents the individual's mindset, characterized as affective and cognitive states; and R represents the individual's approach or avoidance behaviors. Approach behaviors entail positive actions directed toward a particular setting, such as intentions to stay or explore; avoidance behaviors are the opposite (Donovan and Rossiter 1982; Mehrabian and Russell 1974; Sherman and Smith 1987).

Applying this framework to our setting, we posit that users see various stimuli on each web page, some of which are highly relevant to their shopping intentions (e.g., product information, product pictures, pop-up promotions, banner ads), but others may have low relevance or provide a distraction. Thus we can classify the stimuli S into two broad categories: internal and external (Eroglu et al. 2001). The former can be controlled by the retailer, but the latter is beyond its control. According to Montgomery et al. (2004), internal stimuli consist of marketing stimuli (e.g., price, pop-up promotions, banner ads, e-mail solicitation) and web stimuli (e.g., hypertext links, pictures) encountered at the retailer website. Kotler (1974, p. 50) suggests that most purchases are unplanned and that pricing promotions, displays, and packaging information can be designed “to produce specific emotional effects in the buyer that enhance his purchase probability.” User sensitivity to these cues increases with experiential goals, such that the impact of stimuli on purchase intent may be to accelerate or divert the shopping process and purchase. For example, product-related price information may facilitate choice decisions (Moe 2006); an offer of free shipping with a minimum purchase may motivate a user to consider put an additional item into the shopping cart to take advantage of the offer; and pop-up ads may yield significantly higher ad perceptions, click-through rates, and purchase intentions than banner ads (Chatterjee et al. 2003; Cho et al. 2001; Diao and Sundar 2004; Manchanda et al. 2006). To control for the impact of the user’s past purchase and browsing behavior and individual heterogeneity, we also incorporate two categories of variables (Sismeiro and Bucklin 2004): user’s past behavior (visit depth, time spent on last page view, sign-in or not, more items in the cart to earn free shipping, visit during the weekend, made a purchase in last session) and demographic information (age, gender, education, income level). Finally, we define the user’s internal state (O) as his or her latent purchase intention: A greater intention state indicates a stronger purchase orientation. The user’s behavioral responses R include online shopping cart choices, such as continuing to browse without changing the cart, removing or adding items to the cart, purchase, or exiting the site.

In applying the S-O-R framework to predict behavioral responses, previous research has taken a forward reasoning approach, $S \rightarrow O \rightarrow R$, such that it investigates S to find how it affects O , to generate the

corresponding R (Donovan and Rossiter 1982; Eroglu et al. 2001; Sherman and Smith 1987). In contrast, we apply backward reasoning and propose that a person's action reflects her or his unobserved thoughts (Ding 2003; Simon 1957). Because $O \rightarrow R$, we take a user's R as inputs (acts performed during an online visit) to infer unobserved intention O. We build an identification module to analyze each user's shopping cart choices at the page level, so that we can perform this backward reasoning.

Figure 1 contains our framework, where the dashed block represents the core part, $O \rightarrow R$, and the solid arrows are the components of the proposed model system. We use an observation module to capture what each user actually does, page view by page view. The user's internal intent state is unobservable, so the identification module uses the outputs of the observation module to provide an estimate. Our ultimate concern is shopping cart abandonment and purchase conversion; therefore, we check whether our identification module can correctly estimate each user's intent state and predict shopping cart choices, including whether he or she intends to purchase, where he or she abandons the shopping cart, or if he or she engages in pure browsing before exiting the site.

<Insert Figures 1 and 2 Here>

Beyond measuring intentions, retailers hope to encourage users to place items into shopping carts and then make purchases. Therefore, it is necessary to discern how users update their intentions before exiting the site, because intention determines behavior. Furthermore, the stimuli encountered while browsing the site likely influence intentions. Therefore, we also develop a transformation module to make optimal changes to the internal marketing and web stimuli that appear on subsequent pages, before the user exits the site, to alter his or her intentions and encourage actions leading to positive outcomes such as a purchase. In the effort to reduce cart abandonment and increase purchase conversion, the transformation module allows a website interface to optimally reflect concurrent page adaptations to its internal marketing and web stimuli, tailored to each user's real-time intent.

Our theoretical model thus comprises three components: observation, identification, and transformation modules. The core logic for our backward reasoning to infer users' behavioral intentions is depicted in Figure 2. We implement the model as a dynamic adaptive system that enables a retailer to

dynamically improve its own website: It observes each user's shopping/browsing behavior at a page view level (e.g., up to page view t), then performs real-time, rapid learning to identify the user's unobserved intent state ($s = 1, \dots, S$), before automatically displaying contents that optimally match the user's intent by changing the internal marketing and web stimuli at page view $t + 1$.

In extending the S-O-R paradigm, our model captures the dynamics of the user's latent intention over time, as well as infers latent intentions from the actions that the user performs through backward reasoning with dynamic learning. The stimuli we include in the model are ones that any site manager can control, as evidenced in IS and marketing literature. In addition, our dynamic and optimal page adaptation aims to generate desirable outcomes; previous applications of the S-O-R paradigm have often left the consequence of a behavior or response undetermined. Finally, our integrated theoretical model uses direct observations to uncover the dynamics of a user's unobserved intent and generate optimal page transformation for great site performances.

3. Model of Learning Users' Real-Time Intent with Concurrent Optimal Webpage Adaptation

3.1 Observation Module

The observation module records complete visits to a website, including the sequence of pages visited and the links followed, such that it generates a variety of observations during a session. We define a session as a period of interactive information interchange between the user's computer and the web server, so a session starts with web browsing and ends after 20 minutes of inactivity. If a user visits a website and has not clicked any links or viewed any page for 20 minutes, we assume that the viewing session has ended, and the next page view marks the beginning of a new session.

Each observation contains information about the individual user's shopping cart choice behavior, contents of the pages the user views, comparison shopping activities at other sites, and the user's past purchase and browsing behavior, as indicated in the theoretical development section. The page content includes the marketing and web stimuli. Many users do not enter the site's home page directly, because they type a search query or follow external or bookmarked links pointing to the middle of the site, so in addition to recording access counts for pages, this module tracks each user's navigation through the site.

We thus determine where the path begins and analyze precisely what the user does or sees, as well as where he or she goes next.

3.2 Identification Module

As mentioned, we seek to infer each user's real-time intent from shopping cart behavior (i.e., infer O from R). Thus, our identification module uses the outputs of the observation module as inputs to analyze the user's cart choices. From its inference of the user's unobserved intent, the module then makes predictions about the possible outcomes of visiting the website: pure browsing and exit, adding item(s) to the cart but abandoning, or completing a purchase.

3.2.1. Analyzing Shopping Cart Choices

A user can do many things during the shopping process, but decisions related to the shopping cart determine whether a purchase can occur. Therefore, to capture shopping cart choice behavior, in the identification module, we denote the cart choice C_{iqt} ($= 1$ for exit, $= 2$ for browsing without changing the shopping cart, $= 3$ for removing items from the shopping cart, $= 4$ for adding items to the cart, and $= 5$ for purchase) for user i at page view t in session q . According to Simon (1957), intentions determine behavior, and we assume that user i is rational, such that he or she has latent utility V_{ijqt} associated with choice j at page view t in session q . Therefore, in our observational equation,

$$C_{iqt} = j \text{ if } \mathbf{1}'_{ijqt} = 1 \text{ and } V_{ijqt} = \max(\mathbf{V}_{iqt} < \mathbf{1}'_{iqt} >), \quad (1)$$

$\mathbf{V}_{iqt} = [V_{i1qt}, V_{i2qt}, \dots, V_{i5qt}]$ is a 5×1 vector of latent utilities, $\mathbf{1}'_{ijqt}$ is an indicator variable that equals 1 when choice j is available for user i during the t^{th} page view of session q , and $\mathbf{V} < \mathbf{1}' >$ denotes the set of elements from the vector \mathbf{V} whose corresponding indicator operand ($\mathbf{1}'$) is equal to 1. In our identification module, $\mathbf{1}'_{iqt}$ is a vector of values equal to 1, with the exception of the two elements that correspond to deleting items from the cart and purchasing, both of which equal 1 when there is at least one item in the cart, and a third element that refers to adding an item to the cart, which equals 1 when products are available on the web page. With $\mathbf{1}'_{iqt}$, we can eliminate impossible choices (e.g., removing items from an

empty cart; Montgomery et al. 2004). Similar to Bucklin and Sismeiro (2003) and Huberman et al. (1998), we assume the utility of a user's choices on the next page is not certain but rather is stochastically related to the value of prior pages. Therefore, a user's shopping cart choices depend on the stimuli she or he encountered prior to the current page view (see Figure 2). If s represents user i 's cognitive internal state, or unobserved purchase intent ($s = 1, \dots, S$), during session q for page view t , the latent utility associated with choice j in user i 's intent state s is

$$V_{ijqt} = \begin{cases} \beta_{ijs0} + \beta_{ijs1} \mathbf{X}_{iq(t-1)}^1 + \beta_{ijs2} \mathbf{X}_{iq(t-1)}^2 + \beta_{ijs3} Z_{iq(t-1)} + \eta_{ijqt}, & \text{for } j = 2, 3, 4, \text{ or } 5 \\ 0 & \text{for } j = 1 \end{cases}, \quad (2)$$

where β_{ijs0} captures the individual-, choice-, and intent state-specific intrinsic utilities, and β_{ijs1} is a choice-specific $(L + 1) \times I$ parameter vector. For identification, exiting the site serves as the baseline choice, with a utility of 0 (i.e., $V_{i1qt} = 0$). Furthermore, $\mathbf{X}_{iq(t-1)}^1$ is a vector of cumulative marketing and web stimuli received up to $t - 1$ by user i during session q . We take the logarithm of these stimuli variables to capture their potential nonlinear effects and add 1 to each variable to avoid log of zero problem (Manchanda et al. 2004). Similarly, $\mathbf{X}_{iq(t-1)}^2$ and $Z_{iq(t-1)}$ refer to the user's in-session activities (visit depth, time spent on last page view, comparison shopping at competing sites in the category and non-category sites, sign in or not, numbers of items in cart is two or more to receive free shipping) and past behavior variables (purchase from last session, weekend visit or not), respectively. Similarly, we take the logarithm of the user's in-session activity variables to account for potential nonlinear effects, with two exceptions (sign-in and if the number of items in cart is two or more), after adding 1 to the variables. In our empirical test, total users $I = 1,160$, the number of sessions per user q_i ranges from 1 to 17, and the number of page views T_{iq} ranges from 2 to 240.

3.2.2 Potential Endogeneity in the Stimuli

The online retailer may engage in segment- or individual-level targeting when choosing its marketing and web stimuli (\mathbf{X}_{iq}^1), including different price, promotion, banner ad, e-mail solicitation, hypertext links, or

pictures to target different users. To account for potential endogeneity due to individual-level targeting (Manchanda et al. 2004), we allow the marketing and web stimuli to be functions of their lagged values at the user level². The retailer in our study did not adopt any dynamic web page customization according to the user's latent intent state, so the parameters in these endogeneity functions are not affected by the user's latent intent state, nor are they specific to any intent state. Thus,

$$\mathbf{X}_{igt}^1 = \gamma_{0i} + \gamma_{1i} B_{ig(t-1)} + \mathcal{G}_{igt}, \quad (3)$$

where \mathbf{X}_{igt}^1 and $B_{ig(t-1)}$ denote the vector of the log of the cumulative marketing and web stimuli variables on which the retailer can base its individual targeting in the current and previous page views (we add 1 to these variables to avoid zero log issues), respectively. Then γ_{0i} and γ_{1i} are user-specific parameters to be estimated, and \mathcal{G}_{igt} is the random shock.

3.2.3 Simultaneity of Shopping Cart Choices and In-Session Activities

In the meantime, there may exist simultaneity between a user's shopping cart choices and her or his in-session activities (\mathbf{X}_{igt}^2), such as comparison shopping, visit depth, viewing time, sign-in, and if the number of the items in the cart is two or more to receive free shipping. That is, they may be interdependent, due to the impact of the same marketing and web stimuli and unobserved environmental and competition factors. A user's in-session activities also are affected by latent intent states, similar to shopping cart choices. Therefore,

$$\mathbf{X}_{igt}^2 = \phi_{0is} + \phi_{1is} M_{igt} + \tau_{igt}, \quad (4)$$

where \mathbf{X}_{igt}^2 consists of the log of comparison shopping on book category and non-category sites, log of visit depth, log of viewing time, sign-in, and number of items in the cart; M_{igt} denotes a vector of the log of cumulative marketing and web stimuli variables; ϕ_{0is} and ϕ_{1is} are latent, intent state-specific

² In the empirical analysis, we conducted a series of tests which show that the lagged values are valid instruments for the stimuli variables since the former is highly correlated with the latter but has insignificant correlations with the error terms in the cart choice utility function in Equation 2.

parameters to be estimated; and τ_{igt} is the random shock. All the variables in \mathbf{X}_{igt}^2 are continuous except sign-in and if the number of items in the cart is two or more, for which we use binary probit setups.

To account for both potential endogeneity in the marketing and web stimuli and the simultaneity issue, we allow the error terms in Equations 2–4 to correlate, so $\varepsilon_{iqts} = (\eta_{iqts}, g_{igt}, \tau_{igt})' \sim MVN[0, \Sigma]$, where Σ is the variance-covariance matrix to be estimated.

3.2.4 Capturing Individual User Heterogeneity

Different users have different characteristics, so the identification module must account for individual user heterogeneity during the learning process. Consistent with prior research, our identification module employs a hierarchical Bayesian framework to capture user heterogeneity as user demographics (Rossi et al. 1996). With a usual hierarchical Bayesian framework (Rossi et al. 2006), we incorporate heterogeneity across users by assuming that β_{is} , γ_i and ϕ_{is} have random coefficient specifications, and all follow a multivariate regression:

$$\begin{aligned}\beta_{is} &= \bar{\beta}' \mathbf{R}_i + \varsigma_{is}, & \varsigma_{is} &\sim MVN(\mathbf{0}, \Psi_{\varsigma}) \\ \gamma_i &= \bar{\gamma}' \mathbf{R}_i + \xi_i, & \xi_i &\sim MVN(\mathbf{0}, \Psi_{\xi}) \\ \phi_{is} &= \bar{\phi}' \mathbf{R}_i + \omega_{is}, & \omega_{is} &\sim MVN(\mathbf{0}, \Psi_{\omega})\end{aligned}\quad (5)$$

where \mathbf{R}_i is the vector of demographic measures (i.e., age, gender, education, income level), plus an intercept for user i ; $\bar{\beta}$, $\bar{\gamma}$ and $\bar{\phi}$ are parameter matrices; and Ψ_{ς} , Ψ_{ξ} and Ψ_{ω} are all covariance matrices.

3.2.5 Learning Users' Unobserved Intent States

Because a user's unobserved purchase intent states change over time, we propose capturing this latent intent through a first-order, continuous time, discrete-state hidden Markov model (HMM). This approach is consistent with Titus and Everett's (1995) suggestion that way-finding processes and user online navigation paths reflect intents. Similar HMM treatments have helped examine users' unobserved life

stages, competitive promotions, and relationship states (Du and Kamakura 2006; Moon et al.2007; Netzer et al. 2008).

<Insert Figure 3 Here>

The proposed HMM consists of three major components: the starting probabilities of the intent states (\mathbf{v}_{iq}), the transition probability matrix (\mathbf{P}_{igt}) and waiting times of the intent states (w_{igt}). Figure 3 summarizes the HMM working principles: Before the user visits a retailer site, an e-mail solicitation and the duration since his or her last session determine the starting probabilities of the HMM intent states (\mathbf{v}_{iq}), as well as the intent state that exists at the first page view in session q . As the user progresses over the course of the session, the marketing and web stimuli encountered and the user's comparison shopping activities at other sites, up to the last page view ($t - 1$) affect his or her transition probability matrix (\mathbf{P}_{igt}) and waiting times (w_{igt}) for the HMM at the current page view t . The transition probability matrix determines the specific intent state to which the user jumps, if a jump occurs. In the meantime, the waiting time of HMM states determines how long the user stays in the current intent state before jumping to another. This process continues until the end of the user's visit session ($t = T_{iq}$). Note that the user's intent also may persist across sessions, such that he or she could place items in the shopping cart in one session, then come back to purchase them in another session. To account for the interdependence of a user's intent across sessions, we allow the time duration from the end of last session to affect the starting probabilities of HMM states in the current session. Next, we describe the model setup for each of the three HMM components, starting with the waiting times (w_{igt}).

Thus far we have used s to represent a user's unobserved intent state during session q for page view t . But this notation is defined only at integer-time values. Because time is continuous and the user's intent can change at any time, we use a hidden, continuous-time Markov chain D_{igt} to replace s , where D_{igt} equals s at integer values. Waiting time represents how long the user remains in a particular intent state before moving to another. We assume that waiting time between transitions (w_{igt}) from one intent state to

another in our continuous-time domain follows an exponential distribution for user i at page view t in session q , as follows:

$$\Pr[w_{igt} | \lambda_{igt}] = \lambda_{igt, D_{igt}} \exp\{-w_{igt} \lambda_{igt, D_{igt}}\}, \text{ where } \lambda_{igt} = [\lambda_{igt1}, \lambda_{igt2}, \dots, \lambda_{igtS}]^T, \quad (6)$$

where λ_{igts} is an intensity parameter for intent state s for session q at page view t , and the expected waiting time (until transition out of the current intent state) is the inverse of this parameter, $1/\lambda_{igts}$.

Waiting times are intent state-specific and depend on the realization of the first-order HMM. Therefore, the HMM is not memory-less (i.e., the hazard function dynamically changes from page view to page view; Liechty et al. 2003).

Because the user's intent state changes, the transition matrix (\mathbf{P}_{igt}) of the HMM that defines our first-order Markov process is:

$$\Pr[D_{iq, t+w_{igt}} = s | D_{igt} = g, \mathbf{v}_{iq}, \mathbf{P}_{igt}] = P_{igtgs} \text{ if } t > 1, \text{ where } \mathbf{P}_{igt} = \begin{bmatrix} 0 & P_{igt12} & \dots & P_{igt1S} \\ P_{igt21} & 0 & \dots & P_{igt2S} \\ \vdots & \vdots & \ddots & \vdots \\ P_{igtS1} & P_{igtS2} & \dots & 0 \end{bmatrix}, \quad (7)$$

where P_{igtgs} denotes the conditional transition probability that user i switches to intent state s in session q on page view t , given that the previous intent state was g , and the rows sum to 1. The diagonal elements equal 0, because the same state transitions (from s to s) are captured by waiting times. The transition matrix and waiting time process govern the latent intent changes during the visit. Finally, the initial state probability (i.e., that a user is in intent state s at the first page view during a session) is:

$$\Pr[D_{igt} = s | \mathbf{v}_{iq}] = v_{iq, D_{igt}} \text{ if } t = 1, \quad (8)$$

where the probability vector \mathbf{v}_{iq} consists of the initial starting probabilities.

Further, we assume that the row vectors of the transition matrix and the vector of initial probabilities follow a Dirichlet distribution, and the waiting time intensity follows a gamma distribution:

$$\mathbf{P}_{igtj} \sim D(\boldsymbol{\tau}_{igtj}), \quad \mathbf{v}_{iqs} \sim D(\boldsymbol{\alpha}_{iqs}), \quad \lambda_{igts} \sim \Gamma(\hat{\lambda}_{igts}, \check{\lambda}), \quad (9)$$

where \mathbf{P}_{ijt} denotes the j^{th} row of the matrix \mathbf{P}_{it} , and $\hat{\lambda}_{iqts}$ and $\tilde{\lambda}$ denote the shape and scale parameters of the gamma distribution, respectively. Furthermore, we assume that the hyperparameters (τ_{ijt} , α_{iq} , $\hat{\lambda}_{iqts}$) are functions of user comparison shopping activities and the marketing and web stimuli. In turn,

$$\begin{aligned}\log(\tau_{ijt}) &= \bar{\tau}_j \cdot \mathbf{E}_{iq(t-1)} + \mathbf{v}_{ijt} \\ \log(\alpha_{iq}) &= \bar{\alpha}_s \cdot \mathbf{E}_{iq} + \xi_{iq} \\ \log(\hat{\lambda}_{iqts}) &= \bar{\lambda}_s \cdot \mathbf{E}_{iq(t-1)} + \mathbf{p}_{iqts}\end{aligned}\quad (10)$$

where $\mathbf{E}_{iq(t-1)}$ are the stimuli and comparison shopping activities of user i in session q at page view $t - 1$.

For the equation $\log(\alpha_{iq})$, \mathbf{E}_{iq} includes e-mail solicitation and the duration between the current and the last session. Because the hyper-parameters (τ_{ijt} , α_{iq} , $\hat{\lambda}_{iqts}$) should be positive, we assume they follow log-normal distributions with the error terms, such that $\mathbf{v}_{ijt} \sim \text{Normal}(0, \Sigma_v)$, $\xi_{iq} \sim \text{Normal}(0, \Sigma_\xi)$, and $\mathbf{p}_{iqts} \sim \text{Normal}(0, \Sigma_\rho)$.

3.2.6 Model Identification Issues

As we discussed previously, we infer a user's latent purchase intent states from the shopping cart choices at each page view. Because user intent states are unobserved, we must restrict the intent state-specific parameters to identify them. That is, to identify unobserved intent states and parameters in Equation 2, we restrict the average purchase probabilities to be non-decreasing in intent states. Because both the intercepts and the response parameters are state-specific in Equation 2, following Netzer et al. (2008), we impose this restriction at the mean of the vector of covariates. The covariates in Equation 2 thus are mean-centered, and we let the intrinsic utilities of purchase choice β_{i5s0} (where the subscript 5 denotes purchase) increase in state s . Considering the increase in the user's purchase intentions at higher intent states for each page view, a user in a higher intent state is more likely to purchase, suggesting a higher intrinsic utility of making a purchase (i.e., $\beta_{i510} \leq \beta_{i520} \leq \dots \leq \beta_{i5S0}$; Li et al. 2011). Finally, for identification purposes, we set the first and the last two diagonal elements of the variance-covariance

matrix Σ to equal 1, due to the multinomial probit model setup of Equation 1 and the binary nature of the variables of sign-in and if the number of items in the cart is two or more. We present the program estimation in the Web Appendix.

3.3 Transformation Module

Online users dynamically adapt their behaviors to the page-by-page stimuli they encounter, even when they are unaware of their own adaptive behavior (Mandel and Johnson 2002). When it has identified a user's intent state, the site should immediately tailor the content of the next page to match this intent, using the transformation module. The goal of this transformation is to convert users into real buyers.

Our general idea is that, for any user who visits a website, the user's action of placing item(s) into his/her shopping cart marks a progress towards purchase, though, at this moment, it is not known if the user will ultimately complete a purchase. But such an action indicates a potential chance for the site to gain a purchase conversion because putting item(s) into the cart indicates that the user has moved further along his/her purchase process, which is better than pure browsing with exiting the site. Thus, we wish to attract users to place item(s) into the cart if they have not done so. And for those with item(s) in the cart, our goal is to encourage and persuade them to make actual purchases and not abandon the shopping cart.

With this in mind, the transformation module constitutes two steps. As the first step, before the real-time webpage adaptation, we estimate the parameters in the proposed model in the identification module using data of a randomly selected sample of users. At step two, real-time concurrent webpage adaptation is implemented as we show in Figure 2. Specifically, before a new user starts a session, the estimated parameters of the best proposed model identified from step 1 are used as starting values for the user's individual learning model which is the same model defined in the identification module, but for the user only. As the user progresses during the session, the user's model is updated dynamically page view by page view, reflecting his/her intent. Based on this, at each page view, an optimization sub-module (within the transformation module) implements concurrent optimal page adaptation which consists of a two-stage optimization process and one adaption process. The first stage of the optimization process is, if no item is in the cart, to maximize the user's probability of adding an item to the cart since the user

cannot purchase without adding items to cart first. In the second stage, conditional on having some item(s) in the cart, the sub-module will maximize the user's probability of making a purchase. The optimization results will inform the adaptation process about which components on the next page should change, and the corresponding requests that define a set of HTML extensions by using Ajax and server-side scripting techniques will be sent to the website to make such changes. A complete redesign of the structure and contents of a webpage would take time and probably is not feasible; our transformation module only indicates optimal changes to marketing stimuli (i.e., price appearance, pop-up promotion, banner ad, and e-mail solicitation) and web stimuli (i.e., hyperlinks and pictures) on the next page, which are well within each site's control.

<Insert Figures 4a and 4b Here>

To better illustrate how the optimal webpage transformation is implemented for one page view, we provide an example with two screenshots in Figures 4a and 4b. The user in the example is an individual who has viewed two pages on a hypothetically simplified Barnes & Noble's site (bn.com) – the home page and a search results page on books about 'smart marketing'. Next, the user clicks on one particular book to take a close look. The screenshot in Figure 4a shows the product page for the book without page transformation and the one in Figure 4b is the book's page with optimal page transformation based on our proposed model. Specifically, after the user clicks on the book on the search results page with no item in the shopping cart, the transformation module in the proposed model system performs a real-time dynamic updating of the user's individual model parameters and learns that the user is in the low intent state at this point. In the meantime, an optimization on the marketing stimuli (i.e., whether to present price information, a pop-up promotion on free shipping, or a banner ad on another book related to smart marketing) and web stimuli (i.e., the number of hyperlinks and pictures) on this product page is performed in order to maximize the user's probability of adding an item to the cart given no item in the cart yet. The optimally transformed page is shown in Figure 4b. Comparing the two screenshots in the figure with the optimal changes marked by the arrows in Figure 4b, we find that for this user in the low intent state, on the optimally transformed page, the book's price information is not presented, a pop-up

promotion on free shipping and one more hypertext link on textbooks are added compared to the product page without page transformation. Note that the number of pictures (one picture of the book excluding the banner ad picture) and the banner ad presence remain the same for both pages in the figures.

4. Empirical Test and Results

To check the performance of our proposed method, we conducted an empirical test with a real-world data set and then a laboratory experiment, focused on (1) performance in terms of learning and recognizing the user's unobserved purchase intent states while she or he navigates through the website (identification module) and (2) the effectiveness of the proposed intent-based optimal page transformation for reducing shopping cart abandonment and increasing purchase conversion. Using a conventional testing approach for data mining and machine learning, we first investigated a real data set and simulations to test the proposed method, then conducted a laboratory experiment in which we asked participants to visit two experimental sites, one without and one that used the proposed system.

4.1 Data

The data set for our test features user shopping and browsing activities at Barnes & Noble's online bookstore (<http://barnesandnoble.com> and bn.com). It consists of the site visit activity of 1,160 users who visited BN.com during April 1–April 30, 2002, as measured by comScore Media Metrix.³ Table 1 provides their demographic characteristics. Our observation module, written in Perl script, collected the HTML content of each page seen by each user in the month immediately following the data period. We checked that BN.com did not make any major changes to its website or page customization during these two months. In addition, BN.com's online shopping cart appears on every shopping page. The variable descriptions and descriptive statistics for the data at the page view level are in Table 2; a multicollinearity test of these variables revealed no evidence of multicollinearity.

<Insert Tables 1 and 2 about here>

³ Shortly after compiling this data set, Media Metrix was acquired by comScore Networks, which subsequently implemented considerable improvements to both the data collection methodology and the depth of data measured. These improvements included significant increases in panel size to accommodate analyses of home, work, and

The 1,160 users engaged in 1,704 sessions with the site and viewed an average of 8.75 pages per session. The 94 sessions that ended in purchases implied a session conversion rate of 5.5%. In 89.3% of sessions (i.e., 1,522 sessions), users did not have any items in their shopping cart when they exited, which implies pure online browsing.⁴ Of the remaining 182 sessions with at least one item in the cart at some point, about half (48.4%) ended in shopping cart abandonment.

4.2 Results

4.2.1 Learning and Predictive Ability of the Proposed Model System

Because we seek to have the identification module learn from each user’s initial clicks after entering the website, to identify unobserved intent, we adopted a common test method from machine learning literature and randomly divided the users into two groups: 590 users in the training sample for learning and 570 users for out-of-sample predictions, with 9,259 and 7,564 observations, respectively. As we described in Section 3.2, the identification module used a hierarchical Bayesian approach to learn and estimate the possible number of unobserved intent states that each user might exhibit during the course of the visit. That is, given the user’s behaviors from the beginning of the visit up to the current page view t , the identification module starts with one intent state, then tests two, three, and four intent states, and so forth. The best case, or most appropriate number of unobserved intent states, depends on the Bayes factors, or the ratio of posterior marginal densities of any two competing models (Chib and Greenberg 1995; Newton and Raftery 1994) and the prediction performance of the models. In Table 3, along with the main results, we present the learning results from the training sample in the “Estimation Sample” column. To check if our identification module can predict users’ intents for each page view, we applied it to the holdout sample and display the results in the “Holdout Sample” column.

<Insert Table 3 about Here>

university audiences and full measures of transaction details, including products purchased, prices paid, and shopping basket totals. These extended data were not available to us.

⁴ Users virtually never use the proffered wish list capability at bn.com (i.e., 99.9% of sessions), so we did not consider it further in our analyses.

To assess the robustness of the proposed identification module, we also considered four nested benchmark models (see the first column in Table 3). First, we considered an aggregate model without user heterogeneity and intent states as well as endogeneous marketing and web stimuli and user in-session activities. Second, we added individual user heterogeneity to the first model. Third, in a modified Sismeiro and Bucklin (SB 2004) model, we added endogeneity to the second model but did not incorporate unobserved intents. Fourth, following Montgomery et al. (MLSL 2004), we added HMM to account for each user's intent. However, the MLSL model assumes that the user's intent is static during a shopping visit. Both the original SB and MLSL are choice models, so they cannot capture users' decisions about various shopping cart choices. The one-state HMM model thus equals the modified SB model, and the MLSL model is equivalent to a two-state, homogeneous HMM.

We used log marginal density (larger is better), mean absolute errors (MAE, lower is better) of predicted cart choice probabilities, and hit rates (percentage of user cart choices that the model predicts correctly; higher is better) as performance criteria. In terms of in- and out-of-sample MAE and hit rates, the one-state model without endogeneity outperforms the first aggregate-level benchmark model but is outperformed by the one-state SB model, which reveals the importance of incorporating both individual user heterogeneity and the endogeneity of marketing and web stimuli and user in-session activities. However, the two-state MLSL model outperforms the one-state SB model, so we also must account for the user's unobserved intent states. By including heterogeneous HMM to account for changes in unobserved intents, with covariates, our identification module with two intent states outperforms the SB model, the MLSL model, and the three- and four-state models, offering the largest log-marginal density, lowest in- and out-of-sample MAE, and highest hit rates. Therefore, our identification module with two-state HMM is the best option⁵. With a user's observed shopping cart choices up to page view $t - 1$, our identification module can identify unobserved intent at page view t , as either low purchase intention (State 1) or high purchase intention (State 2).

⁵ We also conducted similar model comparison for all the models in Table 3 using the whole data and obtained similar results with the two-state HMM as the best model.

4.2.2 Best Model with Two Intent States

Table 4 summarizes the estimation results for the HMM with two intent states. For BN.com, users were more likely to begin their visit sessions in the low-intent state (75% probability) than in the high-intent state (25%). Most users thus had low purchase orientations at the beginning of the session. Average waiting times for the low- and high-intent states were 2.70 (i.e., inverse of the waiting time intensity estimate λ_{iqts} , given the exponential distribution assumption in Equation 6, or $1/0.37$) and 3.57 ($1/0.28$) page views, respectively. Although each user was likely to start the session in the low-intent state, once reaching the high-intent state, he or she would remain in that state for longer than the duration of the low-intent state. For the two-state HMM, the state transition probability matrix was trivial, with 100% probability of jumping from one state to another if a jump occurs.

<Insert Tables 4 and 5 about Here>

The results for the impact of the four focal sets of factors (marketing stimuli, web stimuli, comparison shopping activities, and user past behavior) on shopping cart choices are in Table 5. From the intercept estimates, we find as expected that high-intent users are intrinsically more likely to add some item to the cart and make a purchase than are low-intent users, but they are less likely to remove an item. More price information displayed in past page views increases the likelihood that high-intent users add or delete items from their cart and make a purchase; this cumulative price information has an insignificant impact on the cart choices of low-intent users (Moe 2006). Pop-up promotions encourage low-intent users to add or delete items, continue browsing, and purchase but have the opposite and thus negative impacts on the cart choices of high-intent users. We posit that pop-up promotions attract low-intent users' attention and interest but distract high-intent users from progressing to their final purchase. More banner ads encourage both low- and high-intent users to continue browsing, add or delete items, and purchase. E-mail solicitations seem to encourage low-intent users to continue browsing, add or delete items, and purchase but have insignificant impacts for high-intent users. In addition, in terms of the interactions among the marketing stimuli, for users in both intent states, we find a positive interaction between price information

and pop-up promotion but a negative interaction of price information with banner ads. Therefore, it may be beneficial for the site to use price information and pop-up promotion together, but not price information and banner ads.

Increasing the number of hypertext links encourages high-intent users to continue browsing, add or delete items, and purchase, with an insignificant impact for low-intent users. Also, increasing the number of pictures on a web page discourages low-intent users from continuing to browse, rather than exiting, but it has insignificant impacts on the cart choices of high-intent users. These findings demonstrate the importance of customizing web stimuli to appeal to different users with different purchase intentions.

The more any user comparison shops at competing bookstore sites during the session, the less likely he or she is to continue browsing, add or delete items, or purchase, which indicates a competition effect. However, for users in both intent states, visiting other non-category sites makes them more likely to continue browsing, add or delete items, and purchase. Thus users' comparison shopping at non-bookstore sites appears complementary to their behaviors at the focal site. Because we take the logarithms of the marketing and web stimuli and comparison shopping variables, the results indicate a decreasing marginal return on the variables' impacts.

In terms of users' past behaviors, when low-intent users view more webpages (visit depth), sign in, or make a purchase in last session, they are discouraged from continuing to browse, add or delete items, and purchase. Earning free shipping status or visiting the site on a weekend instead have positive impacts. Among high-intent users though, time spent in the last page view and making a purchase in the last session discourages them from continuing to browse, add or delete items, or purchase, whereas earning free shipping has a positive impact on their cart choices. Here, we note the importance of the differential impacts of users' past behavior on their cart choices across the two intent states.

Finally, we present the estimates for the demographic variables in the heterogeneity equation, the endogeneity functions, the variance-covariance matrix, and other HMM results in the Web Appendix.

4.3 Optimal Dynamic Webpage Adaptation Performance

As we show in Figure 2, our system generates intent-based page transformations for improving site performance. In accordance with the S-O-R paradigm, we conduct a purchase elasticity analysis of the marketing and web stimuli, to determine how changes to each of them might affect purchases. Mathematically, using the estimates of the two-state proposed model for each variable (i.e., price presence, pop-up promotion, banner ad, e-mail solicitation, number of product links, and number of product pictures), we increase them one at a time by 1%, using the estimation sample and holding everything else constant. The results are in Table 6.

<Insert Table 6 about Here>

The web and marketing stimuli have differential effects across low- and high-intent users. For example, online managers should use pop-up promotions, banner ads, and e-mail promotions to target low-intent users, because doing so will improve their purchase conversion, but they should use price information and banner ads for high-intent users. Increasing the number of product hypertext links on the page for low-intent users and increasing both product links and pictures for high-intent users also can improve their purchase probabilities.

We incorporate the suggestions derived from our elasticity analysis into the transformation module, which then makes corresponding intent-based changes to the next page view, according to each user's intent state. As we have mentioned early, our transformation module runs a two-stage optimization procedure. Specifically, the first stage is, if no item is in the shopping cart, to maximize the user's probability of adding an item to the cart. In the second stage, conditional on have some item(s) in the cart, the module will maximize the user's probability of making a purchase. The optimal results from this procedure inform how marketing and web stimuli on the next page should be changed. To check the performance of this dynamic optimal page transformation, we conducted a simulation test and a lab experiment.

4.3.1. Simulation Test

For this simulation, according to the transformation module, we imagine that BN.com would first use the estimates from the two-state proposed model as starting values for each user's individual learning model. Then, as the user progresses during the session, the user's model is updated dynamically page view by page view with the identification of his or her intent states, and thus performs optimal webpage transformation on the marketing and web stimuli on the page after the first page view, up to the ninth page (average number of page views per session is 9.87 in our data). The goal is to reduce shopping cart abandonment and improve purchase conversion. Note that, as we explicitly model the endogeneity in marketing and web stimuli and the simultaneity in users' in-session activities, the simulation should not alter the user's equilibrium observed in the empirical data.

<Insert Table 7 about here>

We obviously cannot install our proposed system into BN.com, so we developed a simulation site that duplicated the page content of BN.com. We also checked that BN.com did not make any major changes to its website or page customization, so our simulation test should reflect actual consumer shopping processes. In Table 7 we display the session-level predicted shopping cart abandonment and purchase conversion rates for different page views under the proposed system. Users' real shopping data from BN.com showed a 49% abandonment rate, but the proposed system decreased this rate significantly, namely, by 32.4% if all pages after the first page view were optimally transformed, by 33.8% if the pages after the first three page views were customized, and by 29.4% if the pages after the first nine page views were transformed. The purchase conversion rate also improved significantly, by 6.9% relative to the sample rate of 5.9% if all the pages after the first page view were optimally transformed. This improvement declined with increasing delays in the intent-based page transformation. That is, gains due to intent-based page content adaptation were relatively greater for interventions earlier in the user's session, particularly after the first three page views. The propensity to exit the site increases with each additional page browsed in the session, according to Bucklin and Sismeiro (2003). Similarly, in our holdout sample, we found that the shopping cart abandonment rates are 20.3%, 19.5% and 23.4% with 24.1%, 24.9%, and 21.0% reductions (relative to the sample abandonment rate of 44.4%) after the first,

first three, and first nine page views, respectively. The purchase conversion rate improvement followed a similar pattern. Considering BN.com's 49% shopping cart abandonment rate and 5.5% purchase conversion rate, these significant shifts would have key impacts on sales and profitability.

4.3.2 Lab Experiment

To test our proposed model system in real time, we conducted a small laboratory experiment with a cross-subject design. We developed a hypothetical, simple book retailer website in English, with basic functions similar to those on typical retailer websites such as BN.com or Amazon.com (e.g., search box, navigation links, shopping cart, forms, registration, product information). The mock site's checkout process included forms for users to complete but no credit card verification or third-party payment system. If a study participant chose to make a purchase (i.e., click the purchase button), the page containing the order information was recorded in a database, with the user's account and order information.

Forty undergraduate students at a large, Midwestern U.S. university participated in this experiment for course credit. The course previously had introduced students to the uses of advanced information technologies for smart commerce online. Students were randomly assigned to two groups, with 20 students in each: Group 1 used a normal version of the website without the proposed model system (control group), and Group 2 used a version with the proposed model system. The participants were not aware of this difference but instead knew only that the hypothetical website sold books on smart commerce and marketing that were not required by the course. Each participant acted like a normal consumer, visiting a retail website, where he or she could purchase if so motivated or exit the site without purchasing. A purchase required signing in to an account. The experiment lasted for 75 minutes during one class meeting.

When each subject in Group 2 visited the site with the proposed model system, as we describe earlier, the identification module worked as the subject's individual model which first adopted the estimates from the two-state HMM as starting values for his/her individual learning model. Then, as the subject progressed during the session, the user model analyzed the subject's clicks and cart choices, and self-updated dynamically page view by page view. At each page view, as the subject's latent intent state

was identified, the transformation module was triggered to make corresponding intent-based changes to the next page. If the subject had no item in the cart, the transformation module would run the optimization process to suggest optimal changes on the marketing and web stimuli on the next page to maximize the subject's probability of placing item(s) into the cart. Similarly, if the cart had item(s), the transformation module would then run the optimization process to maximize the subject's probability of making a purchase.

<Insert Table 8 about here>

From Table 8, we find that for the control group, for which the site did not identify their latent intention or perform concurrent learning and dynamic webpage transformation, the purchase conversion and cart abandonment rates were 10% and 60%, respectively. The average webpage loading time was 0.40 seconds, with a standard deviation of 0.20. In contrast, for the test group, for whom the site identified latent intentions, we found 61.66% and 38.34% probabilities that they were in low- and high-intent states, respectively. The site also concurrently learned about their intent and preferences and performed dynamic webpage transformation. Their purchase conversion rate increased to 25%, or 15% greater than that in the control group; the cart abandonment rate was 28.57%, or less than half of that of the control group. In addition, the average web page transformation and loading time was 0.48 seconds (standard deviation = 0.32)—slightly longer than that for the control group but still very fast. Thus, the site effectively learned about each person's latent intent state in 2.92 page views, which further confirms the effectiveness of the proposed model system.

5. Conclusions and Further Research

Low purchase conversion and high shopping cart abandonment rates represent a significant problem for online retailers. Understanding online users' real-time intent and dynamic shopping cart choices are critical for retailers' profitability, and the potential to identify each user's real-time intent and then generate optimal intent-based page contents is of utmost interest to online managers. We demonstrate that an individual-level, real-time computer system model can analyze user navigation behavior, as manifested by shopping cart choice decisions, and simultaneously learn unobserved real-time intent with just a few

clicks, to generate optimal dynamic page adaptation before the user exits the site. It thus effectively can turn users into real buyers.

To generate optimal intent-based page adaptation, this proposed model examines individual users' shopping process, identifies the factors that affect shopping cart choices, assesses how users respond to marketing and web stimuli and comparison shop at other sites, and predict how these exposures will change purchase intents during and across sessions. In our study's empirical context, users' unobserved purchase intents consisted of two states, low-intent and high-intent. At the retailer site we investigated, users tended to start in a low-intent state, but they persisted in the high-intent state longer once they reached it. Marketing and web stimuli, as well as users' comparison shopping activities, exerted differential impacts on cart choices, depending on the users' latent intent state.

Our empirical analysis also shows that for online managers to maximize a user's purchase probability, they need to use different optimized tools to target users in different intent states. For example, at BN.com, managers should target low-intent users using pop-up promotions, banner ads, e-mail solicitations, and an optimal number of hypertext links, but for high-intent users, they should use price, banner ads, and an optimal combination of hypertext links and pictures. Our simulation results also indicate that it is a good strategy for retailers to identify users with different intents and intervene early by optimally customizing the web pages each user sees. At BN.com, the optimal timing for an intervention is after three page views, which grants it an efficient means to learn user intent well but still intervene early enough. The results from the lab experiment confirm this finding and demonstrate the superior performance of the proposed model system compared with a site without such a system.

The behavioral measures, such as those that assess individual users' unobserved real-time intent, provide the basis for generating optimal intent-based contents and stimuli on the next page. Our empirical results show that the proposed model system significantly reduces shopping cart abandonment rates, by 32.4%, and improves purchase conversion rates, by 6.9%, if the optimal customization follows the first page view. Users exposed to intent-based content thus appear likely to engage in desired cart choice behavior, including purchasing. As we show, each web page can be optimally customized concurrently to

help each user progress through the shopping process. To the best of our knowledge, this study is the first to incorporate learning of users' unobserved intent with concurrent optimal page adaptation into the shopping process.

We also highlight some limitations and directions for research. First, we conceptualize shopping cart abandonment as a navigational event and thus consider factors that affect the navigation progress through the site. We did not address the effects of price on purchase decisions, due to data unavailability, yet the actual prices for products or shipping and handling offered by the focal retailer and its competitors likely have significant impacts on abandonment rates. These effects could be examined in controlled laboratory experiments, because access to actual clickstream data sets with all competing offers is limited. Second, the marketing mix stimuli of competing websites likely affect a focal firm's performance, but our model does not include these effects explicitly, because we lack data about them. However, we capture some of these effects as the page views of comparison shopping content. If marketing stimuli data from competing websites were available, it would be possible to develop richer mathematical models and account for additional stimuli, such that the focal firm could develop customized offers and stimuli that also mitigate competition effects. Third, data unavailability prevented us from incorporating the quality measures of marketing and web stimuli, word of mouth, customer reviews, or other social network variables. Uncovering the impacts of these variables on online users' shopping and browsing behaviors and their latent shopping intents represents an interesting extension to our research, if such data become available. Fourth, it will also be interesting to examine the cost effectiveness of the site's preventative (i.e., the marketing and web stimuli in the proposed model before the user exits the site) and remedial (i.e., email reminders or retargeting ads after the user abandons the cart and leaves the site) measures in order to improve the purchase conversion rates. Additional research also might seek to develop a model system that exploits cloud computing services, such that it gathers related information from competing websites simultaneously while generating appropriate page adaptations.

REFERENCES

- Adelaar, T., S. Chang, K.M. Lancendorfer, B. Lee and M. Morimoto. 2003. Effects of media formats on emotions and impulse buying intent. *Journal of Information Technology*, 18 (4), 247-266.
- Ajzen, I. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.
- Albert, T. C., P.B. Goes, and A. Gupta. 2004. A model for design and management of content and inter activity of customer-centric web sites. *MIS Quarterly*, 28(2), 161-182.
- Ansari, A., and Mela, C. F. 2003. E-customization. *Journal of Marketing Research* 40 (May), 131-145.
- Baker, J. 1986. The Role of the Environment in Marketing Services: The Consumer Perspective. In *The Services Challenge: Integrating for Competitive Advantage*. Eds. John A. Cepeil et al. Chicago, IL: American Marketing Association, 79-84.
- Bitner, M. J. 1992. Servicescapes: the impact of physical surroundings on customers and employees. *Journal of Marketing*, 56 (April), 57-71.
- Bolin, M., M. Webber, P. Rha, T. Wilson, and R. C. Miller. 2005. Automation and customization of rendered web pages. In 2005 UIST, October 23-27, Seattle, USA.
- Bucklin, R. E., and Sismeiro, C. 2003. A model of web site browsing behavior estimated on clickstream data. *Journal of Marketing Research* 40 (August), 249-267.
- Chatterjee, P., Hoffman, D. L., and Novak, T. P. 2003. Modeling the clickstream: implications for web-based advertising efforts. *Marketing Science* 22 (4), 520-541.
- Chib, S. and E. Greenberg. 1995. Understanding the metropolis Hastings algorithm. *American Statistician* 49, 327-35.
- Cho, C-H, J. Kang, and H. J. Cheon. 2006. Online shopping hesitation. *Cyber Psychology & Behavior*, 9(3), 261-274.
- , J.-G. Lee, and M. Tharp 2001. Different forced-exposure levels to banner advertisements. *Journal of Advertising Research*, 41 (Jul/Aug), 45-56.
- Cho, Y. H., J.K. Kim, and S.H. Kim. 2002. A personalized recommender system based on Web usage mining and decision tree induction. *Expert Systems with Application*, 23, 329-342.
- Currim, I.S., Gurbaxani, V., LaBelle, J., and Lim, J. 2006. Perceptual structure of the desired functionality of Internet-based health information systems. *Health Care Management Science* 9 (2), 151-170.
- Davis, F.D., R.P. Bagozzi, and P.R. Warshaw. 1989. User acceptance of computer technology: a comparison of two theoretical models. *Management Science*, 35(8) 982-1003.
- Diao, F. and S. S. Sundar. 2004. Orienting response and memory for web advertisements: exploring effects of pop-up window and animation. *Communication Research*, 31(5), 537-567.
- Ding, W. 2003. *A Study of Collaborative Scientific Discovery*. Bell & Howell Information and Learning Company.
- Donovan, R. J. and J. R. Rossiter. 1982. Store atmosphere: an environmental psychology approach. *Journal of Retailing*, 58 (1), 34-57.
- Du, R., and Kamakura, W. 2006. Household life cycles and lifestyles in the United States. *Journal of Marketing Research* 43 (February), 121-132.
- eMarketer. 2009. The sad tale of abandoned shopping carts. July 30. Available at <http://www.emarketer.com/Articles/Print.aspx?1007156>.
- Eroglu, S. A., K. A. Machleit and L. M. Davis. 2001. Atmospheric qualities of online retailing: a conceptual model and implications. *Journal of Business Research*, 54 (2), 177-184.
- Everard, A., and Galletta, D. F. 2006. How presentation flaws affect perceived site quality, trust, and intention to purchase from an online store. *Journal of Management Information Systems* 22 (3), 55-95.
- Forrester Research. 2012. The state of retailing online 2012: investments in mobile and tablet commerce. September 25. Available at

<http://www.forrester.com/The+State+Of+Retailing+Online+2012+Investments+In+Mobile+And+Tablet+Commerce/fulltext/-/E-RES82241>.

- Gollwitzer, P. 1999. Implementation intentions: strong effects of simple plans. *American Psychologist*, 54(7), 493-503
- Herlocker, J.L., J. A. Konstan, L. G. Terveen, and J. Riedl. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1), 5-53.
- Hoffman, D. L., and Novak, T. P. 1996. Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations, *Journal of Marketing* (60:3), 50-68.
- Huberman, B. A., P. Pirolli, J. E. Pitkow, and R. M. Lukose. 1998. Strong regularities in World Wide Web surfing. *Science*, 280(3), 95-97.
- Hulten, B. (2012), "Sensory cues and shopper's touching behaviour: the case of IKEA", *International Journal of Retail & Distribution Management*, 40 (4), 273-289.
- Jacoby, J. 2002. Stimulus-organism-response reconsidered: an evolutionary step in modeling (consumer) behavior. *Journal of Consumer Psychology*, 12 (1), 51-7.
- Jarboe, G.R. and C. D. McDaniel. 1987. A profile of browsers in regional shopping malls. *Journal of the Academy of Marketing Science* 15, 46-53.
- Jiang, Z., and I. Benbasat. 2007. Investigating the influence of the functional mechanisms of online product presentations. *Information Systems Research* 18 (4), 454-470.
- Katz, R. N. 2002. *Web portals and higher education: technologies to make IT personal*. San Francisco: Jossey-Bass.
- Kobsa, A. J. Koenemann, and W. Pohl, 2001. Personalized hypermedia presentation techniques for improving online customer relationships. *Knowledge Engineering Review*, 16, 111-155.
- Koufaris, M., A. Kambil and P. A. LaBarbera. 2001. Consumer Behavior in Web-Based Commerce: An Empirical Study, *International Journal of Electronic Commerce*, 6 (2), 115-138.
- Kotler, P. 1974, Atmospherics as a marketing tool, *Journal of Retailing*, (Winter), 48-64.
- Koufaris, M. 2002. Applying the technology acceptance model and flow theory to online consumer behavior. *Information Systems Research* 13 (2), 205-223.
- Lee, L. and D. Ariely. 2006. Shopping goals, goal concreteness and conditional promotions. *Journal of Consumer Research*, 33 (1), 60-70.
- Li, S., B. Sun, and A. L. Montgomery, 2011. Cross-selling the right product to the right customer at the right time. *Journal of Marketing Research*, 48(4), 683-700.
- Liechty, J.C., R. Pieters, and M. Wedel, 2003. Global and local covert visual attention: evidence from a Bayesian hidden Markov model. *Psychometrika* 68 (4), 519-541.
- Manchanda, P., J-P. Dube, K. Y. Goh, and P. K. Chintagunta, 2006. The effect of banner advertising on internet purchasing. *Journal of Marketing Research*, 43(1), 98-108.
- , P. E. Rossi, and P. K. Chintagunta, 2004. Response modeling with nonrandom marketing-mix variables. *Journal of Marketing Research*, 41(4), 467-478.
- Mandel, N., and E.J. Johnson, 2002. When web pages influence choice: effects of visual primes on experts and novices. *Journal of User Research* 29 (2), 235-245.
- Mathieson, K. 1991. Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research*, 2(3) 173-191.
- Mattila, A.S. and Wirtz, J. (2001), "Congruency of scent and music as a driver of in-store evaluations and behaviour", *Journal of Retailing*, 77 (2), 273-289.
- Mazaheri, E., M.O. Richard and M. Laroche. 2012. The role of emotions in online consumer behavior: a comparison of search, experience and credence services. *Journal of Services Marketing*, 26(7), 535-550.
- Mehrabian, A. and J. A. Russell. 1974. *An approach to environmental psychology*. Cambridge, MA: MIT.
- Mithas, S., N. Ramasubbu, M.S. Krishnan, and C. Fornell, 2007. Designing web sites for customer loyalty across business domains: a multilevel analysis. *Journal of Management Information Systems* 23 (3), 97-127.

- Moe, W. W. 2003. Buying, searching, or browsing: differentiating between online shoppers using in-store navigational clickstream. *Journal of Consumer Psychology* 13 (1&2), 29-40.
- . 2006. An empirical two-stage choice model with varying decision rules applied to Internet clickstream data. *Journal of Marketing Research*, 43, 680-692.
- and Fader, P. S. 2004. Dynamic conversion behavior at e-commerce sites. *Management Science* 50 (3), 326-335.
- Montgomery, A. L., S. Li, K. Srinivasan, and J. C. Liechty. 2004. modeling online browsing and path analysis using clickstream data. *Marketing Science* 23 (4), 579-597.
- Moon, S., W. A. Kamakura, and J. Ledolter, 2007. Estimating promotion response when competitive promotions are unobservable. *Journal of Marketing Research* 44 (3), 503-515.
- Nadkarni, S., and Gupta, R. 2007 “A Task-Based Model of Perceived Website Complexity,” *MIS Quarterly* (31:3), 501-524.
- Netzer, O., Lattin, J. M., and Srinivasan, V. 2008. A hidden Markov model of customer relationship dynamics. *Marketing Science* 27 (2), 185-204.
- Newton, M. A., and Raftery, A. E. 1994. Approximate Bayesian inference by the weighted likelihood bootstrap. *Journal of the Royal Statistical Society (Series B)*, 3, 3-48.
- Novak, T. P., Hoffman D. L., and Duhacheck, A. 2003. The influence of goal-directed and experiential activities on online flow experiences. *Journal of User Psychology* 13 (1&2), 3-16.
- Palmer, J. 2002. Website usability, design and performance criteria. *Information Systems Research* 13 (2), 150-167.
- Parboteeah, D. V., J. S. Valacich, and J. D. Wells. 2009. The influence of website characteristics on a consumer's urge to buy impulsively. *Information Systems Research* 20(1), 60-78.
- Parsons, A.G. (2011), “Atmosphere in fashion stores: do you need to change?” *Journal of Fashion Marketing & Management*, 15 (4), 428-445.
- Passani, M. and D. Billsus. 1997. Learning and revising user profiles: the identification of interesting Web sites. *Machine Learning* 27, 313-331.
- Pavlou, P., and D. Gefen. 2004. Building effective online marketplaces with institution-based trust. *Information Systems Research* 15(1), 37-59.
- Putsis, W. P., Jr. and N. Srinivasan. 1994. Buying or just browsing? The duration of purchase deliberation. *Journal of Marketing Research* 31 (August), 393-402.
- Rajamma, R. K., A. K. Paswan, and M. M. Hossain. 2009. Why do shoppers abandon shopping cart? Perceived waiting time, risk, and transaction inconvenience. *Journal of Product & Brand Management* 18(3), 188-197.
- Ramachandran, V., S. Viswanathan and H. C. Lucas. 2010. Clicks to conversion: the impact of product information and price incentives. Working Paper No. RHS-06-141. Robert H. Smith School of Business, University of Maryland.
- Ratchford, B., M. S. Lee and D. Talukdar. 2003. The Impact of the Internet on Search for Automobiles. *Journal of Marketing Research*, 40 (May), 193-209.
- ResearchandMarkets.com. 2004. Shopping cart abandonment and shipping costs. Bringing a personal touch to ecommerce. Available at: <http://www.researchandmarkets.com/reports/423>.
- Retail Week. 2010. Retail technology roundtable: spend and save. *Retail Week*, Sept. 3.
- Ricci, F. 2002. Travel recommender systems. *IEEE Intelligent Systems*, 17, 55-57.
- Rossi, P. E., Allenby, G. M., and McCulloch, R. 2006. *Bayesian statistics and marketing*. UK: Wiley.
- , McCulloch, R. E., and Allenby, G. M. 1996. The value of purchase history data in target marketing. *Marketing Science* 15 (4), 321-340.
- Savitz, E. 2011. Online retailers: fixing shopping cart abandonment. *Forbes.com*, December 2, 2011.
- Schafer, J.B., Konstan, J. A., and Riedl, J. 2001. E-commerce recommendation applications. *Data Mining and Knowledge Discovery* 5, 115-153.
- Sheppard, B. H., J. Hartwick, and P.R. Warshaw. 1988. The theory of reasoned action: a meta analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research* 15, 325-343.

- Sherman, E. and R.B. Smith. 1987. Mood states of shoppers and store image: promising interactions and possible behavioral effects. P. Anderson (Ed.), *Advances in Consumer Research*, 14, Association for Consumer Research, Provo, UT, 251-254.
- Shimp, T. and A. Kavas. 1984. The theory of reasoned action applied to coupon usage. *Journal of Consumer Research* 11, 795-809.
- Simon, H. A. 1957. *Models of man: social and rational*. New York: John Wiley & Sons.
- Sismeiro, C., and Bucklin, R. E. 2004. Modeling purchase behavior at an e-commerce web site: a task completion approach. *Journal of Marketing Research* 41, 306-323.
- Song, J., and Zahedi, F. M. 2005. A theoretical approach to web design in e-commerce: a belief reinforcement model. *Management Science* 51 (8), 1219-1235.
- Spangenberg, E. R., A.E. Crowley, and P.W. Henderson. 1996. Improving the store environment; do olfactory cues affect evaluations and behavior? *Journal of Marketing* 60, 67-80.
- Spangenberg, E.R., Grohmann, B. and Sprott, D.E. (2005), "It's beginning to smell (and sound) a lot like Christmas: the interactive effects of ambient scent and music in a retail setting", *Journal of Business Research*, 58 (11), 1583-1589.
- Straub, D. and R. T. Watson. 2001. Transformational issues in researching IS and net-enabled organizations. *Information Systems Research* 12 (4), 337-345.
- Tam, K. Y., and Ho, S. Y. 2005. Web personalization as a persuasion strategy: an elaboration likelihood model perspective. *Information Systems Research* 16 (3), 271-291.
- Titus, P. A. and P.B. Everett. 1995. The user retail search process: a conceptual model and research agenda. *Journal of the Academy of Marketing Science*, 23(2), 106-119.
- Viswanathan, S, J. N. Kuruzovich, S. Gosain, R. Agarwal. 2007. Online infomediaries and price discrimination: Evidence from the auto-retailing sector. *Journal of Marketing*, 71 (3) 89-107.
- Vividence Corp. 2001. http://www.shop.org/learn/stats_hol2000_general.asp.
- Wellner, A.S. 2000. A new cure for shoppus interruptus. *American Demographics*, 22(8), 44.
- Wolfenbarger, M. and M. Gilly, 2001. Shopping online for freedom, control and fun. *California Management Review* 43(2), 34-55.
- Zettelmeyer, F., F.S. Morton and J. Silva-Risso. 2006. How the Internet Lowers Prices: Evidence from Matched Survey and Auto Transaction Data. *Journal of Marketing Research*, 43 (2), 168-181.

Table 1. User Demographic Characteristics

Variable	Mean	SD	Min	Median	Max
Age	45.81	14.67	11	45	89
Male	.46	.50	0	0	1
Some college education	.81	.40	0	1	1
High income (>\$50,000)	.33	.47	0	0	1
Medium income (\$25,000-\$50,000)	.34	.48	0	0	1

Table 2. Page View Level Descriptive Statistics

Variables	Mean	SD	Min	Med	Max
<i>C_{iq_t} : Shopping Cart Choices</i>					
Purchase	.01	.07	0	0	1
Add item to cart	.02	.15	0	0	1
Delete item from cart	.01	.05	0	0	1
Continue browsing without changing cart	.86	.34	0	1	1
Exit	.10	.30	0	0	1
<i>X_{iq(t-1)}¹ : Marketing Stimuli</i>					
<i>Cumulative price information</i> : number of presences of price information on page	23.97	53.11	0	6	441
<i>Cumulative pop-up promotion</i> : number of pop-up promotions	40.01	72.83	0	13	512
<i>Cumulative banner ad</i> : number of banner advertisements	1.72	4.54	0	0	34
<i>Cumulative email</i> : number of e-mail solicitations from B&N	0.99	5.53	0	0	81
<i>X_{iq(t-1)}¹ : Web Stimuli</i>					
<i>Cumulative hypertext links</i> : number of hypertext links on the webpage	598.7	1604.2	0	108	14207
<i>Cumulative pictures</i> : number of pictures on the webpage	8511	16476	0	127	119818
<i>X_{iq(t-1)}² : Users' In-Session Activities</i>					
<i>Cumulative page views at other competitive bookstores during session</i>	4.36	17.30	0	0	174
<i>Cumulative page views at non-category sites during session</i>	48.72	90.27	0	19	903
<i>Visit Depth</i> : number of webpage viewed	18.74	28.09	1	8	238
<i>Sign-in or not</i> : signed in to the account	.17	.37	0	0	1
<i>Time Duration</i> : time in seconds spent at last webpage	618.6	3453	0	0.4	39068
<i>Free shipping or not</i> : number of items in cart is two or more, to obtain free shipping from the site	.13	.34	0	0	1
<i>Z_{iq(t-1)} : User's Past Behavior</i>					
<i>Last Buy</i> : whether made a purchase in last session	.07	.25	0	0	1
<i>Weekend</i> : whether the visit is on a weekend	.28	.45	0	0	1

Table 3. Learning Performance and Model Comparisons

Model	Log Marginal Density	Estimation Sample		Holdout Sample	
		MAE	Hit Rate (%)	MAE	Hit Rate (%)
Aggregate model	-212147.86	0.36 (0.02)	31.03 (3.46)	0.32 (0.09)	26.30 (17.71)
One-state without endogeneity	-209218.04	0.06 (0.01)	81.77 (0.71)	0.14 (0.01)	67.81 (0.20)
One-state (SB)	-184260.96	0.06 (0.01)	85.73 (0.81)	0.13 (0.01)	75.24 (0.46)
MLSL model	-184086.98	0.06 (0.01)	88.18 (1.13)	0.13 (0.01)	78.43 (0.23)
Two-state	-180834.10	0.05 (0.01)	90.68 (0.77)	0.07 (0.01)	88.82 (0.17)
Three-state	-181908.75	0.06 (0.01)	89.45 (0.92)	0.08 (0.01)	84.34 (1.11)
Four-state	-183757.49	0.06 (0.01)	89.15 (1.25)	0.08 (0.01)	83.14 (1.14)

Table 4: Estimates for the Two-State HMM Model

		Low-Intent	High-Intent
Starting probabilities		0.75 (0.01)	0.25 (0.01)
Average waiting time intensity		0.37 (0.01)	0.28 (0.01)
Transition Matrix	Low-intent state	0	1
	High-intent state	1	0

Table 5. Estimates for the Shopping Cart Choice Model*

Variables	Low-Intent State				High-Intent State			
	Continue browsing	Remove item	Add item	Purchase	Continue browsing	Remove item	Add item	Purchase
Intercept	2.03 (0.10)	-12.91 (0.28)	-7.16 (0.36)	-15.57 (0.30)	2.19 (0.140)	-13.31 (0.23)	-5.96 (0.28)	-8.04 (0.31)
<i>Marketing Stimuli</i>								
Log Price Presence	-0.27 (0.24)	-0.23 (0.24)	-0.22 (0.23)	-0.23 (0.23)	0.54 (0.22)	0.53 (0.22)	0.53 (0.23)	0.54 (0.23)
Log Pop-up promotion	5.71 (0.20)	5.67 (0.20)	5.66 (0.20)	5.67 (0.21)	-7.40 (0.18)	-7.41 (0.18)	-7.41 (0.18)	-7.40 (0.18)
Log Banner ad	3.80 (0.32)	3.80 (0.32)	3.80 (0.32)	3.80 (0.32)	0.81 (0.41)	0.81 (0.41)	0.81 (0.41)	0.81 (0.41)
Log E-mail solicitation	2.49 (0.18)	2.49 (0.19)	2.49 (0.18)	2.49 (0.19)	0.68 (0.45)	0.68 (0.45)	0.68 (0.44)	0.68 (0.45)

<i>Web Stimuli</i>								
Log Product Links	0.12 (0.09)	0.11 (0.10)	0.13 (0.10)	0.10 (0.10)	0.22 (0.10)	0.22 (0.10)	0.23 (0.11)	0.22 (0.10)
Log Pictures	-0.23 (0.11)	-0.17 (0.12)	-0.12 (0.11)	-0.18 (0.12)	0.01 (0.11)	0.03 (0.12)	0.07 (0.12)	0.03 (0.12)
<i>Users' Comparison Shopping Activity</i>								
Log visits to other bookstore sites	-0.31 (0.11)	-0.31 (0.11)	-0.32 (0.11)	-0.32 (0.11)	-1.11 (0.18)	-1.10 (0.18)	-1.10 (0.18)	-1.10 (0.18)
Log visits to other non-bookstore sites	0.34 (0.15)	0.34 (0.16)	0.35 (0.16)	0.35 (0.16)	0.67 (0.20)	0.67 (0.20)	0.67 (0.19)	0.66 (0.20)
<i>Users' Past Behavior</i>								
Log visit depth	-0.29 (0.12)	-0.27 (0.12)	-0.25 (0.11)	-0.22 (0.12)	-0.16 (0.30)	-0.15 (0.31)	-0.16 (0.30)	-0.13 (0.30)
Log time duration	0.01 (0.06)	-0.08 (0.07)	-0.10 (0.08)	-0.08 (0.07)	-0.19 (0.09)	-0.23 (0.09)	-0.27 (0.10)	-0.25 (0.10)
Sign-in	-0.96 (0.20)	-0.97 (0.20)	-1.00 (0.20)	-0.95 (0.20)	0.04 (0.30)	0.04 (0.30)	0.02 (0.30)	0.05 (0.30)
Free shipping	2.44 (0.22)	2.47 (0.22)	2.50 (0.22)	2.49 (0.22)	4.20 (0.24)	4.23 (0.25)	4.25 (0.25)	4.24 (0.24)
Weekend	0.49 (0.17)	0.50 (0.17)	0.50 (0.18)	0.50 (0.18)	-0.19 (0.33)	-0.19 (0.33)	-0.19 (0.32)	-0.19 (0.33)
Last buy	-1.40 (0.24)	-1.41 (0.25)	-1.41 (0.25)	-1.41 (0.25)	-1.11 (0.31)	-1.11 (0.31)	-1.12 (0.31)	-1.12 (0.31)
<i>Interactions among Marketing Stimuli</i>								
Log price \times Log promotion	0.33 (0.05)	0.31 (0.06)	0.31 (0.05)	0.36 (0.05)	0.68 (0.06)	0.53 (0.06)	0.53 (0.06)	0.55 (0.06)
Log price \times Log ad	-1.19 (0.24)	-1.20 (0.25)	-1.20 (0.25)	-1.20 (0.24)	-0.76 (0.20)	-0.77 (0.20)	-0.77 (0.20)	-0.77 (0.20)
Log promotion \times Log ad	0.04 (0.17)	0.02 (0.17)	0.02 (0.17)	0.02 (0.17)	0.26 (0.14)	0.23 (0.15)	0.24 (0.15)	0.23 (0.15)

*Estimates in bold indicate that zero does not lie in the 95% posterior probability interval. For identification, the utility of exiting the site is normalized to 0, as the baseline choice.

Table 6: Purchase Elasticity Analysis

Variables		Overall	Low-Intent State	High-Intent State
Marketing Stimuli	Price	1.04	0.01	1.04
	Pop-up promotion	-2.08	1.04	-31.25
	Banner ad	1.04	1.04	0.41
	E-mail solicitation	7.29	7.29	0.01
Web Stimuli	Product links	1.04	1.04	1.05
	Pictures	0.01	0.01	1.05

Table 7. System Performance under Optimization

Data	Sessions	Sessions with Item in Cart (Purchase)	Metrics	Rates Before System Implementation	New Rates with System Implementation, After Page View*				
					1	3	5	7	9
Est.	906	119 (59)	Abandonment Rate	49.0%	16.6% (0.02)	15.2% (0.02)	15.6% (0.02)	19.4% (0.02)	19.6% (0.02)
			Conversion Rate	5.9%	12.8% (0.01)	13.8% (0.01)	12.6% (0.01)	11.5% (0.01)	11.2% (0.01)
Hold	798	63 (35)	Abandonment Rate	44.4%	20.3% (0.04)	19.5% (0.03)	20.4% (0.04)	21.2% (0.04)	23.4% (0.04)
			Conversion Rate	4.4%	8.3% (0.01)	8.4% (0.01)	8.3% (0.01)	7.9% (0.01)	7.8% (0.01)

*Numbers in the parentheses are standard deviations.

Table 8. Lab Experiment Results

	Control Group	Experimental Group
Number of participants	20	20
Latent intention identified	No	Yes
Low-intent state	N/A	61.66% (0.29)*
High-intent state	N/A	38.34% (0.29)
Dynamic webpage transformation	No	Yes
Shopping cart abandonment rate	60%	28.57%
Purchase conversion rate	10%	25%
Average effective learning time in page views	N/A	2.92 (0.90)
Average page loading time in seconds	0.40 (0.20)	0.48 (0.32)

*Numbers in the parentheses are standard deviations.

Figure 1. System Components and Logic Flow

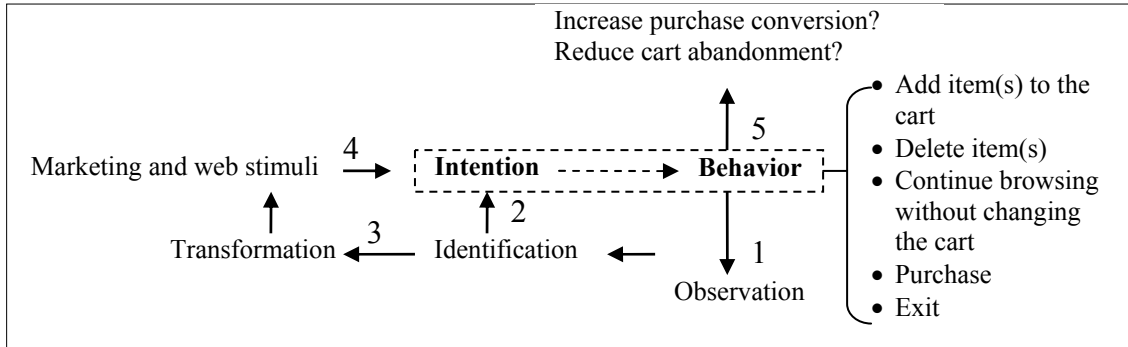
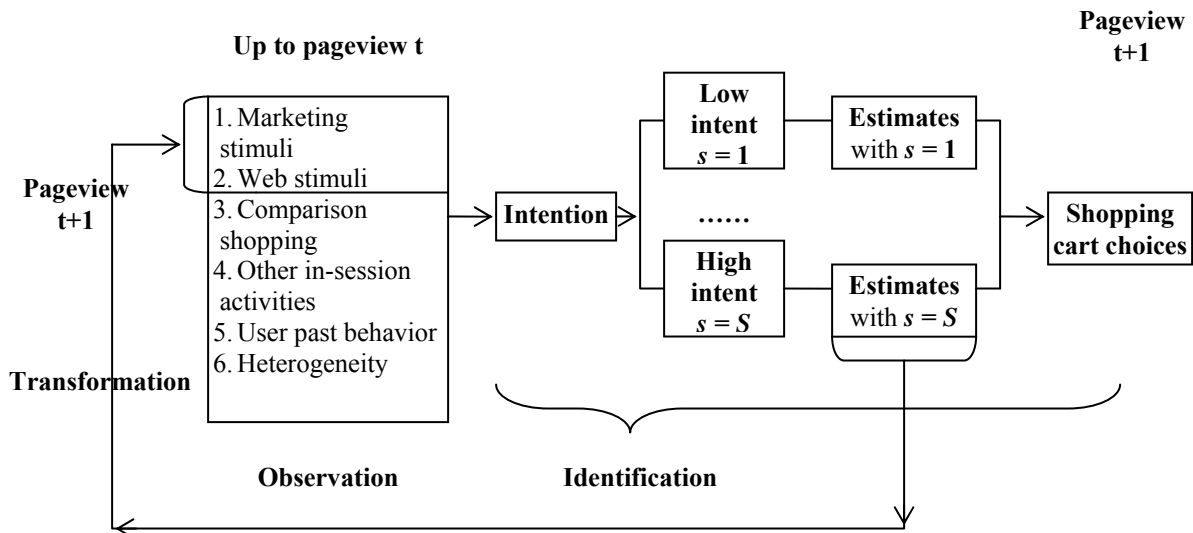


Figure 2. Flow Chart: Learning User Real-Time Intent with Concurrent Webpage Adaptation



Notes: The observation module observes each user's shopping/browsing behavior up to page view t on the site. The identification module performs real-time learning to identify the user's unobserved state of intent ($s = 1, \dots, S$). The parameters for each identified intent state are updated for $s = 1, \dots, S$, which indicates appropriate content adaptation for marketing and web stimuli at page view $t + 1$, which are automatically displayed to the user. The transformed page content at $t + 1$ affects the user's intent state at the next page view.

Figure 3. HMM Flow Chart

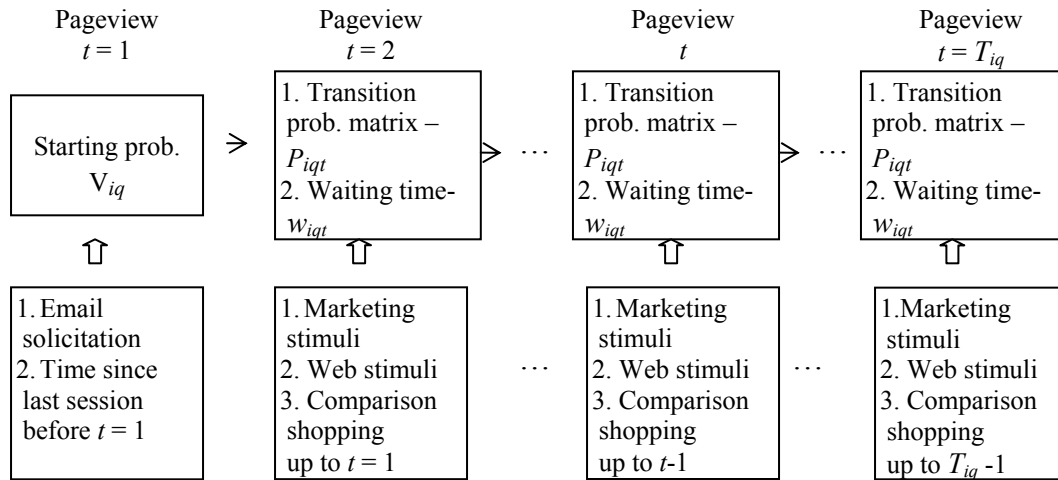


Figure 4a. Example of A Product Page without Page Transformation

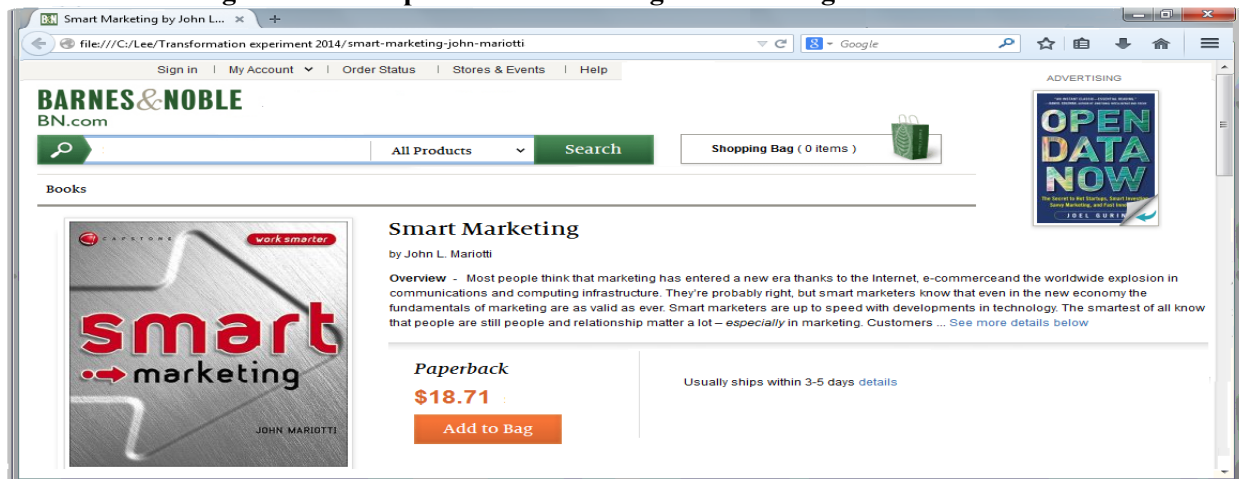


Figure 4b. Example of A Product Page with Optimal Page Transformation

