MARKETING SCIENCE
Vol. 32, No. 5, September–October 2013, pp. 716–732
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2013.0790
© 2013 INFORMS

Information Processing Pattern and Propensity to Buy: An Investigation of Online Point-of-Purchase Behavior

Ofer Mintz
E. J. Ourso College of Business, Louisiana State University, Baton Rouge, Louisiana 70803, omintz@lsu.edu

Imran S. Currim
Paul Merage School of Business, University of California, Irvine, Irvine, California 92697, iscurrim@uci.edu

Ivan Jeliazkov
Department of Economics, University of California, Irvine, Irvine, California 92697, ivan@uci.edu

The information processing literature provides a wealth of laboratory evidence on the effects that the choice task and individual characteristics have on the extent to which consumers engage in alternative-based versus attribute-based information processing. Less attention has been paid to studying how the processing pattern at the point of purchase is associated with a consumer's propensity to buy in shopping settings. To understand this relationship, we formulate a discrete choice model and perform formal model comparisons to distinguish among several possible dependence structures. We consider models involving an existing measure of information processing, PATTERN; a latent variable version of this measure; and several new refinements and generalizations. Analysis of a unique data set of 895 shoppers on a popular electronics website supports the latent variable specification and provides validation for several hypotheses and modeling components. We find a positive relationship between alternative-based processing and purchase, as well as a tendency of shoppers in the lower price category to engage in alternative-based processing. The results also support the case for joint modeling and estimation. These findings can be useful for future work in information processing and suggest that likely buyers can be identified while engaged in information processing prior to purchase commitment, an important first step in targeting decisions.

Key words: information processing; discrete choice; Markov chain Monte Carlo (MCMC); digital strategy
History: Received: November 27, 2012; accepted: April 2, 2013; Preyas Desai served as the editor-in-chief and Barbara Kahn served as associate editor for this article. Published online in *Articles in Advance* July 26, 2013.

1. Introduction

Consumers face daily decisions and trade-offs regarding the products they want to buy. Bettman (1979, p. 178) suggests that when making choices, consumers process information using one of two patterns: (1) choice by processing brands, in which consumers process the available information by examining specific products across attributes (what we call an *alternative-based* processing pattern) and (2) choice by processing attributes, in which consumers process the available information by examining specific product attributes across alternatives (what we call an *attribute-based* processing pattern). Consumers could, of course, process information using any mix of these two basic patterns. Consequently, information processing scholars have developed a key measure, PATTERN (Payne 1976, p. 376), widely used in a variety of experimental studies (see Payne et al. 1993 for a review). This measure takes on values between -1 and $+1$, where values closer to -1 indicate an attribute-based processing pattern and values closer to $+1$ indicate an alternative-based processing pattern.

This raises a logical research question: If consumers process information using one pattern to a greater extent than the other at the point of purchase, would they be more, less, or equally likely to buy a product? In other words, is there a relationship between a consumer's processing pattern and buying behavior, and if so, what is the nature of this relationship? This is the central question that motivates our study and underlies our hypotheses, models, and empirical analysis of a unique data set of processing patterns and purchase decisions of 895 online shoppers. The idea that there may be a systematic relationship between how shoppers process information in online environments and whether they end up making a purchase is inherently intriguing, and given the absence of work on this topic, we take an exploratory approach.

Exploring the relationship between PATTERN and the propensity to buy at the point of purchase has the potential to offer contributions to theory, methodology, and managerial practice. On the theoretical side, the absence of prior work on this topic leads us to explore key hypotheses, detailed in §2, about the relationship between information processing pattern, purchase behavior, and the price category in which consumers are shopping. Our main hypotheses examine whether alternative-based processing patterns are more likely to be associated with purchase than are attribute-based processing patterns and whether shoppers interested in different price categories exhibit systematically different information processing patterns. We also provide a number of generalizations of PATTERN, including a latent specification that overcomes certain limitations of the original observed measure. The new specifications are of theoretical interest in their own right, yet they also serve as valuable robustness checks for our main empirical findings.

Our implementation contributes to four decades of advances in the information processing literature on the antecedent conditions under which consumers are likely to employ alternative-based, attribute-based, or mixed processing patterns. For example, laboratory studies have investigated the effect of individual differences (e.g., novices versus experts), specific properties of the choice task being undertaken (e.g., complexity, as in the number of alternatives and attributes; dissimilarity of options), and the type of choice situation (e.g., whether it involves emotion, time pressure, or a certain type of information display) on information processing patterns (e.g., see the reviews in Bettman 1979, Payne et al. 1993, and Bettman et al. 1998). However, because all subjects in such experiments were typically required to choose among products in forced choice scenarios after accessing and processing the information provided, these studies are unable to link the information processing pattern to *whether* the subject will buy. In contrast, in our point-of-purchase setting, we are able to study this missing link as we observe both buyers and nonbuyers. By examining this relationship and testing its empirical relevance, this paper offers a foundation for connecting the aforementioned studies on drivers of processing patterns with propensity to buy. To the best of our knowledge, this paper is one of the first attempts to bridge the large literature on information processing patterns that flourished in laboratory settings during the 1970s, 1980s, and 1990s and continues through today (e.g., Dhar and Nowlis 2004, Valenzuela et al. 2009) to purchasing behavior in a contemporary online shopping environment. We comment on the value of this connection in §5.

On the methodological side, we pursue several objectives. One question that arises quite naturally, particularly in our online point-of-purchase setting, is whether modeling and estimation of information processing patterns and purchase behavior can be approached separately or should instead be accomplished jointly in a system of equations. Because the measures of information processing pattern we consider are censored to the interval $[-1, 1]$ and purchase decisions are binary, we develop a joint (Tobit and probit) discrete data model to analyze the interactions between those two outcomes and discuss a simulation-based algorithm for parameter estimation. In our baseline model, the information processing pattern is allowed to endogenously determine propensity to buy based on the recognition that unobserved consumer background variables, in addition to the price category in which a consumer shops, may influence both the processing pattern and the eventual purchase decision. For example, shoppers may arrive at the point of purchase with different backgrounds, levels of experience, and prior information about the product category, which can result in better awareness or focus on particular products or features (e.g., Hauser and Wernerfelt 1990, Ratchford 1982). This, in turn, could influence information processing (Hong and Sternthal 2010, Simonson et al. 1988) and affect the propensity to purchase.

In this context, an unusual and important challenge arises in tackling our main goal of properly formulating the dependence of the propensity to buy on PATTERN because the latter is censored to the interval $[-1, 1]$ and exhibits point mass at both endpoints. To address this challenge, §4 considers two types of econometric models: in the first category, propensity to buy depends on the observed value of PATTERN, widely employed in experimental studies of consumer behavior; in the second category, modeling involves the latent variable (random utility) underlying that outcome, which allows unobserved consumer variables (background, level of experience, prior information, etc.) to affect propensity to buy. There is little a priori theoretical guidance on which type of model may be more appropriate, and indeed, there exist conditions related to the presence or absence of unobserved variables wherein either of the two may be sensible. In showing that a latent measure of an information processing pattern, which can be inferred from the data, is found to be more powerful than its traditional observed counterpart, we offer a methodological contribution that may prove useful in future information processing studies.

This and other modeling issues motivate our next main methodological objective, which is to address the problem of model uncertainty to arrive at a preferred model specification. We do so by estimating alternative specifications of the relationship between the information processing pattern, purchase, and price category and performing formal model comparisons to distinguish among them. The results support several key hypotheses about the relationship between these variables. Although our proposed model has not been employed in the marketing literature, its main contribution is to basic research in information processing; it provides new and important substantive evidence on the connection between information processing pattern and propensity to buy. The results hold regardless of whether we work with the traditional PATTERN or examine generalizations that dispense with some limitations of the original observed measure. Finally, our methodological framework provides a foundation for future modeling extensions. Because we link a shopper's processing pattern to *whether* that shopper purchases, our techniques differ from commonly used multinomial models in marketing, which focus on *which* product is chosen conditionally on purchase and product attributes. In our discussion, we suggest how future research could embed the model presented in this paper in a larger hierarchical model of consumer choice and describe the associated methodological challenges.

The managerial implications of our results relate to a key challenge for commercial websites—namely, the conversion of visitors to buyers. Many well-known websites such as Dell, CNET, Amazon, and Apple offer shoppers a choice to process information using alternative-based, attribute-based, or a mix of alternative- and attribute-based patterns. For example, Apple's website allows consumers interested in information on iPods to get the information in a format that enables alternative-based, attribute-based, or a mix of alternative- and attribute-based processing (see Figure 1). Likewise, shoppers on Amazon or CNET can examine a single product alternative or can alternatively select a set of products for side-by-side comparisons in a format similar to Figure 1. In addition, new tools based on clickstream and mouse-tracking technology can now capture consumer navigation across different pages at a site. Consequently, understanding the relationship between processing patterns and propensity to buy has the potential to allow management to identify shoppers who are more likely to purchase while they are engaged in information processing prior to a purchase commitment. Because website visitors can be stratified by their processing pattern, managers can direct incentives more efficiently, e.g., by prioritizing likely buyers for follow-up communication if they abandon their shopping carts. The significance of recognizing and exploiting this channel for identifying likely buyers is even greater in settings where consumers may be anonymous, as in many Web-based environments, and management has little additional information to rely on.

The remainder of this paper is organized as follows. Section 2 examines hypotheses about the relationship between the information processing pattern, propensity to buy, and price category. Section 3 discusses our data, and §4 presents our econometric methodology and offers generalizations of the traditional PATTERN. Section 5 contains our results, whereas §6 offers a discussion of their implications, limitations, and suggestions for future research.

2. Background and Hypotheses

In this section, we provide background literature for three purposes. First, we develop a hypothesis between consumers' information processing patterns and propensity to buy, which underscores the need for studying the two constructs jointly and highlights the limited scope of prior studies that focus only on a subsample of consumers who arrive at a point of purchase. As noted in §1, processing patterns can be dynamic and context driven, e.g., the consequence of strategic attempts to cope with processing limitations, and may involve unobserved factors that are also relevant to purchase decisions. Second, we develop a hypothesis between the price category in which a customer shops and his or her information processing pattern because the former is easily observed and could serve as a useful segmentation variable for managers seeking to prioritize consumers who abandon their shopping carts for follow-up communications aimed at persuading them to purchase. Third, we briefly indicate how our study differs from others that have investigated consumer conversion behavior on the Internet.

2.1. Relationship Between Information Processing and Propensity to Buy

Consumers who assess products individually, in isolation, or one at a time are better able to judge whether a product's features meet their goals or purchasing criteria with less distraction about whether another competitive product is better (Dhar and Nowlis 2004). Such consumers therefore are expected to develop more accurate representations of the products they assess (Payne et al. 1993) and are better able to judge the overall suitability and determine whether a product will satisfy their requirements (Simon 1956). In contrast to such alternative-based processing, consumers who process information in attribute-based patterns aim to determine which alternative is best on each attribute. If only one attribute is important, such a lexicographic processing strategy can be useful in identifying which product to purchase (Valenzuela et al. 2009). When multiple attributes are important, however, choice is typically complicated and made more difficult by the need to make trade-offs between multiple products that are superior on one attribute but not another, because buying a certain product implies that other superior features of competitive products must be sacrificed. This can be true in many product categories, but it can be particularly relevant in durable goods categories, where the trade-offs of a

Figure 1 Apple's Presentation of Alternative-Based and Attribute-Based Information

Comparison Chart

	 iPod shuffle		 iPod nano		 iPod classic	 iPod touch		
Storage	1GB	2GB	8GB	16GB	120GB	8GB	16GB	32GB
Songs	240	500	2,000	4,000	30,000	1,750	3,500	7,000
Price	$49	$69	$149	$199	$249	$229	$299	$399
Color	●●●●●	●●●●●	●●●●●●●●●	●●●●●●●●●	●●	●	●	●
Battery life	Up to 12 hours of music playback		Up to 24 hours of music playback; up to 4 hours of video playback		Up to 36 hours of music playback; up to 6 hours of video playback	Up to 36 hours of music playback; up to 6 hours of video playback		
Display			2-inch (diagonal) color LCD with LED backlight		2.5-inch (diagonal) color LCD with LED backlight	3.5-inch (diagonal) widescreen Multi-Touch display		
Ports	Stereo minijack		Dock connector, stereo minijack		Dock connector, stereo minijack	Dock connector, stereo minijack		
Connectivity	USB through included dock		USB through dock connector; component and composite video through dock connector (with AV cables, sold separately); audio through headphone jack		USB through dock connector; component and composite video through dock connector (with AV cables, sold separately); audio through headphone jack	USB through dock connector; component and composite video through dock connector (with AV cables, sold separately); audio through headphone jack		
Wireless data						Wi-Fi (802.11b/g) Nike + iPod support built in Maps location-based service		
Charge time	About 4 hours (2-hour fast charge to 80% capacity)		About 3 hours (1.5-hour fast charge to 80% capacity)		About 4 hours (2-hour fast charge to 80% capacity)	About 4 hours (2-hour fast charge to 80% capacity)		
Audio support	AAC (8 to 320 Kbps), Protected AAC (from iTunes Store), MP3 (8 to 320 Kbps), MP3 VBR, Audible (formats 2, 3, and 4), WAV, and AIFF		AAC (16 to 320 Kbps), Protected AAC (from iTunes Store), MP3 (16 to 320 Kbps), MP3 VBR, Audible (formats 2, 3, and 4), Apple Lossless, WAV, and AIFF		AAC (16 to 320 Kbps), Protected AAC (from iTunes Store), MP3 (16 to 320 Kbps), MP3 VBR, Audible (formats 2, 3, and 4), Apple Lossless, WAV, and AIFF	AAC (16 to 320 Kbps), Protected AAC (from iTunes Store), MP3 (16 to 320 Kbps), MP3 VBR, Audible (formats 2, 3, and 4), Apple Lossless, WAV, and AIFF		
Photo support			Syncs iPod-viewable photos in JPEG, BMP, GIF, TIFF, PSD (Mac only), and PNG formats		Syncs iPod-viewable photos in JPEG, BMP, GIF, TIFF, PSD (Mac only), and PNG formats	Syncs iPod-viewable photos in JPEG, BMP, GIF, TIFF, PSD (Mac only), and PNG formats		
Video support			H.264 video, up to 1.5 Mbps, 640 by 480 pixels, 30 frames per second, Low-Complexity version of the H.264 Baseline Profile with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats; H.264 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Baseline Profile up to Level 3.0 with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats; MPEG-4 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Simple Profile with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats		H.264 video, up to 1.5 Mbps, 640 by 480 pixels, 30 frames per second, Low-Complexity version of the H.264 Baseline Profile with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats; H.264 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Baseline Profile up to Level 3.0 with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats; MPEG-4 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Simple Profile with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats	H.264 video, up to 1.5 Mbps, 640 by 480 pixels, 30 frames per second, Low-Complexity version of the H.264 Baseline Profile with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats; H.264 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Baseline Profile up to Level 3.0 with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats; MPEG-4 video, up to 2.5 Mbps, 640 by 480 pixels, 30 frames per second, Simple Profile with AAC-LC audio up to 160 Kbps, 48kHz, stereo audio in .m4v, .mp4, and .mov file formats		
Size	1.07 x 1.62 x 0.41 inches (27.3 x 41.2 x 10.5 mm) including clip		3.6 x 1.5 x 0.24 inches (90.7 x 38.7 x 6.2 mm)		4.1 x 2.4 x 0.41 inches (103.5 x 61.8 x 10.5 mm)	4.3 x 2.4 x 0.33 inches (110 x 61.8 x 8.5 mm)		
Weight	0.55 ounce (15.6 grams)		1.3 ounces (36.8 grams)		4.9 ounces (140 grams)	4.05 ounces (115 grams)		
Included accessories	Earphones, USB dock		Earphones, USB cable, dock adapter		Earphones, USB cable, dock adapter	Earphones, USB cable, dock adapter, polishing cloth		

Source. http://www.apple.com/ipod/compare-ipod-models/, accessed March 2009.

given choice have wider-ranging consequences over longer periods of time. We therefore propose the following hypothesis.

Hypothesis 1 (H1). *Shoppers who process information using alternative-based patterns to a greater extent than attribute-based patterns are more likely to buy; i.e., PATTERN is positively associated with propensity to buy.*

This proposition also relates to the two-stage choice literature, where research has suggested that consumers first employ a simplifying heuristic-based decision rule to reduce the number of alternatives in their choice set followed by a more careful evaluation of the remaining alternatives before selecting a product. This idea dates back to early work on information display boards (Bettman and Park 1980, Jacoby et al. 1978, Payne 1976), which was subsequently picked up by modelers (Gensch 1987, Gilbride and Allenby 2004, Hauser and Wernerfelt 1990, Liu and Arora 2011). As a result, one might expect that an alternative-based processing rule might be more temporally related to purchase; i.e., alternative-based processing is the strategy that people use after they have developed some initial expertise about the choice set, either by initial attribute-based screening in early phases of the choice or by prior search. However, the model-based studies model *only* buyers' behavior and focus on *which* product(s) will be selected, similar to the experimental studies that require subjects to choose a product with no option to abstain. Consequently, these works do not investigate nonbuyers or *whether* the shopper will buy any product, thus limiting their scope to a subsample of all consumers that arrive at a point of purchase. The possibility that customers could decide not to make a purchase after completing either of the two stages highlights the importance of modeling the decision of *whether* to buy for strengthening the managerial implications of such models. And because prior experimental and nonexperimental work has not allowed an option to abstain, there is little, if any, evidence regarding H1 in a point-of-purchase setting where shoppers have that option.

A key component of assessing the relationship between information processing patterns and propensity to buy is to recognize that because both are driven by common unobserved factors such as shoppers' background, experience, and prior information about the product category, one must carefully account for potential endogeneity. Shoppers who arrive at the point of purchase with significant experience may process information via a preference validation pattern to ensure that the product does not contain any negative properties that discourage purchase (Iyengar and Lepper 2000, Moorthy et al. 1997), and hence they may be more likely to process information via an alternative-based pattern. In related research, Dhar and Nowlis (2004) find that when subjects are in a buy or no-buy decision response mode (versus unconditional brand choice response mode), decision processes will likely be characterized by alternative-based evaluations. Less experienced shoppers, on the other hand, may exhibit an exploratory attribute-based processing pattern in accordance with the two-stage decision process. Because such unobserved individual-specific factors would affect shoppers' processing patterns and purchase decisions, those patterns and decisions must be studied jointly. Our results suggest that such joint modeling and estimation is supported by the data.

2.2. Relationship Between Price Category and Information Processing Pattern

One of our goals in this paper is to investigate the relationship between the price category that shoppers access and their information processing pattern. There are a number of reasons to expect that consumers who limit themselves to the low price category will be more likely to employ alternative-based information processing. Such a hypothesis is sensible in the context of the two-stage choice literature, because these shoppers first use attribute-based processing to identify the lowest-priced alternatives, followed by more careful alternative-based processing to ensure that an alternative does not have features that deter purchase. More generally, consumers who limit themselves to the lowest price category typically employ a price-aversion strategy to make choices, whereas consumers who do not limit themselves to the lowest price category typically use the best-value strategy to make choices (Tellis and Gaeth 1990). Price aversion involves choosing the lowest-priced alternative to minimize immediate cost while guarding against the presence of any nonprice features that preclude purchase. As a result, price-averse consumers are more likely to process information in an alternative-based pattern on the lowest-priced alternative and exhibit higher values of PATTERN. In contrast, the best-value strategy involves making trade-offs between price and nonprice features to identify the alternative that is the best value. Therefore, best-value consumers are more likely to process information in an attribute-based pattern to more easily determine what value they receive (forgo) from superior (inferior) attribute levels when they pay more (less) and exhibit lower values of PATTERN.[1] Consequently, we propose the following hypothesis.

Hypothesis 2 (H2). *Shoppers in a lower* (*higher*) *price category are more likely to process information using alternative* (*attribute*)*-based information processing patterns; i.e., price is negatively associated with PATTERN.*

[1] The reader is referred to Tellis and Gaeth (1990) for further theoretical details on price aversion and best-value strategies.

An empirical validation of the relationship between price category and consumer information processing pattern can be very useful because, if selection of a particular price category is found to lead to a certain processing pattern and subsequent buying behavior, managers could use price category as another important proxy (in addition to information processing pattern, as argued earlier) for segmenting and prioritizing. Segmentation on the basis of processing pattern may have information content advantages, but it requires tracking and computations, whereas segmentation by price category, despite its potential crudeness, may be easier to implement because the choice of price category is easily observed.

2.3. Studies on Conversion Behavior on the Internet

Although our main contribution is to basic research in information processing, this paper relates to econometric studies that investigate conversion behavior on the Internet, i.e., converting visits into purchases (Moe 2006, Moe and Fader 2004, Montgomery et al. 2004, Sismeiro and Bucklin 2004). These efforts typically focus on determining how customer and search characteristics relate to purchase behavior but do not study the information processing pattern and its relationship to propensity to buy at the point of purchase.

3. Data

Our data set consists of observations on a sample of shoppers who visited the website of a popular global electronics retailer and were unaware that their actions were to be analyzed. We are required to keep the product category and retailer's identity confidential. The Decision Board Platform (http://www.decisionboard.org; see Mintz et al. 1997), a computerized decision process tracing program, was installed on the retailer's website. The Decision Board Platform is conceptually similar to Mouselab (e.g., Payne et al. 1993) and has been employed in a wide variety of research fields, including political science, engineering safety, and business studies. Customers who visited the website over a 50-hour period examined products based on different price categories with links for each price category listed visibly on the center of the main product category home page. These links stated "compare [electronic products] at a glance, see only the features that matter most to you for [the price range selected]." Shoppers then entered the Decision Board Platform and were able to compare three alternatives at a glance (given in columns) on 11 product attributes (presented in rows) with the model identifier and the price of the alternative listed as column headers. The attribute values in the corresponding cells were hidden, and consumers were instructed to click on cells that were important to them to access those values. A "Customize and Buy" option was visibly placed at the bottom of each product column. A mock-up of the Web page is given in Figure 2 using an iPod example.

The Decision Board Platform kept track of the accessed cells, the sequence in which they were accessed, and the customize and buy decisions of each shopper. Our analysis includes $n = 895$ consumers who accessed more than one cell and exhibited a processing pattern during the 50-hour period.

3.1. Measures

3.1.1. Price Category. Shoppers started by entering one of two price categories to conduct their processing: low (less than \$999) and high (\$999 or more). Price category was the only theoretically relevant variable recorded by the Decision Board Platform prior to the initiation of information processing. Such a dearth of background information tends to be the norm in Web-based environments and is representative of the information that management typically has about a website visitor.

3.1.2. Information Processing Pattern. A widely used measure, PATTERN (Payne et al. 1993), was employed to quantify the extent to which consumers process information via alternative- or attribute-based patterns. The measure is defined as a ratio—the numerator is the difference between the number of alternative- and attribute-based transitions, N_{ALT} and N_{ATT}, respectively, and the denominator is the sum of the two; i.e.,

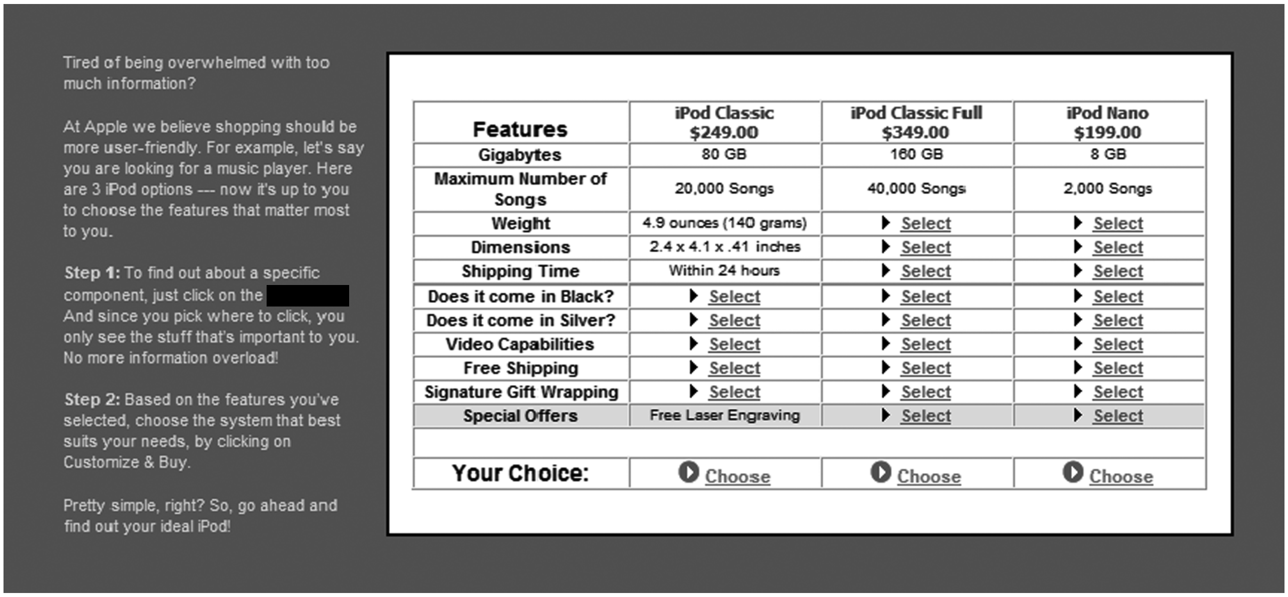
$$P = \frac{N_{ALT} - N_{ATT}}{N_{ALT} + N_{ATT}}. \tag{1}$$

The resulting scores are censored to the range $[-1, 1]$ and exhibit point mass at both -1 and 1. Lower values indicate more attribute-based processing, whereas higher values represent alternative-based patterns. The use of this measure links our results to earlier experimental information processing studies and facilitates future connections between the antecedent conditions for processing patterns and propensity to buy at the point of purchase.

In §4, we examine refinements and generalizations of the metric in Equation (1) and discuss their theoretical implications. We conduct extensive additional analyses employing the new measures and report our findings in §5. The results offer broad validation of our main conclusions across all versions of the processing pattern, serving as useful robustness checks of the results based on the traditional measure in Equation (1).

3.1.3. Propensity to Buy. Propensity to buy is recorded as a binary variable, 0 or 1, with 1 indicating that a shopper chose a product. If a consumer clicked

Figure 2 Mock-Up of the Matrix Presented to Shoppers by the Decision Board Platform Using iPod as an Example



on the "Customize and Buy" link at the bottom of an alternative's column, the product's purchase page appeared, and shoppers could customize the product or purchase it as is. It is unknown if all consumers clicking on "Customize and Buy" subsequently procured the product, since this required the completion of additional tasks (as in Sismeiro and Bucklin 2004). Final purchases were recorded on a secure server to which we do not have access for legal reasons. Yet consumers selecting "Customize and Buy" demonstrate a greater inclination to buy than if they did not click on that link. In either case, our model can be viewed as a first stage in a hurdle model and, an interpretation of H1 and H2 would be valid for that stage (analysis of the second stage of the hurdle model would obviously require that actual purchase data be available).

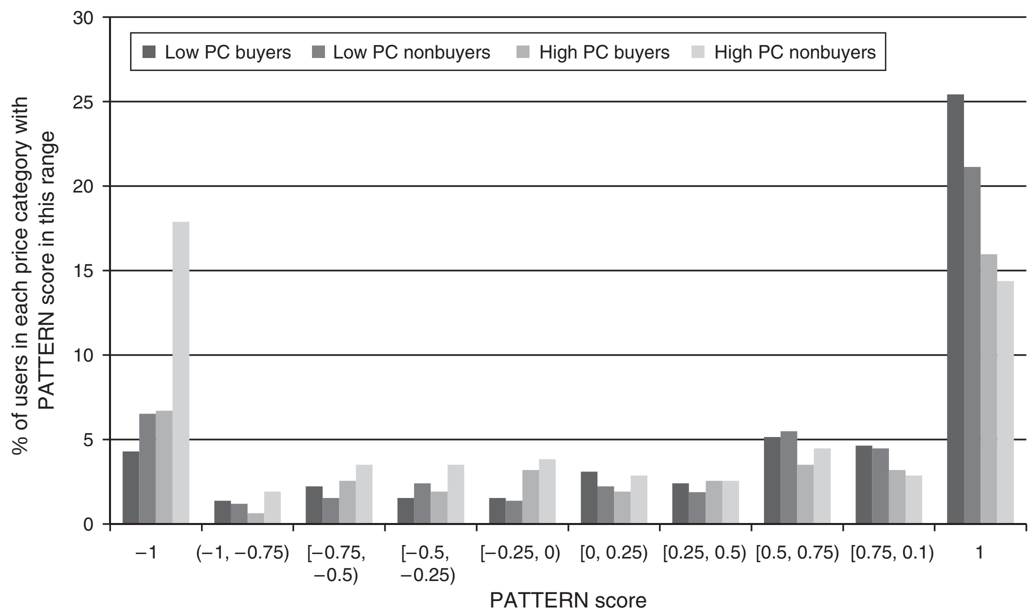
3.2. Descriptive Statistics

Among the 895 shoppers in our sample, 594 (66%) exhibited an alternative-based processing pattern (had a positive PATTERN score), 284 (32%) exhibited an attribute-based processing pattern (had a negative PATTERN score), and 17 (2%) exhibited neither alternative- nor attributed-based processing patterns (had a PATTERN score of 0). Figure 3 gives a detailed distribution of PATTERN for buyers and nonbuyers in both price categories and indicates the mass points at -1 and $+1$ and the variation between these two end points.

Of the 895 consumers, 433 (48%) proceeded to customize and buy and 462 (52%) did not. Among the shoppers who processed via an alternative-based processing pattern, 314 (53%) proceeded to customize and buy. Among the consumers who processed via an attribute-based processing pattern, 112 (40%) proceeded to buy. In addition, 582 (65%) consumers chose to shop in the low price category, whereas 313 (35%) consumers shopped in the high price category. Of the 582 consumers who chose to shop in the low price category, 430 (74%) used alternative-based processing; among the 313 shoppers in the high price category, 164 (52%) used alternative-based processing.

As the retailer required that the product alternative and attribute information remain confidential, in Table 1, we provide ordinal information about the attribute in each price category. Table 2 provides descriptive analysis of the information processing revealed by shoppers in the low and high price categories. Summaries are presented for "pure" alternative-based ($\text{PATTERN} = 1$), "pure" attribute-based ($\text{PATTERN} = -1$), and mixed ($1 > \text{PATTERN} > -1$) processing. Three variants of mixed processing are examined: (a) two-stage theory (consumers begin with attribute-based processing and then switch to alternative-based processing), (b) reverse two-stage theory (consumers begin with alternative-based processing and then switch to attribute-based processing), and (c) random/mixed strategies (consumers process information without discernible patterns). According to H1, alternative-based processing is more likely to be associated with purchase than attribute-based processing. Although H1 will be tested formally based on models presented in the next section, Table 2 offers some preliminary support for that hypothesis: 57% of buyers reveal either pure alternative-based (46%) or two-stage theory-based (11%) processing, whereas 11% of buyers reveal pure attribute-based processing, and 32% of buyers reveal non-two-stage theory information processing. Surprisingly, a total of 43% of

Figure 3 **Frequency Distribution of PATTERN for Buyers and Nonbuyers in Low and High Price Categories (PCs)**



buyers do not conform to traditional two-stage theory, which provides an important additional rationale for studying the processing patterns of all shoppers, not just the subsample of buyers.

We emphasize three additional points before proceeding. First, our data set exhibits excellent variability over the constructs being investigated, which is important for estimation purposes. Second, there is no dominant or clearly superior product in the low or high price categories in Table 1 that could have led to alternative-based processing (also recall that information remains hidden until consumers access a cell). And third, the high "Customize and Buy" rate in our sample suggests that at least some customers must have come to the point of purchase well prepared by processing additional information prior to arrival or possessing more experience in the product category. Because such customer characteristics can impact both the information processing pattern and propensity to buy, they underscore the importance of joint modeling and suitable formulation of dependencies, which we address next.

Table 1 **Descriptive Information on Products in the Low and High Price Categories**

	Level		
Attribute	Product 1	Product 2	Product 3
	Low price category		
Price	1	2	2
1	2	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	2	1	3
6	1	1	2
7	1	1	1
8	2	1	2
9	1	1	1
10a/b	1/2	2/1	3/3
11	1	2	1
	High price category		
Price	1	2	1
1	1	2	1
2	1	1	1
3	1	2	1
4	1	1	1
5	2	3	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	1	1
10a/b	1/2	3/3	2/1
11	1	1	2

Notes. Level 1 indicates the lowest value and level 3 indicates the highest value for an attribute. Prices and model names and numbers were available to shoppers without clicking a cell. Attribute 10a (height) and attribute 10b (weight) were accessible by clicking the same cell. Level 1 for price indicates a cheaper alternative, level 1 for height (attribute 10a) indicates a shorter alternative, and level 1 for weight (attribute 10b) indicates a lighter alternative; for the remaining attributes, lower levels correspond to less desired values.

4. Empirical Methodology

4.1. Overview

We now present an econometric model that is specifically tailored to the setting considered in this paper. Recall that our main motivation is to examine the relationship between the processing pattern and propensity to purchase in the presence of unmeasured exogenous factors (e.g., familiarity or expertise) that affect both the purchase propensity and processing pattern. The model is intended to accommodate

Table 2 **Descriptive Statistics by Type of Information Processing for Buyers**

	% of all shoppers			% for subsample of buyers		
Type of processing	Both PCs	Low PC	High PC	Both PCs	Low PC	High PC
"Pure" alternative	41	47	30	46	49	38
"Pure" attribute	16	11	25	11	8	16
Mixed processing	43	43	45	44	43	46
Two-stage theory	11	8	17	11	9	16
Reverse two-stage theory	19	22	13	18	20	13
Random/mixed strategies	14	13	15	14	13	17

Notes. PC, price category. The average number of clicks (with standard deviations in parentheses) for each type of processing are as follows: "pure" alternative, 9.98 (5.83); "pure" attribute, 7.29 (8.32); mixed, 18.12 (9.38); two-stage theory, 20.54 (9.34); reverse two-stage theory, 19.66 (9.01); and random/mixed strategies, 14.12 (8.67).

three particular aspects of the problem at hand. First, our model accounts for the discrete nature of the dependent variables—in particular, propensity to buy is a binary indicator variable, whereas our measure of the information processing pattern is censored on the interval $[-1, 1]$ and exhibits point mass at both end points. To deal with this aspect, our modeling and estimation approach relies on data augmentation (Albert and Chib 1993, Chib 1992), which employs the latent threshold-crossing utility representation of the model to facilitate estimation. A second issue we address is the potential for the information processing pattern to be endogenous in its relationship with the propensity to buy due to the possibility of unmeasured factors that affect both variables. If these potential features of the theory are not accounted for in the model, they could render it misspecified. Models with endogeneity, however, have been difficult to estimate when the response variables of interest are not continuous because standard two-stage estimators are inapplicable in this context. Moreover, an important modeling question arises in our setting: Should endogeneity of the information processing pattern be modeled in terms of observed or latent variables? Although this question does not arise in continuous data problems, it is a point worth emphasizing in discrete data settings. Third, we specifically account for model uncertainty by performing formal model comparisons based on marginal likelihoods and Bayes factors. These techniques allow us to consider the extent to which the data support the main hypotheses we consider here and allow us to distinguish among several competing specifications.

4.2. Model Specification

For consumer $i = 1, \ldots, n$, the general specification we consider is given by the bivariate system:

$$\begin{aligned} y^*_{iIP} &= x'_{i1}\beta_1 + \varepsilon_{i1}, \\ y^*_{iPB} &= x'_{i2}\beta_2 + y^*_{iIP}\theta + \varepsilon_{i2}, \end{aligned} \tag{2}$$

where y^*_{iIP} and y^*_{iPB} are the latent random utilities underlying the information processing pattern and propensity to buy, respectively, and x_{i1} and x_{i2} are exogenous covariates with corresponding parameter vectors β_1 and β_2. The observed information processing pattern, y_{iIP}, relates to the latent measure y^*_{iIP} through the two-sided censored (Tobit) relationship

$$y_{iIP} = \begin{cases} -1 & \text{if } y^*_{iIP} \le -1, \\ y^*_{iIP} & \text{if } y^*_{iIP} \in (-1, 1), \\ 1 & \text{if } y^*_{iIP} \ge 1, \end{cases} \tag{3}$$

whereas y_{iPB} relates to the latent y^*_{iPB} through the indicator link function $y_{iPB} = 1\{y^*_{iPB} > 0\}$. In the foregoing equations, the latent y^*_{iIP} and y^*_{iPB} have the customary random utility interpretation underlying the theory on discrete choice analysis in econometrics. In particular, the latent y^*_{iIP} represents the net value or utility of alternative-based processing; because of observed factors (e.g., covariates) and unmeasured characteristics (e.g., familiarity or expertise) that cannot a priori be restricted, the measure is unbounded and can take on values on the entire real line, with larger values implying greater net utility of alternative-based processing. For example, if consumers are more familiar, have more expertise, or have conducted more research prior to arriving at the point of purchase—all of which is unobserved—thus resulting in them having a strong intention to purchase a particular alternative, the latent net value or utility of alternative-based processing will be high. In contrast, if consumers do not have a strong intention to purchase a particular alternative prior to arriving at the point of purchase, but they intend to engage in trade-offs between price and benefits at the point of purchase, the latent net value or utility of alternative-based processing will be low, thereby leading to attribute-based processing. Similarly, y^*_{iPB} can be viewed as the net utility of purchase, which, if positive, results in purchase.

Therefore, even though the observed data y_{iIP} can only take values in the range $[-1, 1]$ and $y_{iPB} \in \{0, 1\}$, the latent variables that determine those outcomes are unrestricted. In the Tobit equation, the modeling framework and the mapping in (3) relate, in a theoretically coherent way, the *discrete* point mass at the end points to the covariates and parameters employed in modeling the *continuous* outcomes $y_{iIP} \in (-1, 1)$. Moreover, if y^*_{iIP} is far outside the interval $(-1, 1)$ (or analogously, when y^*_{iPB} is far from 0), a much larger shock would be needed to observe any change in observed behavior. The fact that the latent utilities can change without necessarily inducing a corresponding change in y_{iIP} or y_{iPB} is a key distinction of this model relative to regressions involving only observed processing and purchase decisions.

In the system of equations in (2), the errors follow

$$\begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \sim \mathrm{N}(0, \Omega), \quad \text{where } \Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & 1 \end{pmatrix};$$

i.e., Ω is a symmetric positive definite matrix that incorporates the usual unit variance restriction in probit models. For the purposes of estimation, the model in (2) can be written as

$$y^*_i = X_i \beta + \varepsilon_i, \tag{4}$$

where

$$y^*_i = \begin{pmatrix} y^*_{iIP} \\ y^*_{iPB} \end{pmatrix}, \quad X_i = \begin{pmatrix} x'_{i1} & 0 & 0 \\ 0 & x'_{i2} & y^*_{iIP} \end{pmatrix},$$

$$\beta = (\beta'_1, \beta'_2, \theta)', \quad \text{and } \varepsilon_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix}.$$

In the specific application that we consider, the vector of exogenous covariates x'_{i1} contains an intercept and a dummy variable for the high price category, and x'_{i2} contains an intercept term.

In (2), y^*_{iPB} is affected by the latent y^*_{iIP}, yet another sensible specification would be to allow the y^*_{iPB} to depend on the observed y_{iIP}; i.e.,

$$\begin{aligned} y^*_{iIP} &= x'_{i1}\beta_1 + \varepsilon_{i1}, \\ y^*_{iPB} &= x'_{i2}\beta_2 + y_{iIP}\theta + \varepsilon_{i2}. \end{aligned} \tag{5}$$

Equation (5) can be written in matrix form as Equation (4), however, with X_i given by

$$X_i = \begin{pmatrix} x'_{i1} & 0 & 0 \\ 0 & x'_{i2} & y_{iIP} \end{pmatrix}.$$

This suggests that despite their differences, both models can be estimated similarly because they contain analogous components.

Let $y_i = (y_{iIP}, y_{iPB})'$, $y = (y'_1, \ldots, y'_n)'$, and $y^* = (y^*_1, \ldots, y^*_n)'$; and let $\psi = (\beta, \theta, \omega)$ represent the vector of model parameters, where ω contains the unique unrestricted elements of Ω. The likelihood function $f(y \mid \psi) = \prod_i f(y_i \mid \psi)$ for this model requires multivariate integration to obtain each likelihood contribution $f(y_i \mid \psi) = \int_{S_i} f(y^*_i \mid \psi)\, dy^*_i$, where S_i is the feasible region defined by the mapping between y^*_i and y_i. This feature complicates estimation by maximum likelihood; however, Bayesian simulation-based estimation is easy to implement. Details on our Markov chain Monte Carlo (MCMC) estimation approach are presented in the appendix. The Bayesian approach is also useful because it provides finite-sample inferences about parameters and model probabilities, and it easily enables comparisons of nested and nonnested models (which will be discussed shortly).

To differentiate between models (2) and (5), we note that Equation (2) implies that values of the latent y^*_{iIP}, even those outside the observable range $[-1, 1]$ in (3), matter for purchase behavior. In other words, the magnitude of latent factors such as consumer background, experience, and prior information about a product category that affect information processing patterns determine propensity to buy. On the other hand, Equation (5) implies that purchase behavior depends only on the observed processing pattern y_{iIP}, as has typically been done in experimental studies of consumer behavior. In this case, y^*_{iIP} influences y_{iPB} only to the extent that it affects y_{iIP} through the nonlinear censoring mechanism in Equation (3).

To gauge the empirical relevance of the specifications involving observed or latent drivers of behavior, cast light on the propositions presented in §2, and study the case for joint modeling and estimation relative to simpler alternatives, we examine several competing models. These models are formally compared based on their marginal likelihoods, as discussed next.

4.3. Model Comparison

For any two models M_j and M_k, Bayesian model comparison proceeds on the basis of the ratio of posterior model probabilities, known as the posterior odds:

$$\frac{\Pr(M_j \mid y)}{\Pr(M_k \mid y)} = \frac{\Pr(M_j)}{\Pr(M_k)} \frac{f(y \mid M_j)}{f(y \mid M_k)}.$$

The first fraction on the right-hand side is known as the prior odds ratio, and the second is called the Bayes factor. Of central importance in determining the Bayes factor is the marginal likelihood $f(y \mid M_l)$, defined as the integral of the likelihood function $f(y \mid \psi_l, M_l)$ with respect to the prior $\pi(\psi_l \mid M_l)$; i.e.,

$$f(y \mid M_l) = \int f(y \mid \psi_l, M_l)\pi(\psi_l \mid M_l)\, d\psi_l. \tag{6}$$

Well-known properties of marginal likelihoods and Bayes factors are that they lead to finite-sample model probabilities, do not require competing models to be nested (unlike Wald, likelihood ratio, or Lagrange multiplier tests), and have appealing asymptotic properties that give rise to the information criterion of Schwarz (1978) (see Greenberg 2008, Chapter 3; O'Hagan 1994, Chapter 3).

A less known, yet very important, point is that marginal likelihoods provide a measure of *sequential out-of-sample predictive fit*. This can be seen by writing

$$\begin{aligned} f(y \mid M_l) &= \prod_{i=1}^{n} f(y_i \mid \{y_j\}_{j<i}, M_l) \\ &= \prod_{i=1}^{n} \int f(y_i \mid \{y_j\}_{j<i}, \psi_l, M_l)\pi(\psi_l \mid \{y_j\}_{j<i}, M_l)\, d\psi_l, \quad (7) \end{aligned}$$

where the second line uses the law of total probability to represent the marginal likelihood as the product of n one-step-ahead sequential predictive densities. The third line of (7) makes it explicit that the adequacy of a model, as captured by its marginal likelihood, corresponds to its cumulative out-of-sample predictive record, where the fit of observation i is measured with respect to the posterior density based only on data $\{y_j\}_{j<i}$ up to the ith data point, without conditioning on $\{y_j\}_{j\geq i}$. In contrast, in-sample measures of fit condition on the entire data set.

There are also important advantages of the model comparison framework presented here relative to customary out-of-sample comparisons in which a researcher would estimate the model using part of the data and then examine how successfully that model can predict the remainder of the data. The sequential out-of-sample fit measure provided by marginal likelihoods overcomes key difficulties (both general and context specific) of traditional out-of-sample comparisons.[2] The fast and efficient computation of (6) is afforded by the method of Chib (1995); the MCMC sampler is also employed for evaluating the likelihood function (Jeliazkov and Lee 2010). The construction of proper priors for use in (6) is done by the method of training samples and is discussed in the appendix.

[2] First, note that the marginal likelihood in (6) is invariant to permutation of the indices of the data: the same value will be obtained if the data were rearranged. This invariance is desirable because, in contrast, typical out-of-sample comparisons depend on what part of the data is used in estimation and what is retained for the purpose of comparison. Second, as a practical matter, the conditional mean of y_{iPB} in (2) depends on the latent value of y^*_{iIP}; thus, it is unavailable and must be averaged over the latent data. Third, since y_{iIP} is censored, traditional metrics of predictive accuracy such as mean squared error or R^2 are inapplicable. Finally, we must account for the fit in both equations in our bivariate system, and hence comparisons based only on the binary outcome y_{iPB} would be inappropriate.

4.4. Alternative Information Processing Measures

The original PATTERN measure in Equation (1) was modified in several ways to deal with several possible drawbacks. First, transitions that occur after a particular column or row has been exhausted may be misclassified. This is because customers must switch to a new row or column, which may affect their PATTERN score even if they wish to continue the same browsing behavior. To avoid this problem, we constructed a new measure that does not include transitions immediately following the completion of a row or column of clicks.[3] We call this measure *rational PATTERN* because it accounts for a key constraint to browsing behavior.

Our second modification generalizes (1) by weighing each move m_t, $t = 1, \ldots, T$, that a customer makes. In particular, let the indicator functions $1\{m_t = ATT\}$ and $1\{m_t = ALT\}$ take the value of 1 whenever transition m_t is attribute- or alternative-based, respectively. The measure of information processing pattern we propose is

$$\tilde{P} = \frac{\sum_{t=1}^{T} w_t 1\{m_t = ALT\} - \sum_{t=1}^{T} w_t 1\{m_t = ATT\}}{\sum_{t=1}^{T} w_t 1\{m_t = ALT\} + \sum_{t=1}^{T} w_t 1\{m_t = ATT\}}, \quad (8)$$

where w_t is a suitably defined weight function. Our application examines weights of the form

$$w_t = \exp(-\kappa|t - \tau|) \quad (9)$$

for scalars $\kappa \in \Re$ and $\tau \in [1, T]$. With constant weights, for example, when $\kappa = 0$, $\tilde{P}$ reduces to the PATTERN score in (1); otherwise, the two differ. When $\kappa > 0$, letting $\tau = T$ gives weights that emphasize consumer behavior toward the end of a browsing session—in other words, it introduces a *recency-based* weighting of clicks. Conversely, $\tau = 1$ implies that behavior at the beginning of a session is weighted more heavily, so we have a *primacy-based* weighting of clicks. Choosing τ to be the index of the median transition also enables us to weight the first half of clicks differently from the second half of clicks, or the middle of a customer's search differently than the two ends. This type for analysis can be performed similarly at other quantiles as well.

We also examined two alternatives that only consider runs of transitions, thereby focusing on overtly "determined" information processing patterns. Specifically, let $N_{ALT,R}$ represent the number of alternative "run" clicks (two or more consecutive alternative clicks), and let $N_{ATT,R}$ be the number of attribute run clicks (two or more consecutive attribute clicks). Then, a run-based measure can be defined as

$$R = \frac{N_{ALT,R} - N_{ATT,R}}{N_{ALT,R} + N_{ATT,R}}. \quad (10)$$

[3] We thank a reviewer for suggesting this measure.

We also examine a related measure that accounts only for the longest attribute and alternative runs, or the "maximum run" measure, which is defined as

$$M = \frac{L_{ALT} - L_{ATT}}{L_{ALT} + L_{ATT}}, \tag{11}$$

where L_{ALT} and L_{ATT} measure the number of clicks in the longest alternative and attribute runs, respectively.

In §5.3, we study several versions of the metrics and weight functions discussed here to examine whether they provide support for the main conclusions that emerge from our baseline specification using traditional PATTERN.

5. Results

Our results are presented in three steps. In §5.1, we present model comparisons between the latent and observed versions of the information processing pattern, followed by comparisons to determine whether joint modeling of outcomes and endogeneity are supported by the data. In §5.2, we interpret the parameter estimates of the best-fitting model to examine whether H1 and H2 are supported. Then, in §5.3, we introduce several alternative measures of information processing, and we use them to test whether H1 and H2 are robust across specifications.

5.1. Model Comparisons

Estimated log-marginal likelihoods for our baseline models M_1 and M_2, given in Equations (2) and (5), respectively, are presented in Table 3. A direct comparison between these models reveals that the data strongly support the latent specification. The difference in log-marginal likelihoods implies posterior odds in favor of the latent variable model M_1 versus M_2 of roughly 8,519:1. Given this set of three models, the log-marginal likelihoods suggest that the posterior probability that M_1 is the correct model is 0.99873, whereas the corresponding probability for M_2 is 0.00012. This is strong evidence that the relationships in our setting are better captured at the underlying latent utility level than by the observed processing pattern. The key distinction between the observed data model M_2 and the latent data specification M_1 is the way in which y^*_{iIP} enters the model for y_{iPB}.

The indirect relationship in M_2, where the latent y^*_{iIP} first determines the observed y_{iIP} through the nonlinear censoring mechanism in (3) and only then affects propensity to buy y_{iPB}, is not supported by the results in Table 3. Instead, the data strongly favor M_1, where the full magnitude of y^*_{iIP}, even values whose extent is driven by observed covariates and unobserved factors outside the $[-1, 1]$ interval, is relevant for customers' decisions. An important practical implication for future empirical work on information processing and purchase behavior is that including the observed PATTERN as a regressor may not be fully adequate and that a random utility specification may be preferable. This possibility has not been studied in the literature previously, but it appears to be strongly supported by the model comparison results.

Table 3 also presents the log-marginal likelihood for model M_3, which is a simplification of M_1 intended to test whether joint modeling of the outcomes and endogeneity are indeed important features of the setting. Model M_3 captures the notion that the information processing pattern and propensity to buy can be modeled independently because the model does not include y^*_{iIP} or y_{iIP} as regressors in the equation for y^*_{iPB}, and it rules out correlation in the errors. Model M_3 performs worse than model M_1, with posterior odds of 871:1 in favor of M_1 (implying that the posterior probability of M_3 given the data is 0.00115), a result that strengthens the case for joint modeling and estimation and supports key modeling aspects such as accounting for endogeneity, modeling at the level of the underlying latent variables, and allowing for correlation in the errors. These are new results in the literature and provide a platform for understanding the interaction pathways at the point of purchase.

Table 3 Model Comparisons

Model	X_i	Ω	$\ln f(y \mid M_k)$	$\Pr(M_k \mid y)$[a]
		Latent data endogeneity model		
M_1	$X_i = \begin{pmatrix} x'_{i1} & 0 & 0 \\ 0 & x'_{i2} & y^*_{iIP} \end{pmatrix}$	$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & 1 \end{pmatrix}$	−1,471.80 (0.02)	0.99873
		Observed data endogeneity model		
M_2	$X_i = \begin{pmatrix} x'_{i1} & 0 & 0 \\ 0 & x'_{i2} & y_{iIP} \end{pmatrix}$	$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & 1 \end{pmatrix}$	−1,480.85 (0.02)	0.00012
		Independent equations model		
M_3	$X_i = \begin{pmatrix} x'_{i1} & 0 \\ 0 & x'_{i2} \end{pmatrix}$	$\Omega = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 1 \end{pmatrix}$	−1,478.57 (0.01)	0.00115

[a]Given this set of models, the posterior model probabilities are computed from the posterior odds ratios discussed in §4.3, assuming that the models are a priori equiprobable.

We also estimated a model in which y_{iPB} depends not only on the type of processing y^*_{iIP} but also on the quantity of search, i.e., on the number of clicks customer i has completed. The marginal likelihood for that model was estimated to be $-1{,}474.35$, which was lower than that of model M_1. Quantity was not significant in the propensity to buy equation. Taken together, these results indicate that the presence of a quantity variable is not supported by the data and that the original specification is more appropriate.

5.2. Parameter Estimates

Parameter estimates for the benchmark model M_1 are presented in Table 4. Inferences are based on an MCMC simulation run of length 10,000 draws, following a burn-in cycle of 1,000 iterations. Consistent with H1, we find that shoppers who employ alternative-based information processing patterns are likely to have a higher propensity to purchase than shoppers who employ attribute-based patterns. This relationship is supported by the data and found to have a robust effect across specifications. In the baseline model, the effect of y^*_{iIP} has a posterior mean of 0.27 with posterior standard deviation of 0.05. The values of β and the error covariance ω_{12} provide strong evidence, as do the model comparisons, that the equations should be treated jointly and that correlation in the errors must be properly accommodated.

The results in Table 4 also support H2, the conjecture that customers in the low price category tend to process information in alternative-based processing patterns. Specifically, the coefficient on price category has a posterior mean of -0.79 with a posterior standard deviation of 0.12. Because of the correlation in the errors and the nonlinearity in the relationship between y^*_i and y_i, interpretation of the magnitude of that coefficient in practical terms is not straightforward. However, the simulation techniques presented in Chib and Jeliazkov (2006) and Jeliazkov et al. (2008) allow for an uncomplicated simulation-based evaluation of the marginal effect. In particular, covariate effect estimation proceeds as a forecasting problem in which, given a draw Ψ from the posterior, a value of y^*_i is generated and converted to y_i for both the high and low price categories. Performing this simulation multiple times and averaging the resulting differences between the high and low price categories' simulated values of y_i gives an estimate of the average effect of the exogenous covariate price category and is a useful way to interpret the magnitude of the coefficient on price category. Using this simulation approach, we have been able to determine that y_{iIP} increases by approximately 0.35 when the price category is changed from 1 (high) to 0 (low), suggesting that shoppers who select the lower price category are more likely to employ alternative-based processing patterns (consistent with price aversion). Because of the endogeneity in the model, a change in this covariate also has an effect on y_{iPB}, which increases by approximately 0.09. Both of these effects are of sufficiently high magnitudes such that they should be of interest to website marketing managers.

Furthermore, Table 4 lists the inefficiency factors for the model parameters, which show that the MCMC algorithm exhibits good mixing and convergence properties. The inefficiency factors approximate the ratio of the numerical variance of the posterior mean from the correlated MCMC draws relative to that from hypothetical independent draws (the latter quantity can be obtained by the method of batch means). Values close to 1 indicate approximate independent and identically distributed sampling. Researchers familiar with MCMC sampling in latent data models will recognize that the chain is performing very well—mixing is better for parameters that rely less on latent data (as in the first equation of the model) and tends to be somewhat slower for parameters that depend on latent data (e.g., the elements of β in the second equation and ω_{12}). The performance of the MCMC algorithm was further verified in a set of MCMC runs with different lengths and starting points.

Table 4 **Parameter Estimates for Model M_1**

Parameter	Covariate	Mean	SD	95% interval	Inefficiency
β	Eq. (1) intercept	0.96	0.08	(0.81, 1.11)	2.02
	Price category	−0.79	0.12	(−1.03, −0.56)	1.31
	Eq. (2) intercept	−0.22	0.05	(−0.32, −0.11)	4.10
	y^*_{iIP}	0.27	0.05	(0.17, 0.35)	7.36
ω_{11}		2.51	0.18	(2.18, 2.90)	3.50
ω_{12}		−0.45	0.11	(−0.66, −0.24)	14.11

5.3. Results Based on Alternative Measures of Information Processing

Results from the original PATTERN measure of Equation (1) are reproduced in row 1 of Table 5, followed by results from models using the alternatives discussed in §4.4. Results for rational PATTERN are given in row 2 of Table 5. The modification affected few individuals, and hence the correlation between the two measures, was very high (0.996), leading to nearly identical conclusions.

Table 5 further presents results from a variety of weighted measures constructed as in Equation (8) for various settings of κ and τ in (9). Weights that emphasize different degrees of recency-based behavior ($\kappa > 0$, $\tau = T$) are studied in rows 3–6, whereas results from primacy-based weights ($\kappa > 0$, $\tau = T$) are given in rows 7–10. (Recall that when $\kappa = 0$,

Table 5 Results for Models Employing New Processing Pattern Measures

Description of y_{ilP} variable	Coefficient on y^*_{ilP}	Price category coefficient	Supported? H1	Supported? H2	Description of y_{ilP} variable	Coefficient on y^*_{ilP}	Price category coefficient	Supported? H1	Supported? H2
1. Traditional PATTERN	0.27 (0.17, 0.36)	−0.78 (−1.01, −0.54)	Yes	Yes	10. Primacy weighted ($\kappa = 1$, $\tau = 1$)	0.16 (0.09, 0.24)	−1.11 (−1.41, −0.81)	Yes	Yes
2. Rational PATTERN (excludes clicks after completing a row or column)	0.21 (0.13, 0.29)	−0.92 (−1.20, −0.65)	Yes	Yes	11. Inverted U-shaped weights ($\kappa = 1$, $\tau =$ median)	0.21 (0.12, 0.29)	−0.94 (−1.22, −0.66)	Yes	Yes
3. Recency weighted ($\kappa = 0.25$, $\tau = T$)	0.21 (0.13, 0.29)	−0.87 (−1.15, −0.59)	Yes	Yes	12. U-shaped weights ($\kappa = 1$, $\tau =$ Nearest end point)	0.21 (0.13, 0.30)	−0.90 (−1.18, −0.63)	Yes	Yes
4. Recency weighted ($\kappa = 0.5$, $\tau = T$)	0.21 (0.12, 0.28)	−0.85 (−1.14, −0.56)	Yes	Yes	13. Differential weight (1st half) ($w_t = 2$ for 1st 1/2 of clicks, $w_t = 1$ for 2nd 1/2 of clicks)	0.21 (0.12, 0.29)	−0.94 (−1.22, −0.68)	Yes	Yes
5. Recency weighted ($\kappa = 0.75$, $\tau = T$)	0.20 (0.12, 0.28)	−0.84 (−1.14, −0.54)	Yes	Yes	14. Differential weight (2nd half) ($w_t = 1$ for 1st 1/2 of clicks, $w_t = 2$ for 2nd 1/2 of clicks)	0.21 (0.13, 0.29)	−0.91 (−1.18, −0.63)	Yes	Yes
6. Recency weighted ($\kappa = 1$, $\tau = T$)	0.21 (0.13, 0.28)	−0.85 (−1.17, −0.54)	Yes	Yes	15. Differential weight (1st 1/3) ($w_t = 2$ for first 1/3 of clicks, $w_t = 1$ for remaining 2/3)	0.21 (0.12, 0.29)	−0.94 (−1.22, −0.67)	Yes	Yes
7. Primacy weighted ($\kappa = 0.25$, $\tau = 1$)	0.20 (0.11, 0.28)	−1.00 (−1.28, −0.73)	Yes	Yes	16. Differential weight (2nd 2/3) ($w_t = 1$ for first 1/3 of clicks, $w_t = 2$ for remaining 2/3)	0.21 (0.13, 0.29)	−0.91 (−1.18, −0.64)	Yes	Yes
8. Primacy weighted ($\kappa = 0.5$, $\tau = 1$)	0.19 (0.11, 0.27)	−1.05 (−1.34, −0.77)	Yes	Yes	17. "Runs" measure	0.14 (0.09, 0.19)	−1.64 (−2.12, −1.16)	Yes	Yes
9. Primacy weighted ($\kappa = 0.75$, $\tau = 1$)	0.18 (0.10, 0.26)	−1.08 (−1.37, −0.79)	Yes	Yes	18. "Maximum run length" measure	0.28 (0.18, 0.38)	−0.74 (−0.96, −0.53)	Yes	Yes

Equation (8) reduces to (1).) Inverted U-shaped and U-shaped weights are applied in rows 11 and 12, whereas results for several differential weights are given in rows 13–16. Finally, regressions from the runs measures of Equations (10) and (11) are presented in rows 17 and 18, respectively. An examination of the evidence in all 18 rows of Table 5 suggests that the different metrics offer unanimous support for H1 and H2 and support the main conclusions that emerged from the baseline specification involving PATTERN.[4]

6. Discussion

Our results support the key hypotheses motivating this paper—that alternative-based processing is associated with higher propensity to buy and that shoppers in the low price category tend to process information in alternative-based patterns. The results also support the latent version of PATTERN over its traditional counterpart and reveal that joint modeling of the information processing pattern and buying behavior is important at the point of purchase. These results have important implications for basic research in information processing and management.

[4] The main conclusions were subjected to an additional robustness check; we reestimated model M_1 after eliminating shoppers with fewer than five clicks. The results obtained from the subsample were consistent with those from the full data sample and supported both H1 and H2.

To the best of our knowledge, our main finding that alternative-based processing is associated with an increased propensity to buy is the first empirical result on this relationship, particularly in a field setting. This result offers a link between the experimental information processing studies over the last four decades and propensity to buy in a field setting. By connecting these studies with propensity to buy, our results can generate a set of managerial strategies whose effectiveness can be explored at the point of purchase. For example, one antecedent condition for the processing pattern studied in the information processing literature is the complexity of the task environment (e.g., the number of alternatives, attributes, dissimilarity of the options). Posting additional products on a commercial website may help to boost sales by serving a wider audience; it can also increase the complexity of the choice task and may promote attribute-based processing patterns and lower propensity to buy. Our work provides a first step toward quantifying these trade-offs that could aid management in its pursuit of an optimal website design. An example of such facilitation is occasionally, although not often, found in business practice. For instance, Apple's website recommends particular iPods for certain types of use—iPod Shuffle is for those who want to clip a lightweight model to a sleeve, belt, or gym clothing (for ultraportability); the iPod Nano is for those who want to shake, shuffle, and roll (for music lovers); the iPod Classic is for those who enjoy for music, movies, podcasts, and audio books; and the iPod Touch is for those interested in Internet browsing, games, videos, and songs.

Similar managerial implications can be derived from our second finding—namely, that consumers in the low price category are more likely to employ alternative-based processing and have a higher propensity to buy than shoppers in the high price category. The practical appeal of using this finding as a targeting device is the ease with which a price category can be observed, as it does not require additional tracking or the sophisticated computation required for obtaining a pattern of information processing. Prices and other product attributes should be highly visible to enable low price category shoppers to quickly find an acceptable product and employ an alternative-based evaluation that ensures that the product does not have any negative facets that deter purchase. In the high price category, visible prices and other product attributes would allow customers to easily identify similarly priced alternatives so that they can efficiently conduct an attribute-based evaluation that allows them to understand the trade-offs required to attain the best value.

The implications of our findings are relevant not only to websites but to brick-and-mortar stores as well. Imagine two similar consumers who are interested in products displayed in a particular product category; however, one consumer is more interested in a specific product and has some questions about its features (alternative-based processing), whereas the other is interested in how three different products compare on several features (attribute-based processing). Based on our results, a salesperson should, ceteris paribus, expect that the customer who is interested in one product will be more likely to buy and that the other customer should be asked clarifying questions (intended use, expected quality, etc.) and given recommendations so as to quickly focus on just one or two alternatives and help transition to alternative-based processing. We expect that our results will generalize across product categories, retailers, and consumers, although further empirical work is warranted to understand the idiosyncrasies of each specific setting.

Methodologically, our conclusions are that joint modeling is important and, equally if not more importantly, that the latent variable specification of PATTERN is better than conditioning on its observed counterpart. This finding will be important for future information processing studies that study processing patterns in lab-based and online shopping settings wherein there may be unobserved consumer background variables, which can influence information processing patterns and propensity to buy. Because our data collection setting involves shoppers who arrived online to either purchase or not purchase in a durable product category, we have one observation per shopper; consequently, we are unable to allow for heterogeneity across shoppers in the impact of information processing patterns on propensity to buy or the impact of price category on information processing patterns. The focus of the current work is to establish the basic relationships between PATTERN and propensity to buy as well as price categories and PATTERN. Important future advances would be to consider segments of consumers whose processing strategies or cognitive costs vary, e.g., those in Table 2, and longitudinal studies designed to understand and better account for heterogeneity across shoppers and dynamic aspects of decision making such as state dependence, evolution of tastes, habit formation, or learning. One benefit of the model considered herein is that it can be inserted as a component in a larger hierarchical model. In particular, embedding our current model as a layer in a hierarchical multinomial choice model can provide guidance not only about whether a customer will buy but also, conditionally upon buying, about what product will be purchased. We hope that future research on these issues will build on our efforts.

Acknowledgments
The authors thank the Internet Technology Group, Inc. for providing the data. They also thank Rick Andrews, Aradhna Krishna, Ron Niedrich, Connie Pechmann, and S. Siddarth, as well as seminar participants at Bar Ilan University, Hebrew University, Interdisciplinary Center (IDC) Herzliya, Tel Aviv University, and conference participants at the 2011 International Behavioral Decision Making Conference, the 2010 Direct/Interactive Marketing Research Summit, and the Paul Merage School of Business Brown Bag Series, for helpful comments. The authors also thank the editor, two associate editors, and two reviewers for helpful suggestions on positioning, options to the PATTERN measure, and the methodological framework. This research was partially supported by the Institute of Mathematical Behavioral Studies at the University of California, Irvine and the Dean's Office of the Merage School of Business.

Appendix. MCMC Estimation Algorithm and Prior Specification

Estimation Approach

Under the prior distributions $\pi(\beta) = \mathrm{N}(\beta \mid \beta_0, B_0)$ and $\pi(\Omega) \propto \mathrm{IW}(r_0, R_0)1\{\Omega_{22} = 1\}$, we develop an MCMC estimation algorithm, which recursively samples the full-conditional distributions of β, Ω, and the latent data $\{y_i^*\}$. The MCMC algorithm includes the latent data explicitly so as to facilitate estimation. The approach proceeds as follows.

Algorithm 1 (MCMC Estimation)

1. Sample $[\beta \mid y^*, \Omega] \sim \mathrm{N}(\widehat{\beta}, \widehat{B})$,

$$\text{where } \widehat{B} = \left(B_0^{-1} + \sum_i X_i' \Omega^{-1} X_i\right)^{-1} \quad \text{and}$$

$$\widehat{\beta} = \widehat{B}\left(B_0^{-1}\beta_0 + \sum_i X_i \Omega^{-1} y_i^*\right).$$

2. Sample $[\Omega \mid y^*, \beta]$ by drawing $\omega_{11\cdot 2} \sim \mathrm{IW}(r_0 + n, Q_{11})$ and $\omega_{12} \sim \mathrm{N}(Q_{22}^{-1} Q_{21}, \omega_{11\cdot 2} Q_{22}^{-1})$, where $\omega_{11\cdot 2} \equiv \omega_{11} - \omega_{12}\omega_{22}^{-1} \cdot \omega_{21}$ and $Q = R_0^{-1} + \sum_i (y_i^* - X_i\beta)(y_i^* - X_i\beta)'$, from which Ω can be recovered directly.

3. For $i = 1, \ldots, n$, sample $[y_{iIP}^* \mid y_{iPB}^*, y_{iIP}, \beta, \Omega] \sim \mathrm{TN}_{S_i}(\mu_{1|2}, V_{1|2})$, where the region S_i of the truncated normal distribution is implied by the censoring of y_{iIP}, and $\mu_{1|2}$ and $V_{1|2}$ are the usual conditional moments; at each step, also sample $[y_{iPB}^* \mid y_{iIS}^*, y_{iPB}, \beta, \Omega] \sim \mathrm{TN}_{S_i}(\mu_{2|1}, V_{2|1})$, where S_i is the region $(0, \infty)$ if $y_{iPB} = 1$ and $(-\infty, 0)$ otherwise.

The first step in Algorithm 1 follows the sampling of seemingly unrelated regression models (see Chib and Greenberg 1995), the second step follows from the properties of the inverse Wishart distribution (Drèze and Richard 1983, Chib et al. 2009), and the final step exploits the data augmentation techniques proposed in Chib (1992) and Albert and Chib (1993).

Training Sample Priors

The prior densities in our application are determined through a training sample approach (for an overview, see Greenberg 2008 or O'Hagan 1994). We take 150 observations (roughly one-sixth of our data) as a training sample, and the remainder is retained as a comparison sample. The data in the training sample are used to construct a first-stage posterior distribution that, in turn, will be used as a proper informative (training sample) prior when analyzing the comparison sample. This approach takes advantage of the Bayesian updating principle and uses all available information in the sample—the eventual posterior distribution combines information from the training sample prior with information from the comparison sample, which is embodied in the likelihood. This technique is also appealing because of its neutrality on the signs and magnitudes of covariate effects. To be cautious, however, we have conducted local sensitivity analysis to ensure that the size of the training sample does not alter the model rankings, parameter estimates, or the substantive conclusions of our paper. This was done by varying the size of the training sample by up to one-half its original size and examining the resulting model rankings to ensure that they do not change. Varying the setup in this way did not change model rankings, with M_1 always leading M_2 and M_3 with a log-marginal likelihood difference of at least 6.

References

Albert JH, Chib S (1993) Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* 88(422):669–679.

Bettman JR (1979) *An Information Processing Theory of Consumer Choice* (Addison-Wesley, Reading, MA).

Bettman JR, Park CW (1980) Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *J. Consumer Res.* 7(3):234–248.

Bettman JR, Luce MF, Payne JW (1998) Constructive consumer choice processes. *J. Consumer Res.* 25(3):187–217.

Chib S (1992) Bayes inference in the Tobit censored regression model. *J. Econometrics* 51(1/2):79–99.

Chib S (1995) Marginal likelihood from the Gibbs output. *J. Amer. Statist. Assoc.* 90(432):1313–1321.

Chib S, Greenberg E (1995) Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models. *J. Econometrics* 68(2):339–360.

Chib S, Jeliazkov I (2006) Inference in semiparametric dynamic models for binary longitudinal data. *J. Amer. Statist. Assoc.* 101(474):685–700.

Chib S, Greenberg E, Jeliazkov I (2009) Estimation of semiparametric models in the presence of endogeneity and sample selection. *J. Comput. Graphical Statist.* 18(2):321–348.

Dhar R, Nowlis SM (2004) To buy or not to buy: Response mode effects on consumer choice. *J. Marketing Res.* 41(4):423–432.

Drèze JH, Richard J-F (1983) Bayesian analysis of simultaneous equation systems. *Handbook of Econometrics* (North-Holland, Amsterdam), 517–598.

Gensch DH (1987) A two-stage disaggregate attribute choice model. *Marketing Sci.* 6(3):223–239.

Gilbride TJ, Allenby GM (2004) A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* 23(3):391–406.

Greenberg E (2008) *Introduction to Bayesian Econometrics* (Cambridge University Press, New York).

Hauser JR, Wernerfelt B (1990) An evaluation cost model of consideration sets. *J. Consumer Res.* 16(4):393–408.

Hong J, Sternthal B (2010) The effects of consumer prior knowledge and processing strategies on judgments. *J. Marketing Res.* 47(2):301–311.

Iyengar SS, Lepper MR (2000) When choice is demotivating: Can one desire too much of a good thing? *J. Personality Soc. Psych.* 79(6):995–1006.

Jacoby J, Chestnut RW, Fisher WA (1978) A behavioral process approach to information acquisition in nondurable purchasing. *J. Marketing Res.* 15(4):532–544.

Jeliazkov I, Lee E (2010) MCMC perspectives on simulated likelihood estimation. *Adv. Econometrics* 26:3–39.

Jeliazkov I, Graves J, Kutzbach M (2008) Fitting and comparison of models for multivariate ordinal outcomes. *Adv. Econometrics* 23:115–156.

Liu Q, Arora N (2011) Efficient choice designs for a consider-then-choose model. *Marketing Sci.* 30(2):321–338.

Mintz A, Geva N, Redd SB, Carnes A (1997) The effect of dynamic and static choice sets on political decision making: An analysis using the Decision Board Platform. *Amer. Political Sci. Rev.* 91(3):553–566.

Moe WW (2006) An empirical two-stage choice model with varying decision rules applied to Internet clickstream data. *J. Marketing Res.* 43(4):680–692.

Moe WW, Fader PS (2004) Dynamic conversion behavior at e-commerce sites. *Management Sci.* 50(3):326–335.

Montgomery AL, Li S, Srinivasan K, Liechty JC (2004) Modeling online browsing and path analysis using clickstream data. *Marketing Sci.* 23(4):579–595.

Moorthy S, Ratchford BT, Talukdar D (1997) Consumer information search revisited: Theory and empirical analysis. *J. Consumer Res.* 23(4):263–277.

O'Hagan A (1994) *Kendall's Advanced Theory of Statistics: Bayesian Inference* (John Wiley & Sons, New York).

Payne JW (1976) Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ. Behav. Human Performance* 16(2):366–387.

Payne JW, Bettman JR, Johnson EJ (1993) *The Adaptive Decision Maker* (Cambridge University Press, Cambridge, UK).

Ratchford BT (1982) Cost-benefit models for explaining consumer choice and information seeking behavior. *Management Sci.* 28(2):197–212.

Schwarz G (1978) Estimating the dimension of a model. *Ann. Statist.* 6(2):461–464.

Simon HA (1956) Rational choice and the structure of the environment. *Psych. Rev.* 63(2):129–138.

Simonson I, Huber J, Payne JW (1988) The relationship between prior brand knowledge and information acquisition order. *J. Consumer Res.* 14(4):566–578.

Sismeiro C, Bucklin RE (2004) Modeling purchase behavior at an e-commerce web site: A task-completion approach. *J. Marketing Res.* 41(3):306–323.

Tellis GJ, Gaeth GJ (1990) Best value, price-seeking, and price aversion: The impact of information and learning on consumer choices. *J. Marketing* 54(2):34–45.

Valenzuela A, Dhar R, Zettelmeyer F (2009) Contingent response to self-customization procedures: Implications for decision satisfaction and choice. *J. Marketing Res.* 46(6):754–763.