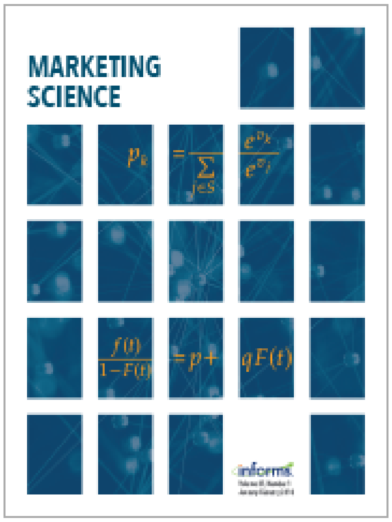


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Satisficing Choice Model

Peter Stüttgen, Peter Boatwright, Robert T. Monroe,

MARKETING SCIENCE
Vol. 31, No. 6, November–December 2012, pp. 878–899
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.1120.0732

A Satisficing Choice Model

Peter Stüttgen
Business Administration, Carnegie Mellon University in Qatar, Doha, Qatar, pstuettg@andrew.cmu.edu

Peter Boatwright, Robert T. Monroe
Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
{boatwright@cmu.edu, bmonroe@cmu.edu}

Although the assumption of utility-maximizing consumers has been challenged for decades, empirical applications of alternative choice rules are still very new. We add to this growing body of literature by proposing a model based on the idea of a "satisficing" decision maker. In contrast to previous models (including recent models implementing alternative choice rules), satisficing depends on the order in which alternatives are evaluated. We therefore conduct a visual conjoint experiment to collect search and choice data. We model search and product evaluation jointly and allow for interdependence between them. The choice rule incorporates a conjunctive rule for the evaluations and, contrary to most previous models, does not rely on compensatory trade-offs at all. The results strongly support the proposed model. For instance, we find that search is indeed influenced by product evaluations. More importantly, the model results strongly support the satisficing stopping rule. Finally, we perform a holdout prediction task and find that the proposed model outperforms a standard multinomial logit model.

Key words: noncompensatory choice; eye tracking; visual conjoint experiment
History: Received: July 15, 2011; accepted: June 12, 2012; Preyas Desai served as the editor-in-chief and Michel Wedel served as associate editor for this article. Published online in *Articles in Advance* September 20, 2012.

1. Introduction

The large majority of choice models in the marketing literature focus on understanding the influence of product attributes or marketing variables on consumer choice. Implicitly, these models assume that the consumer has all information, or at least enough information, needed to form a consideration set according to some rule. But in reality, a consumer needs to *acquire* product information before that information can, in turn, influence his choice decision. Thus, Pieters and Warlop (1999) suggest that visual attention may help understand consumer choice. Using eye-tracking methodology, they find that consumers do not gather all information about the alternatives. Moreover, they show that the percentage of people that look at the different alternatives is predictive of the alternatives' choice shares. In addition, they find directional support that skipping pieces of information (brand information, ingredient information, and/or the pictorial) for an alternative is negatively related to its choice share. In our data, the percentage of people who skipped some information for an alternative is an even better predictor of choice shares (Pearson's $r = -0.75$) than the percentage of people who looked at a given alternative ($r = 0.36$). As discussed in more detail in §2, the phenomenon of skipped information within a product cannot be satisfactorily explained by standard utility-maximizing models, even if allowing for search cost. Yet the high correlation between choice shares and the percentage of skipped information suggests that being able to understand *why* people skip some attribute information may be important for understanding consumer choice behavior. The satisficing model proposed in this paper does not only fully explain such behavior but also responds to the repeated call for models with better "ecological rationality" (Netzer et al. 2008, p. 346), i.e., the call for models that are "better representations of decision processes" (Johnson et al. 1989, p. 268).

The assumption that consumers are utility maximizers has been criticized for a long time on grounds of the unrealistically high cognitive burden these rules impose on the decision maker (Simon 1955) and because of common violations of basic axioms of utility-maximizing rules (e.g., transitivity; see Loomes et al. 1991). Instead, consumers are believed to use simplified choice heuristics when making their choices (Gigerenzer and Todd 1999). Although several simple alternative choice rules were proposed several decades ago (e.g., Coombs 1951), such rules have been incorporated into empirical choice models only recently (e.g., Gilbride and Allenby 2004). We contribute to this new stream of research in the marketing literature by proposing a choice model based on Herb Simon's idea of "satisficing" choice (Simon 1955).

Satisficing is a simple choice rule in which the first alternative that is good enough according to some criterion is chosen. The outcome of a satisficing choice then is search path-dependent. Any model of product choice therefore must also account for the search path. Moreover, the search path may be dependent on what "good enough" means. This interdependence between search and evaluation allows us to parsimoniously explain the phenomenon of skipped information in the framework of a simple choice rule. To do so, we model the search path and the resulting evaluation jointly. We collect search path data in what we term a "visual conjoint experiment." Using realistic shelf images for which product attributes vary according to a conjoint design, we use eye-tracking technology to record consumers' search paths in addition to the standard choice data.

The satisficing choice rule implies a distinct stopping rule for the search process. Our results strongly confirm this stopping rule, providing support that consumers may in fact use a satisficing choice rule rather than a utility-maximizing model in the product category that we studied. The satisficing choice rule is further supported by the results of a holdout prediction task in which the proposed satisficing model comfortably outperforms a multinomial logit model.

The remainder of this paper is organized as follows. We will first briefly review the relevant streams of literature and then describe the experiment and the data before explaining the proposed model and estimation. Finally, we present the results and conclude with a general discussion.

2. Literature Review

We first review the traditional approach to choice models and its merits, but we also argue why that approach cannot convincingly explain the observed data with skipped information. Then we review the literature on bounded rationality and show that a satisficing choice rule in contrast can explain the observed data very parsimoniously. Finally, given our use of an eye-tracking experiment, we briefly review the relevant work on eye-tracking–based search models.

2.1. Maximizing Choice

At least since Guadagni and Little (1983) introduced the multinomial logit model to the marketing literature in their seminal paper, the idea of a compensatory utility-maximizing choice has been the predominant framework for empirical analyses of consumer choice. The theory of utility-maximizing choice has its foundations in the tenets of microeconomics. Typically, utility is specified as a linear combination of the alternative's attributes, thereby making it a compensatory process (i.e., a "bad" value for one attribute can be compensated for by a "good" value for another attribute). The approach has proven to be straightforward to implement and can yield valuable managerial insights—for instance, enabling managers to segment the market (e.g., Kamakura and Russell 1989) or to understand the impact of marketing decisions (e.g., Gupta 1988). Building on this framework, the more recent advent of structural models has allowed researchers to examine consumers' strategic and forward-looking behavior (e.g., Sun 2005).

However, the assumption of a rational consumer with unlimited cognitive capabilities, as theoretically appealing as it may be from a normative standpoint, has long been challenged as an appropriate representation of actual human decision makers (e.g., Simon 1955, Kahneman and Tversky 1979). Even proponents of the utility-maximizing approach typically agree that decision makers may not actually make decisions following the rules of the model, but rather they act *as if* they did. The models then are seen as a description of the outcome rather than the process. Nonetheless, following Shugan's call for incorporating a "cost of thinking" (Shugan 1980) to allow for more realistic models, the literature on choice models has started to account for limited consumer search and introduced cognitively less demanding (for the decision maker) models.

Importantly, Hauser and Wernerfelt (1990) and Roberts and Lattin (1991) incorporated consideration sets into choice models; i.e., they proposed a two-stage process in which only a subset of the available alternatives is selected in the first stage for a utility-maximizing choice in the second stage. The formation of the consideration set was originally dependent on cost–benefit trade-offs for including an additional brand into the consideration set or, with the rise of structural models, an explicit trade-off of search cost and expected benefit (Mehta et al. 2003). Yet these models of constrained utility maximization, in fact, typically *increase* the computational burden of the decision maker rather than decrease it. In the words of Gigerenzer and Todd (1999, p. 11), "The paradoxical approach of optimization under constraints [i.e., optimization including a search or other sort of cost] is to model 'limited' search by assuming that the mind has essentially unlimited time and knowledge with which to evaluate the costs and benefits to further information search."

Yet even if we embrace constrained utility maximization as an appropriate framework to model consumer decision making, we encounter serious problems in trying to explain the search patterns observed in our visual conjoint experiment (to be described in §3). Incorporating the cost of thinking (i.e., search cost and/or cost for evaluating information) in the model framework can easily explain why consumers may not evaluate all available alternatives. If consumers use a stopping rule for search based on cost–benefit trade-offs, it may be optimal to not search all available options. However, search costs do not plausibly explain why consumers may start evaluating some alternatives but not collect all information about them (in the standard framework with compensatory utility), a frequent pattern in the search process. One explanation might be that search cost within a product is higher than across products. However, this does not seem likely because (a) search cost should be very low within a product, as the shopper only needs to move his eyes minimally, and (b) integration of information should not be very difficult, as the product category (ramen noodles) does not involve difficult trade-offs between attributes. Alternatively, incomplete search within a product could occur in a search cost framework if a given consumer cares a lot more about one attribute (say, flavor) than another (say, price). In that case, knowing that a product has a flavor he really dislikes may make it unprofitable for him to acquire the other attributes for this product, because even the lowest price may not offset the disutility caused by the flavor. However, in our data we find that it is not always the same attribute that is missing (within a person), so therefore this cannot be the correct explanation.

As we will show below, though, both limited search across alternatives as well as limited search within alternatives can be easily explained by a model that is not based on a utility-maximizing framework but instead uses a satisficing choice rule.

2.2. Bounded Rationality

In his previously mentioned critique of the rational utility-maximizing agent, Simon says that "the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information to the computational capacities that are actually possessed by . . . man" (Simon 1955, p. 99).[1] This is the foundation of what has come to be known as "bounded rationality." In this view, decision makers are aware of their cognitive limits and therefore rely on simplified choice rules (depending on the task). The best known of these simplified choice rules are the lexicographic rule, the conjunctive and disjunctive rules, and elimination by aspects.

In the lexicographic rule (von Neumann and Morgenstern 1947), a decision maker focuses on the attribute that is most important to her and simply chooses the alternative that is best on that particular attribute. If there is a tie, she compares the tied alternatives on her second most important attribute and chooses the alternative that is preferred according to that attribute. The process continues until a unique choice is found (or until all attributes are exhausted).

In the conjunctive and disjunctive rules (Coombs 1951, Dawes 1964), the decision maker has individual threshold levels for all attributes. In the conjunctive rule, every product that passes *all* of these thresholds is acceptable to him, whereas in the disjunctive rule, all products that pass *at least one* threshold are acceptable.

Finally, elimination by aspects (Tversky 1972) is essentially a combination of the lexicographic and the conjunctive rules in which a decision maker first focuses on the most important attribute. Then, not only does the best alternative make it to the second round, but so do all alternatives passing the threshold for the particular attribute.

Notice that all of these choice rules are noncompensatory; i.e., a bad value for one attribute may be enough for not choosing a particular product, regardless of how good it may be on other attributes. Thus, these decision rules tremendously simplify the decision process, because the decision maker does not have to evaluate any trade-offs between attributes. Despite the fact that these models were proposed several decades ago, and despite more and more behavioral evidence that central assumptions of the utility-maximizing framework seem to be violated in reality (see Bettman et al. 1991 for a review of consumer decision making), most empirical applications of noncompensatory models stem only from the last decade (but see Fader and McAlister 1990).

Several of these applications have extended the linear utility framework to be able to capture screening rules based on these simplified rules (e.g., Swait 2001, Elrod et al. 2004), and others have directly modeled the simplified rules (see Gilbride and Allenby 2004 and Jedidi and Kohli 2005 for the conjunctive and disjunctive rules, Kohli and Jedidi 2007 for the lexicographic rule, and Gilbride and Allenby 2006 for elimination by aspects). However, there are two main issues with these alternative choice rules. First, these simplified choice rules are *attribute based*; i.e., to make a choice, the decision maker needs to know the values of at least one attribute for *all* alternatives. In our data (see §3) it is easy to see that this assumption is clearly violated in almost all instances. In contrast, satisficing is an *alternative-based* choice rule, meaning that alternatives are evaluated sequentially rather than simultaneously across one or more attributes. Second, the conjunctive and disjunctive rule only result in a set of acceptable products. Thus, the applications of the

[1] Following the rules of the well-known children's game, we attempt to do exactly as "Simon says" in this paper.

disjunctive and conjunctive rules so far have either been used to predict a whole set (like the acceptable MBA candidates in Jedidi and Kohli 2005) or are followed by, or combined with, a compensatory choice rule (Swait 2001, Elrod et al. 2004, Gilbride and Allenby 2004).

"Satisficing," a term coined by Simon that combines "satisfactory" and "sufficing," in contrast results in a unique choice *and* can explain limited search. The decision process is very simple: Start by evaluating one alternative. If it is satisfactory (according to a criterion to be defined), choose that product and stop searching. If not, evaluate the next alternative. Continue this process until you have found a satisfactory alternative.

It is obvious why people following a satisficing choice rule may not search all alternatives. However, depending on the satisfaction criterion, satisficing can also explain incomplete information acquisition within a product. For instance, say the satisfaction criterion is given by a conjunctive rule. Then, once the decision maker knows that the product fails to meet the threshold on one attribute (be it price, flavor, or brand), there is no reason to continue the search within this product.

One difficulty for the empirical application of a satisficing rule is that the choice outcome depends on the sequence of evaluation. If there is more than one satisfactory product, the decision maker will choose whichever she comes across first. Thus, it is essential to know the search sequence. We therefore use eye-tracking technology to record the sequence of information acquisition in our visual conjoint experiment.

2.3. Eye Tracking

Eye-tracking hardware has improved tremendously over the last 15 years, allowing for unobtrusive observation of a person's eye fixations. In an early application, Russo and Leclerc (1994) relied on human coders to code the location of the fixations based on a video of each participant's face. They identified three processing stages in consumer choice: orientation, evaluation, and verification. In the verification stage, most relevant to our work for reasons that will become clear later, participants continue to search and acquire information despite already having made a choice. In the study by Pieters and Warlop (1999), the data collection was automated, but consumers had to keep their heads fixed to the apparatus. Recent improvements in software and hardware now allow participants to move freely in about a $25 \times 25 \times 25$-inch box, and the location of their eye fixations is determined based on the eyes' reflection of infrared signals.

Applications of eye tracking in marketing research have brought valuable insights in consumers' processing of print ads (e.g., Pieters et al. 2002), the optimal design of TV commercials (Teixeira et al. 2010), and the effect of in-store marketing activities (Chandon et al. 2009), among others. For a review of the findings from eye-tracking applications in marketing, see Pieters (2008). Most relevant to our application, though, is the research explaining consumer search patterns. van der Lans et al. (2008a) showed that consumer search is influenced both by features of the stimulus, so-called bottom-up effects, and by strategic or intentional strategies, so-called top-down effects. Liechty et al. (2003) proposed a hidden Markov model to capture two distinct types of search: local and global. In the local search state, "stimuli are explored in detail by extracting information from specific and adjacent locations," whereas the global state "is characterized by longer saccades" (i.e., movements between fixations) and "stimuli are explored to identify locations to extract information" (Liechty et al. 2003, p. 520).

3. Visual Conjoint Experiment

As discussed above, the empirical application of a satisficing model requires knowledge of both the search path and product choice. Moreover, because search and product evaluations may be interdependent, we model the search path and the evaluations jointly. To collect the data needed to do this, we conduct what we term a *visual conjoint experiment*, i.e., we develop a standard conjoint design, but then translate the resulting choice sets into realistic images of shelves from which participants make their selections.

Our experimental setup is similar to the one used by Reutskaja et al. (2011). However, there are several major differences. First, Reutskaja et al. (2011) added time pressure (three seconds per decision) and a severe punishment ($3) for failing to make a choice. The introduction of the punishment may explain why the authors failed to find consistent support for a satisficing rule, because participants were forced to make a choice even if no product was in fact satisfactory. Second, the analysis presented in Reutskaja et al. (2011) is on the aggregate level. Yet aggregate analysis may fail to find discontinuities in individual-level choice rules if there is heterogeneity in the individual points of discontinuity (e.g., see Williams and de Dios Ortuzar 1982). Finally, Reutskaja et al. (2011) treated the search path as exogenous (and on the product level), whereas we endogenously model the search path on the attribute level. This allows us to gain more insight into the underlying processes.

3.1. Participants

The experiment was conducted at the Doha, Qatar campus of Carnegie Mellon University. Participants were 75 undergraduate students from the Doha, Qatar campuses of Carnegie Mellon University, Texas A&M University, Georgetown University, Northwestern

Figure 1 Example Stimulus

University, and Cornell Medical College. Eleven of these participants were excluded from the analysis because of calibration problems and/or incomplete eye recordings, leaving a total of 64 students (29 female, 35 male) in the sample.[2] Participants' ages ranged from 17 to 23 years, with a mean of 19.88 years. Participants' nationalities were predominantly (~55%) South Asian (e.g., Indian, Pakistani, Bangladeshi), and participants from Middle Eastern countries combined for a total of 18 participants (28%).[3] Six out of the 64 participants were U.S. citizens. Subjects were paid approximately \$14 (depending on their choices), and sessions lasted between 30 and 60 minutes.

3.2. Stimuli and Procedure

We chose instant noodles (also known as ramen noodles) as the product category. Products vary by price, flavor, and brand. We used four brands, five equidistant price levels (ranging from ~\$1.10 to ~\$1.90 for a five-pack of noodles), and 10 flavors. The brands and flavors were selected from brands and flavors present in the local market. Similarly, the price levels spanned the price range found in the local market. The conjoint design is orthogonal and consists of 15 choice sets with 15 alternatives each. We translated each choice set into an image of three shelves with five alternatives each. To approximate a realistic amount of clutter on the shelves, each alternative has four facings. See Figure 1 for an example.[4]

Subjects participated in the experiment in individual sessions. Figure 2 shows the setup of the experiment. Participants were standing in front of a screen with the eye tracker (Tobii Technology X60) in between them and the screen. We used a 50-inch, high-definition television (1920×1080 pixels) to display the stimuli, which allowed the products to be approximately life-sized and made all information easily readable. The eye tracker also allowed free movement of the head, recording eye positions at a frequency of 60 Hz and with high precision (the deviation of measured and true gaze direction is at most 0.5°).

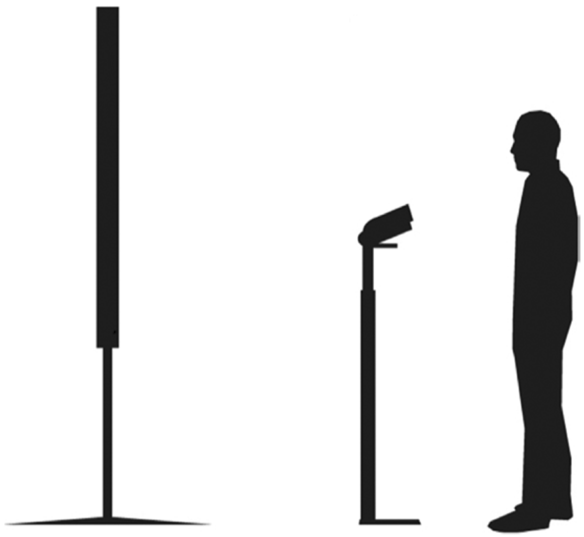
After reading the instructions, including a list of the available brands and flavors as well as an example shelf image, the eye-tracking software was calibrated. For the individual-specific calibration, subjects were asked to follow a dot moving around the screen with their eyes to "teach" the software how eye movements relate to location on the screen. Calibration was repeated after one-third and after two-thirds of the experiment had been conducted to ensure high-quality data. After calibration, the first shelf image appeared on the screen, and participants could take as long as they needed to make a decision. Once they reached a decision, they clicked a button on a

[2] The calibration procedure is explained in §3.2.

[3] Many of the participants lived in Qatar for most, if not all, of their lives.

[4] Note that the shelf positions are randomized, whereas in reality, products are usually grouped by brand. Our design may lead to more search in the experiment than in reality, which provides us with more information on consumer preferences.

Figure 2 Experimental Setup



Source. ©Tobii Technology, Sweden. Reprinted with permission.

presentation clicker that caused the screen to blur, and the products were overlayed with letters from A to O. This was done to prohibit acquisition of additional information after a choice had been made. Note, however, that the verification state of Russo and Leclerc (1994), if present, then inherently becomes a part of the recorded search path; we therefore allowed for the presence of a verification stage in our model. Subjects then indicated their choice by announcing the corresponding letter to the experimenter, or they said that they chose not to buy anything from this particular choice set (Pieters and Warlop 1999). After the last choice, participants completed a questionnaire to collect, among other things, explicit measures of their preferences.

To ensure that the task was incentive compatible, one of the choice sets was selected at the end of the experiment and the corresponding purchase realized (i.e., participants received their chosen item and paid the respective price from their participation fee).

3.3. Data

For each participant, we then have the 15 choice outcomes, the questionnaire responses, and the sequence of the locations of eye fixations for each choice set. Because our interest lies mainly in information acquisition, we aggregate the pixel-level data into meaningful areas of interest (AOI)—namely, the price tag, the flavor information, and the rest of the package for each of the alternatives, plus fixations on the background (Pieters and Warlop 1999, Shi et al. 2012). All fixations in our data are longer than 50 milliseconds, the minimum length required for information acquisition in complex visual scene perception (van Diepen et al. 1995). We thus assume that a fixation on an AOI leads to acquisition of the respective information. Following Shi et al. (2012), we exclude fixations on the background (7.7% of all fixations) as well as consecutive repeat fixations on the same AOI (20.2%), because they are not informative about a consumers information acquisition process (consecutive repeat fixation may occur from moving the eyes slightly to read flavor or price information, for instance). Thus, we have 45 AOIs (15 products with 3 AOIs each) that provide an exhaustive and mutually exclusive partition of each shelf image. Because the packaging distinguishes brands, and brands are well known, we assume that participants learn a product's brand by looking anywhere on the packaging (including the flavor AOI), whereas they have to fixate on the corresponding AOI to learn the flavor or price.

Figure 3 shows the median number of fixations per choice set. It is obvious that participants tend to search longer in the first few shelf images, most likely to get used to the task. For the effect of the number of fixations on the likelihood of termination (see §4.4.4), we therefore normalize the number of fixations by the median number of fixations for the respective choice set.[5] The number of fixations within a subject varies greatly across choice sets; even when only considering the last 11 choice sets (i.e., when median fixations have stabilized), the mean (across participants) standard deviation (for one participant across choice sets) is 13.9 fixations. This suggests that participants do not simply follow a fixed-search stopping rule but employ a more variable stopping rule depending on the information acquired in a particular search.

Tables 1 and 2 provide a summary of brand and flavor choices, respectively, giving a first indication of consumer preferences. The Maggi brand as well as the chicken and onion chicken flavors are clear consumer favorites. On average, a person chose 3.03 out of the 4 brands ($SD = 0.96$) and 5.09 of the 10 flavors ($SD = 2.08$) across the 15 choice tasks. Almost 70% of people chose more than half of the brands, and almost 60% chose at most half of the available flavors. Overall, participants decided not to buy in 7.0% of choices. As mentioned in §1, the percentage of people looking at a given alternative is positively correlated with its choice share ($r = 0.36$), and the percentage of people who gather all three attributes for an alternative is very strongly correlated with its choice share ($r = 0.75$).

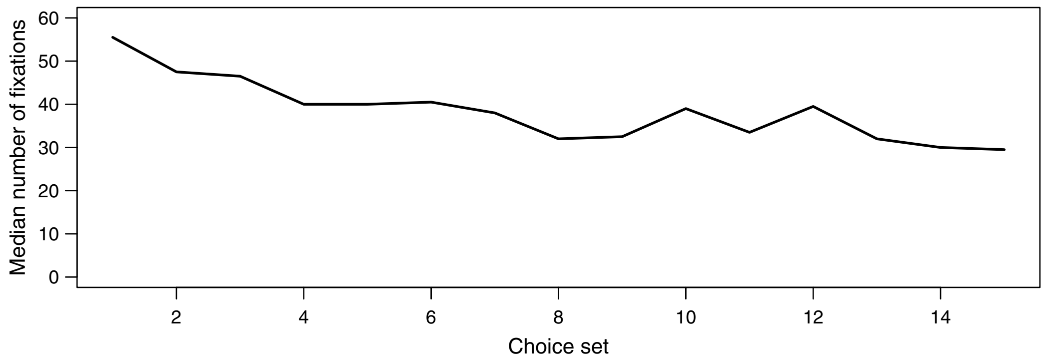
4. Model

4.1. Notation

In this section, we briefly introduce the notation for the rest of the paper. Lowercase letters are variables to

[5] Results are qualitatively equivalent if we instead exclude the first three choice sets. We prefer the normalization as to not lose the information contained for other parts of the model.

Figure 3 Median Number of Fixations per Choice Set



be estimated, with Greek letters (e.g., ϕ, γ) referring to model parameters and Latin letters (e.g., l, g) referring to unobserved states of the search process. Capital Latin letters (e.g., L, R) are generally observed explanatory variables. Scripted Latin letters (e.g., $\mathscr{S}$, $\mathscr{U}$) are variables that depend on the parameter estimates but are deterministic conditional on a set of estimates and the observed data. Finally, we use $\mathbb{I}_y$ as the indicator function for condition y being true.

We use the following subscripts: $i = 1, \ldots, 64$ for participants, $j = 1, \ldots, 15$ for choice sets, $f_{ij} = 1, \ldots, F_{ij}$ for participant i's fixations in choice set j, and $x_j = 1, \ldots, 15$ for the products in choice set j. However, for readability we typically suppress subscripts i and j in the presentation of the model. Except for the parameters of the hierarchy level, all parameters to be estimated are individual specific but constant across choice sets. In contrast, explanatory variables are typically individual specific and choice set specific.

We let $A \in \{BR, FL, PR\}$ (brand, flavor, and price, respectively) denote the available sets of the three different attributes, and a (a_{x_j}) refers to one particular attribute level (the attribute level of product x_j). The AOIs that make up each image are identified by $h = 1, \ldots, 45$. In turn, $P(h)$ and $A(h)$ then indicate to which product h belongs (in order to specify subscript x) and which attribute information is contained in AOI h. Similarly, $br_j(h)$ (short for $br(p_j(h))$) denotes the brand (in choice set j) of the product to which h belongs. Finally, I_{ijf} denotes the information set of person i in choice set j after fixation f; i.e., it specifies which pieces of information the person knows at that point in time.

Table 1 Brand Choice Shares

Brand	%
Fantastic	19.0
Indomie	22.4
Koka	12.7
Maggi	45.9

Table 2 Flavor Choice Shares

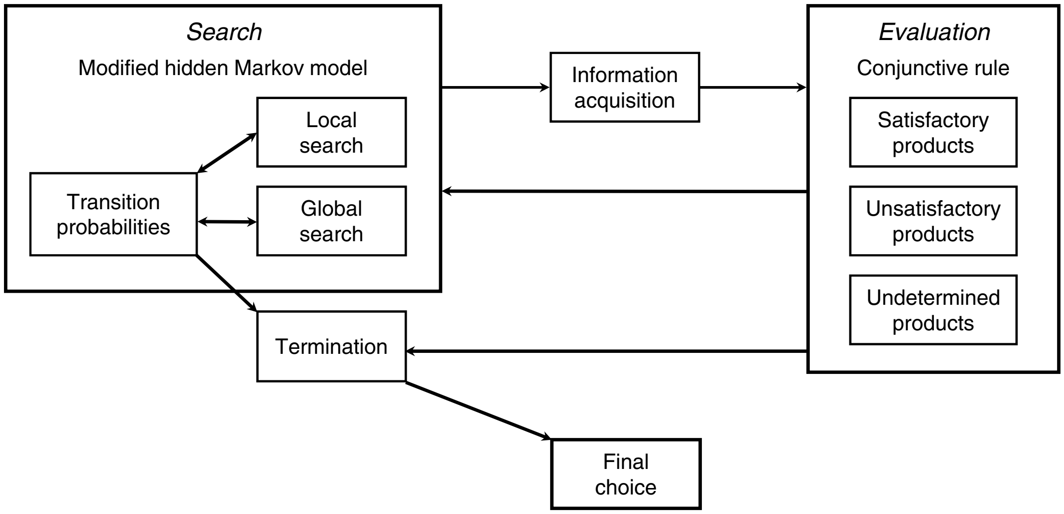
Flavor	%
Beef	6.6
Cheese	11.8
Chicken	23.1
Curry	12.0
Lobster	4.7
Mushroom	4.6
Onion chicken	18.0
Shrimp	6.9
Tomato	5.8
Vegetable	6.5

4.2. General Model

In recent years, there has been a growing literature on memory-based versus stimulus-based choices (e.g., Lee 2002, Rottenstreich et al. 2007). Following Lee (2002), stimulus-based decisions are decisions based on information available in the physical environment, whereas memory-based decisions are decisions based on information retrieved from memory. Of course, many decisions in real life are a mixture of these two extremes. However, choice models in marketing have traditionally focused on the impact of the stimuli (i.e., the attributes of the available products), sometimes with some reduced-form variables like the loyalty variable of Guadagni and Little (1983) capturing at least one aspect of memory-based influences. (But see Mehta et al. 2004 for a recent structural model of memory-based choices and the impact of forgetting.)

Because it is reasonable to believe that consumers may be more likely to use a satisficing choice rule in categories with which they are already very familiar (see §7.1.3 for a wider discussion on when people may be more likely to use a satisficing versus a maximizing choice rule), a satisficing choice model certainly needs to allow for both memory-based and stimulus-based influences on the decision process. We therefore present such a general model in this subsection. However, as we will argue below, our experimental setup unfortunately does not allow us to estimate any memory-based effects in this paper. We therefore restrict the estimation to stimulus-based influences.

Figure 4 Overview of the Model



One may think of the model as consisting of two interrelated parts: search and evaluation (see Figure 4).[6] Though we present them separately, satisficing truly combines both. During search, a person continuously acquires more information. Based on the information gathered up to that point, the person also continuously updates his evaluations of the alternatives, grouping them into satisfactory products, unsatisfactory products, and undetermined products (i.e., products for which he does not have enough information yet to make a judgment). Crucially, we allow for these evaluations to in turn influence the continued search. Of course, the evaluations will also influence the final decision once a person moves to the termination stage.

Both evaluation and search are likely affected by the stimulus *as well as* memory. As for the search, the eye-tracking literature has demonstrated that the stimulus can help guide the consumer's attention as well as distract it. Yet the search path may also be heavily influenced by consumers' memory of the shelf layout. For instance, a consumer may remember the shelf location of the alternative she last bought and start her search close to the recalled location. Similarly, the evaluation of the products will clearly be influenced by the stimulus (e.g., what are the actual product attributes), but it may also be influenced by a consumer's memory of, say, a brand's quality.

If participants did not have the verification stage identified by Russo and Leclerc (1994), then they should stop immediately after finding the first satisfactory alternative. Thus, if people follow a satisficing choice rule, any continued search after finding a satisfactory alternative constitutes a verification stage *by definition*. In almost 80% of the observed choices, the participants continue to search after fully evaluating the alternative she will end up choosing (which, also by definition, is satisfactory).[7]

To accommodate for the verification stage, we relax Simon's strict satisficing rule to allow for continued search after encountering the first satisfactory alternative. Nonetheless, even in the relaxed version, the probability of stopping the search should increase significantly after encountering the first satisfactory alternative. The results for the parameter associated with this stopping rule then provide information on whether participants in fact used a satisficing choice rule or not. In a satisficing choice rule *without* verification stage, this parameter (λ_{t3} in Equation (20)) should go to infinity, resulting in a stopping probability of 1 immediately after seeing the first satisfactory option. In the presence of a verification stage, it will not be infinity but should still be positive and large if a person uses a satisficing choice rule.

4.3. Evaluation

Evaluation refers to the process of determining whether a given product is satisfactory or not, i.e., whether it passes a so-called "aspiration level" or not. If utility is assumed to follow the standard linear additive form, the aspiration level is simply given by a minimum acceptable utility level. However, we use a conjunctive rule as the satisfaction criterion to be used within the satisficing choice, as it can

[6] It may be tempting to think of the evaluation part as the choice model. However, because satisficing is a search path-dependent choice rule, both parts together make up the choice model.

[7] In addition, all participants tend to return their gaze to the alternative they end up choosing right before pressing the button. This is most likely an artifact of the experimental setup asking them to say the corresponding letter after the picture is blurred. (We thank an anonymous reviewer for suggesting this explanation.) We will refer to this later as the end-of-sequence effect.

parsimoniously explain the above-mentioned correlation between the amount of information acquisition per alternative and choice shares. For each of the attributes, an individual has a set of acceptable levels and only if she (1) has learned all three attribute levels for a given product (or can determine the acceptability of an attribute based on memory) and (2) all three attribute levels are acceptable is the product judged to be satisfactory to her.[8] As such, the aspiration level is defined by the combination of the attribute-level acceptability judgments. In contrast, as soon as at least one unacceptable attribute level has been found for a product, the product is judged to be unsatisfactory. The difference in the amount of information needed to judge an option to be satisfactory (i.e., *all* attributes) versus unsatisfactory (potentially only one attribute) parsimoniously explains the above-mentioned correlation. Finally, if no unacceptable attribute level has been encountered yet and not all attribute levels for a given product have been learned, the product status is undetermined. Note that this implies that all products are undetermined at the beginning of the search and can change status at any point during the search.[9]

More formally, let $\gamma_{ia} \in \{0, 1\}$ indicate whether attribute-level a is acceptable to person i. The terms $\mathcal{S}_{ijxf}$, $\mathcal{U}_{ijxf}$, and $\mathcal{UD}_{ijxf}$ are indicators for whether product x has been judged to be satisfactory or unsatisfactory, or is still undetermined, respectively, as of fixation f in choice set j by individual i.

We can then write the conjunctive rule above as follows (suppressing i and j):

$$\mathcal{S}_{xf} = \prod_A \gamma_{a_x} \cdot \mathbb{I}_{a_x \in I_f}, \tag{1}$$

$$\mathcal{U}_{xf} = \max_A ((1 - \gamma_{a_x}) \cdot \mathbb{I}_{a_x \in I_f}), \tag{2}$$

and

$$\mathcal{UD}_{xf} = 1 - \mathcal{S}_{xf} - \mathcal{U}_{xf}. \tag{3}$$

The product over all attributes in Equation (1) reflects the notion that *all* attributes of a product have to be known and acceptable for the product to be satisfactory. In contrast, a product is unsatisfactory if *at least one* attribute is known and unacceptable, as captured by the maximum operator in Equation (2). Finally, a product is undetermined if and only if it is neither satisfactory nor unsatisfactory yet.

Which attribute levels are acceptable may depend on the decision maker's previous experience and memory. For categorical attributes such as brand and flavor, these judgments are independent, conditional on the respective recalled evaluations of these brands and flavors. Thus, we use the binary logit probability for each attribute level:

$$\Pr(\gamma_{ia} = 1) = \frac{e^{\hat{\gamma}_{a0} + \hat{\gamma}_1 Memory_{ia}}}{1 + e^{\hat{\gamma}_0 + \hat{\gamma}_1 Memory_{ia}}}. \tag{4}$$

The $Memory_{ia}$ variable in this context may simply be the recency of the last purchase or some smoothed function of previous purchases (akin to Guadagni and Little's 1983 loyalty measure). On the other hand, one could also specify some structure on a consumer's process of evaluating—say, a brand's quality as in Erdem and Keane (1996)—or include the effect of forgetting as in Mehta et al. (2004).[10]

However, because all decisions in our experiment were made in one session without any consumption experience in between, all memory-related factors affecting the evaluation of products stay constant across all choices. Moreover, because we do not have the purchase history for each participant, Equation (4) reduces to the same probability of a given attribute level being acceptable for all participants. Yet this still allows for consumer heterogeneity (which may be driven, among other things, by different past experiences) by a hierarchical structure. For brands and flavors, the individual-level acceptability then reduces to a Bernoulli distribution; i.e.,

$$\gamma_{ia} \sim \text{Bern}(\hat{\gamma}_a) \quad \text{for } A \in \{BR, FL\}. \tag{5}$$

For ordered attributes such as price, we instead estimate the highest acceptable price (of the prices used in the experiment) for each person ρ_i, and the indicators for the price levels then follow easily. Equation (4) is then extended to an ordered logit model for ρ_i. However, for the same reasons as mentioned above, we neither have enough information to calculate a $Memory_{i,\,\text{Price}}$ variable for the individuals nor have any variation in this variable across the experiment. Thus, we cannot estimate the parameters of the ordered logit model. Instead, it reduces to a multinomial distribution, at least allowing for heterogeneity across consumers:

$$\rho_i \sim \text{MN}(1, \vec{\rho}) \tag{6}$$

[8] To avoid confusion, we will use "(un)satisfactory" for the product level and "(not) acceptable" for the attribute level for the remainder of the paper.

[9] See Appendix §A.1 for an extension in which consumers may make cross inferences between attributes, thereby relaxing the assumption that a person needs to learn all three attributes directly.

[10] Any potential correlation between judging, say, chicken flavor and onion chicken flavor acceptable can also be captured in the definition of what exactly memory means in a particular context; i.e., $Memory_{i,\,\text{Chicken}}$ could be a function of the past experience of not only chicken flavor but also onion chicken flavor. Finally, Simon (1955) also allowed for the aspiration level to be adjusted *during* choice occasions, if the search also provides information about the distribution of attributes. This could be incorporated in the model by extending (4) to be conditional on a person's information set. However, because the instructions in our experiment included a list of the available brands and flavors, the information set does not change over time, and we therefore do not explicitly model changing aspiration levels in this paper.

and

$$\gamma_{pr} = 1 \quad \text{if and only if } pr \leq \rho_i. \tag{7}$$

Because we allow for continued search after finding the first satisfactory alternative (to accommodate the verification stage), it is possible that a decision maker finds more than one satisfactory alternative before terminating the search. To keep the choice model as close to Simon's original idea of satisficing (i.e., all a consumer cares about is passing a certain threshold rather than some relative ranking), we posit random choice between all satisfactory alternatives. We also introduce the notion of a "trembling hand" to allow for unobserved error. In game theory, a trembling hand allows for nonzero probabilities of actions off the equilibrium path (Osborne and Rubinstein 1994). In our context, it allows for nonzero choice probabilities for undetermined alternatives as well as for no-choice. Although undetermined alternatives should never be chosen, we observe that the chosen option is undetermined in about 4.7% of cases and therefore allow for this error.[11] Similarly, no-choice should never be chosen in the presence of a satisfactory alternative; however, in the absence of a satisfactory option, no-choice is the predicted outcome given a satisficing choice rule. In contrast, we posit a choice probability of exactly 0 for all unsatisfactory options as well as all options that were never fixated on at all.

Let F denote the last fixation in a given choice set by a given person. We then have the following choice probabilities for unsatisfactory, satisfactory, and undetermined alternatives (suppressing i and j):

$$\Pr(x \mid \mathcal{U}_{xF} = 1) = 0, \tag{8}$$

$$\Pr(x \mid \mathcal{S}_{xF} = 1) = \frac{1}{\mathcal{C}}, \tag{9}$$

$$\Pr(x \mid \mathcal{UD}_{xF} = 1) = \frac{\tau}{\mathcal{C}} \quad 0 \leq \tau \leq 1, \tag{10}$$

$$\Pr\Big(\text{no choice} \,\Big|\, \max_x \mathcal{S}_{xF} = 1\Big) = \frac{\tau}{\mathcal{C}}, \tag{11}$$

$$\Pr\Big(\text{no choice} \,\Big|\, \max_x \mathcal{S}_{xF} = 0\Big) = \frac{1}{\mathcal{C}}, \tag{12}$$

where $\mathcal{C}_{ij}$ is the appropriate normalizing constant such that all probabilities sum to 1; τ is restricted to be between 0 and 1 and should in fact be fairly close to 0. If choosing an undetermined alternative (or choosing no-choice in the presence of a satisfactory alternative) is truly an error, its choice probability should be much lower than that of a satisfactory option. (This is similar to having very small but nonzero choice probabilities for "very unlikely choices" in the standard logit or probit models.)

Finally, τ_i follows a Beta distribution; i.e.,

$$\tau_i \sim \text{Beta}(\hat{\tau}_1, \hat{\tau}_2).$$

4.4. Search

Search refers to the information acquisition process. In particular, the model describes the sequence of the locations a consumer is fixating on while gathering the information. Following the literature based on Liechty et al. (2003), we model consumer information search with a modified hidden Markov model. There are two unobserved search states, local and global search, as well as a termination state, defined by the button press of the participant at the end of his search.

4.4.1. Initial Fixation. The first fixation in the search process is typically assumed to be exogenous (and in the global state) to initialize the Markov model (van der Lans et al. 2008a). However, in a choice context in which consumers are familiar with the category, understanding where they start their search is actually crucial for marketers. The first fixation could be influenced by shelf location (e.g., eye level versus not eye level), by the consumer's direction of traffic down the aisle, or by a consumer's memory of the shelf layout. In the latter case, the initial fixation would likely be close to the recalled fixation of the last purchased (or most preferred) product. However, because neither of these factors come into play in our experiment (the "shelf" is not large enough, participants do not move, and shelf location is random across shelfs), we revert back to treating the first fixation as exogenous but note that this is an important area for future research in real-world settings.

4.4.2. Global Search. Recall that global search consists mainly of large "jumps" across the shelf image to explore different areas (Liechty et al. 2003). Because this implies moving to an area that has previously been only in peripheral vision, targeting a specific location for the next fixation is difficult for the participant. Therefore, we assume product-level effects for the probabilities of moving, and the probability of the exact location within that product (i.e., which AOI) is proportional to the size of the respective AOI.

In the global state, eye movements are largely influenced by the saliency and/or luminance of image areas (van der Lans et al. 2008b). As such, participants may use the different color patterns of the packagings to guide their search behavior (Wolfe and Horowitz 2004, Wolfe et al. 1990). Because these are confounded with brand through brand-specific packaging, we use separate brand intercepts to capture

[11] See the discussion in §7.1.2 for an extension allowing for attributes not impacting evaluations, which could explain choices of undetermined options. However, even for participants who choose an undetermined option at least once, they typically do so only once or twice across the 15 choices. Thus, this truly seems to be an error either in the decision-making process or in the eye-tracking data.

this effect. Moreover, the status of the alternative may influence the probability of moving to a respective product. If a product has already been judged to be satisfactory or unsatisfactory, there is no reason from an information acquisition viewpoint to return to that product later in the search. Notice that by incorporating the status into the search probabilities, these probabilities become path-dependent and vary with time.

To formalize these considerations, let $\vec{\mathcal{S}}_{ij,f-1}$ and $\vec{\mathcal{U}}_{ij,f-1}$ be the vectors containing the all product-level indicators $\mathcal{S}_{ijx,f-1}$ and $\mathcal{U}_{ijx,f-1}$. Also, define $r_j(h)$ as the ratio of the size of h in choice set j relative to the size of $P_j(h)$. Suppressing i and j again, the probability of moving to AOI h at fixation $f>1$ in the global state g conditional on the statuses of all alternatives (summarizing the previous search) is then given by

$$\eta_{fg}(h \mid \vec{\mathcal{S}}_{f-1}, \vec{\mathcal{U}}_{f-1}) = \frac{\psi_{fg}(h \mid \cdot)}{\Psi_{fg}}, \tag{13}$$

where

$$\psi_{fg}(h \mid \cdot) = r(h) \cdot \exp(\phi_{g0br(h)} + \phi_{g1}\mathcal{S}_{p(h),f-1} + \phi_{g2}\mathcal{U}_{p(h),f-1}) \quad \forall\, h \neq h_{f-1}, \tag{14}$$

and $\psi_{fg}(h_{f-1}) = 0$ (because of the exclusion of consecutive repeat fixations). The appropriate normalizing constant is $\Psi_{fg} = \sum_h \psi_{fg}(h \mid \cdot)$. Although Equation (13) has the same form as a standard logit probability, we want to stress that we do not assume that a person is moving to the AOI with the maximum value for ψ_{fg} subject to unobserved extreme value error. Instead, eye fixations are believed to be at least partly not under conscious control and therefore to have a truly random component (Shi et al. 2012).

4.4.3. Local Search. In contrast to global search, local search is aimed at gathering specific pieces of information in the same area of the image as the previous fixation (Liechty et al. 2003). Thus, refixating on the same product or moving to an adjacent product should be most likely. Moreover, staying in the same area of the image allows for targeted search for specific attribute information. In the local state, people may also use systematic search strategies such as search by attribute. Finally, similar to the global state, we expect people to be less likely to return to a product that has already been determined to be unsatisfactory.

Let $R_{ijf}(h \mid h_{ij,f-1})$, $N_{rijf}(h \mid h_{ij,f-1})$, and $N_{lijf}(h \mid h_{ij,f-1})$ be dummy variables for whether moving to AOI h would constitute a refixation on the same product, a move to the neighboring product to the right, or a move to the neighboring product to the left, respectively.[12] In case the product last fixated on was on the left (right) edge of the shelf, $N_{lijf}(h)$ ($N_{rijf}(h)$) equals 1 for the AOIs corresponding to the product(s) directly above and/or below that product. Further, let $L_{ijf}(h \mid h_{ij,f-1})$ (or short $L_{ijf}(h)$) be the sum of these three indicators, thereby defining the "local" area around the last fixation as the same plus the neighboring products (Shi et al. 2012).

Finally, to capture targeted search for attribute A (TS_A) within the local area (e.g., looking specifically for prices) and search by attribute (SBA_A) within the local area (e.g., moving from price tag to price tag), define the following indicators

$$TS_A(h) = \mathbb{I}_{A(h)=A} \cdot L_f(h)$$

and

$$SBA_A(h) = \mathbb{I}_{A(h)=A} \cdot \mathbb{I}_{A(h)=A(h_{f-1})} \cdot L_f(h);$$

i.e., $TS_A(h)=1$ if moving to AOI h is a local move to attribute A, and $SBA_A(h)=1$ if moving to AOI h is a local move and both the current AOI and h correspond to attribute A.

Similar to the specification for the global state and once again suppressing i and j, we then have (for $f>1$)

$$\eta_{fl}(h \mid h_{f-1}, \vec{\mathcal{U}}_{f-1}) = \frac{\psi_{fl}(h \mid \cdot)}{\Psi_{fl}}, \tag{15}$$

where

$$\psi_{fl}(h \mid \cdot) = \exp\begin{pmatrix} \phi_{l0}L_f(h) + \phi_{l1}R_f(h) + \phi_{l2}N_{rf}(h) \\ +\,\phi_{l3}\mathcal{S}_{p(h),f-1} + \phi_{l4}\mathcal{U}_{p(h),f-1} \\ +\sum_A \phi_{l5A} TS_A(h) + \sum_A \phi_{l6A} SBA_A(h) \end{pmatrix} \quad \forall\, h \neq h_{f-1}, \tag{16}$$

and $\psi_{fl}(h_{f-1}) = 0$ (because of the exclusion of consecutive repeat fixations). Again, $\Psi_{fl} = \sum_h \psi_{fl}(h \mid \cdot)$ is the appropriate normalizing constant. The first line represents the effects of staying in the local area, the second line captures the impacts of previous evaluations, and the third line allows for targeted search and search by attribute. Notice that whereas staying in the local area should be most likely in the local search state, moving outside of the local area is not impossible because of the partially random nature of fixations.

4.4.4. Transition Probabilities. To complete the hidden Markov model for the search, we need to specify the transition probabilities between the states. Letting $s_{ijf} \in \{g, l, t\}$ be the state of the fth fixation of individual i in choice set j, the transition probabilities take the general form

$$\Pr(s_{ijf} = s^* \mid f, s_{ij,f-1}, h_{ij,f-1}, \vec{\mathcal{S}}_{ij,f-1}, \vec{\mathcal{U}}_{ij,f-1}) = \frac{\pi_{ijf}(s^* \mid \cdot)}{\sum_s \pi_{ijf}(s \mid \cdot)}. \tag{17}$$

[12] For readability, we will omit the conditioning term for the remainder of the paper as it can be easily derived from the first set of subscripts.

For the transition to the global and to the local state, we again allow transitions to depend on the status of the product last fixated on. If the status is already determined to be either satisfactory or unsatisfactory, transition to the global state may be more likely in order to move to a different area of the image. Moreover, it should depend on the previous state, allowing for autocorrelation between states. Again suppressing i and j, we have

$$\pi_f(g \mid \cdot) = \exp(\lambda_{g0} + \lambda_{g1}\mathbb{I}_{s_{f-1}=g} + \lambda_{g2}\mathcal{S}_{p(h_{f-1}),\,f-1} + \lambda_{g3}\mathcal{U}_{p(h_{f-1}),\,f-1}) \tag{18}$$

and

$$\pi_f(l \mid \cdot) = \exp(\lambda_{l0} + \lambda_{l1}\mathbb{I}_{s_{f-1}=g} + \lambda_{l2}\mathcal{S}_{p(h_{f-1}),\,f-1} + \lambda_{l3}\mathcal{U}_{p(h_{f-1}),\,f-1}). \tag{19}$$

Most interesting, though, explicitly modeling the transition to the termination state allows for better insights into what causes consumers to quit searching (Liechty et al. 2003). This is of particular interest in our application, because the satisficing choice rule has a very distinct stopping rule that we can directly model. Though the stopping rule is not deterministic, as it allows for the verification stage, satisficing implies that transitioning to the termination state should be significantly more likely after finding the first satisfactory alternative. In addition, people are more likely to quit searching the longer the search has been. With the understanding that the transition to the termination state occurs after the last recorded fixation, we then let

$$\pi_f(t \mid \cdot) = \exp(\lambda_{t0} + \lambda_{t1}\mathbb{I}_{s_{f-1}=g} + \lambda_{t2}f^* + \lambda_{t3}\max(\vec{\mathcal{S}}_{f-1})), \tag{20}$$

where f_j^* is the running count of fixations normalized by the median number of fixations for the respective choice set, as explained in §3.3. If consumers use a satisficing rule, λ_{t3} (capturing the implied stopping rule of the satisficing model) should be positive. Moreover, its magnitude relative to λ_{t2} determines how important that effect is relative to the effect of the number of fixations (which can be interpreted as a proxy for search cost and/or fatigue).

4.4.5. Heterogeneity. Similar to the evaluation part of the model, all parameters in the search part of the model allow for interindividual heterogeneity through a normal hierarchical structure; i.e., we have

$$\phi_{i\cdot} \sim \mathrm{N}(\bar{\phi}_{\cdot}, \sigma_{\cdot})$$

and

$$\lambda_{i\cdot} \sim \mathrm{N}(\bar{\lambda}_{\cdot}, \varsigma_{\cdot}).$$

5. Estimation and Identification

The model is completed by a set of uninformative priors (see Appendix §A.2 for details) and estimated with a Markov chain Monte Carlo (MCMC) algorithm in a Bayesian framework (Gelfand and Smith 1990, Casella and George 1992). The MCMC allows for efficient integration over the inherent discontinuities of the model caused by the indicator nature of acceptability and satisfaction judgments. See Appendix §A.3 for details of the conditional posterior distributions used in the MCMC. We use 100,000 draws after discarding 30,000 draws as a burn-in period. Convergence of the estimated parameters is assessed using the the Heidelberger and Welch convergence diagnostic as implemented in the Bayesian Output Analysis package for R (Heidelberger and Welch 1983, Smith 2007).

Several restrictions have to be placed on this model for identification. Analogous to standard choice models, the brand intercepts in Equation (14) for the global search are not separately identified. We therefore normalize $\phi_{g0i,\,\text{Fantastic}}$ to 0. Equation (16) for the local search is also not uniquely identified. In particular, as the brand, flavor, and price AOIs collectively make up the complete product, ϕ_{l0i} is not separately identified from the set $\{\phi_{l5Bi}, \phi_{l5Fi}, \phi_{l5Pi}\}$ and from the set $\{\phi_{l6Bi}, \phi_{l6Fi}, \phi_{l6Pi}\}$. We therefore normalize ϕ_{l5B} and ϕ_{l6B} to 0.

Once again analogous to intercepts in standard models, only two out of the three λ_{s0} and the three λ_{s1} parameters for the transition probabilities are identified. (Note that $\lambda_{s0} + \lambda_{s1}$ is the intercept conditional on the last fixation being in the global state.) We thus normalize λ_{g0} and λ_{g1} to 0. The identification of λ_{g2} and λ_{l2} as well as of λ_{g3} and λ_{l3} relies on fitting the relative probability of termination. However, as reported below, we find that the probability of termination is so low (at least before finding a satisfactory option) that only the difference of these parameters is well informed, as the termination probability changes only minimally for a wide range of absolute values of these parameters. We therefore set $\pi_{ijf}(g \mid \cdot) = 1$ for all i, j, and f. However, this still allows us to investigate the relative impact of an alternative's status on transitioning to the global versus the local state as well as, more importantly, the use of the stopping rule implied by the satisficing choice.

In addition to the "full model" described above, we also estimate an "independent model" in which all search model parameters corresponding to the status of the alternatives (ϕ_{g1i}, ϕ_{g2i}, ϕ_{l3i}, ϕ_{l4i}, λ_{l2i}, λ_{l3i}, and λ_{t3i}) are set to 0. This allows us to analyze the influence of modeling the two parts jointly on the obtained results.

6. Results

Analogous to the presentation of the model, we separate the presentation of the results into search and evaluation. We focus on the results of the full model and discuss the results of the independent model wherever they allow extra insight into the model.[13] Finally, we present a holdout prediction analysis and compare the result against standard logit models.

Although we will analyze all parameter results in detail, it is worth highlighting the results or the parameters corresponding to the alternatives' statuses. Because estimating these indicators and having search depend on them is one of the main features of the proposed model, it is important to check whether they do in fact affect search in reasonable ways. As discussed below, we find that all but one of the parameters corresponding to the alternatives' statuses—including the one for the stopping rule—are nonzero (i.e., the 95% highest density interval of the posterior does not include 0), and of those, all but two have the expected sign. The exceptions are easily explained by a closer look at the data. Thus, the results suggest that these (un)satisfactory judgments are in fact meaningful and have a real impact on continued search.

6.1. Search

Table 3 presents the posterior means and standard deviations for the population-level hierarchies for the search model parameters.

6.1.1. Global Search. The brand intercepts in the global search capture effects of packaging on consumer search. Because packaging is confounded with the brands in our experiment, packaging might help consumers find their preferred brands quickly (van der Lans et al. 2008a). If this was the case, the brand intercepts should be correlated with brand preferences. Given the large standard deviations of the population-level hierarchies (e.g., $\sigma_{g0,\,\text{Indomie}} = 0.92$), we need to perform this analysis at the individual level rather than at the aggregate level.[14] We therefore correlate the individual-level estimates for the brand intercepts with the explicit measures of brand preference collected during the experiment. Across all participants, the correlation is 0.27, suggesting that consumers may use colors/packaging to guide their search process, but that this process is far from perfect. We also calculate the corresponding correlation for each individual (i.e., based on only four data points; the correlation could not be computed for five participants because of a lack of variation in the explicit data). We find a wide variety of values (mean = 0.41). Further analysis reveals that these correlations (i.e., the extent to which packaging is used to guide search in the global state) are positively related to explicit preference of Maggi ($r = 0.27$) but negatively related to explicit preferences for

Table 3 Posterior Means (and Standard Deviations) for the Search Parameters

Parameter	Variable	M (SD)	Variable	M (SD)
		Global search		
Brand intercepts	$\bar{\phi}_{g0,\,\text{Indomie}}$	−0.35 (0.14)**	$\sigma_{g0,\,\text{Indomie}}$	0.92 (0.24)
	$\bar{\phi}_{g0,\,\text{Koka}}$	0.11 (0.07)*	$\sigma_{g0,\,\text{Koka}}$	0.17 (0.05)
	$\bar{\phi}_{g0,\,\text{Maggi}}$	0.48 (0.09)**	$\sigma_{g0,\,\text{Maggi}}$	0.41 (0.10)
Satisfactory	$\bar{\phi}_{g1}$	0.84 (0.11)**	σ_{g1}	0.44 (0.14)
Unsatisfactory	$\bar{\phi}_{g2}$	−0.49 (0.08)**	σ_{g2}	0.21 (0.07)
		Local search		
Local area	$\bar{\phi}_{l0}$	3.12 (0.08)**	σ_{l0}	0.36 (0.08)
Same product	$\bar{\phi}_{l1}$	1.73 (0.04)**	σ_{l1}	0.09 (0.02)
Right	$\bar{\phi}_{l2}$	−0.04 (0.04)	σ_{l2}	0.06 (0.01)
Satisfactory	$\bar{\phi}_{l3}$	0.32 (0.05)**	σ_{l3}	0.12 (0.03)
Unsatisfactory	$\bar{\phi}_{l4}$	−0.75 (0.05)**	σ_{l4}	0.10 (0.02)
Targeted search	$\bar{\phi}_{l5,\,\text{Flavor}}$	−0.58 (0.04)**	$\sigma_{l5,\,\text{Flavor}}$	0.09 (0.02)
	$\bar{\phi}_{l5,\,\text{Price}}$	−1.33 (0.08)**	$\sigma_{l5,\,\text{Price}}$	0.34 (0.07)
Search by attribute	$\bar{\phi}_{l6,\,\text{Flavor}}$	0.17 (0.05)**	$\sigma_{l6,\,\text{Flavor}}$	0.09 (0.03)
	$\bar{\phi}_{l6,\,\text{Price}}$	1.99 (0.08)**	$\sigma_{l6,\,\text{Price}}$	0.28 (0.07)
		Transition probabilities		
Intercept	λ_{l0}	1.63 (0.11)**	ς_{l0}	0.70 (0.15)
Last state: Global	λ_{l1}	0.27 (0.12)**	ς_{l1}	0.50 (0.15)
Satisfactory	λ_{l2}	−0.04 (0.06)	ς_{l2}	0.10 (0.03)
Unsatisfactory	λ_{l3}	−0.48 (0.05)**	ς_{l3}	0.08 (0.02)
Intercept	λ_{t0}	−5.03 (0.21)**	ς_{t0}	1.31 (0.36)
Last state: Global	λ_{t1}	0.10 (0.14)	ς_{t1}	0.29 (0.15)
Fixations	λ_{t2}	1.54 (0.14)**	ς_{t2}	0.72 (0.23)
Satisficing stopping rule	λ_{t3}	2.40 (0.18)**	ς_{t3}	0.45 (0.18)

*The 90% highest density interval does not include 0. **The 95% highest density interval does not include 0.

Fantastic ($r = -0.25$), Koka ($r = -0.48$), and Indomie ($r = -0.56$). The yellow color of the Maggi packaging is fairly distinctive on the shelf, allowing people to use packaging to find their favorite brand. However, the light blue color of the Indomie packaging truly stands out from the other products on the shelf. The strong negative correlation of preference for Indomie and the Indomie intercept in the global search then suggests that it might indeed be so eye-catching that people tend to unintentionally look at it, despite not liking the brand that much. This speaks to the at least partially subconscious nature of eye fixations. In summary, this suggests that using colors to guide search in a top-down process, as demonstrated by van der Lans et al. (2008a), may work well if packages are distinct enough to be distinguishable in peripheral vision, yet it can be overcome by marketers if packages are eye-catching enough to guide eye fixations inadvertently. However, further research is needed to validate this hypothesis.

As expected, a consumer is less likely to return to an alternative already judged to be unsatisfactory ($\bar{\phi}_{g2} = -0.49$). In contrast, a consumer is more likely to return to an alternative already judged to be satisfactory ($\bar{\phi}_{g1} = 0.84$). (However, on the individual level, the highest density interval of the posterior

[13] None of the results of the independent model differs qualitatively from the full model unless mentioned.

[14] We thank an anonymous reviewer for this suggestion.

for ϕ_{g1i} includes 0 for more than 50% of the participants. Because there tend to be very few fixations after finding the first satisfactory alternative, there simply is not much information on this parameter in the data.) We did not expect this result, because there should be no reason to return to an already determined alternative from an information acquisition viewpoint. One possible explanation might be that consumers make explicit comparisons between different satisfactory alternatives, going back and forth between them. However, a look at the search paths suggests that this result may mainly be driven by an end-of-search effect, as almost everyone finishes her search by returning to the product she chose, which of course is satisfactory.

6.1.2. Local Search. Similarly to the global search, fixating on an alternative that is already determined to be unsatisfactory is less likely in local search ($\bar{\phi}_{l4} = -0.75$), as consumers seem to be more likely to return to an alternative already judged to be satisfactory ($\bar{\phi}_{l3} = 0.32$). Again, this is due to the end-of-sequence effect discussed previously. Moreover, the highest density interval of the posterior for ϕ_{l3i} includes 0 for *all* participants.

Confirming the characterization of the local search, the probability of staying in the "local" area is about 90% based on the results for $\bar{\phi}_{l0}$ to $\bar{\phi}_{l2}$. Refixating on the same product is most likely ($\bar{\phi}_{l1} = 1.73$), and moving to the left is somewhat more likely than moving to the right ($\bar{\phi}_{l2} = -0.04$). The remaining parameters have to be interpreted relative to the normalized brand parameters and with the previous ones in mind. The negative signs of $\bar{\phi}_{l5,\,\text{Flavor}}$ and $\bar{\phi}_{l5,\,\text{Price}}$ are probably due to the relative differences in size of the AOI relative to the remaining package. However, when combined with $\bar{\phi}_{l1}$, the sum is still positive, suggesting that targeted information search within the same product is still more likely than moving to the next product. Finally, we find support for strategic search by attribute for search by price ($\bar{\phi}_{l6,\,\text{Price}} = 1.99$) but not by flavor ($\bar{\phi}_{l6,\,\text{Flavor}} = 0.17$). The much weaker effect for flavor is probably caused by the considerable clutter on the shelves, which makes it difficult to move from one flavor AOI to another flavor AOI, in particular because the location of the flavor information on the package differs by brand. For price, this is a lot easier because all prices are in one line at the bottom of each shelf. However, there is considerable heterogeneity across participants in how strong this systematic search effect is ($\sigma_{l4,\,\text{Price}} = 0.28$), suggesting that some people may be more likely to use search by attributes than others.

6.1.3. Transition Probabilities. As the final part for the search model, let us move to the transition probabilities. In general, transitioning to the local state is a lot more likely than transitioning to the global state ($\bar{\lambda}_{l0} = 1.63$). This is reflected in the finding that about 80% of all fixations are estimated to occur in the local state. Moreover, on the population level, people are more likely to switch back to the local state after a fixation in the global state ($\bar{\lambda}_{l1} = 0.27$); i.e., a move to a different part of the shelf is typically followed by local examination in that area. However, the high level of heterogeneity across participants ($s_{l1} = 0.50$) suggests that there may be interindividual differences; looking at individual-level estimates, we find that 4.7% of participants are more likely to stay in the global state if the last fixation was also in the global state (relative to moving to the global state if the last fixation was in the local state), whereas 14.1% are more likely to move to the local state. (For the remaining participants, the highest density interval of the posterior distribution includes 0.)

Participants are equally likely to quit their search after the local state as after the global state ($\bar{\lambda}_{t1} = 0.10$), with the initial stopping probability not surprisingly being extremely low ($\bar{\lambda}_{t0} = -5.03$). More interestingly, however, is a look at how the stopping probability changes over time. As one might expect, people become more likely to stop their search the longer they have already searched ($\bar{\lambda}_{t2} = 1.54$); this effect holds true for 87.5% of the participants, where the others do not seem to be affected by the length of their search. In contrast, we find support for the increase in stopping probability implied by the satisficing choice rule for *all* participants; i.e., having found the first satisfactory alternative significantly increases the stopping probability ($\bar{\lambda}_{t3} = 2.40$). To understand the relative magnitudes of these two effects, recall that the number of fixations is relative to the median number of fixations for a given choice set. Thus, the impact of finding the first satisfactory alternative on the stopping probability is 1.56 times larger than the effect of having searched the median search length (or alternatively, more than 50 times larger than the impact of one additional fixation).

In the independent model, the satisficing stopping rule is not part of the model because λ_{t3i} is set to 0. We find that in this case the initial probability to stop searching immediately is significantly higher ($\bar{\lambda}_{t0} = -3.78$); i.e., the intercept picks up some of the effect that is now missing from the model. More interestingly, though, we also find that the effect of the length of the search increases significantly (i.e., no overlap of the 95% highest density intervals) to $\bar{\lambda}_{t2} = 2.47$, also picking up some of the missing effect. The fact that the probability of having found at least one satisfactory alternative increases with the length of the search is trivially true, as by definition there is no satisfactory alternative at the beginning of the search and almost

Table 4 Log-Marginal Likelihood

	Bare	Brand	TAS	Indep. model	NoStop	Full model
Empirical GD ($R = 100{,}000$)	−128,841	−128,254	−125,240	−123,748	−121,529	−121,366
Asymptotic GD	−135,793	−135,123	−132,239	−130,585	−129,960	−129,328

always at least one at the end of the search. Taken together with the results from the full and independent models, this essentially constitutes a Bayesian test for mediation (Zhang et al. 2009). The results can then be interpreted as evidence that the effect of the length of search is partially mediated by the indicator for having found at least one satisfactory alternative.

Taken together, these findings lend strong support to the hypothesis that consumers do, in fact, use the stopping rule implied by the satisficing choice rule.

6.1.4. Analysis of Relative Contribution of Model Components. Because this is the first model of consumer search using eye-tracking data in a choice context, we examine the explanatory contributions of the components of the search model. To do so, we follow Guadagni and Little (1983) and build up the search part of the model element by element. Thus, in addition to the full and independent models described above, we also estimate the following four models: Model 1 (referred to as "Bare" in Table 4) is a bare-bone benchmark of the search model in which the location fixation in the global state is random (i.e., relative to AOI size), and in the local state only the indicators for the local area are used as explanatory variables. Transition probabilities are also independent of the evaluations. Thus, only the following individual-level search parameters are estimated (all other parameters are set to 0): $\phi_{l,0-2}$, $\lambda_{l,0-1}$, and $\lambda_{t,0-2}$. For Model 2, the brand intercepts for the global search ($\phi_{g0,\,\text{Indomie/Koka/Maggi}}$) are added, capturing the additional explanatory power of the packaging effects in the global search ("Brand"). Model 3 adds the targeted attribute search (TAS) in the local state ($\phi_{l5,\,\text{Flavor/Price}}$). Adding the search-by-attribute parameters in the local state ($\phi_{l6,\,\text{Flavor/Price}}$) results in the independent model. Finally, Model 4 ("NoStop") is identical to the full model except for omission of the satisficing stopping rule (i.e., $\lambda_{t4} = 0$).

Table 4 reports the log of the Gelfand–Dey (GD) estimator of the marginal likelihood of the search data for each of the models (Gelfand and Dey 1994).[15] This estimator does not suffer from the same simulation pseudo-bias favoring more complex models as the usual harmonic mean estimator (Duan et al. 2011, Lenk 2009). The GD estimator is given by

$$\hat{p}(y \mid M_k) = \left\{ \frac{1}{R} \sum_{r=1}^{R} \frac{q(\Theta_k^r)}{L(y \mid \Theta_k^r, M_k)\, p(\Theta_k^r \mid M_k)} \right\}^{-1},$$

where Θ_k^r represents parameter estimates from the rth iteration of the MCMC for model k, $q(\Theta_k^r)$ and $p(\Theta_k^r)$ are a proposal density and the prior density of Θ_k^r, and $L(y \mid \Theta_k^r, M_k)$ is the likelihood of the data given model k and parameter draw Θ_k^r. In essence, the GD estimator is the harmonic mean of the likelihoods weighted by an importance sampling weight.

In our application, the vector Θ_k consists of all population-level parameters of the search model contained in model k. Following the recommendation of Lenk (2009), we use a multivariate normal distribution with mean and covariance calculated from the MCMC draws as proposal density (we use the log of the draws of the variance parameters for the multivariate normal distribution).

The likelihood of the data $L(y \mid \Theta_k^r, M_k)$ is calculated as the product of the individual-level search path likelihoods (as defined in Equations (13)–(20)) and the likelihoods of the individual-level parameters (including the unobserved states), conditional on the population-level draw:

$$L(y \mid \Theta_k^r, M_k) = \prod_i L(y_i \mid \theta_i^r)\, p(\theta_i^r \mid \Theta_k^r). \tag{21}$$

The first row of Table 4 presents the log marginal likelihoods for the different models using 100,000 draws for the GD estimator. However, because of the vast amount of data (more than 40,000 eye fixations) and the complexity of the model, the empirical estimates are affected by infrequent "outliers" (Rossi et al. 2005, p. 167).[16] We therefore also report an asymptotic estimate of the GD estimator derived from the empirically observed value. The summands in Equation (21) empirically turn out to be log-normally distributed for all of the submodels (where the log is tested to be normal with the Kolmogorov–Smirnov test). We can then use the closed-form solution for the mean of a log-normal distribution with the empirically observed parameters for the different models as the asymptotic value for the GD estimator.

[15] In this section, we are concerned only with the explanatory power of the components of the search model in explaining the search data. In §6.2.1, we analyze the effect of allowing for the interdependence of search and evaluation on the estimation of participants' preferences.

[16] Differences in log-likelihoods are in the hundreds, compared to even the log of one million being less than 15 (which would be the adjustment if we took one million draws).

The resulting values for the log-GDs are far below the empirically observed values, as taking the full tails into account includes even more drastic outliers. However, the order of GDs does not change; i.e., the conclusions about model choice remain the same.

For both the empirical and the asymptotic GD estimate, the largest improvement is observed when introducing more explanatory variables to the local search state (i.e., for the TAS and independent models). This is to be expected, given the above-reported finding of about 80% of the fixations being in the local search. More interestingly, there still is a sizable improvement when adding the interdependence between search and evaluation into the model, over and above the already detailed explanatory variables included in the independent model. (The difference is smaller for the asymptotic GD estimate than for the empirically observed one, as the variance of the log-normal distribution of the summands is larger for the NoStop model than for the independent model.) Finally, adding the satisficing stopping rule further increases the marginal likelihood. The increase of the marginal likelihood is especially impressive given the partial mediation of the effect of the number of fixation by the satisficing stopping rule. Once again, this suggests that the satisficing stopping rule is a very meaningful addition to the model, suggesting that participants incorporate this stopping rule in their search behavior.

6.2. Evaluation

Table 5 presents the posterior means and standard deviations for the population-level hierarchies for the evaluation model parameters.

By and large, these estimates conform roughly to the choice shares presented in Tables 1 and 2. *Every* participant chose Maggi at least once, which is reflected in the high probability of acceptability ($\hat{\gamma}_{\text{Maggi}} = 0.98$). At first glance, it might be surprising that the population-level probability of the highest price being available is almost 50%. However, a closer look at the data reveals that, in fact, 48% of the participants chose a product priced at QR7.00 at least once (implying that that price is acceptable to them), so the estimate is perfectly on target. Keeping in mind that even this highest price is only \$1.90 for a five-pack of noodles, this is not all that surprising. In contrast, only one person never chose a product that cost more than QR4.00. Finally, the posterior means of $\hat{\tau}_1$ and $\hat{\tau}_2$ are 1.29 and 38.6, respectively, combining to a Beta distribution with mean 0.032 and standard deviation 0.028 for the trembling hand parameters τ_i.

To further test the face validity of our results, we correlate the individual-level results with the explicit measures of brand and flavor preference collected in the questionnaire. Across all participants, the correlation is 0.47 for brands and 0.57 for flavors. These correlations are strong considering the numerous ties in the explicit measures as a result of using a five-point Likert scale and, more importantly, the numerous ties in the model estimates as a result of its deterministic nature (if a person chose several different flavors, they all have a probability of being acceptable of 1). For the individual-level correlations based on only 4 and 10 values, respectively (to avoid scale issues across participants), the mean correlation is 0.60 for brands and 0.61 for flavors.[17]

Table 5 Posterior Means (and Standard Deviations) for the Choice Parameters

Parameter	M (SD)
Brand	
$\hat{\gamma}_{\text{Fantastic}}$	0.80 (0.05)
$\hat{\gamma}_{\text{Indomie}}$	0.70 (0.06)
$\hat{\gamma}_{\text{Koka}}$	0.58 (0.06)
$\hat{\gamma}_{\text{Maggi}}$	0.98 (0.01)
Flavor	
$\hat{\gamma}_{\text{Beef}}$	0.47 (0.06)
$\hat{\gamma}_{\text{Cheese}}$	0.62 (0.06)
$\hat{\gamma}_{\text{Chicken}}$	0.89 (0.04)
$\hat{\gamma}_{\text{Curry}}$	0.66 (0.06)
$\hat{\gamma}_{\text{Lobster}}$	0.35 (0.06)
$\hat{\gamma}_{\text{Mushroom}}$	0.45 (0.06)
$\hat{\gamma}_{\text{OnionChicken}}$	0.84 (0.05)
$\hat{\gamma}_{\text{Shrimp}}$	0.45 (0.06)
$\hat{\gamma}_{\text{Tomato}}$	0.45 (0.06)
$\hat{\gamma}_{\text{Vegetable}}$	0.46 (0.07)
Price	
$\vec{p}_{4.00}$	0.03 (0.01)
$\vec{p}_{4.75}$	0.04 (0.02)
$\vec{p}_{5.50}$	0.22 (0.05)
$\vec{p}_{6.25}$	0.22 (0.05)
$\vec{p}_{7.25}$	0.49 (0.06)

Finally, to gain additional insight into whether consumers may or may not be using the proposed satisficing choice rule, we analyze the number of satisfactory options (defined as the modal number across MCMC draws) a person has found before stopping his search. Across all participants and choice sets, the median set size of satisfactory options is 1 (mean = 1.74), with almost 80% of sets consisting of at most two satisfactory options. (For comparison, the median sizes of the set of unsatisfactory options and the set of undetermined options are 6 and 4, respectively.) On an individual level, more than 70% of participants have on average less than two satisfactory options across choice sets before terminating their search. Once again this suggests that having found one satisfactory alternative is sufficient for many people to stop their search very soon after, lending further

[17] Individual-level correlations could not be calculated for 27 individuals for brands and for 1 individual for flavors because of no variation in the explicit and/or estimated preference measures.

support to the hypothesis that they follow a satisficing choice rule. On the other hand, 8% of the participants have on average more than three satisfactory options before making their final choice, suggesting that a satisficing choice model may not be appropriate for them.

6.2.1. Impact of the Evaluation-Search Interdependence on the Estimation of Preferences. As discussed in §6.1.4, allowing for the interdependence of evaluations and search helps tremendously in explaining search paths. In this section, we show that allowing for this interdependence also improves the estimation of consumers' preferences. In addition to providing more information on the continued search, including the evaluations in the search model also allows the model to derive information from the continued search about whether a given alternative had already been judged to be satisfactory or unsatisfactory. For instance, if every time a person encountered, say, chicken flavor, she moved on to another product, that is quite strong evidence that chicken flavor may have been enough for her to not like the alternative. Thus, chicken flavor would not be acceptable. If, however, the person continued looking at the product to find out price for instance, this suggests that chicken flavor may have been acceptable to her (even if she did not choose chicken flavor in any of the choice sets).

To check this, we identify three vegetarians in our data (defined by never choosing a nonvegetarian flavor and giving the lowest possible explicit rating to all nonvegetarian flavors). The full model, which uses the information on the continued search in the estimation of the acceptability of attributes, assigns very low probabilities for them to find any meat flavor acceptable.[18] The independent model, in contrast, does not use the sequence information for the estimation of preferences but instead relies solely on final choices (out of the encountered alternatives). In turn, the independent model performs very poorly at identifying the vegetarians. Although the nonvegetarian flavor acceptability probabilities for the vegetarians are consistently below the respective hierarchy levels, they are far from identifying vegetarians as such. On average, nonvegetarian flavors are estimated to be acceptable for vegetarians with a probability of 56%, with several estimates even being over 90%.

The same principle holds true when applied to flavors/brands not chosen by a given participant (vegetarian or not). We consistently find that the added search information in the full model allows (but does not force) the estimates for those flavors/brand to be drawn further from the population hierarchy than in the independent model. Thus, incorporating the link between search and evaluations allows us to get better insight into individual-level preferences.

6.3. Holdout Prediction

We conduct two different holdout analyses. In the first, we reestimate the model using only 12 of the 15 choice sets and use the remaining 3 choice sets for prediction, where the goal is to evaluate holdout fit using individual-level estimates. (We hold out choice sets 3, 8, and 13.) In the second, we holdout both participants and choice sets, estimating the model on 12 choice sets and 44 participants, to evaluate the predictive ability of the satisficing model relative to a standard multinomial logit model.

6.3.1. Holdout Fit. Because the model provides probabilities of acceptability for each level of each attribute, we can check how well the model fits the holdout choices on an attribute level. The holdout choices conform extremely well with the model results. Recall that we estimate the individual-level posterior probability that an attribute is acceptable. We define the individual-level acceptable sets for a given attribute as those that are acceptable with a probability of at least 95%. The median acceptable set includes five flavors, three brands, and the lowest four prices. We find that 86% of the flavors, 96% of the brands, and 98% of the prices chosen in the holdout choices are within the respective acceptable sets.[19] Although this is consistently higher than the relative size of the acceptable sets, the model strongly benefits from its deterministic nature. If a flavor chosen in the holdout choices was chosen by the same individual in one of the estimation choices, the probability of it being acceptable is necessarily 1. The somewhat lower hit rate for flavors is then mainly caused by the greater number of flavors to choose from and the resulting higher probability that a flavor chosen in the holdout choices may not have been chosen in the estimation choices.

Next, we move to the product level. Using the results for which attribute levels are acceptable as well as the data on which pieces of information participants looked at for the holdout choices, we can calculate the product-level probabilities of each product for being satisfactory, unsatisfactory, or undetermined for each participant. To integrate over the

[18] The single exception out of the 15 estimates is for a participant who did not see any chicken flavored option in 6 of the 15 choice sets. Although that person's probability for finding chicken flavor acceptable is one of the lowest, it is still higher than one would like because it is driven mainly by the population mean, $\hat{\gamma}_{\text{Chicken}} = 0.89$, as a result of the scarcity of individual-specific information for this person.

[19] We exclude the no-choice instances that occur in the holdout choices for the analyses in this as well as in the next paragraph because the analyses are not applicable to them.

joint distribution of the acceptance probabilities, we determine the status of each product for each draw in the MCMC and then average across the draws (rather than averaging the acceptance probabilities across draws and then determining the probabilities, which would lose information on the correlations between acceptable attribute levels). Examining the products chosen in the holdout choices, we find that 74% of the choices have a probability above 95% of being satisfactory for the respective participant (compared with 10% of all possible products). The chosen product has the highest probability of being satisfactory in 85% of all cases; however, in more than half (54%) of those cases, it is tied for first place with at least one more product. Once again, this is due to the fairly deterministic nature of the model.

Finally, we compute choice probabilities for the holdout choices to examine the hit rate, defined as the probability that the chosen option has the highest predicted choice probability. Choice probabilities depend on the number of satisfactory and undetermined options at the time of decision as well as the trembling hand parameter. As above, we calculate choice probabilities (according to Equations (8)–(12)) for each draw in the MCMC and then average across draws.[20] The resulting hit rate is a a very impressive 72.4%. However, the extremely high hit rate does not take into account that in many of these correct predictions, the chosen product is tied with one or more other products for the highest choice probability (as would be expected given the ties in the probabilities of being satisfactory reported in the previous section). So although it is a very encouraging result that the model picks the chosen option to be among the top choices in more than 70% of the cases, hit rates for models with ties may not be as informative as they are for models without ties.

6.3.2. Predictive Ability. In the previous section, we used all participants in the estimation sample as well as the holdout sample to evaluate the individual-level fit. In contrast, we follow the recommendation of Elrod (2002) for model validation and use only 44 participants and 12 choice sets as the estimation sample and the remaining 20 participants and three choice sets as holdout sample. The rationale is that for predictive purposes, a model needs to generalize not only to different choice sets but also to different members of the same population. Moreover, this allows us to circumvent the issue of the deterministic nature of the model that is the main driver for the good performance of the model in the previous section. Using a different set of participants for estimation and prediction eliminates the deterministic connection between estimation choices and holdout choices. Thus, this offers a stronger test for the predictive ability of the proposed model.

To evaluate the predictive performance of the proposed model, we compare it to a hierarchical multinomial logit (MNL) model. We construct individual information sets for each participant such that the estimation of the MNL model only uses the information available to the participant at the time of the decision. As for the satisficing model, we posit zero-choice probability for all alternatives that were never fixated on. For undetermined options, the model assumes a risk-neutral decision based on expected utility (based on rational expectations of the underlying attribute distributions). Finally, the model allows for heterogeneity across participants in the preference parameters.

For both models, we compare the observed choices from the holdout participants to the predicted choice probabilities from the respective model. To calculate the predicted choice probabilities, we need to integrate over the population-level preference estimates from the estimation sample. (We use the posterior means of the estimates for this analysis.) As this is a $4 + 10 + 1 = 15$ (brands + flavors + price)-dimensional integral, we numerically integrate by simulating 500,000 realizations (hypothetical consumers) from the population hierarchy. For each realization, we calculate the likelihood of the three holdout choices for each of the 20 holdout participants. We then average across realizations and finally multiply out the likelihoods across the holdout participants. (The order is important to correctly account for the statistical dependence of the three holdout choices per participant.) Because we have no information on the information sets of the hypothetical consumers, the choice probabilities are calculated with all products in the respective information set. This analysis is purely concerned with predicting choices; i.e., we do not use the observed search paths of the holdout participants, nor do we simulate search paths for the hypothetical consumers.

Following Elrod (2002), we use the log-likelihood (LL) of the holdout choices as measure of predictive ability.[21] The LL for the MNL model is -153.9, whereas the LL for the proposed satisficing model is -137.6. Thus, the satisficing model generalizes better in terms of predictive ability to other choice sets and other consumers. Whereas we use no information on search in the prediction task, we do use the

[20] Note that we only use the data on which information was acquired, not the sequence in which it was acquired; i.e., we do not use the search part of the model in the predictions.

[21] Given the Bayesian framework, one may want to use the Bayes factor instead (i.e., using the integral of likelihood times prior). We choose to focus on the likelihood because it is primarily the likelihood that differentiates the models, seeing that we use uninformative priors.

search in the estimation of the satisficing model (we use individual information sets for the estimation of the MNL model also; i.e., the added information is information on the search *sequence*). Thus, one might think that this additional information is the cause for the better predictive ability. We conduct the same holdout prediction task using the independent model to test for the effect of adding search information to the estimation sample. The LL for the independent model is −141.5. Thus, using search path information in the estimation improves holdout prediction. Yet even when using the same information in the estimation of the parameters, the proposed model predicts holdout choices better than a standard MNL model.

7. Discussion

The proposed model continues the line of research started by Gilbride and Allenby (2004) and Jedidi and Kohli (2005). This line of research truly brings a paradigm shift to the empirical choice model literature in marketing, a shift away from compensatory utility maximizing and toward a quest for more realistic models of consumer choice. Most models in this new line of research employ a two-stage approach in which the simple heuristic is used to form a consideration set in the first stage, followed by a compensatory utility-maximizing choice in the second stage. In contrast, the proposed model does not rely on compensatory trade-offs at all. This is possible thanks to a search stopping rule based on Simon's idea of a satisficing decision maker (Simon 1955). In a satisficing choice rule, the sequence in which products are evaluated is essential. We therefore collect choice and eye-tracking data in a visual conjoint experiment and jointly model search and evaluation.

The results lend significant support to the proposed model. Most importantly, the stopping rule implied by the satisficing rule is strongly supported by the parameter estimates. In addition, the distinction between satisfactory and unsatisfactory products is meaningful in explaining the search pattern. We also show that the joint model of search and evaluation informs the parameters of the evaluation model much better than the independent model. The model performs extremely well in a holdout prediction task. It has very good holdout fit on the individual level with a hit rate of almost 80%, and it clearly outpredicts a MNL model in a holdout prediction task.

It has long been accepted that consumers do not really calculate the compensatory utilities implied by the standard models. Our results show that it is possible to estimate choice models that conform more closely to the actual decision-making process—and that it may be worthwhile to do so. We therefore fully agree with Netzer et al. (2008) that is is time to improve what they call the "ecological fit" of the choice models to the respective task.

7.1. Extensions

Being the first model (to our knowledge) of its kind, the proposed model, of course, can and should be extended to match real choice situations even better. In the following, we will briefly discuss three of the many possible extensions.

7.1.1. Probabilistic Conjunctive Rule. Jedidi and Kohli (2005) introduced a probabilistic version of the conjunctive rule, which could be used to replace the deterministic version used in our model. In terms of the behavioral interpretation of these rules, the difference lies in when a consumer decides what is acceptable to him: In the deterministic version, the consumer decides before the decision process starts what is acceptable to him (leading to thresholds that are constant across choice sets). In contrast, in the probabilistic version, the consumer has a probability of whether he finds a certain attribute level acceptable whenever he sees it (i.e., independent decisions are made on the spot whenever he encounters the attribute level).

To incorporate the probabilistic conjunctive rule into the model, Equations (1) and (2) will be replaced by equations calculating the probability of an alternative to be satisfactory by multiplying the attribute-level probabilities (conditional on the attributes having been fixated on). Notice that, although behaviorally still a noncompensatory rule (if one attribute is judged to be unacceptable, the product will be unsatisfactory), this formulation is mathematically compensatory (a decrease in the probability of a certain attribute to be acceptable can be made up for by an increase in another attribute's probability). Although this may or may not be deemed desirable, the probabilistic version would give the model an added degree of freedom to determine which products were acceptable to a decision maker in a given choice set (bringing the average number of satisfactory products at the end of search even closer to the number of one (which would be predicted by the theory) than as reported above for the deterministic version).

7.1.2. Important vs. Unimportant Attributes. Another potential extension addresses the question whether all attributes are used in the evaluation of a product or maybe only a subset of them. For instance, seeing how even the highest price is acceptable to almost half the participants in our experiment, one may wonder whether such participants actually use price in the decision process. In theory, this could explain choices for what we termed *undetermined* alternatives. Based on the search paths, the model should be able to distinguish between people who incorporate price but all levels are acceptable and people who do not incorporate price in their decisions. The former will look at prices before making a decision, whereas the latter typically will not. We chose not to include this in our model because, as mentioned in footnote 11, the choice of undetermined alternatives truly seems to be an error in our data for almost everybody. (However, for three participants who chose undetermined products in 4 out of the 15 choices, this may well explain their behavior.)

To include this distinction into the model, a latent indicator for whether an attribute was important for the decision would have to be estimated for each attribute (or for the attribute(s) for which one expects that it may not be important to everybody). Equations (1) and (2) would then be modified such that the evaluations depend only on the important attributes.

7.1.3. Applicability. Of course, we do not intend to imply that all consumers always follow a satisficing decision rule. Heterogeneity across people in their tendency to use simple choice heuristics (often imprecisely called "satisficing") versus maximizing decision rules have been well documented (e.g., Schwartz et al. 2002). Moreover, the same person is likely to employ different choice rules when buying instant noodles versus a car, for instance. And even for the same task, choice rules have been found to vary depending on time pressure, fatigue, etc. (e.g., Swait and Adamowicz 2001). Thus, the question is when and by whom a satisficing rule is more likely to be used.

In general, satisficing is expected to be more prevalent in low-involvement categories like most grocery items. Based on the above-mentioned research, we also expect people to be more likely to use a simple choice rule such as satisficing when under time pressure or when fatigued, both of which may be affected by the category complexity. Moreover, consumers may use sampling behavior as predicted by a forward-looking maximizing choice rule when facing a new category but then switch to a satisficing rule after preferences have been formed. Thus, the probability of satisficing choice might increase with familiarity with the category. Finally, the type of choice rule employed may depend on the search environment. For instance, the presentation of information in online search bots by product versus by attribute influences consumers' evaluation strategies (Shi et al. 2012). Presentation by product facilitates a product-based search process and should thereby make satisficing more likely. Also, the ability to filter alternatives by minimum or maximum attribute levels may lead to more noncompensatory decision making in the style of the conjunctive model used in our model.

Clearly, future research is needed to address the questions of when we should prefer to model consumer choice by a satisficing rule over modeling it by a maximizing choice. However, it should not come as a surprise that for frequently purchased (at least for the subject pool) and fairly inexpensive goods like instant noodles, consumers employ simpler choice rules such as the satisficing rule estimated in this paper. And if they do, our models should reflect that. Or so Simon says.

Appendix

A.1. Cross Inferences in Evaluation

As discussed in §4.3, the satisfactory-or-not decision can be influenced by memory in two ways: (1) which attribute levels are acceptable may be based on prior experience and the memory thereof, and (2) a consumer may make inferences about unseen attributes based on already examined attributes and his memory. In the following, we propose some potential model modifications to capture these effects.

Suppose that consumers are very familiar with a category, knowing one attribute level may give them information about another attribute. For instance, average prices may differ between brands. Learning the brand then also updates the expected price. Of course, this effect could also be incorporated in the model. We chose not to include it in the main body of the paper, because brand, flavor, and price are independent in our conjoint design (and we made it clear to the participants in the instructions). However, we now present a brief overview on how to model this effect.

To do so, the model needs to be extended to include an additional layer of information in the consumer's mind—namely, beliefs about the distribution of the attribute levels. In the proposed model, the consumer knows a given attribute level either with certainty or not at all. Yet if the consumer is familiar with the category (and the attribute levels are not randomly combined as in our experiment), she may have an idea about the distribution of, say, prices in her mind. This belief will then be a function of the consumer's information set at a given time. Before learning anything about a product, it will be equal to her prior beliefs about price for the category. However, once the brand, for instance, is learned, the distribution is updated to the belief about prices of that brand. Finally, once the price tag is fixated on, the distribution reduces to the true price level.

How exactly one attribute may or may not be informed by the knowledge of another attribute depends on the consumer as well as the particular category/application in question. However, once this dependency has been specified, one can simply use these distributions (or the expectations thereof) for the conjunctive rule rather than having to keep track of which attributes have already been fixated on and which have not. Thus, Equations (1) and (2) become, respectively,

$$\mathcal{S}_{xf} = \prod_{A} \gamma^*_{D[a_x \mid I_f]} \tag{22}$$

and

$$\mathcal{U}_{xf} = \max_{A}(1 - \gamma^*_{D[a_x \mid I_f]}), \tag{23}$$

where $D[a_x \mid I_f]$ denotes the belief about the distribution of attribute a for product x based on the information set after fixation f. The definition of γ would have to change slightly (denoted by γ^*) to account for the fact that it is now a function operating on a distribution. For interval-scaled attributes such as price, one could simply let $\gamma^*_{D[\cdot]} = \gamma_{E[a_x \mid D]}$,

Table A.1 Priors

Choice	Search
$\hat{\gamma}_a \sim$ Beta(1, 1)	$\bar{\phi}_{\cdot} \sim$ Normal(0, 100)
$\bar{\rho} \sim$ Dirichlet(1, 1, 1, 1, 1)	$\sigma_{\cdot} \sim$ Inverse Gamma(0.5, 0.5)
$\hat{\tau}_1 \sim$ Gamma(2, 1)	$\bar{\lambda}_{\cdot} \sim$ Normal(0, 100)
$\hat{\tau}_2 \sim$ 10·Gamma(2, 1)	$\varsigma_{\cdot} \sim$ Inverse Gamma(0.5, 0.5)

i.e., reduce the distribution to its expectation and employ the original γ defined in the main body of the text. For categorical attributes such as brand and flavor, the distributions cannot be collapsed as easily. A consumer may find a brand or flavor acceptable without having fixated on it, if she is sure enough that it will be a brand/flavor she likes. So, for instance, we could let

$$\gamma^*_{D[\cdot]} = \mathbb{I}_{\sum_a[\Pr(a\,|\,D)\gamma_a]\geq 0.8},$$

where $\Pr(a \mid D)$ is the probability of attribute level a given distribution D. Thus, in this case, a consumer would find a brand or flavor acceptable if she is at least 80% sure that it is an acceptable brand or flavor.

A.2. Priors

To complete the hierarchical Bayesian setup, a set of priors is needed. We choose largely uninformative priors, as shown in Table A.1. We scale the prior for $\hat{\tau}_2$ to be 10 times the prior for $\hat{\tau}_1$ to reflect the idea that the trembling hand probability should be fairly small. Nonetheless, the priors are wide enough to allow for a wide spectrum of Beta distributions on the trembling hand probabilities.

A.3. Conditional Posterior Distributions

To implement the MCMC for estimating the model, we need the conditional posterior distributions from which to draw. All parameters for the population hierarchies follow a standard conjugate distribution. Here, we present the conditional distributions for the individual-level parameters (including for the unobserved search states, as we augment the data with the search states in the estimation procedure).

First, define the following notation:

$\nu_{ij}(x) = \Pr(c_{ij} = x \mid \tau_i, \vec{\mathcal{S}}_{ijF}, \vec{\mathcal{U}}_{ijF}, \vec{\mathcal{UD}}_{ijF})$ as defined in Equations (8)–(12), where c_{ij} is defined as the outcome of person i's choice in choice set j.

Let $\theta_{ijf}(s^*)$ be the probability that state $s_{ijf} = s^*$ given in Equation (17) (suppressing the conditioning arguments and including Equations (18)–(20)).

Recall from Equations (13) and (15) that $\eta_{g,\,ijf}(h)$ and $\eta_{l,\,ijf}(h)$ (suppressing the conditioning arguments) denote the probabilities that fixation f_{ij} is on AOI h in the global state and local state, respectively. Also recall that following Equations (1) and (3) $\vec{\mathcal{S}}_{ijF}$, $\vec{\mathcal{U}}_{ijF}$, and $\vec{\mathcal{UD}}_{ijF}$ are deterministically determined given $\vec{\gamma}_{i,\,BR}$, $\vec{\gamma}_{i,\,FL}$, ρ_i, and the observed search path of person i in choice set j.

We then have the following conditional distributions (for all i): For the acceptability of brands and flavors, i.e., for $A \in \{BR, FL\}$,

$$\gamma_{ia} \propto \mathrm{Bern}(\hat{\gamma}_a) \cdot \prod_{j=1}^{15} \nu_{ij}(c_{ij} \mid \gamma_{ia}) \cdot \prod_{j=1}^{15} \left\{ \prod_{f=2}^{F_{ij}} [\eta_{s_{ijf},\,ijf}(h_{ijf} \mid \gamma_{ia}) \cdot \theta_{ijf}(s_{ijf} \mid \gamma_{ia})] \cdot \theta_{ij,\,F+1}(t \mid \gamma_{ia}) \right\}.$$

For the maximum acceptable price level,

$$\rho_i \propto \mathrm{MN}(\vec{\rho}) \cdot \prod_{j=1}^{15} \nu_{ij}(c_{ij} \mid \rho_i) \cdot \prod_{j=1}^{15} \left\{ \prod_{f=2}^{F_{ij}} [\eta_{s_{ijf},\,ijf}(h_{ijf} \mid \rho_i) \cdot \theta_{ijf}(s_{ijf} \mid \rho_i)] \cdot \theta_{ij,\,F+1}(t \mid \rho_i) \right\}.$$

For the trembling hand parameter,

$$\tau_i \propto \mathrm{Beta}(\hat{\tau}_1, \hat{\tau}_2) \cdot \prod_{j=1}^{15} \nu_{ij}(c_{ij} \mid \tau_i).$$

For the parameters in the global and the local search, i.e., for $m \in \{g, l\}$,

$$\phi_{im\cdot} \propto \mathrm{N}(\bar{\phi}_{m\cdot}, \sigma_{m\cdot}) \cdot \prod_{j=1}^{15} \left\{ \prod_{\{f:\,1<f\leq F_{ij} \text{ and } s_{ijf}=m\}} \eta_{m,\,ijf}(h_{ijf} \mid \phi_{im\cdot}) \right\}.$$

For the transition probability parameters, i.e., for $m \in \{l, t\}$,

$$\lambda_{im\cdot} \propto \mathrm{N}(\bar{\lambda}_{m\cdot}, \varsigma_{m\cdot}) \cdot \prod_{j=1}^{15} \left\{ \prod_{f=2}^{F_{ij}} \theta_{ijf}(s_{ijf} \mid \lambda_{im\cdot}) \cdot \theta_{ij,\,F+1}(t \mid \lambda_{im\cdot}) \right\}.$$

For the unobserved search states, recall that the first fixation for each choice set is assumed to be in the global state. Also recall that the termination state by definition is the final state (and only the final state). Therefore, the remaining fixations are either in the global state or in the local state. Below is the conditional posterior probability for one fixation to be in the global state (conditional, among other things, on the states of the other fixations). The conditional posterior probability for the local state is then simply the remaining probability such that the two probabilities sum to 1:

$$\Pr(s_{ijf} = g) = \frac{\theta_{ijf}(g \mid s_{ij,\,f-1}) \cdot \theta_{ij,\,f+1}(s_{ij,\,f+1} \mid s_{ijf} = g) \cdot \eta_{g,\,ijf}(h)}{\sum_{s\in\{g,\,l\}}[\theta_{ijf}(s \mid s_{ij,\,f-1}) \cdot \theta_{ij,\,f+1}(s_{ij,\,f+1} \mid s_{ijf} = s) \cdot \eta_{s,\,ijf}(h)]} \quad \forall\, i, j, \text{ and for } f = 2, \ldots, F_{ij}.$$

References

Bettman JR, Johnson EJ, Payne JW (1991) Consumer decision making. Robertson T, Kassarjian H, eds. *Handbook of Consumer Behavior* (Prentice Hall, New York), 50–84.

Casella G, George EI (1992) Explaining the Gibbs sampler. *Amer. Statistician* 49(4):327–335.

Chandon P, Hutchinson JW, Bradlow ET, Young SH (2009) Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *J. Marketing* 73(6):1–17.

Coombs CH (1951) Mathematical models in psychological scaling. *J. Amer. Statist. Assoc.* 46(256):480–489.

Dawes RM (1964) Social selection based on multidimensional criteria. *J. Abnormal Soc. Psych.* 68(1):104–109.

Duan JA, McAlister L, Sinha S (2011) Reexamining Bayesian model-comparison evidence of cross-brand pass-through. *Marketing Sci.* 30(3):550–561.

Elrod T (2002) Recommendations for validation of choice models. *Proc. Sawtooth Software Conf.* (Sawtooth Software, Sequim, WA), 225–243.

Elrod T, Johnson RD, White J (2004) A new integrated model of noncompensatory and compensatory decision strategies. *Organ. Behav. Human Decision Processes* 95(1):1–19.

Erdem T, Keane MP (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* 15(1):1–20.

Fader PS, McAlister L (1990) An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data. *J. Marketing Res.* 27(3):322–332.

Gelfand AE, Dey DK (1994) Bayesian model choice: Asymptotics and exact calculations. *J. Roy. Statist. Soc. Ser. B* 56(3):501–514.

Gelfand AE, Smith AFM (1990) Sampling based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* 85(410):398–409.

Gigerenzer G, Todd PM (1999) Fast and frugal heuristics: The adaptive toolbox. Gigerenzer G, Todd PM, ABC Research Group, eds. *Simple Heuristics That Make Us Smart* (Oxford University Press, New York), 3–36.

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

Gilbride TJ, Allenby GM (2006) Estimating heterogeneous EBA and economic screening rule choice models. *Marketing Sci.* 25(5):494–509.

Guadagni PM, Little JDC (1983) A logit model of brand choice calibrated on scanner data. *Marketing Sci.* 2(3):203–238.

Gupta S (1988) Impact of sales promotions on when, what, and how much to buy. *J. Marketing Res.* 25(4):342–355.

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

Heidelberger P, Welch PD (1983) Simulation run length control in the presence of an initial transient. *Oper. Res.* 31(6):1109–1144.

Jedidi K, Kohli R (2005) Probabilisitc subset-conjunctive models for heterogeneous consumers. *J. Marketing Res.* 42(4):483–494.

Johnson EJ, Meyer RJ, Ghose S (1989) When choice models fail: Compensatory models in negatively correlated environments. *J. Marketing Res.* 26(3):255–270.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.

Kamakura WA, Russell GJ (1989) A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* 26(4):379–390.

Kohli R, Jedidi K (2007) Representation and inference of lexicographic preference models and their variants. *Marketing Sci.* 26(3):380–399.

Lee AY (2002) Effects of implicit memory on memory-based versus stimulus-based brand choice. *J. Marketing Res.* 39(4):440–454.

Lenk PJ (2009) Simulation pseudo-bias correction to the harmonic mean estimator of integrated likelihoods. *J. Comput. Graphical Statist.* 18(4):941–960.

Liechty J, Pieters R, Wedel M (2003) Global and local covert visual attention: Evidence from a Bayesian hidden Markov model. *Psychometrika* 68(4):519–541.

Loomes G, Starmer C, Sugden R (1991) Observing violations of transitivity by experimental methods. *Econometrica* 59(2):425–439.

Mehta N, Rajiv S, Srinivasan K (2003) Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* 22(1):58–84.

Mehta N, Rajiv S, Srinivasan K (2004) Role of forgetting in memory-based choice decisions: A structural model. *Quant. Marketing Econom.* 2(2):107–140.

Netzer O, Toubia O, Bradlow ET, Dahan E, Evgeniou T, Feinberg FM, Feit EM, et al. (2008) Beyond conjoint analysis: Advances in preference measurement. *Marketing Lett.* 19(3):337–354.

Osborne MJ, Rubinstein A (1994) *A Course in Game Theory* (MIT Press, Cambridge, MA).

Pieters R (2008) A review of eye-tracking research in marketing. Malhotra NK, ed. *Review of Marketing Research*, Vol. 4 (Emerald Group Publishing, Bingley, UK), 123–147.

Pieters R, Warlop L (1999) Visual attention during brand choice: The impact of time pressure and task motivation. *Internat. J. Res. Marketing* 16(1):1–16.

Pieters R, Warlop L, Wedel M (2002) Breaking through the clutter: Benefits of advertising originality and familiarity for brand attention and memory. *Management Sci.* 48(6):765–781.

Reutskaja E, Nagel R, Camerer CF, Rangel A (2011) Search dynamics in consumer choice under time pressure: An eye-tracking study. *Amer. Econom. Rev.* 101(2):900–926.

Roberts JH, Lattin JM (1991) Development and testing of a model of consideration set composition. *J. Marketing Res.* 28(4):429–440.

Rossi PE, Allenby GM, McCulloch R (2005) *Bayesian Statistics and Marketing* (John Wiley & Sons, Hoboken, NJ).

Rottenstreich Y, Sood S, Brenner L (2007) Feeling and thinking in memory-based versus stimulus-based choices. *J. Consumer Res.* 33(4):461–469.

Russo EJ, Leclerc F (1994) An eye-fixation analysis of choice processes for consumer nondurables. *J. Consumer Res.* 21(2): 274–290.

Schwartz B, Ward A, Monterosso J, Lyubomirsky S, White K, Lehman DR (2002) Maximizing versus satisficing: Happiness is a matter of choice. *J. Personality Soc. Psych.* 83(5):1178–1197.

Shi SW, Wedel M, Pieters R (2012) Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Management Sci.* Forthcoming.

Shugan SM (1980) The cost of thinking. *J. Consumer Res.* 7(2):99–111.

Simon HA (1955) A behavioral model of rational choice. *Quart. J. Econom.* 69(1):99–118.

Smith BJ (2007) boa: An R package for MCMC output convergence assessment and posterior inference. *J. Statist. Software* 21(11):1–37.

Sun B (2005) Promotion effect on endogenous consumption. *Marketing Sci.* 24(3):430–443.

Swait J (2001) A non-compensatory choice model incorporating attribute cutoffs. *Transportation Res. Part B* 35(10):903–928.

Swait J, Adamowicz W (2001) The influence of task complexity on consumer choice: A latent class model of decision strategy switching. *J. Consumer Res.* 28(1):135–148.

Teixeira TS, Wedel M, Pieters R (2010) Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing. *Marketing Sci.* 29(5):783–804.

Tversky A (1972) Elimination by aspects: A theory of choice. *Psych. Rev.* 79(4):281–299.

van Diepen PMJ, de Graef P, d'Ydewalle G (1995) Chronometry of foveal information extraction during scene perception. Findlay JM, Walker R, Kentridge RW, eds. *Eye Movement Research: Mechanisms, Processes and Applications* (Elsevier, Amsterdam), 349–362.

van der Lans R, Pieters R, Wedel M (2008a) Competitive brand salience. *Marketing Sci.* 27(5):922–931.

van der Lans R, Pieters R, Wedel M (2008b) Eye-movement analysis of search effectiveness. *J. Amer. Statist. Assoc.* 103(482):452–461.

von Neumann J, Morgenstern O (1947) *Theory of Games and Economic Behavior*, 2nd ed. (Princeton University Press, Princeton, NJ).

Williams HCWL, de Dios Ortuzar J (1982) Behavioural theories of dispersion and the mis-specification of travel demand models. *Transportation Res. Part B* 16(3):167–219.

Wolfe JM, Horowitz TS (2004) What attributes guide the deployment of visual attention and how do they do it? *Nature Rev. Neurosci.* 5(6):1–7.

Wolfe JM, Yu KP, Stewart MI, Shorter AD, Friedman-Hill SR, Cave KR (1990) Limitations on the parallel guidance of visual search: Color $\times$ color and orientation $\times$ orientation conjunctions. *J. Experiment. Psych.: Human Perception Performance* 16(4):879–892.

Zhang J, Wedel M, Pieters R (2009) Sales effects of attention to feature advertisements: A Bayesian mediation analysis. *J. Marketing Res.* 46(5):669–681.