




3-2004

Dynamic Conversion Behavior at E-Commerce Sites

Wendy W. Moe

Peter S. Fader
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/marketing_papers

 Part of the Behavioral Economics Commons, Business Administration, Management, and Operations Commons, Business Analytics Commons, Business Intelligence Commons, E-Commerce Commons, Marketing Commons, Sales and Merchandising Commons, and the Technology and Innovation Commons

Recommended Citation

Moe, W. W., & Fader, P. S. (2004). Dynamic Conversion Behavior at E-Commerce Sites. *Management Science*, 50 (3), 326-335. <http://dx.doi.org/10.1287/mnsc.1040.0153>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/marketing_papers/267
For more information, please contact repository@pobox.upenn.edu.

Dynamic Conversion Behavior at E-Commerce Sites

Abstract

This paper develops a model of conversion behavior (i.e., converting store visits into purchases) that predicts each customer's probability of purchasing based on an observed history of visits and purchases. We offer an individual-level probability model that allows for different forms of customer heterogeneity in a very flexible manner. Specifically, we decompose an individual's conversion behavior into two components: one for accumulating visit effects and another for purchasing threshold effects. Each component is allowed to vary across households as well as over time. *Visit effects* capture the notion that store visits can play different roles in the purchasing process. For example, some visits are motivated by planned purchases, while others are associated with hedonic browsing (akin to window shopping); our model is able to accommodate these (and several other) types of visit-purchase relationships in a logical, parsimonious manner. The *purchasing threshold* captures the psychological resistance to online purchasing that may grow or shrink as a customer gains more experience with the purchasing process at a given website. We test different versions of the model that vary in the complexity of these two key components and also compare our general framework with popular alternatives such as logistic regression. We find that the proposed model offers excellent statistical properties, including its performance in a holdout validation sample, and also provides useful managerial diagnostics about the patterns underlying online buyer behavior.

Keywords

stochastic models, e-commerce, online purchasing conversion, buyer behavior

Disciplines

Behavioral Economics | Business | Business Administration, Management, and Operations | Business Analytics | Business Intelligence | E-Commerce | Marketing | Sales and Merchandising | Technology and Innovation

**Which Visits Lead to Purchases?
Dynamic Conversion Behavior at e-Commerce Sites**

Wendy W. Moe
Peter S. Fader

February 2001

Wendy W. Moe (wendy.moe@bus.utexas.edu) is Assistant Professor of Marketing at the University of Texas at Austin. Peter S. Fader (faderp@wharton.upenn.edu) is Associate Professor of Marketing at the Wharton School of the University of Pennsylvania. The authors would like to thank WeBI (the Wharton e-Business Initiative) and Media Metrix, Inc., for generously providing the data used in this study as well as Bruce Hardie for his very useful comments and suggestions. This paper stems from Wendy Moe's dissertation work, and she extends special thanks to the other members of her dissertation committee (Eric Bradlow, Barbara Kahn, Don Morrison, and David Schmittlein) and the Marketing Science Institute for supporting this research through the 1999 Alden Clayton Dissertation Award.

Which Visits Lead to Purchases? Dynamic Conversion Behavior at e-Commerce Sites

ABSTRACT:

This paper develops a model of conversion behavior (i.e., converting store visits into purchases) that predicts each customer's buying probability based on an observed history of visits and purchases. We offer an individual-level probability model that allows for cross-customer heterogeneity in a very flexible manner. We also allow visits to have a variety of effects in the purchasing process. For example, some visits are motivated by planned purchases while others are simply browsing visits -- requiring a model that would accommodate different visit-to-purchase relationships. Finally, consumers' shopping behavior may evolve over time as a function of past experiences, and we capture these effects as well. Our proposed Conversion Model decomposes an individual's purchasing conversion behavior into separate visit effects and purchasing threshold effects. Each component varies across households as well as over time. We apply this model to the problem of "managing" visitor traffic at a leading e-commerce site (Amazon.com). By predicting purchasing probabilities for a given visit, the Conversion Model can identify those visits that are likely to result in purchases. These premium visits may be re-directed to a server that will provide a better shopping experience while those visits that are less likely to result in a purchase may be considered as targets for promotions.

Which Visits Lead to Purchases? Dynamic Conversion Behavior at e-Commerce Sites

1. Introduction

As it is virtually costless for Internet shoppers to visit an e-commerce site, many online merchants experience large volumes of visitor traffic. Amazon, for example, had approximately 21 million unique visitors at their site just in the month of December 2000, according to Media Metrix, while Victoria's Secret occasionally experiences huge surges in traffic -- as many as 1000 visitors per second (Quick 1998). Consequently, an important issue facing e-commerce managers is how best to handle the large numbers of shoppers given limited server capacity.

One solution is to invest in server capacity, since too little capacity results in congested server traffic and provides users with a poor shopping experience. But too much capacity is expensive, and large investments in server capacity are poorly rewarded, since conversion rates of visits to purchases are uniformly low across the industry. Forrester (1999) reported that over 70% of online retailers experienced less than a 2% overall purchase conversion rate. In other words, over 98% of the visits that e-commerce sites must serve do not result in purchase transactions. Should e-commerce sites invest to ensure that the 2% of the visits resulting in purchases will be positive shopping experiences or invest less and hope that 2% of the visits will still result in a purchases despite any potential negative shopping experiences?

One way e-commerce sites have chosen to deal with this dilemma is to "manage" visitor traffic. For example, Victoria's Secret re-directs purchasing visits to faster servers while mere browsing visits are hosted by their less efficient and often congested servers (Quick 1998). But purchasing visits are only identified *after* the shopper places an item in the shopping cart. The shopping experience leading up to that event is still hosted by the slower server. The ability to

identify purchasing visits at the start of the session rather than after the customer has finished shopping and begun to complete the transaction, would offer a better overall shopping experience for the more promising visits, and therefore might increase the number of visits that proceed towards a successful purchase transaction.

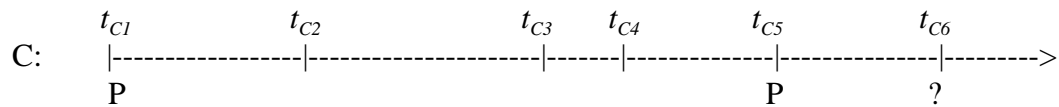
Because of the low returns to serving the large volume of visitor traffic, a related issue is how best to allocate marketing resources in order to improve conversion rates. Many e-commerce sites hold the philosophy that every visit is a buying opportunity and therefore try to induce purchases with promotions and discounts. However, offering promotions and discounts for all visits is inefficient. Many purchasers are likely to initiate a transaction without the added incentive of a promotion; offering these shoppers a promotion would be a poor use of resources. A more efficient option is to limit promotional offerings to those who were unlikely to buy rather than to extend an offer to all shoppers and reducing margins for the likely buyers.

The key in both of these applications is the ability to predict purchasing probabilities for a given visit. Those visits that are likely to result in a purchase need to be identified and the visitors possibly re-directed to a server that will provide a better shopping experience and increase the visitors' likelihood of buying. Those visits that are less likely to result in a purchase, without any added incentive, may be identified as targets for promotion.

With these (and other) resource allocation decisions in mind, we develop a model of conversion behavior that predicts each customer's probability of purchasing at a given visit based on that individual's observed history of visits and purchases. We offer an individual-level probability model that allows for customer heterogeneity in a very flexible way. We discuss and allow for the fact that visits may play very different roles in the purchasing process. For example, some visits are motivated by planned purchases while others are simply browsing

indicates that his purchasing probability is higher than individual A's. Therefore, when modeling conversion probabilities, we must consider the potential effect of visits in conjunction with any Bayesian updating of a customer's latent purchasing tendencies as we observe a series of visits and purchases over time.

In addition to the role past visits have on purchasing, past purchasing behavior may also affect future conversion behavior. Consider individual C with a visiting history identical to A.



The only difference is the timing of past purchases. How will this affect C's purchasing probability at t_{C6} ? One could argue that since a purchase just occurred at t_{C5} , person C is less likely to purchase at t_{C6} , implying that some sort of purchase cycle exists. However, individual C could be more likely to purchase if recent purchases are somehow more memorable and influential in reducing purchasing-related anxiety. For example, as a shopper repeatedly purchases from a site, loyalty (or inertia) may build, increasing future conversion probabilities. Furthermore, what would be the effect of a third purchase (at, say, t_{C3}) on future behavior?

The Conversion Model developed in this research will address three important issues that the above scenarios help illustrate. First, what is the effect of the number and timing of past visits on future purchasing behavior? Second, what is the effect of the number and timing of past purchases on future behavior? And third, do these effects change over time as customers gain more experience at a particular website? Past research has examined almost exclusively the issue of predicting future purchases with past purchasing behavior. This research will integrate observed visiting behavior into our model of purchasing while allowing for evolving behavior.

3. Model Development

Conversion Model Framework. There are four key components of the Conversion Model that allow us to accommodate the variety of nonstationary visit and purchase effects discussed in the previous section:

1. *Baseline probability of purchasing.* For each individual, there is a baseline probability of purchase at each visit. This baseline reflects the extent to which visits are purchase-directed and likely to result in purchasing regardless of any effects from past visits or purchases.

2. *Positive visit effect on purchasing.* Each visit has its own stochastic impact (assumed to be non-negative), and as the effects of these visits accumulate, the probability of purchase increases over time.² In other words, as a shopper makes more visits, she will be increasingly likely to purchase in subsequent visits, depending on the magnitude of these visit effects on purchasing. If and when a purchase occurs, this “bank” of visit effects is reset to zero, and visit effects begin to build up again upon the next visit opportunity. However, incremental visit effects may also be zero. In this case, there is no accumulated visit effect, and purchasing is driven primarily by the baseline effect.

3. *Negative purchasing threshold effect.* Purchasing propensity is negatively affected by an individual’s level of purchase-related anxiety toward a given retailer. For example, shoppers new to a site may be risk averse and reluctant to provide personal information, such as credit card numbers, home addresses, etc., to an unknown vendor as part of the transaction process. In

²This is similar in spirit (but quite different methodologically) to a model developed by [Lenk and Rao \(1995\)](#) for coupon redemption, in which the shopper visits the store a number of times before a coupon is redeemed. They conceptualize their model as store visits having effects on the shopper that are stochastically determined and accumulate until coupon redemption occurs. Although their conceptual framework can be applied to the relationship between store browsing and purchasing, it does not allow for a single shopper to undergo multiple transitions (i.e., more than one coupon redemption or, in our case, more than one purchase event), nor does it allow for evolving visit/purchase effects over time.

effect, shoppers have a threshold that must be overcome before a purchase will occur. Therefore, as a shopper visits a store, the associated visit effects are measured against a purchasing threshold construct which varies across individuals. When the accumulated effect of visits becomes high relative to this threshold, purchase occurs.

4. Evolving effects over time. The expected magnitudes of both visit effects and purchasing threshold may evolve over time, as the customer gains experience with the shopping environment. For example, subsequent visits to a website may have smaller effects on purchasing as the shopper gets used to the environmental stimuli and becomes less persuaded by content that has been seen repeatedly in the past. Therefore, our Conversion Model will allow for and provide a measure to characterize the trends that may be affected by past visiting. Additionally, purchasing thresholds may also evolve as a function of past purchasing experiences. For example, as shoppers make repeated purchases, the original reluctance to purchase from an unknown site may decrease, making future purchasing more likely. Many researchers (e.g., Beatty and Ferrell 1998) have found that customers are less resistant to purchase from a particular store in the future if they have already purchased from them in the past.

Alternatively, shoppers may be more likely to buy early on as it is a novel experience. As the novelty wears off with repeated purchases, visit effects may need to overcome an increasingly higher hurdle to induce a purchase. Again, the Conversion Model will allow for and provide a measure to characterize the trends that may be affected by past purchasing.

Beta-Bernoulli Model. We begin the formal specification of our Conversion Model by first introducing the simple beta-Bernoulli process, which is a natural benchmark model for us to build upon. The beta-Bernoulli, when applied to visiting and purchasing behavior, assumes that

at any given visit, customer i has a probability, p_i , of purchasing. Additionally, it assumes that these individual purchasing probabilities are beta distributed across the population to account for cross-customer heterogeneity.

$$f(x_i; p_i) = p_i \quad \text{and} \quad g(p_i) = \frac{1}{B(a, b)} p_i^{a-1} (1-p_i)^{b-1} \quad (1)$$

where $B(a, b)$ is the beta function, $B(a, b) \sim \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. Therefore, the unconditional probability of an individual buying at a given visit is then:

$$f(x_i) = \int_0^{\infty} p_i \cdot g(p_i) dp_i = E[p_i] = \frac{a}{a+b} \quad (2)$$

If a customer makes a sequence of visits, the probability of buying at each visit is typically assumed to be the same across visits, i.e., purchasing probability is a latent trait that may vary across customers but is stationary over time. If this assumption of stationarity were to hold (an assumption that our Conversion Model will later relax) purchasing behavior would not evolve over time, and past visits and purchases would have no impact on latent purchasing probabilities. But even in this stationary environment, our estimate of a customer's purchasing probability would be updated from visit to visit, using Bayes theorem:

$$E[p_{ij} | x_{ij}, n_{ij}] = \frac{a + x_{ij}}{a + b + n_{ij}} \quad (3)$$

where x_{ij} is the number of prior visits resulting in a purchase and n_{ij} is the total number of prior visits.

An important characteristic of a beta random variable is that it can be expressed as a ratio of gamma random variables. Specifically, if X and Y are both gamma-distributed random variables such that $X \sim \text{gamma}(a, \gamma)$ and $Y \sim \text{gamma}(b, \gamma)$, then the fraction $X / (X+Y)$ is by

definition beta distributed with parameters a and b (Johnson, Kotz, and Balakrishnan 1995, p.350). This is a fundamental relationship that will help us develop our Conversion Model.

Conversion Model. Starting with the basic beta-Bernoulli structure, we bring in the various elements of the aforementioned conceptual framework in the following manner. Let p_{ij} be the probability of customer i purchasing at visit j . This probability is a function of both visit impacts as well as a purchasing threshold. That is, each time a customer goes to the site, there is a non-negative stochastic impact that will accumulate toward a purchasing threshold. It follows that the customer's purchasing probability is determined by the net effect of these visit impacts relative to this threshold.

$$p_{ij} = \frac{\text{Net Effect of Visits Since Last Purchase } (V_{ij})}{\text{Net Effect of Visits Since Last Purchase } (V_{ij}) + \text{Purchasing Threshold } (\tau_{ij})} \quad (4)$$

We assume that the net visit effect, V_{ij} , consists of two components; a baseline propensity to buy (v_{i0}) that applies at every visit and the incremental effects (m_{ij}) from the visits that have occurred since the last purchase. Thus, V_{ij} is the sum of the baseline effect and all qualifying visit effects. In a customer's first purchasing cycle, for example, V_{ij} is equal to $v_{i0} + m_{i1} + \dots + m_{ij}$ for customer i who has made j visits.

If customer i has a purchasing threshold, τ_{ij} , her purchasing probability can then be written as:

$$p_{ij} = \frac{v_{i0} + m_{i0} + m_{i1} + \dots + m_{ij}}{v_{i0} + m_{i0} + m_{i1} + \dots + m_{ij} + \tau_{ij}} \quad (5)$$

To accommodate heterogeneity, let us assume that the baseline purchasing propensity v_{i0} is gamma distributed across the population with shape parameter r_v and scale parameter γ . Let us

also assume that the visit impacts, as well as purchasing thresholds, are heterogeneous and gamma distributed such that $m_{ij} \sim \text{gamma}(\mu, \gamma)$ and $\tau_{ij} \sim \text{gamma}(r_{\tau}, \gamma)$ ³. The resulting purchasing probability is then a ratio of two gamma random variables which, as noted earlier, is a beta-distributed random variable.

Evolving Visiting Effects. Thus far, we have assumed that the gamma distribution governing the impact of visits remains stationary over time with parameters μ and γ . In other words, while the impact of successive visits may vary over time, we have not allowed for any trends in this stochastic process. However, it is likely that, as customers familiarize themselves with the store site, visit effects (m_{ij}) might change systematically. We therefore extend the model to allow for the possibility that the influence of store visits will increase, decrease, or stay the same depending on how familiar the individual is with the site. We approximate the customer's experience with the site with the number of times she has previously visited the site and implement the evolutionary trend through the shape parameter governing the incremental visit effects, m_{ij} (Schmittlein and Morrison 2000). Therefore, we assume:

$$m_{ij} \sim \text{gamma}(\mu_j, \gamma) \quad \text{where} \quad \mu_j = \mu_0 k^j \quad (6)$$

and thus the net effect of visits for the first purchase cycle then becomes:

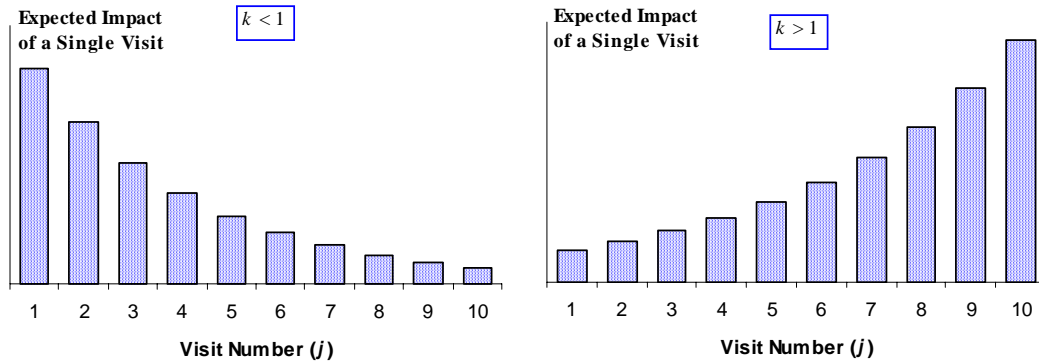
$$V_{ij} \sim \text{gamma}(r_v, \gamma) + \text{gamma}(\mu_0 k^1, \gamma) + \text{gamma}(\mu_0 k^2, \gamma) + \dots + \text{gamma}(\mu_0 k^j, \gamma) \quad (7)$$

The parameter k ranges from 0 to infinity and characterizes how visit impacts evolve as consumer familiarity increases. If k equals 1.0, there is no evolutionary effect; the stochastic process governing m_{ij} is a simple stationary one. If k is less than 1.0, visits tend to become less influential over time, while if k is greater than 1.0, visits tend to become more influential as

³Since the sum of the baseline effect and the visit impacts are compared against the threshold to determine purchasing propensity, it is reasonable to assume that they are measured on the same scale, and thus the distributions governing these three effects share the same scale parameter γ .

consumers evolve (Figure 1). But despite the upward or downward trend on the shape parameter, each successive draw of m_{ij} is still a random variable, thereby allowing any particular visit to have an unusually high or low impact. This allows for the possibility of impulse purchases, albeit with different probabilities, at any given visit.

Figure 1. Evolving Visit Effects



Across multiple purchases, (7) generalizes to:

$$V_{ij} \sim \text{gamma} \left(r_v + \sum_{u=lp+1}^i \mu_0 k^u, \gamma \right) \quad (8)$$

where lp indicates the visit during which the last purchase occurred. If customer i has not yet been observed to make a purchase, then all past visits would contribute to V_{ij} , or $lp = 0$. In summary, we assume that the net visit impact is driven by a sum of gamma-distributed random variables yielding a net visit impact that is itself gamma distributed.⁴

Figure 2 (a&b) illustrates how the net effect of visits can accumulate over a customer's history. Assuming that visit effects evolve such that $k > 0$, incremental visit effects will stochastically vary about a mean, which itself is evolving (see Figure 2a). From visit to visit,

⁴A discount factor was also tested to differentially weight past visit effects depending on the amount of time elapsed (i.e., days) between visits. However, discounting proved to offer only negligible improvements in model performance and was therefore omitted from the model.

these incremental effects accumulate until a purchase is made, at which time the net effect is reset to zero (see Figure 2b).

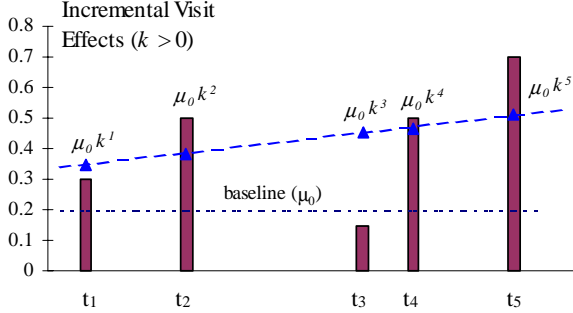


Figure 2a. Incremental Visit Effects

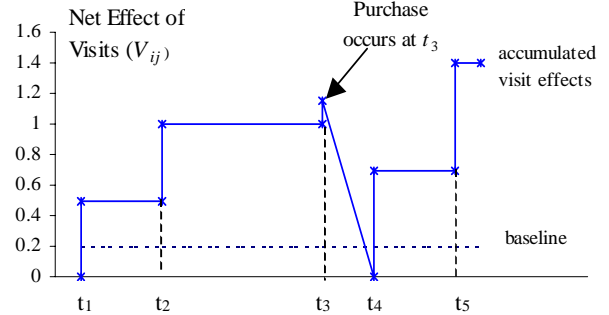


Figure 2b. Net Effect of Visits

Evolving Purchasing Thresholds. Having completed our discussion of how we model visit effects, i.e., the numerator of equation (4), we now turn our attention to the remaining term in the denominator regarding purchasing thresholds. Initially, the purchasing threshold for customer i is τ_{i0} , which we specify as a gamma-distributed random variable with shape parameter r_τ and scale parameter γ to account for cross-customer heterogeneity. However, over time, a customer's purchasing threshold may evolve as a function of her observed behavior, specifically, her past purchasing experiences. In much the same way that visit impacts can evolve over time, we implement the evolution of purchasing thresholds through the shape parameter:

$$\tau_{ij} \sim \text{gamma}(r_\tau \exp\{\psi \cdot \text{PUR}_{ij}\}, \gamma) \quad (9)$$

where r_τ captures the initial purchasing threshold, ψ is a parameter that governs the magnitude and direction of the dynamic process, and PUR_{ij} represents the number of purchases that customer i has made up to (but not including) visit j . This specification for the shape parameter allows the threshold to either increase, decrease, or remain constant depending on the direction of the evolutionary parameter, ψ . If ψ equals zero, the purchasing threshold distribution remains stationary with a shape parameter of r_τ regardless of past purchasing experiences. If, however, ψ

is less than zero, thresholds decline as the customer gains purchasing experience with the retailer, and she becomes more likely to buy at future visits.⁵

Figure 3 (a&b) illustrate how the purchases threshold can evolve as a function of past purchasing events. Notice that depending on ψ , this threshold may either increase or decrease (Figure 3a). Assuming $\psi < 0$ such that purchasing increases future conversion probabilities, the threshold changes after each purchase as seen in Figure 3b. Whether or not a purchase occurs is the stochastic result of a conversion probability derived from the combination of visit effects (Figure 2b) and the purchasing threshold (Figure 3b).

Figure 3a. Threshold Dynamics

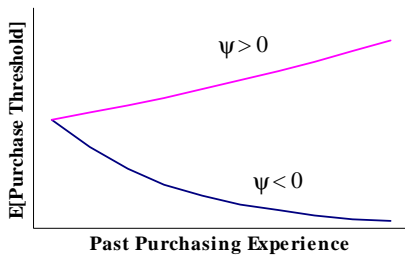
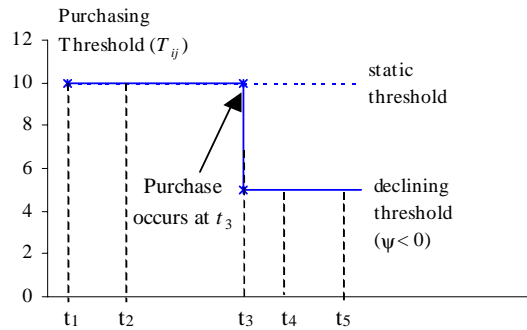


Figure 3b. Evolving Purchasing Threshold



4. Estimation

Two key components determine an individual's purchasing probability at a given visit, (1) the net effect of past visits and (2) purchasing thresholds. Both of these components are assumed to be gamma-distributed random variables to account for heterogeneity. Consequently, the probability of purchase is calculated as $V_{ij} / (V_{ij} + \tau_{ij})$, where $V_{ij} \sim \text{gamma}(Q, \gamma)$ and $\tau_{ij} \sim \text{gamma}(R, \gamma)$, where Q and R represent the expressions for the shape parameters shown in

⁵We also tested a model that included a discount factor to incorporate the actual timing of past purchases. Like the inclusion of the discount rate for visits, discounting past purchases did not add to our model's fit or predictive power.

equations (8) and (9), respectively. As noted earlier, this combination of gamma random variables can be reexpressed as a beta distribution with parameters Q and R .

To help illustrate the complete model and the parameter estimation process, consider customer A described in §2. With no prior information, an initial estimate of the buying probability at visit t_{A1} would simply be:

$$p_{A1} \sim \text{beta}(Q_{A1}, R_{A1}) \quad \text{and} \quad E[p_{A1}] = \frac{Q_{A1}}{Q_{A1} + R_{A1}} \quad (10)$$

$$\text{where} \quad \begin{aligned} Q_{A1} &= r_v + \mu_0 k \\ R_{A1} &= r_\tau \end{aligned}$$

But after observing a purchase at that time, we can update the buying probability much like we do in (3) for the beta-Bernoulli model. But in addition to Bayesian updating, which occurs even in a stationary environment, the shopper's fundamental behavior is also evolving from visit to visit. This is captured by changes in the beta parameters, Q_{ij} and R_{ij} , according to (8) and (9).

The probability that consumer A buys during her second visit then becomes:

$$p_{A2} \sim \text{beta}(Q_{A2}+1, R_{A2}+1) \quad \text{and} \quad E[p_{A2}] = \frac{Q_{A2}+1}{Q_{A2}+R_{A2}+1} \quad (11)$$

$$\text{where} \quad \begin{aligned} Q_{A2} &= r_v + \mu_0 k^2 \\ R_{A2} &= r_\tau \exp\{\psi\} \end{aligned}$$

If the second visit ends without a purchase, then the probability of buying in the third visit, after having observed the two previous visit outcomes, is updated as follows:

$$p_{A3} \sim \text{beta}(Q_{A3}+1, R_{A3}+2) \quad \text{and} \quad E[p_{A3}] = \frac{Q_{A3}+1}{Q_{A3}+R_{A3}+2} \quad (12)$$

$$\text{where} \quad \begin{aligned} Q_{A3} &= r_v + \mu_0 k^2 + \mu_0 k^3 \\ R_{A3} &= r_\tau \exp\{\psi\} \end{aligned}$$

Since a purchase occurred at the third visit,

$$p_{A4} \sim \text{beta}(Q_{A4}+2, R_{A4}+3) \quad \text{and} \quad E[p_{A4}] = \frac{Q_{A4}+2}{Q_{A4}+R_{A4}+3} \quad (13)$$

$$\text{where } Q_{A4} = r_v + \mu_0 k^4$$

$$R_{A4} = r_\tau \exp\{2\psi\}$$

From this inductive illustration, we jump right to the complete likelihood function, recognizing that, for notational convenience (as shown in (2)), we can use the expressions for $f(x_i)$ and $E[p_i]$ interchangeably.

$$L = \prod_{i=1}^N \int \int \dots \int \prod_{j=1}^{J_i} p_{ij}^{I_{ij}} \cdot (1-p_{ij})^{(1-I_{ij})} dp_{ij} = \prod_{i=1}^N \prod_{j=1}^{J_i} E[p_{ij}]^{I_{ij}} \cdot (1-E[p_{ij}])^{(1-I_{ij})} \quad (14)$$

where I_{ij} is a 0/1 variable indicating whether or not customer i actually makes a purchase at visit j .

The final component of the model allows for a segment of “hard-core never buyers.” In many situations, there may be a set of shoppers who visit a store to look around but have absolutely no intention of ever buying anything there. We will assume that these shoppers comprise a fraction, $(1-\pi)$, of the population. This parameter will affect our expectation of p_{ij} such that if a consumer’s history contains no purchases ($x_{ij} = 0$),

$$E[p_{ij} | x_{ij} = 0] = (1 - \pi) + \pi \frac{Q_{ij}}{Q_{ij} + R_{ij} + n_{ij}} \quad (15)$$

If, however, consumer i had purchased in the past and therefore proven herself not to be a never buyer, then the expected probability of buying for all visits following that purchase would be:

$$E[p_{ij} | x_{ij} > 0] = \frac{Q_{ij} + x_{ij}}{Q_{ij} + R_{ij} + n_{ij}} \quad (16)$$

Parameter estimation is performed using ordinary maximum likelihood procedures. We utilize the MATLAB programming language on a standard desktop PC. In this setting, the complete model requires less than ten minutes to obtain optimal estimates for its parameters. This estimation procedure is quite robust; we have seen no evidence of local optima or any other irregularities.

5. Data

We use clickstream panel data collected by Media Metrix, Inc. Media Metrix maintains a panel of approximately 10,000 households whose Internet behavior is recorded over time. The data contains information regarding when each household visits a given site. Additionally, the data include the precise day and time when a specific URL was viewed. To consolidate the data, we aggregate visits to the daily level. Any session in which the individual views a URL with the on-line store's domain name is considered a visit to that store. If a given household visited a store multiple times in a single calendar day, that is coded as one visit on the day in which the session began. For our purposes, we examine the panel's shopping behavior at a leading online bookstore, Amazon.com. From March 1, 1998 through October 31, 1998, there were 4,379 panelists who made at least one visit making a total of 11,301 visits.

Purchase is defined as any visit during which a purchase occurred. For many online stores, buyers are linked to a specific web page that acts as a receipt or purchase confirmation after an order has been submitted, including "one-click" purchases. Those visits in which the panelist saw the "confirm-order" page of the store's website were identified as purchase visits. The number of units purchased and the total amount spent were not considered in this analysis.

Of the 4,379 panelists who visited the store in our data period, only 851 households bought during that time, but they made a total of 1,573 purchases. This results in an overall conversion rate of 13.92% (1,573/11,301) -- a very high conversion rate for online retailers who tend to have conversion rates below 5% in most circumstances. However, this overall conversion rate says nothing about how buying propensities differ across shoppers or how it changes over time for a given shopper. Because average conversion rates tend to be so low, it is vital to identify those visits (and visitors) with higher likelihoods of conversion.

Table 1 summarizes the visiting and purchasing behavior at the online bookstore and how it appears to be changing over time. All measures seem to indicate that site performance is improving from the first four-month period to the next. The conversion rate is increasing, the number of visitors as well as buyers is increasing, and so on. But these aggregate measures are misleading as the inflow of new shoppers and the dropout of existing shoppers may mask the true underlying dynamics that are occurring at the individual level (Moe and Fader 2001). If we examine the same statistics while accounting for the entering and exiting of shoppers, a very different pattern emerges. Table 2 shows the conversion rate statistics for only those shoppers who appeared to be active at the store throughout the entire data period, namely those who made a visit to the store in both the first two months and the last two months of the data period. This illustrative sub-sample avoids any problems due to censoring and thus provides a better view of individual-level dynamics. The behavior of these households during the four intermediate months are then compared.

Table 1. *Summary of Bookstore Visiting and Purchasing*

	All 8 months	Months 1-4	Months 5-8
Number of visitors	4,379	2,693	2,717
Number of buyers	851	468	531
Number of visits	11,301	5,402	5,899
Number of purchases	1,573	705	868
Conversion rate	13.9%	13.1%	14.7%
Visits/visitor	2.58	2.01	2.17
Purchases/buyer	1.85	1.51	1.63
Purchases/visitor	0.36	0.26	0.32

Table 2. *Conversion Rate Summary for Active Bookstore Shoppers Only*

	Months 3-4	Months 5-6
Conversion rate	26.0%	20.8%
Number of visits	757	472
Number of purchases	197	98

In general, conversion rates for this subset of the population are much higher than those seen in the population as a whole as more active shoppers tend also to be more likely buyers (Moe and Fader 2001). But, contradicting the increasing conversion rates for the entire sample as seen in Table 1, this group's conversion rates are actually decreasing over time. Therefore, without modeling behavior at the individual level, e-commerce managers may be drawing incorrect conclusions. The Conversion Model is able to capture these individual-level patterns,

and therefore provides a better indication of differences across households as well as the dynamics over time.

6. Results

As a start, the full six-parameter Conversion Model was estimated on the entire eight months of bookstore data (see Table 3, row 1). From the model results, there are three main dynamics we are looking to identify. First is the influence of visits. Do visits have some effect on conversion probabilities or is most purchasing driven by a strong baseline effect? The model seems to indicate that visits do have effects, above and beyond the baseline, that accumulate and increase purchasing probabilities as indicated by a relatively large $\mu_0 = 0.276$ when compared to the baseline, $r_v = 0.062$.

Table 3. *Parameter Estimates for Conversion and Nested Models at Bookstore*

	r_v	μ_0	k	r_τ	ψ	π	LL	BIC
1. Full model*	0.062 (0.021)	0.276 (0.026)	0.932 (0.014)	2.314 (0.086)	0.117 (0.016)	0.790 (0.024)	-4264.25	8578.81
2. Visit effect accumulation only	0.099 (0.038)	0.436 (0.205)	1	4.047 (0.018)	0	0.904 (0.029)	-4279.06	8600.04
3. Visit effect accumulation and evolution	0.155 (0.037)	0.317 (0.035)	0.918 (0.016)	3.159 (0.177)	0	0.777 (0.025)	-4266.17	8574.26
4. Threshold dynamics only	1.348 (0.072)	0	1	8.782 (0.120)	0.234 (0.032)	0.868 (0.040)	-4299.44	8632.42
5. Beta-Bernoulli	0.613 (0.015)	0	1	4.436 (0.040)	0	1	-4308.03	8632.83

NOTE: The numbers in bold indicate the values at which the parameters were fixed, and standard errors for the parameter estimates are presented in parentheses.

* Selected Model

Second is the adaptation effect. That is, does the incremental effect of each visit systematically evolve as the shopper gains experience? In this case, k is less than one, suggesting that subsequent visits have a diminishing (but still positive) impact on purchasing behavior as the shopper makes more visits to the site.

Changes in purchasing thresholds, or the effect of past purchases, is the final dynamic that we examine. From the full model estimated, it seems that purchasing thresholds increase as a function of discounted purchasing experiences ($\psi=0.117$), and thus a consumer is less likely to re-purchase soon after a transaction occurs. This result should be rather very discouraging to many online retailers who believe that online purchasing experiences will be so positive that future purchasing will be more likely over time.

Overall, conversion probabilities are driven largely by the accumulation of visit effects suggesting that there are also interesting dynamics occurring among the panelists over time. Visits have a smaller effect on purchasing probabilities as consumers adapt to the shopping environment. Also, purchasing thresholds are increasing as a function of past purchasing experiences. Both of these dynamics suggest that conversion probabilities are decreasing over time, as a function of past visit experiences and as a result of increased purchasing experiences -- contradicting the trends suggested by the aggregate conversion rate measures (but supporting the type of data pattern shown in Table 2). This bodes very poorly for the future of this website if such dynamics were to continue.

Table 3 also shows the parameter estimates and log-likelihoods of several nested models, including the static, two-parameter, beta-Bernoulli model⁶. Additional likelihood ratio tests were conducted to identify the model that fit significantly better than the others given the number of

⁶A beta-Bernoulli model that allowed for a segment of hard-core never-buyers was also estimated but did not improve upon the fit of the simple two-parameter beta-Bernoulli.

parameters it used. Though two models were comparable in terms of BIC (the full model and the specification that ignores threshold dynamics), a likelihood ratio test shows that the inclusion of the purchasing threshold evolution does add to model fit in a statistically significant manner ($p < 0.05$).

For completeness, we also compare the Conversion Model results to those of a logistic regression incorporating recency and frequency measures as explanatory variables. Specifically, we model purchasing in each session as a function of (1) the number of past visits, (2) the number of past purchases, (3) the number of visits since the last purchase, (4) time elapsed (in days) since the last visit, and (5) time elapsed (in days) since the last purchase. The results of the logistic regression are presented in Table 4. The fit of the model is vastly inferior to that of the Conversion Model.

We also estimated a latent-segment logistic-regression to better accommodate customer heterogeneity. When we expand the model to include multiple segments, we find only two distinct segments: one large segment (72%) that responds very little to past purchases and one (28%) that is very sensitive to past purchase behavior (i.e., if you have purchased in the past, you are more likely to purchase again in the future). Adding a third segment does not add appreciably to the fit of the model. Although the latent-segment model improves upon the single segment logistic regression, it still provides a far poorer fit than the Conversion Model.

Table 4. Logistic Regression Results

	Single Segment Logistic Regression	Two-Segment Logistic Regression $w = 0.717 (0.071)$	
Constant	1.944 (0.037)	2.864 (1.041)	1.126 (0.249)
# of past visits	0.066 (0.011)	0.314 (0.147)	0.240 (0.031)
# of past purchases	-0.369 (0.031)	-0.685 (0.118)	-2.173 (0.228)
# of visits since last purchase	0.003 (0.019)	-0.241 (0.451)	-0.224 (0.051)
Time (days) since last visit	-0.000 (0.001)	-0.003 (0.018)	0.001 (0.002)
Time (days) since last purchase	-0.007 (0.001)	0.003 (0.292)	-0.011 (0.003)
LL	-4367.79		-4285.50
BIC	8791.55		8692.30

*Standard errors are presented in parentheses

7. Predicting Purchases and Model Validation

As a validation and forecasting exercise, we estimated a variety of models on the first four months of data. We then identify those individuals who continued to visit at least once in the next four months. Given the model results, we predict the probability of purchase for these individuals' first actual visit after the model estimation period. This is a challenging basis for model testing, since the use of a single holdout purchase does not allow over- and under-predictions to "average out" across visits at the individual level, as would be the case for more aggregate measures such as a "hit rate." Besides our Conversion Model, we considered four viable benchmarks:

Benchmark 1 - Historical Conversion Rates. The first benchmark is a simple projection of historical (static) conversion rates -- the current “state of the art” among e-commerce managers. An individual’s predicted buying probability in future visits is simply the total number of past purchases in the estimation period divided by the total number of visits in the same period. Though this method is very simple and easy to implement, it is severely limited in its ability to accurately predict behavior for relatively inactive individuals. For example, an individual who made one purchase in the one visit observed in the estimation period would be said to have a predicted conversion rate of 100% for future visits, and unlikely outcome.

Benchmark 2 - Beta-Bernoulli. The second benchmark model is the beta-Bernoulli which is more sophisticated than the historical conversion rate as it allows for stochastic outcomes of a latent conversion process. In other words, the individual who made one purchase out of one visit would not have a 100% conversion rate; instead, each household’s projected future conversion probability is a Bayesian shrinkage estimate (equation 3) that combines the parameters of the population’s beta heterogeneity distribution and each household’s own distinct visiting and purchasing history.

Benchmark 3 - Logistic Regression. Both the one- and two-segment models are used to predict future buying probabilities. Forecasts based on the one-segment logistic regression are straightforward. Two-segment logistic regression forecasts involve calculating the posterior segment weights for each individual given behavior in past sessions, using a direct application of Bayes rule. The predicted buying probability for a given visit is then a weighted average of the estimated buying probabilities under each segment.

Conversion Model. Finally, we compare these benchmarks to our dynamic Conversion Model. Like the beta-Bernoulli model, the Conversion Model accommodates variance around

each shopper's latent tendencies, and utilizes Bayesian updating to provide more accurate estimates as we observe more data for each shopper. But additionally, the Conversion Model incorporates nonstationarity in conversion probabilities. As a result, those shoppers who may have been "written off" by the competing benchmark models because of low observed conversion probabilities in the past would not be so easily dismissed by the Conversion Model. Instead, depending on the pattern of past visits and purchases, some of these shoppers may be building up to a future purchase. For example, of the 1022 panelists who made visits in both the model estimation period and the forecasting period, 759 of them did not buy at all in the first four months. Of these 759 shoppers, 10.7%, or 81 individuals, actually made purchases in their next visits. Using historical conversion rates, all of these 759 shoppers would have been dismissed as having 0% conversion probabilities in the future. The beta-Bernoulli model, with its more lenient shrinkage estimates, predicts that 9.1% of the observed non-buyers will buy in their next visit. But by allowing for the accumulation of visit effects, the Conversion model mirrors the actual future conversion most closely, estimating that 11.1% of the historical non-buyers will actually buy in the next visit. This is significantly different from the prediction offered by the beta-Bernoulli ($p < 0.001$) but not significantly different from the actual percent of future buyers ($p = 0.365$).

Overall, of the 1022 next visits for which we predict purchasing, 160 visits actually resulted in a purchase, leading to a 15.7% conversion rate. The average predicted purchasing probability across these 1022 visits as calculated by the Conversion Model is 14.7%, only a 6.3% error. Compared to the average purchasing probabilities according to the historical conversion rates (13.8%), the beta-Bernoulli (13.0%), the single-segment logistic regression (19.2%), and the two-segment logistic regression (14.2%), the Conversion Model provides the most accurate

overall conversion predictions. But it appears to be inferior to the two-segment logistic regression, which predicts an overall conversion rate of 15.9%.

However, these conversion rates are an average of purchasing probabilities across all active panelists' next visits, while our objective is to accurately predict the outcome of each individual visit. Therefore, to better evaluate forecasting accuracy across individual visits, we calculate a weighted root mean squared error (RMSE) that compares predicted buying probabilities at the individual-level to actual conversion rates. Specifically, for all integer values of y , we identify all sessions that were predicted to have a $y\%$ of buying, where $y\%$ is represented by the range between $(y-0.5)\%$ and $(y+0.5)\%$. We then calculate the actual conversion rate among the visits at each percentage level. In other words, what percent (z_y) of those n_y visits that were predicted to have a $y\%$ purchasing probability actually bought? But since the conversion rate at each percentage level is derived from an uneven number of visits, we calculated a weighted RMSE, weighted by the number of visits, n_y , observed at each percentage level, y .

$$\text{Weighted RMSE} = \sqrt{\frac{\sum_{y=0}^{100} n_y \cdot (y - z_y)^2}{\sum_{y=0}^{100} n_y}} \quad (17)$$

Comparing across models, we find, as expected, that the weighted RMSE is poor when using the historical conversion rate to predict future buying (RMSE = 14.4%). The beta-Bernoulli is a significant improvement over the historical method with an RMSE of 9.9%. The logistic regressions, while faring rather well in predicting overall conversion rates, are far worse when we evaluate the predictions in this manner (single-segment RMSE = 24.5%; two-segment

RMSE = 17.7%). The Conversion Model, in sharp contrast, offers the best predictive accuracy with an RMSE of 8.9%.

Having shown that the proposed Conversion Model consistently outperforms the other models on a *relative* basis, we now provide one final statistical test to demonstrate that its estimated probabilities are valid on an *absolute* basis as well. This test, adapted from Barnett, et al (1981), provides a formal assessment of whether or not the probabilities match up properly with the observed outcomes.

Barnett, et al (1981) sought to determine the predictive accuracy of a diagnostic model of Hodgkin's disease. In their setting, patients were predicted as being in one of four stages of the disease. Therefore, assignment of patients into each of these four stages is discrete while the predictive probabilities are continuous, similar to comparing a continuous buying probability to a discrete event of purchase or no purchase. Because we are dealing with only two possible outcomes (stages) we can simplify the Barnett test as follows:

1. Sort the predicted purchasing probabilities for each visit in non-decreasing order. Let us define the largest as p_1 , the second as p_2 , and the smallest as p_N , etc.
2. Now the visits are divided into groups within which numerical predictions are close together. That is, begin forming a first group with the visit with prediction p_1 , the visit with prediction p_2 , etc. Stop upon reaching K , the smallest number for which $\sum_{i=1}^K p_i$ is at least five.
3. Starting with the next lowest prediction, construct a second group in an analogous way using the same stopping rule. Proceed similarly until as many groups as possible are formed.⁷
4. For each group, calculate the total number of actual visits that resulted in a purchase. This number can then be compared to the predicted number of purchases for the group, $\sum p_i$ in a standard chi-squared test.

⁷If the final (remainder) set of observations is not large enough to form a full group they should be added into the last group formed.

When we apply this test to the Conversion Model predictions, we get a chi-squared value of 15.17, or a p -value of 0.232, indicating that the individual-level predictions offered by the Conversion Model provide very good indicators of actual future purchasing. In general, to obtain such a close fit for a set of probabilities is a good achievement, but to do so for a set of probabilities in a *holdout* period is a strong testimonial for the model's capabilities.

Having established the validity of the Conversion Model in so many different ways, we are comfortable to rely upon it as the methodological basis to perform the type of resource allocation task mentioned at the outset of the paper.

8. Identifying High Purchasing Probability Visits

Having conceptualized, developed, and estimated a model that accommodates the different dynamics occurring across visits, we now return to the managerial decisions that originally motivated this modeling exercise. Our goal in this section is to evaluate the model's ability to accurately identify likely purchase visits based only on the visitor's behavioral history to date. This will help the e-commerce manager target the appropriate customers who may benefit from having access to faster servers as well as those who may be good (or bad) targets for short-term promotional efforts.

To test the value of using the Conversion Model to identify high purchasing probability visits, we rank visits by their conversion probabilities as predicted by the historical conversion rate, beta-Bernoulli model, logistic regression (one- and two-segment models), and the Conversion Model. Specifically, we use each model to separately identify the top-ranking 10% of customers who have the highest probabilities of purchase at that point in time. Across all five models, a total of 138 different panelists are included in one or more of these top-ranked subsets.

It is interesting to note that *none* of the panelists were categorized as being in the top 10% by all five methods and 86 of them were included in the top 10% by only one method. In other words, the five methods are very distinct in identifying the valuable panelists.

Table 6 provides the *actual* conversion rate of the top 10% of customers, according to each method. The top 10%, according to projections based on historical conversion rates, resulted in only 29.3 purchases per 100 households. Though this conversion rate is better than the 15.7% for the population as a whole, it is inferior to the conversion rates that result from using either the beta-Bernoulli or the Conversion Model. Specifically, 33.0% of the top customers, according to the beta-Bernoulli, actually made a purchase while 37.0% of the top Conversion Model customers actually purchased. In other words, if an e-commerce manager uses the Conversion Model to identify the high purchasing probability customers as premium visitors, the return in terms of conversion rates is 12.8% higher than targeting those identified by the static beta-binomial model. The Conversion Model also outperforms both logistic regression models in its ability to identify the more promising visits.

Table 6. *The Top 10%'s Actual Conversion Rates at Bookstore*

Top 10% according to...	Actual Conversion Rate
Proposed Conversion Model	37.0%
Beta-Bernoulli model	33.0%
Logistic regression (1 seg)	33.0%
Historical conversion rates	29.3%
Logistic regression (2 seg)	28.0%

9. Conclusions

The Internet has provided e-commerce managers with an over-abundance of data and metrics. The objective of this paper was to investigate one of these metrics – conversion rates. Specifically, the model presented in this paper allows for a more careful and useful examination of conversion behavior than can be provided by a simple aggregation of the number of visits and purchases. We illustrated that aggregate summary measures can offer highly misleading results and conclusions. The Conversion Model avoids this mistake by directly addressing heterogeneity across consumers as well as dynamics over time. Because consumers have different reasons for visiting a web store, it is important to understand and account for various patterns in the relationship between visiting and purchasing. These patterns are often overlooked but are addressed carefully in our Conversion Model.

Understanding how conversion probabilities vary across consumers and change from visit to visit is valuable information that can allow e-commerce managers to better treat each visitor based on his/her past behavior. The patterns of visits and purchases may reveal how best to serve a given visitor. E-commerce managers should ensure that the high purchasing probability visits are positive shopping experiences. Those visits less likely to result in a purchase may be better targets for a promotion.

The research problem of examining conversion probabilities is very complex, and several issues are still unexplored. For example, we have ignored the activity *within* each visit. The sequence of pageviews (e.g., duration, type of pages examined, etc.) could have a great influence on changing conversion probabilities and the likelihood that a consumer will buy in any given visit. Once the process of evolving conversion behavior within a single visit is better understood,

we can then begin to investigate the effectiveness and efficiency of promotional tools offered at different moments of the consumer's store visit.

Methodologically, the Conversion Model allows for heterogeneity and various different effects of store visits. However, it is unable to characterize each individual by their dominant shopping behavior, i.e., searcher, browser, directed buyer, etc. Therefore, one area for potential future research is to extend this model to utilize hierarchical Bayesian methods. This would allow marketers to differentiate among different types of shoppers and thus design different marketing campaigns to better suit each consumer.

Additionally, this paper has focused exclusively on conversion probabilities without addressing the issue of modeling visiting patterns. Such a model is presented in Moe and Fader (2001), but a promising future project would be to integrate these two approaches to obtain a more complete picture of online visit-purchase behavior. In combining these two models, we can begin to investigate the efficiency of marketing dollars aimed at increasing visits versus encouraging conversion behavior. Such insights would be of tremendous value to practitioners. But before any comprehensive optimization schemes can be achieved, it is important that we first gain a complete, theoretically grounded, understanding of each of the processes that underlie the observable behaviors. We believe that this model is a useful first step in this direction, and we encourage future researchers to build upon it.

REFERENCES

- Alba, Joseph W. and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (March), 411-454.
- Babin, Barry J., William R. Darden, and Mitch Griffin (1994), "Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value," *Journal of Consumer Research*, 20 (4), 644-656.
- Barnett, A.I., Cynthia J. Rutherford, Jane F. Desforges, Nancy Gutensohn, and B. Davies (1981), "Evaluating the Validity of a Bayesian Program for Predicting Stage in Hodgkin's Disease," *Methods of Information in Medicine*, 20 (3), 174-8.
- Bettman, James R. (1979), *An Information Processing Theory of Consumer Choice*, Reading, MA: Addison-Wesley.
- Bloch, Peter H., Daniel L. Sherrell, and Nancy M. Ridgway (1986), "Consumer Search: An Extended Framework," *Journal of Consumer Research*, 13 (June), 119-126.
- Brucks, Merrie (1985), "The Effects of Product Class Knowledge on Information Search Behavior," *Journal of Consumer Research*, 12 (1), 1-16.
- Hirschman, Elizabeth C. (1984), "Experience Seeking: A Subjectivist Perspective of Consumption," *Journal of Business Research*, 12 (March), 115-136.
- Jarboe, Glen R. and Carl D. McDaniel. (1987). "A Profile of Browsers in Regional Shopping Malls," *Journal of the Academy of Marketing Science*, 15 (Spring): 46-53.
- Janiszewski, Chris (1998), "The Influence of Display Characteristics on Visual Exploratory Search Behavior," *Journal of Consumer Research*, 25 (3), 290-301.
- Johnson, Norman L., Samuel Kotz, and N. Balakrishnan (1995), *Continuous Univariate Distributions*, 2nd edition, Vol. 1, John Wiley & Sons, Inc., New York, New York.
- Lenk, Peter J. and Ambar G. Rao (1995), "Transition Times: Distributions Arising from Time Heterogeneous Poisson Processes," *Management Science*, 41 (7), 1117-1129.
- Moe, Wendy W. and Peter S. Fader (2000), "Capturing Evolving Visit Behavior in Clickstream Data," *Wharton Marketing Department Working Paper #00.003*.
- Punj, Girish N. and Richard Staelin (1983), "A Model of Consumer Information Search Behavior for New Automobiles," *Journal of Consumer Research*, 9 (March), 366-80.
- Putsis, William P., Jr. and Narasimhan Srinivasan (1994), "Buying or Just Browsing? The Duration of Purchase Deliberation," *Journal of Marketing Research*, 31 (August), 393-402.
- Quick, Rebecca (1998), "Gawkers or Shoppers? Selling Bras on the Web," *The Wall Street Journal*, December 29, B1.
- Sherry, John F., Jr., Mary Ann McGrath, and Sidney J. Levy (1993), "The Dark Side of the Gift," *Journal of Business Research*, 28 (November), 225-245.
- Schmittlein, David C. and Donald G. Morrison (2000), "A Live Baby or Your Money back: The Marketing of in vitro Fertilization Procedures," *Wharton Marketing Department Working Paper*.
- Wilkie, William L. and Peter R. Dickson (1985), "Shopping for Appliances: Consumers' Strategies and Patterns of Information Search," in *Perspectives in Consumer Research*, ed. Harold H. Kassirjian and Thomas R. Robertson, Englewood Cliffs, NJ: Prentice Hall, 1-25.