Attention and Preference Measurement

Liu (Cathy) Yang

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

Attention and Preference Measurement

Liu (Cathy) Yang

This dissertation contains two essays examining the role of attention and information processing in stated choices under choice-based preference measurement tasks.

While choice experiments have long been used in marketing as a way to measure consumer preferences, full rationality of consumers is always assumed, meaning consumers are able to process all the choice relevant information before making a decision. Moreover, conditioned on the premise that consumers process all the choice-relevant information, incentive-alignment mechanism introduced in choice experiments are considered the gold standard for inducing consumers to choose as they would in real-life situations. However, if consumers are boundedly rational and processing information is costly, we expect consumers to maximize not only the utility derived from the option they choose but also the utility derived from the process. Therefore, given a certain incentive structure, the amount of information processed by consumers is endogenized by individual preference toward the focal product in a choice experiment. Furthermore, research has shown that varying incentives in experiments might also result in changes in attention, which implies that the amount of attention paid in real-life choice situations (the probability of realizing a choice is 1) is different than the attention paid to choices paired with smaller incentives in most preference-measurement tasks (the probability of realizing a choice is strictly greater than 0 but lower than 1). In this dissertation, we first focus in Chapter 1 on the link between information processing and stated choices in an incentive-alignment choice experiment by developing a new preference measurement. We explore the impact of incentives on attention, information processing, and stated choices by conducting an experiment described in Chapter 2.

In Chapter 1, we develop a dynamic discrete choice model of information search and choice under bounded rationality, that we calibrate using a combination of eye-tracking and choice data. Our model extends the directed cognition model of Gabaix et al. (2006) by capturing fatigue,

proximity effects, and imperfect memory encoding and by estimating individual-level parameters and partworths within a likelihood-based, hierarchical Bayesian framework. We show that modeling eye movements as the outcome of forward-looking utility maximization improves out-of-sample predictions, enables researchers and practitioners to use shorter questionnaires, and allows better discrimination between attributes.

In Chapter 2, we empirically investigate whether incentives impact attention, information processing, and stated choices. We vary the probability that the respondent's choice will be realized from 0 (hypothetical) to 0.01, 0.50, 0.99, and 1 (deterministic) and collect data on both response times and eye tracking. We find a U-shaped relationship between the probability that the choice will be realized and the level of attention. Hypothetical questions and deterministic questions induce similar attention and information processing but different choices.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank the members of my dissertation committee, Eva Ascarza, Raghu Iyengar, Martijn de Jong, Kamel Jedidi, and Olivier Toubia for their time and comments on various parts of the dissertation. In particular, I am very grateful for the tremendous help and guidance from my advisor, Olivier Toubia, throughout my PhD program. I would also like to give special thanks to Martijn de Jong for his intelligent input on both chapters of my dissertation.

For my family and friends.

Essay 1: A Bounded Rationality Model of Information Search and Choice in Preference Measurement

Chapter 1

A Bounded Rationality Model of Information Search and Choice in Preference Measurement

1.1 Introduction

Choice experiments are used routinely in marketing, economics, and psychology. One common example is choice-based conjoint analysis. An implicit assumption typically made in the choice-based preference measurement literature is that respondents are fully rational and thus can systematically process all the choice-relevant information (i.e., attribute levels of all alternatives) and choose the alternative that provides the greatest utility. However, the bounded rationality literature (Simon, 1955) suggests that this assumption is not necessarily valid. Instead, consumers likely balance the utility of the option they choose with the (cognitive) utility derived from the choice process itself (e.g., Payne et al. (1988, 1992, 1993)).[1]

Trading off the costs of processing information with the benefits from the choice should result

[1] In doing so, they may revert to noncompensatory decision rules, e.g. disjunctive, conjunctive (Gilbride and Allenby, 2004, 2006; Jedidi and Kohli, 2005), lexicographic and elimination-by aspects (Johnson et al., 1989; Payne et al., 1988; Tversky et al., 1988; Yee et al., 2007). From a modeling perspective, we leverage the fact that noncompensatory decision rules are nested within additive utility models and we do not model them directly (Jedidi and Kohli, 2005, 2008; Yee et al., 2007).

in some choice-relevant information not being processed at all. This phenomenon has long been recognized in the marketing literature (e.g., Hagerty and Aaker (1984), Hauser et al. (1993), and Meyer (1982)) and has been documented recently using eye-tracking evidence (Shi et al., 2013; Stüttgen et al., 2012; Toubia et al., 2012). Recent models have endogenized the information acquisition process (Shi et al., 2013; Stüttgen et al., 2012) but not in a way that explicitly captures the strategic, dynamic tradeoff between the effort spent acquiring information and the benefits of making better-informed decisions.

This chapter attempts to close this gap. We develop a dynamic discrete choice model of information processing and choice under bounded rationality that we calibrate using a combination of eye-tracking and choice data. To the extent that consumers are strategic in their information acquisition process and that information acquisition is motivated by utility maximization, the information acquisition process should contain valuable information about consumers' preferences. Indeed, we show that complementing choice data with eye-tracking data and modeling eye movements as outcomes of forward-looking utility maximization improve out-of-sample performance, enable practitioners and researchers to use shorter questionnaires, and allow better discrimination between attributes.

Although we collected our eye-tracking data in a dedicated lab, commercial solutions are available, such as Eyetrackshop (`www.eyetrackshop.com`) and YouEye (`www.youeye.com`), that allow for collection of eye-tracking data in an online environment using the consumer's webcam. Therefore, we believe that the model developed in this chapter and the data on which it relies will be widely accessible in the near future and that market researchers will be able to collect eye-tracking data systematically to augment traditional choice data.

The rest of the chapter is organized as follows. We review some relevant prior literature in Section 1.2, present our model in Section 1.3, describe our data in Section 1.4 and the estimation results in Section 1.5, and conclude in Section 1.6.

1.2 Prior Literature

Our model bridges the literatures on dynamic discrete choices and eye-tracking. Before reviewing these literatures, we briefly introduce the context and type of data considered in our model. We consider a consumer who makes a series of choices in which alternatives are described by attributes that may have several levels. We assume that the choice-relevant information is presented to the consumer in a matrix such as the one shown in Figure 1.1 with one column per alternative and one row per attribute (alternative formats could be modeled as well, as in Shi et al. (2013)). We also assume that we observe, for each choice question, a series of eye fixations that end when the consumer chooses one of the alternatives. On each search opportunity, the consumer makes a choice between acquiring some choice-relevant information (by visiting a cell in the matrix) or ending the search and choosing one of the alternatives based on the information acquired up to that point. In the latter case, the consumer moves on to the next choice question.

1.2.1 Dynamic Models of Search

The decisions made by consumers in the process of acquiring choice-relevant information and choosing one alternative are an example of a typical dynamic choice setting, in which each decision may affect the utility offered by various future possible decisions. For example, acquiring a new piece of information on one alternative may change the identity of the alternative in the choice set with the highest expected utility. Such choice problems can be modeled using dynamic discrete choice models (e.g., Ching et al. (2012), Chintagunta et al. (2012), Dube et al. (2012), Hartmann and Nair (2010), Huang et al. (2012), Misra and Nair (2011), Rust (1987), Toubia and Stephen (2013), Yao and Mela (2011)). However, the standard approach to dynamic discrete choice modeling poses at least two challenges in our case. First, the state space is likely to be too large to allow estimation of a traditional dynamic discrete choice model using tools and computers available today. For example, suppose there are four alternatives per choice question described by six attributes. In this typical scenario, simply keeping track of which cells of the matrix were visited by the consumer would require 2^{24} possible states. Second, such an approach would not fit well with the assumption that consumers trade off decision accuracy and decision cost. In particular, once it is assumed that processing information and making decisions is potentially costly, search models that require solving dynamic programs suffer from the "infinite regress problem" – agents should optimize how they will optimize their decisions and optimize how they will optimize the way they optimize their decisions, and so on and so forth (Gabaix et al., 2006). More generally, several researchers have argued that models based on dynamic programming, while being normative, should be adjusted to capture the behavior of boundedly rational consumers (e.g., Assuncao and Meyer (1993) and Hutchinson and Meyer (1994)).

In light of these issues, we base our model on the directed cognition (DC) model proposed and validated by Gabaix et al. (2006). According to this model, on each search occasion t, the participant chooses as if this search occasion is the last one in that question. In other words, if the consumer decides to acquire some information, he or she does so as if he or she would be making a choice immediately after acquiring this new piece of information.

The DC model offers several benefits in addition to being computationally tractable for complex search problems like ours. First, although not fully forward-looking, the DC model is not myopic either and does capture the basic tradeoff in search problems (i.e., search more and choose later using more information vs. choose now based on the currently available information). Second, there is evidence that this model describes the actual behavior of human agents better than traditional search models based on optimal solutions to dynamic programs. Using a simple experimental setting, Gabaix et al. (2006) showed that the DC model predicts behavior better than a search model based on the optimal strategy, which in that case was available in closed form, using the Gittins-Weitzman index (Gittins, 1979; Weitzman, 1979). In another experiment, the authors showed that the model predicts behavior well in a more complex setup that shared some similarities with choice-based conjoint (CBC) analysis. In particular, Gabaix et al. (2006) used a setting in which choices were presented in a matrix format similar to Figure 1.1. However, each cell contained monetary payoffs and the value of each alternative was the sum of the amounts in the corresponding cells (i.e., subjects received the monetary value of the chosen alternative). Gabaix et al. (2006) tested the DC model using a Mouselab paradigm (see Payne et al. (1993)) in which most information was hidden and subjects could "open" only one cell at a time.

Although the DC model provides us with a framework that informs our modeling efforts, our specific model differs significantly from the model used by Gabaix et al. (2006). We will compare our implementation of the DC model to theirs in Section 1.3.3 after describing our model in more detail.

We also note that the DC model of Gabaix et al. (2006) is related to earlier studies in the marketing literature. For example, Hagerty and Aaker (1984) considered a similar context in which information is presented to consumers in a matrix form. In their model, at each search opportunity consumers evaluated the expected gain from visiting each cell in the matrix. That gain is linked to the probability that visiting a piece of information will change the identity of the option that provides the greatest expected utility. Similarly to Gabaix et al. (2006), Hagerty and Aaker (1984) assumed that consumers select the cell in the matrix that will maximize the expected gain in utility in the next period. See Meyer (1982) for a related model.

1.2.2 Eye-tracking Research in Marketing

Eye tracking data are composed of fixations and saccades (Wedel and Pieters, 2000). Fixations represent the time periods in which participants fix their eyesight to a specific location; saccades represent the eye movements between two fixations.

As a way to directly measure attention and involvement, eye-tracking studies have been conducted in numerous marketing settings, including branding (Pieters and Warlop, 1999; van der Lans et al., 2008a), advertising (Pieters et al., 1999, 2002; Pieters and Wedel, 2004; Rosbergen et al., 1997; Wedel and Pieters, 2000, 2008; Wedel et al., 2008), search effectiveness (van der Lans et al., 2008b), and brand display on supermarket shelves (Chandon et al., 2009).

Other studies have used eye tracking in preference measurement settings. Toubia et al. (2012) used eye tracking in a purely descriptive manner to measure the impact of "gamifying" a preference measurement task on the amount of attention paid by consumers. Musalem et al. (2013) used eye tracking to explore how consumers' preferences for each level of an attribute relate to the amount of attention paid to that attribute level and to alternatives that contained it. Shi et al. (2013) used eye-tracking data to study and model how consumers switch back and forth between attribute-based and alternative-based strategies when acquiring information about products described in a matrix format. The paper closest to ours is probably Stüttgen et al. (2012), which developed and estimated a model of search and choice in which consumers were assumed to use a satisficing rule; that is, they evaluated alternatives one after another and chose the first alternative that was deemed satisfactory. The authors further assumed that consumers used a conjunctive rule to decide whether an alternative was satisfactory (i.e., all attributes of the alternative need to be acceptable).

The standard approach for modeling eye-tracking data among these papers was either to treat eye fixations as exogenous (e.g., Musalem et al. (2013)) or to endogenize eye fixations using hidden Markov models (Liechty et al., 2003; Shi et al., 2013; Stüttgen et al., 2012; van der Lans et al., 2008a,b). The states in hidden Markov models of eye movements typically capture various information acquisition strategies or modes of search. For example, Stüttgen et al. (2012) followed Liechty et al. (2003) and assumed that consumers move back and forth between a "local search" state and a "global search" state that involve eye movements in the periphery of the current eye position (local) and in different areas (global). Their model captures how the transition probabilities between these states are influenced by the consumer's ongoing evaluations of the various alternatives (i.e., which alternatives have already been classified as satisfactory and which as unsatisfactory based on the information processed up to that point).

Our model takes a different approach. The key differentiator of our model (compared to extant models based on hidden Markov processes) is that we allow consumers to be *strategic and forward-looking* in how they acquire information. We model information acquisition as the result of forward-looking utility maximization; the utility derived by a consumer comes not only from the chosen product but also from the information acquisition process itself. Another key feature of our model is that we allow for imperfect memory encoding; in other words, a consumer may need multiple fixations in a region of interest before remembering the information it contains.

1.3 Model

In this section we develop a dynamic discrete choice model in which the amount of information processed by consumers is endogenized and modeled as the result of forward-looking utility maximization in which the consumer derives (positive or negative) utility both from his or her final choice and from the search process itself. The model is designed to be calibrated using a combination of eye-tracking and choice data.

1.3.1 Specification

For ease of exposition, we focus on one consumer when describing our model. We index choice questions by k, and each choice question consists of selecting one of J alternatives that are described by I attributes. For ease of presentation, we assume without loss of generality that all attributes have the same number of levels, L. The choice-relevant information is presented in a matrix such as the one shown in Figure 1.1 with one column per alternative and one row per attribute.

For simplicity, we assume that the choice questions come from a random experimental design. In this case, attributes vary independently across alternatives and there is no need to model inferences consumers may make across attributes and alternatives. Our approach could easily be extended to nonrandom experimental designs.

Each time period in our model is a search occasion, t, in which the consumer chooses between acquiring some choice-relevant information (by visiting a cell in the matrix that contains the level of one attribute for one choice alternative) and ending the search and choosing one of the alternatives based on the information acquired up to that point. In the latter case, the consumer moves to the next choice question.

As mentioned earlier, our model is inspired by the DC model proposed and validated by Gabaix et al. (2006). According to that model, on each search occasion t, the participant chooses as if the search occasion is the last one in that question. In other words, if the consumer decides to acquire some information, he or she does so as if he or she is going to make a choice immediately after acquiring this new piece of information.

We develop a likelihood-based implementation of the DC model that allows for heterogeneity in preferences. Like any dynamic model, our implementation specifies an action space, a set of state variables, a utility function, and state-transition probabilities. We next define each of these components and the resulting likelihood function.

Actions

We denote the current position of the eyes in the $I \times J$ matrix that contains the choice-relevant information by $p = (i, j)$. On each search occasion, the consumer may move his or her eyes to a different location (i', j') in the matrix or end the information acquisition process and choose one of the alternatives (j'), thereby moving to the next choice question.[2]

States

Although the attribute levels for each alternative in a choice question are known to the researcher, they are unknown to the consumer at the beginning of the question. Take Figure 1.1 as

[2]We only consider fixations within the regions of interest that contain choice-relevant information. In the first search occasion in each question, the number of possible cells to move to is $I \times J$ instead of $I \times (J-1)$ (there is no "current" position). We collapse consecutive fixations within the same cell as one fixation since they are likely to be caused by participants randomly moving their eyes in a very small range due to blinking (nonconsecutive fixations in the same cell are recorded as distinct fixations).

an example. Consumers learn the level of each attribute in each alternative when they move their eyes to the relevant cell in the matrix. In our data, we found that consumers revisited 62.22% of the cells they visited at least once. Therefore, it would be unreasonable to assume that consumers learn the level of attribute i for alternative j with certainty after only one fixation in cell (i, j). Instead, we assume an imperfect memory encoding process in which consumers form a set of beliefs about the true value of each cell. These beliefs are updated after each fixation, and they converge to the truth as the number of fixations increases.

Our two observed state variables are p, which captures the current eye position, and a set of numbers $\{n_{i,j}\}$ in which $n_{i,j}$ is the number of times cell (i, j) is visited (i.e., number of fixations in the cell that contains information on attribute i for alternative j).

We follow Wedel and Pieters (2000) and assume that consumers extract a chunk of information with each fixation on a cell. We denote as η the amount of information extracted per fixation. Again following Wedel and Pieters (2000), we further assume that the total amount of information stored by the consumer related to cell (i, j) is the sum of the activation levels of all memory traces: $\eta \times n_{i,j}$. Suppose the true level in cell (i, j) is l_0. After $n_{i,j}$ fixations in that cell, the total amount of information in support of l_0 being the true level is $\eta \times n_{i,j}$. The total amount of information in support of any other level being the true level is 0. If we assume some error in memory retrieval $(\delta_{i,j,l})$ (Wedel and Pieters, 2000), then the probability that the consumer believes l_0 is the true level is $Prob(\eta n_{i,j} + \delta_{i,j,l_0} > \delta_{i,j,l},\ \forall\ l \neq\ l_0)$. If we assume that $\delta_{i,j,l}$ follows a double exponential distribution, the probability weight associated with each level l, $w_{i,j,l}$, becomes

$$w_{i,j,l}(\eta, n_{i,j}) = \begin{cases} \frac{exp(\eta n_{i,j})}{L-1+exp(\eta n_{i,j})} & \text{if } l \text{ is the true level} \\ \frac{1}{L-1+exp(\eta n_{i,j})} & \text{if } l \text{ is not the true level} \end{cases} . \tag{1.1}$$

We denote the $1 \times L$ array of probability weights corresponding to all possible levels in cell (i, j) as $w_{i,j}(\eta, n_{i,j})$ which equals $[w_{i,j,1}(\eta, n_{i,j}), ..., w_{i,j,L}(\eta, n_{i,j})]$. Before the first visit to a cell, the consumer starts with a uniform belief, $w_{i,j}(\eta, 0) = [\frac{1}{L}, ..., \frac{1}{L}]$ that reflects the random experimental design (nonrandom designs would potentially give rise to different initial beliefs). The set of weights corresponding to a cell is updated after each visit to the cell and converges to a vector that has a weight of 1 on the true level and 0 on all the other levels. Appendix A.1 illustrates this process using a simple example.

We further assume the existence of unobserved state variables in the form of idiosyncratic shocks $\epsilon(a)$ that capture information unobservable to the econometrician. This addition allows us to write a likelihood function for our model following a standard distributional assumption (see Rust (1987)).

Utility Function

The utility derived by the consumer at each search occasion is a function of the current state and the consumer's action. We make a distinction between product-related utility derived by the consumer (i.e., the utility that comes from the alternative (j) chosen by the consumer) and search-related utility (i.e., the utility that comes from the search process, which may be positive or negative). The consumer derives product-related utility only upon ending the search.

We first describe product-related utility. For ease of exposition, we do not include a subscript for the consumer in our equations although all of the parameters are estimated at the individual level. We assume effects coding; that is, the partworth for the last level of an attribute is equal to minus the sum of the partworths for the other levels. Let β_i be the $(L-1)\times 1$ vector containing a consumer's partworths for attribute i under effects coding. The $L \times 1$ vector containing all partworths for attribute i is $\widetilde{\beta}_i = I_i^0 \beta_i$ where I_i^0 is an $L \times (L-1)$ coding matrix:

$$I_i^0 = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ 0 & 0 & ... & 1 \\ -1 & ... & ... & -1 \end{bmatrix}. \tag{1.2}$$

With perfect memory encoding, the consumer would know the true level contained in cell (i, j) after one fixation, and the partworth corresponding to that cell would simply be the appropriate element of $\widetilde{\beta}_i$. However, with imperfect memory encoding, the consumer assigns a set of probability weights, $w_{i,j}(\eta, n_{i,j})$, to each possible level in cell (i, j), and the *expected value* of the partworth corresponding to that cell is a weighted average of the partworths for attribute i: $w_{i,j}(\eta, n_{i,j})\widetilde{\beta}_i$.

This expression converges to the appropriate element of $\widetilde{\beta}_i$ as the number of fixations on the cell increases. The product-related utility is specified as

$$u_{product}(a|\{n\}, \beta) = \begin{cases} 0 & \text{if } a = \textit{move to } (i', j') \\ \sum_i w_{i,j'}(\eta, n_{i,j'})\widetilde{\beta}_i & \text{if } a = \textit{choose } j' \end{cases}. \tag{1.3}$$

Appendix A.1 illustrates computation of product-related utility using a simple example.

In a case in which the consumer decides not to choose any alternative and continues searching instead, he or she derives search-related utility. Given the finding that the amount of information processed by participants tends to decrease as they progress through the questionnaire (Stüttgen et al., 2012; Toubia et al., 2012), we allow for fatigue effects by modeling search-related utility as a function of the question number, k.

Shi et al. (2013) documented that a large proportion of consecutive fixations are to contiguous cells when information about products is displayed using matrices as in our case. This is consistent with eye movements between cells being more cognitively costly when the cells are more distant. Moreover, Shi et al. (2013) identified an asymmetry in consumers' propensity to make horizontal vs. vertical eye movements. To capture these physiological factors, we also allow search-related utility to be a function of the distance between the current location of the eyes, p, and the next cell visited and allow for different weights on horizontal and vertical movements. In particular, we model search-related utility as[3]

$$u_{search}(a|p,k,\theta) = \begin{cases} \theta_0 + \theta_1 k + \theta_2 d(a,p,\theta_3) & \text{if } a = \textit{move to } (i',j') \\ 0 & \text{if } a = \textit{choose } j' \end{cases} \tag{1.4}$$

where $d(a,p,\theta_3)$ is a weighted Euclidean distance between the current cell (i,j) and the next cell (i',j') defined as $\sqrt{(i-i')^2 + \theta_3(j-j')^2}$. The parameter θ_3 captures asymmetries between vertical and horizontal eye movements (this parameter is constrained to be non-negative). Note that we do not restrict the signs of the parameters θ_0, θ_1, and θ_2 .

Transition Probabilities

The state variables capture eye position and the number of fixations in each cell. The transitions between states are deterministic from the perspective of the researcher based on the consumer's actions. However, from the perspective of the consumer, there is some uncertainty regarding the true value of each cell so the transitions between states are probabilistic.

Suppose that the consumer's action, a, is to visit cell (i',j'). From the perspective of the consumer, for each possible level l there is a probability (given by $w_{i',j',l}(\eta, n_{i',j'})$) that level will

[3]We also tested a version of the model with only the parameter θ_0 in the search-related utility function. The deviance of information criteria favored the more complex version, and out-of-sample performance was not greatly affected. Details are available from the author.

be found in cell (i', j'). Therefore, the value function is based on a set of transition probabilities given by the current set of probability weights, $w_{i',j'}(\eta, n_{i',j'})$, which are a function of the current state variable, $n_{i',j'}$. These weights represent the probability of seeing each level in each cell based on the consumer's current beliefs. (Note that we assume that the consumer knows which cell he or she is visiting; the only uncertainty is related to the level contained in the cell.)

Suppose that l_0 is the true level in cell (i', j'). To specify the Bellman equation, we need to model all possible state transitions from the perspective of the consumer. In particular, we need to model what would happen when the consumer sees a level in cell (i', j') that is different from l_0. Although this never happens, it must be addressed because the weights $(w_{i',j',l}(\eta, n_{i',j'}))$ are positive for all levels, not just the true level. If level $l \neq l_0$ is found in cell (i', j'), the amount of information in support of that level being the true level would increase from 0 to η. As a result, the probability weight associated with the level would be updated to $\frac{exp(\eta)}{L-2+exp(\eta n_{i',j'})+exp(\eta)}$. The weight associated with the true level would be updated to $\frac{exp(\eta n_{i',j'})}{L-2+exp(\eta n_{i',j'})+exp(\eta)}$ and the weights associated with the other levels to $\frac{1}{L-2+exp(\eta n_{i',j'})+exp(\eta)}$. With a slight abuse of notation, we denote the product-related utility based on the new beliefs that are formed if level l is to be found in cell (i', j') as $u_{product}(a'|\{n'_{i',j',l}\}, \beta)$.

After the fixation to cell (i', j'), the state variable corresponding to that cell, $n_{i',j'}$, is incremented by 1 and the set of probability weights corresponding to that cell is updated to $w_{i',j'}(\eta, n_{i',j'} + 1)$ based on Equation (1.1).

There is no need to specify transition probabilities for the other state variables (p (current fixation position) and $\epsilon(a)$) because the former evolves deterministically and the latter is assumed to satisfy the conditional independence assumption (Rust, 1987).

Likelihood Function

The DC model assumes that consumers act on each search occasion as if this search occasion is their last opportunity to acquire new information. Mathematically, this implies that consumers behave on each search occasion as if they are solving the following optimization problem:

$$\begin{aligned}&\max\{\max_{a=\{j'\}}\{u_{product}(a|\{n\}, \beta) + \epsilon(a)\},\\&\max_{a=\{i',j'\}}\{u_{search}(a|p,k,\theta) + \epsilon(a) + \sum_l w_{i',j',l}(\eta, n_{i',j'}) \max_{a'=\{j\}}\{u_{product}(a'|\{n'_{i',j',l}\}, \beta) + \epsilon(a')\}\}\}\end{aligned} \quad (1.5)$$

The first term, $\max_{a=\{j'\}}\{u_{product}(a|\{n\},\beta)+\epsilon(a)\}$, is the maximum utility the consumer can derive by ending the search and choosing one of the alternatives given the current state variables $\{n\}$ and the consumer's partworths β. The second term is the maximum utility the consumer can derive by continuing the search, and $u_{search}(a|p,k,\theta)+\epsilon(a)$ is the search-related utility, $w_{i',j',l}(\eta,n_{i',j'})$ captures the state-transition probabilities (from the perspective of the consumer), and $\max_{a'=\{j\}}\{u_{product}(a'|\{n'_{i',j',l}\},\beta)+\epsilon(a')\}$ is the maximum utility derived from choosing one of the alternatives in the next period given that level l is found in cell (i',j').

Assuming that the idiosyncratic shocks, ϵ, satisfy the conditional independence assumption and follow a double-exponential distribution gives rise to the following likelihood function in which $\Theta=\{\beta,\theta,\eta\}$.

$$P(a|\{n\},p,\Theta)=\frac{\exp(V_a(\{n\},p|\Theta))}{\sum_{a'}\exp(V_{a'}(\{n\},p|\Theta))} \tag{1.6}$$

where:

$$V_a(\{n\},p|\Theta)=$$
$$\begin{cases} u_{search}(a|p,k,\theta)+\sum_l w_{i',j',l}(\eta,n_{i',j'})\log\sum_{a'=\{j\}}\exp(u_{product}(a'|\{n'_{i',j',l}\},\beta)) & \text{if } a=\textit{move to } (i',j') \\ u_{product}(a|\{n\},\beta) & \text{if } a=\textit{choose } j' \end{cases} \tag{1.7}$$

1.3.2 Identification and Estimation

The parameters to be estimated in our proposed dynamic discrete choice model are $\Theta=\{\beta,\theta,\eta\}$. As with a standard CBC analysis, the partworths, β, are identified at the individual level through the choices consumers make between various alternatives. The parameters θ_0, θ_1, θ_2, and θ_3 capture search-related utility. Given the value of the partworths, the intercept θ_0 is identified because we observe consumers choosing either to continue the search or to stop the search and select one of the alternatives. The parameter θ_1 captures the effects of fatigue (through the question number) and is identified because we observe multiple questions per consumer. The parameters θ_2 and θ_3 capture the effect of distance on search utility and are identified because the information in each cell varies randomly across questions. We estimate both β and θ at the individual level. Finally, η is a parameter that captures the amount of information extracted per fixation (i.e., it may be interpreted as capturing the speed with which consumers learn the content of a cell). This parameter is identified primarily through the common occurrence of revisits to cells that were previously visited by the same consumer in the same question. However, we have found that this parameter is only weakly identified. Therefore, we estimate it only at the aggregate level using a grid search. In particular, we fix the parameter η and estimate all the other parameters given that value of η for multiple values of η. We keep the value of η that gives rise to the lowest deviance of information criteria (DIC)[4] . We estimate our model using a hierarchical Bayes method (Atchadé and Rosenthal, 2005). The first stage prior for $\{\theta_n, \beta_n\}$ (where n indexes consumers) is normal with $\{\theta_n, \beta_n\} \sim N(\mu_0, \Lambda)$. The second stage priors are $\mu_0 \sim N(0, 1000*I)$ and $\Lambda^{-1} \sim Wishart(I, 23+3)$ where 23 is the number of heterogeneous parameters in the model. A total of 150,000 Markov chain Monte Carlo (MCMC) iterations are performed using the first 100,000 as burn-in. We apply a grid search method for the learning parameter η; we estimate the model with $\eta = 0 - 5$ with a step of 1 and select the best-fitting model based on DIC.

We confirmed the identification of our model and tested our estimation approach using a simulation study. Appendix A.2 provides the details of our simulation study. We generated a synthetic data set using a set of parameters inspired by the estimates from our study reported in Section 1.5. We found that the parameters were recovered adequately.

1.3.3 Comparison with Gabaix et al.

We used the DC model of Gabaix et al. (2006) as a basic framework for our model, but our model differs significantly in several important ways. First, the Gabaix et al. model was applied to a context in which each cell contained a monetary amount and the payoff from the chosen alternative was the sum of the monetary values of its cells. Product-related utility in Gabaix et al.'s model was simply the amount of money earned in the game. We apply our model to a context in which each cell contains an attribute level and product-related utility is parameterized by a set of partworths. Gabaix et al. assumed that the value in each cell was drawn from a continuous normal distribution, while in our case the values are drawn from a discrete uniform distribution. As a result, Gabaix et al. was able to derive a closed-form expression for the expected benefit from each possible action

[4] The DIC is defined as $-4E_\Theta[\log f(a|\{n\}, p, \Theta)|a] + 2\log f(a|\{n\}, p, \hat{\Theta}(a, \{n\}, p))$, where $\hat{\Theta}(a, \{n\}, p) = \arg\max_\Theta \{f(a|\{n\}, p, \Theta)\}$ (Celeux et al., 2006).

(see Equation (3) in Gabaix et al. (2006)) while we use the general Bellman equation. Gabaix et al. assumed that search-related utility is constant (the opportunity cost of time) while we allow search-related utility to be affected by fatigue and proximity effects (i.e., consumers may search less and less over time and may be more likely to move their eyes to nearby cells). Gabaix et al. also assumed perfect memory encoding (a consumer learns the content of a cell perfectly after one visit) while we allow for imperfect memory encoding. While Gabaix et al. calibrated the one parameter in their model (opportunity cost of time) by fitting moments of the data (average amount of search in the game), we develop a likelihood-based, hierarchical Bayesian framework. And Gabaix et al. calibrated their model at the aggregate level; we allow for heterogeneity across consumers.

1.4 Data

1.4.1 Set-up

We collected the choice-based conjoint (CBC) data in the context of Dell laptop computers and used six attributes ($I = 6$) with four levels each ($L = 4$): processor speed (1.6 GHz, 1.9 GHz, 2.7 GHz, and 3.2 GHz), screen size (26 cm, 35.6 cm, 40 cm, and 43 cm), hard drive capacity (160 GB, 320 GB, 500 GB, and 750 GB), Dell support subscription (1 year, 2 years, 3 years, and 4 years), McAfee antivirus subscription (30 days, 1 year, 2 years, and 3 years), and price (350, 500, 650, and 800 euros).

In the main task, each participant answered twenty choice questions, where each offering four alternatives ($J = 4$). The questions were generated randomly (once for all participants, i.e., all participants saw the same set of questions). Before answering the twenty questions, participants completed one training question designed to familiarize them with the interface. Figure 1.1 provides a screenshot of one of choice questions.

Figure 1.1: Screenshot from the first question in the main task.

PART I - QUESTION 1

Please indicate your favorite product from the set below.
One out of 100 respondents will be selected as a winner and will receive 800 euros, which will be used to purchase a laptop automatically.
If you are selected as a winner, with 50% probability you will receive your preferred laptop from Part II. With probability 50%, you will receive your preferred laptop from one randomly selected question from Part I. All questions from Part I are equally likely to be selected.
In all cases, you will receive both the laptop and the difference between 800 euros and its price.

	A	**B**	**C**	**D**
Processor speed	2.7 Ghz	2.7 Ghz	1.6 Ghz	3.2 Ghz
Screen size	40 cm	43 cm	40 cm	35.6 cm
Hard drive	160 GB	500 GB	320 GB	320 GB
Dell support	3 years	2 years	2 years	4 years
McAfee subscription	2 years	3 years	2 years	2 years
Price	500 euro	650 euro	500 euro	800 euro
	○	○	○	○

submit

In addition to the main task, participants completed an external validity task. We used a typical setting in which the external validity task consisted of one choice task with eight alternatives that were chosen randomly (once for all participants; all participants saw the same set of alternatives) subject to the constraint that each level of each attribute would be present in at least one of the alternatives. This task also was preceded by one training question to familiarize participants with the interface.

We randomized the position of the external validity task relative to the main task so that half of the participants completed the external validity task first and half the main task first. This difference was our only between-subject variation.

Our study followed an incentive-alignment scheme typical of CBC studies (Ding et al., 2005; Ding, 2007). One participant was selected randomly as a winner and received 800 euros, which were to be used to automatically purchase a laptop based on his or her answers to the survey. The winner received the alternative chosen in the external validity task with probability 50% and the alternative chosen in each question in the main task with probability 2.5%. The winner received the preferred laptop along with the difference between 800 euros and the price of the laptop.

Our participants were recruited at a large European university. They all participated in the survey in the university's behavioral lab using the online platform developed by the author.

Eye-Tracking Data

Participants completed the survey while being monitored by a free-standing nonintrusive Tobii® 2150 eye tracker that sampled infrared corneal reflections at 50Hz with a $0.35°$ spatial resolution and an accuracy of $0.5°$. The stimuli were presented on a 21-inch LCD monitor with a display resolution of 1,600 by 1,200 pixels. The position of the left eye and right eye were recorded separately (van der Lans et al., 2011). Fixations and saccades were differentiated using van der Lans et al. (2011)'s velocity-based algorithm. We defined the region of interest (ROI) for each piece of information as the area within the boundary of the cell that contained the information (see Figure 1.1).

1.4.2 Descriptive Statistics

We collected complete eye-tracking data for 70 participants. Of those, 33 completed the external validity task before the main task and 37 completed it after the main task.

We now provide a descriptive analysis of our eye-tracking data for the twenty questions in the main task. The average proportion of cells visited at least once (with at least one fixation) across all questions and participants was 69.65%. The proportion differs slightly with the order in which the main task and the external validity task were completed: 67.79% for the main task first and 71.74% for the external validity task first (p-value=0.13). We accommodated this difference by adding a parameter to our search-utility specification (see Section 1.5.1.). Figure 1.2 plots the average proportion of information visited in each choice question. The downward trend in this graph confirms the need to control for question position in our model and is consistent with previous findings (Stüttgen et al., 2012; Toubia et al., 2012). Figure 1.3 shows the distribution of the proportion of information visited across all choice questions and participants. Figure 1.4 shows the distribution of the number of visits per piece of information (each "piece of information" consists of the level of one attribute for one alternative) for all pieces of information, choice questions, and participants. This chart shows that information that is processed is likely to be visited multiple times by the same consumer in the same question, which confirms the need to model memory encoding as imperfect; it would not be reasonable to assume that visiting a cell once is enough for a consumer to completely memorize its content. Figure 1.5 shows the distribution of the distances between two consecutive eye fixations across all choice questions and participants[5]. We see that consumers are much more likely to move their eyes to an adjacent (distance = 1) cell in the 6 by 4 matrix containing all choice-relevant information than they are to move a more distant cell. This is consistent with previous studies (Shi et al., 2013; Stüttgen et al., 2012) and confirms the need model search-related utility as a function of the distance between cells.

[5]If the respondent moves his or her eyes between cell (i, j) and (i', j'), the distance is defined as $\sqrt{(i-i')^2 + (j-j')^2}$.

Figure 1.2: Average proportion of information visited per choice question vs. question number.

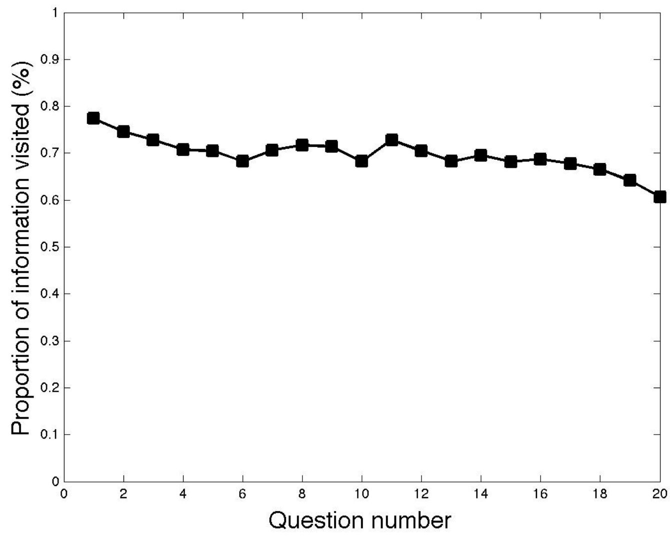


Figure 1.3: Distribution of the proportion of information visited per choice question.

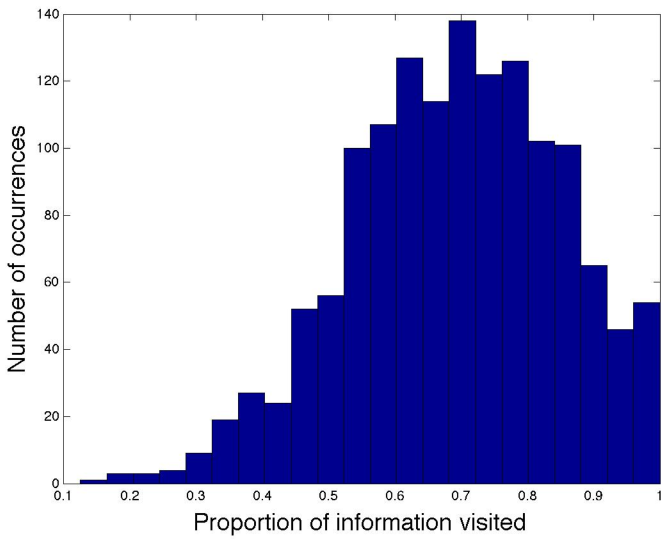


Figure 1.4: Distribution of the number of visits per piece of information.

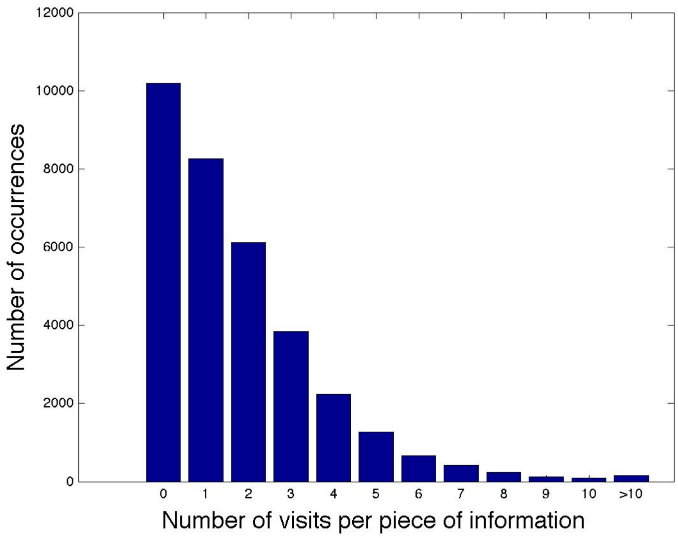
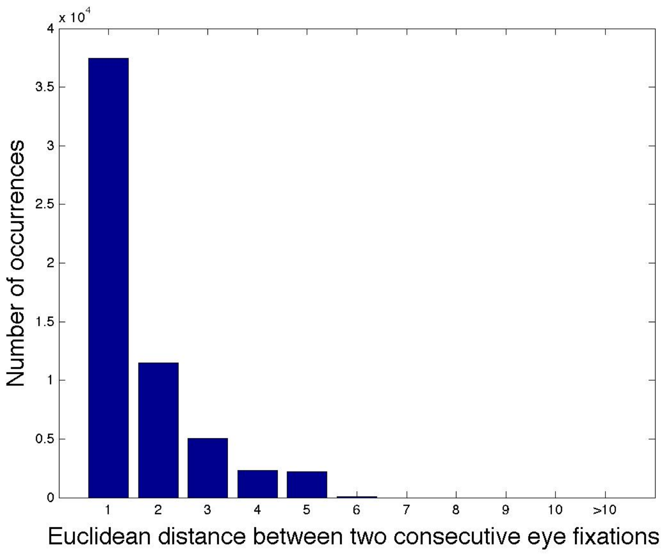


Figure 1.5: Distribution of the Euclidean distance between successive eye fixations.



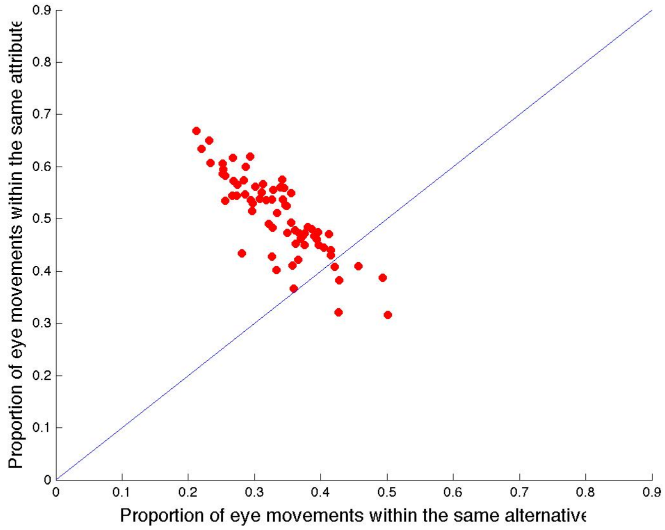
Note: If the respondent moves his or her eye from the cell (i, j) to cell (i', j'), the distance is defined as $\sqrt{(i-i')^2 + (j-j')^2}$.

Table 1.1 shows the proportion (across all questions and participants) of eye movements that were to a different choice alternative within the same attribute, to a different attribute within the same alternative, and to a different attribute in a different alternative. While most movements were either within the same alternative or within the same attribute, there is no evidence that either alternative-based processing or attribute-based processing dominates. To explore the possibility that each type of processing dominates for subsets of consumers, we present a scatter plot of the proportion of within-attribute and within-alternative eye movements at the participant level in Figure 1.6 (each dot represents one participant). We see that most participants use a hybrid of attribute-based and alternative-based searches, although attribute-based searches were more prevalent on average. To further investigate the existence of attribute-based and alternative-based searches, we report the distribution of the number of attributes visited per alternative (across all alternatives, respondents, and choice questions) in Figure 1.7 and the number of alternatives visited per attribute (across all attributes, respondents, and choice questions) in Figure 1.8. Attribute-based search would lead to some attributes not being visited at all, and alternative-based search would lead to some alternatives not being visited at all. We find that an alternative (attribute) is completely ignored only 1.13% (4.33%) of the time. This further suggests no evidence for purely attribute-based or alternative-based searches, and confirms the need for a model to be flexible enough to allow for any type of eye movements.

Table 1.1: Overall proportion of eye movements to a different alternative within the same attribute, to a different attribute within the same alternative, and to a different attribute in a different alternative.

Type of search	**Proportion**
Different alternative within the same attribute	0.51
Different attribute within the same alternative	0.34
Different attribute in a different alternative	0.16

Figure 1.6: Scatter plot of the proportion of eye movements to a different alternative within the same attribute versus a different attribute within the same alternative.



Note: Each dot corresponds to one respondent.

Figure 1.7: Distribution of the number of attributes visited per alternative (across all alternatives, respondents and choice questions).

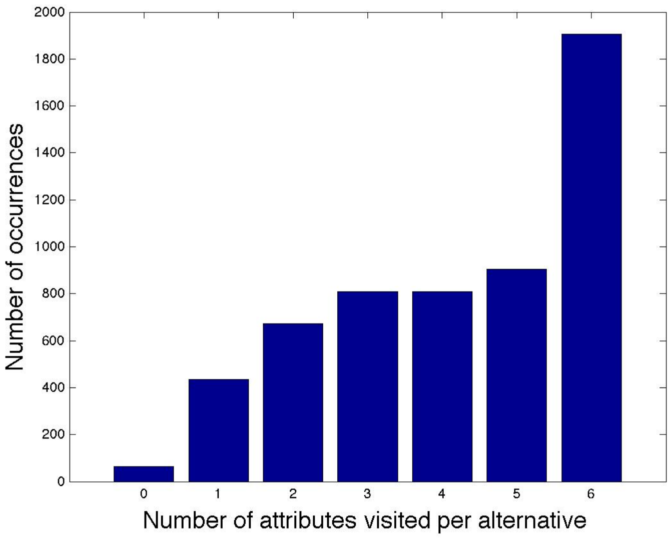
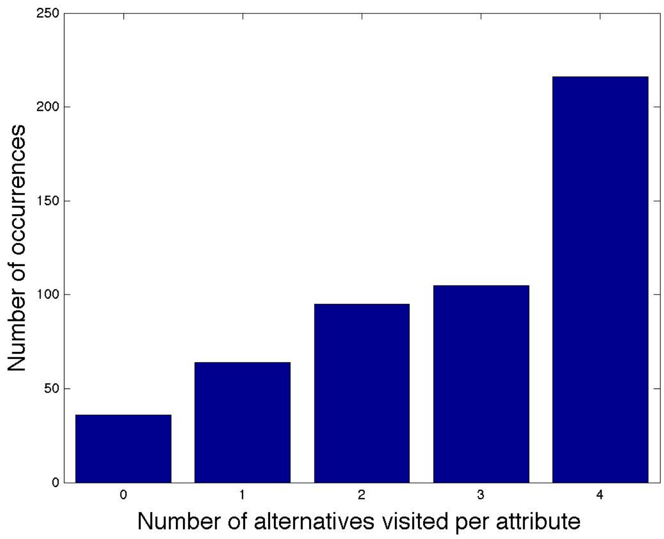


Figure 1.8: Distribution of the number of alternatives visited per attribute (across all attributes, respondents and choice questions).



1.5 Estimation Results

We use the last four questions of the main task as holdouts. We vary the number of questions used for estimation between the first eight and sixteen to assess the benefits of the proposed model when the number of choice questions is increased.

1.5.1 Proposed Model

We estimate the proposed model described in Section 1.3 with one small adjustment to capture the change in position of the external validity task and its effect on the propensity to search. Thus, we add one term to the search-related utility equation:

$$u_{search}(a|p,k,\theta) = \begin{cases} \theta_0 + \theta_{11}k + \theta_{12}\mathbb{1}(ext\ val) + \theta_2 d(a,p,\theta_3) & \text{if } a = move\ to\ (i',j') \\ 0 & \text{if } a = choose\ j' \end{cases} \tag{1.8}$$

where $\mathbb{1}(ext\ val)$ is an indicator function equal to 1 if the participant completed the external validity task before the main task. The additional parameter θ_{12} captures the effect of completing the external validity task before the main task on the propensity to search for information in the main task. Recall that we omit consumer subscripts for ease of exposition, but the partworths and search-related utility parameters are all estimated at the individual level. In addition, we constrain the partworths for price to be monotonic using rejection sampling (Allenby et al., 1995).

1.5.2 Benchmark Models

We use several benchmark models to test various components of the proposed model. All benchmarks are estimated using the same Bayesian approach and the same prior specifications as the proposed model.

Our first set of benchmarks does not model searches and focuses only on choice. We refer to this set as the "choice-only" benchmarks. The first is a standard multinomial logit (MNL) choice model that assumes that participants have full knowledge of the alternatives in each choice question (i.e., the information in all of the 24 cells in Figure 1.2 is assumed to be known). The second benchmark is an MNL model that takes into account information on the specific cells visited (cells with at least one fixation) by each participant in each question and assumes that participants only use the information contained in the cells that they visited when they evaluate the alternatives. The likelihood function in these benchmarks is based only on the choice data. Therefore, we cannot compare them to the proposed model based on measures of in-sample fit, and we use out-of-sample predictions instead.

Our second set of benchmarks models both the choices made by consumers and their eye movements and therefore may be compared to our proposed model based on in-sample fit statistics (deviance of information criteria (DIC)). We refer to this set as "search+choice" benchmarks. In each search+choice benchmark, the same imperfect memory encoding process from the proposed model (Equation 1.1) is assumed, and the same specification is used for product-related utility (Equation 1.3) and search-related utility (Equation 1.4). The only difference is in specification of the forward-looking term in the value function (Equation 1.7). The first benchmark in this set (labeled "future product-related utility unanticipated") assumes that consumers only take search-related utility into consideration when deciding whether and how to search for information and that they ignore future product-related utility. This benchmark assumes that the value function from Equation (1.7) takes the following form.

$$V_a(\{n\},p|\Theta) = \begin{cases} u_{search}(a|p,k,\theta) & \text{if } a = \textit{move to } (i',j') \\ u_{product}(a|\{n\},\beta) & \text{if } a = \textit{choose } j' \end{cases} \tag{1.9}$$

Our second benchmark in this set explores the possibility that, while consumers may take future utility into account when deciding whether to continue searching for information, they may not take into account how the results of the search will impact their future beliefs, which will impact their future expected utility. This benchmark assumes that consumers behave on each search occasion as if they will not update their beliefs after the search. We label this benchmark as "future belief updating unanticipated." Note that this benchmark assumes that subjects ignore future updating of beliefs at the time of the decision but updates beliefs after subjects acquire the new information. The value function takes the following form.

$$V_a(\{n\},p|\Theta) = \begin{cases} u_{search}(a|p,k,\theta) + \log\sum_{a'=\{j\}} \exp(u_{product}(a'|\{n\},\beta)) & \text{if } a = \textit{move to } (i',j') \\ u_{product}(a|\{n\},\beta) & \text{if } a = \textit{choose } j' \end{cases} \tag{1.10}$$

We estimate η, the learning parameter, for each benchmark separately using a grid search of the same set of values as in the proposed model. Table 1.2 reports the DICs for the proposed model and the second set of benchmarks when the number of questions used for calibration varies from eight to sixteen. We find that the proposed model has a better fit than both benchmarks. Thus, it is reasonable to assume that consumers take future product utility into account when deciding whether and how to search and that they anticipate how searching will impact their beliefs about the various alternatives.

1.5.3 Posterior Check

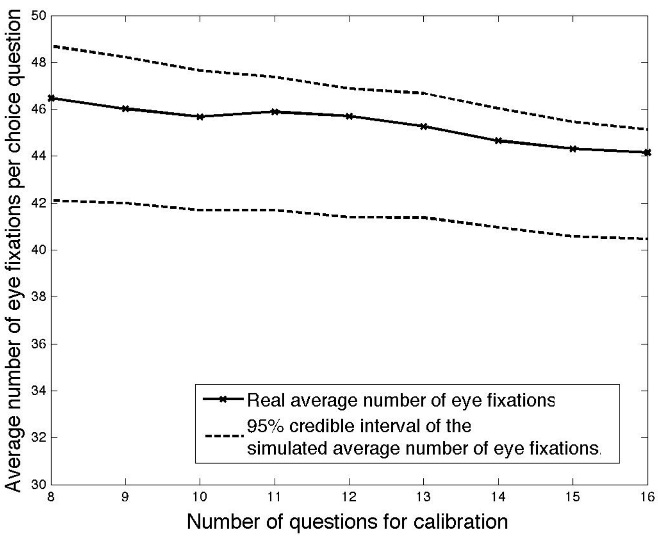
In addition to computing the DICs, we measure in-sample fit for the proposed model by evaluating how well it recovers some key statistics of the data (Gelman et al., 1996). At each iteration of the Gibbs sampler, we use the parameter estimates in that iteration to simulate the number of eye fixations for each respondent and for each choice question in the calibration set and repeat this analysis when the number of questions used for calibration varies from eight to sixteen. Figure 1.9

Table 1.2: Comparison of proposed model with search+choice benchmarks based on deviance of information criteria (DIC).

	Number of questions used for calibration								
Model	8	9	10	11	12	13	14	15	16
Proposed model	140758.518	156895.061	172595.803	190780.442	207433.841	222558.975	236862.551	252007.345	267657.965
Future belief-updating unanticipated	142849.949	159134.888	175173.667	193607.851	210737.014	226183.093	240554.609	255871.439	271813.857
Future product-related utility unanticipated	143390.124	159746.143	175816.279	194222.653	211265.653	226766.771	241243.021	256665.679	272617.374

shows the real average number of eye fixations across choice questions and respondents together with the 95%-credible interval of this statistic across iterations of the Gibbs sampler as the number of questions used for calibration is varied. In all cases, the true values falls within the 95%-credible interval.

Figure 1.9: Posterior check of the average number of eye fixations per choice question vs. number of questions used for calibration.



1.5.4 Parameter estimates

Table 1.3 shows the posterior means and 95%-credible intervals of the first-stage prior parameters defined in Section 1.3.2., i.e., the population mean of the partworths and search-utility parameters

(μ_0) and the population variance of those parameters (diagonal elements of Λ). The estimates presented in Table 1.3 are based on the proposed model using the first sixteen questions for calibration. (Recall that all parameters in Table 1.3 are estimated at the individual level; we report only the population means here.) All results are based on $\eta = 3$, the value suggested by the grid search for this parameter. As previously mentioned, we used effects coding such that the partworths sum up to 0 within each attribute.

The signs of θ_{11} and θ_{12} are consistent with the descriptive statistics reported earlier that showed that search decreases as the questionnaire progresses and that placing the external validity first slightly increases the amount of attention spent in the main task. The sign of θ_2 is consistent with the finding that consumers tend to move to cells that are close to the one they are currently visiting. The fact that θ_3 is less than 1 is consistent with Table 1.1, which shows that horizontal eye movements within the same attribute are more frequent than vertical ones within the same alternative. The positive sign of θ_0 suggests that search-related utility may be positive in certain situations.

The posterior means of the importance of each attribute based on the proposed model and the set of choice-only benchmarks are presented in Table 1.4. The proposed model extracts more information about each attribute from each choice question. Therefore, we expect this model to give rise to more discrimination across attributes; for each consumer, there should be more variance in attribute importance across attributes. Table 1.4 also reports posterior means and credible intervals of the averages (across consumers) of the variance (across attributes) of the partworth importances. That is, at each iteration of the MCMC, we compute the variance of the attribute importances for each consumer and calculate the average of this variance across consumers. As predicted, the proposed model gives rise to much more variation across attributes than the benchmarks (the 95%-credible intervals do not overlap with the benchmarks). In other words, complementing choice data with search data and modeling the information acquisition process as the result of forward-looking utility maximization increases discrimination across attributes.

1.5.5 Out-of-sample Predictions

We compare out-of-sample prediction performance between the proposed model and the two sets of benchmarks (search+choice and choice-only), based on the hit rate on both the holdout questions

Table 1.3: Population estimates from the proposed model.

		Posterior Population Mean	95% Credible Interval	Posterior Population Variance
	θ_0	2.33	[2.13, 2.51]	1.05
	θ_{11}	-0.01	[-0.03, 0.00]	0.36
Search-related parameters	θ_{12}	0.28	[-0.07, 0.61]	1.3
	θ_2	-1.07	[-1.08, -1.05]	0.45
	θ_3	0.70	[0.66, 0.73]	0.75
	1.6 Ghz	-8.80	[-9.11, -8.51]	17.03
Processor speed	1.9 Ghz	-3.21	[-3.44, -2.97]	6.54
	2.7 Ghz	3.94	[3.69, 4.15]	5.07
	26 cm	-0.31	[-0.54, -0.09]	17.06
Screen size	35.6 cm	1.42	[1.16, 1.77]	4.36
	40 cm	0.31	[-0.01, 0.51]	4.14
	160 GB	-3.36	[-3.87, -2.99]	4.27
Hard drive	320 GB	-0.62	[-1.12, -0.28]	2.57
	500 GB	1.20	[0.98, 1.51]	2.68
	1 year	-1.51	[-1.75, -1.27]	1.93
Dell support	2 years	0.90	[0.69, 1.11]	1.81
	3 years	0.16	[-0.07, 0.41]	1.55
	30 days	-0.83	[-1.03, -0.57]	4.37
Anti virus	1 year	0.05	[-0.23, 0.39]	1.44
	2 years	0.58	[0.31, 0.82]	2.13
	350 euro	4.98	[4.68, 5.28]	14.18
Price	500 euro	1.49	[1.34, 1.66]	1.97
	650 euro	0.04	[-0.13, 0.17]	1.29

Note: The first sixteen questions are used for calibration. We used effects coding so the partworth for the last level of each attribute is minus the sum of the other three partworths.

Table 1.4: Average attribute importances and average variance of attribute importances.

		Proposed Model	Choice-only Benchmarks	
			Assume consumers fully informed	Use knowledge of which cells visited
Average attribute importance	Processor speed	0.335	0.281	0.232
	Screen size	0.145	0.140	0.140
	Hard drive	0.140	0.199	0.189
	Dell support	0.076	0.091	0.125
	Anti virus	0.081	0.090	0.108
	Price	0.222	0.200	0.206
Average variance of attribute importances		0.010	0.006	0.003
95% Credible interval		[0.009, 0.011]	[0.004, 0.007]	[0.002, 0.003]

Note: The first sixteen questions are used for calibration.

(the last four questions in the main task) and the external validity task. For each consumer and out-of-sample question, we measure the hit rate by computing the estimated choice probability of the chosen alternative at each MCMC iteration and then computing the average across MCMC iterations.[6] We plot how the average performance of each model evolves as the number of questions used for calibration varies between eight and sixteen. Results are reported in Figures 1.10 through 1.13.

[6]We computed the utility of each alternative in each out-of-sample choice question by multiplying the characteristics of the alternatives by the consumer's partworths. This standard approach implicitly assumes that consumers take into account the full description of all the alternatives in the choice set. We also tried making out-of-sample predictions for the search+choice models based on counterfactual simulations. In particular, we could simulate the consumer's search process in the out-of-sample questions and estimate the resulting choice probabilities. Predictive performance was slightly worse using this approach. Details are available from the author.

Figure 1.10: Proposed model vs. choice-only benchmarks - average holdout hit rate vs. number of questions used for calibration.

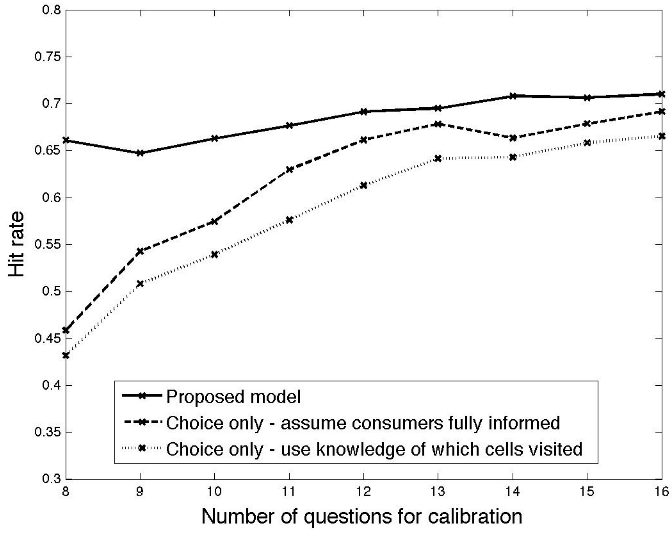


Figure 1.11: Proposed model vs. search+choice benchmarks - average holdout hit rate vs. number of questions used for calibration.

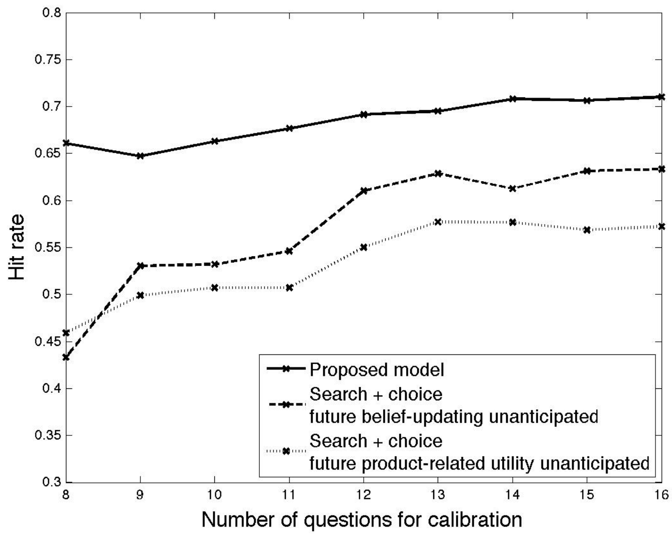


Figure 1.12: Proposed model vs. choice-only benchmarks - average external validity hit rate vs. number of questions used for calibration.

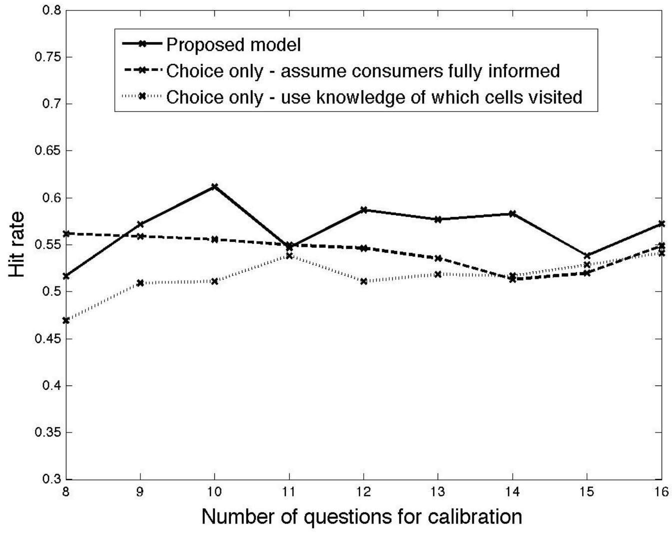
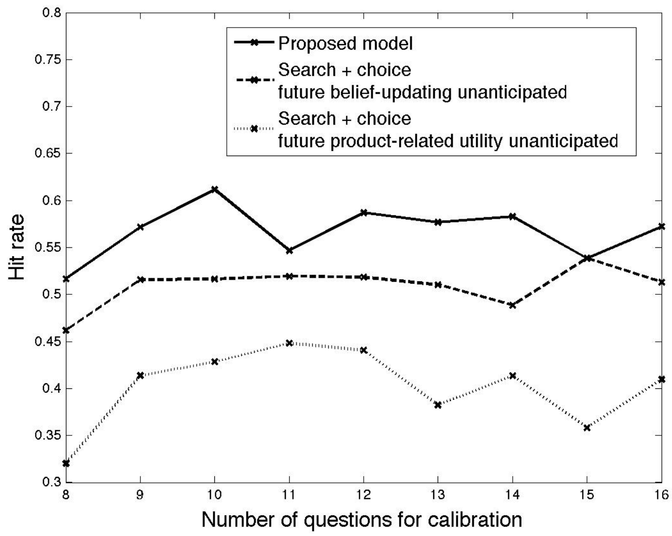


Figure 1.13: Proposed model vs. search+choice benchmarks - average external validity hit rate vs. number of questions used for calibration.



We see in Figure 1.10 that the hit rate on the holdout questions is systematically higher in the proposed model than in the choice-only benchmarks and that the difference becomes less and less

Table 1.5: Proposed model vs. choice-only benchmarks - regression results.

	Holdout questions		**External Validity**	
	Coefficient	P-value	Coefficient	P-value
Intercept	0.918	0.000	0.618	0.000
Choice only - assume consumers fully informed dummy	-0.187	0.000	-0.518	0.002
Choice only - use knowledge of which cells visited dummy	-0.426	0.000	-0.679	0.000
q	0.040	0.000	0.006	0.843
q^2	0.001	0.728	-0.017	0.184
Choice only - assume consumers fully informed dummy $\times$ q	0.098	0.000	-0.027	0.521
Choice only - use knowledge of which cells visited dummy $\times$ q	0.088	0.000	0.048	0.259
Choice only - assume consumers fully informed dummy $\times$ q^2	-0.025	0.000	0.018	0.342
Choice only - use knowledge of which cells visited dummy $\times$ q^2	-0.018	0.001	0.009	0.645

Note: The number of observations in each regression is the number of respondents $\times$ 9 (number of questions used for calibration varies from eight to sixteen) $\times$ 3 (number of models being compared). The dependent variable is the logit of the hit rate. We include respondent fixed effects. The variable q is the (mean-centered) number of questions used for calibration.

Table 1.6: Proposed model vs. search+choice benchmarks - regression results.

	Holdout questions		**External Validity**	
	Coefficient	P-value	Coefficient	P-value
Intercept	0.918	0.000	0.618	0.000
Search+choice - future belief-updating unanticipated dummy	-0.491	0.000	-0.789	0.000
Search+choice - future product-related utility unanticipated dummy	-0.716	0.000	-1.342	0.000
q	0.040	0.000	0.006	0.835
q^2	0.001	0.730	-0.017	0.163
Search+choice - future belief-updating unanticipated dummy $\times$ q	0.062	0.000	0.037	0.362
Search+choice - future product-related utility unanticipated dummy $\times$ q	0.024	0.055	-0.007	0.861
Search+choice - future belief-updating unanticipated dummy $\times$ q^2	-0.017	0.002	0.017	0.346
Search+choice - future product-related utility unanticipated dummy $\times$ q^2	-0.010	0.078	-0.002	0.888

Note: The number of observations in each regression is the number of respondents $\times$ 9 (number of questions used for calibration varies from eight to sixteen) $\times$ 3 (number of models being compared). The dependent variable is the logit of the hit rate. We include respondent fixed effects. The variable q is the (mean-centered) number of questions used for calibration.

pronounced as the number of questions grows. This is consistent with the fact that the proposed model is able to extract more information from each choice question. Therefore, it is able to achieve greater predictive performance with fewer questions. It takes approximately twelve choice questions for the best choice-only benchmark to reach the performance achieved by the proposed model after eight choice questions, and the performance of the best choice-only benchmark after sixteen questions is similar to the performance of the proposed model after only twelve questions. In addition, the choice-only benchmark that uses eye-tracking data to account for which cells were visited in each question performs worse than the standard choice-only benchmark that assumes that all cells are visited. Thus, to extract valuable information from eye-tracking data, it is not enough to merely capture which information was processed by the consumer and it is better to endogenize the information search process.

To compare the performance of these models statistically, we run an ordinary least squares regression with the logit of the hit rate as the dependent variable.[7] The number of observations in each regression is the number of respondents $\times$ 9 (number of questions used for calibration varies from eight to sixteen) $\times$ 3 (number of models being compared). We include respondent fixed effects to capture the panel structure of the data. We use the proposed model as the baseline and include dummy variables for each of the two benchmarks. We also include covariates that capture the increasing trend in performance as the number of questions increases. Results are reported in Table 1.5. We find that the main effects corresponding to the two benchmarks are significantly negative; the proposed model on average performs significantly better than either benchmark.

Figure 1.11 compares the proposed model to the search+choice benchmarks, and Table 1.6 reports the results of the regression. The proposed model performs better than either of the benchmarks. The worst performing benchmark is the one that assumes that consumers ignore future product-related utility when deciding whether and how to acquire information. The benchmark that assumes that consumers take future product-related utility into account but ignores how their beliefs will be updated performs better but still not as well as the proposed model. This suggests that the gain from the proposed model comes from assuming both that consumers take future utility into account and that they take the impact that the additional search will have on their future decisions into account.

[7] We take the logit because the hit rate is bounded between 0 and 1.

Figures 1.12 and 1.13 compare the models' performance on the external validity task. Although the proposed model still performs better than the benchmarks and the difference remains significant on average (see Tables 1.5 and 1.6), the comparisons are a little noisier for at least two reasons. First, the external validity task was a single choice question while the holdout comparisons are based on the last four questions, which reduces the variance in performance across consumers. Second, it turns out that the choice shares of the eight alternatives in the external validity question were very unevenly distributed; the three most popular alternatives were chosen by 95.0% of the respondents (with respective shares of 55.7%, 28.6%, and 11.4%).

1.6 Conclusions

In this chapter, we develop a joint model of information processing and choice that explicitly captures the strategic, dynamic tradeoff between search-related utility and product-related utility. We find that the proposed model offers better out-of-sample predictions than a wide range of benchmarks that either do not leverage data on the information search process or do so without endogenizing searches as the outcome of forward-looking utility maximization. Our model also allows for greater discrimination between various attributes than the benchmarks, which only model choice and not search.

Our contribution is both methodological and managerial. Methodologically, our model extends Gabaix et al. (2006)'s DC model in several important ways. That model had a single parameter (opportunity cost of time) that was estimated at the aggregate level by matching moments of the data. In contrast, our model specifies a rich search-related utility function that captures fatigue and proximity effects and a product-related utility function parameterized by a set of partworths. Moreover, our model allows consumers to have imperfect memory encoding. We estimate our model within a likelihood-based, hierarchical Bayesian framework that allows for heterogeneity across consumers.

Managerially, as discussed in the introduction, there are commercial solutions available today that allow for collection of eye-tracking data in an online environment using the consumer's webcam. We expect such solutions to be increasingly common as large companies such as Facebook acquire such capabilities in 2012, (Facebook acquired GazeHawk, a startup that provides webcam eye-tracking services, (ZDNet, 2012)) and with the development of open-source solutions such as ITU Gaze Tracker (`www.gazegroup.org/downloads/23-gazetracker`). Therefore, we believe that the approach developed in this chapter will be increasingly accessible to market researchers. We show that complementing choice data with eye-tracking data and modeling eye movements as the outcome of forward-looking utility maximization improve out-of-sample performance, enable practitioners and researchers to use shorter questionnaires, and allow greater discrimination between attributes. We envision eye-tracking data being collected systematically in online market research to augment and improve the responses given by consumers.[8]

Finally, we believe that the present research offers several directions for future research. First, our model may be extended to account for risk aversion, loss aversion, regrets, and other behavioral phenomena (Hauser et al., 1993). Second, our model could provide a framework for developing and testing new theories related to information search and choice. Third, additional physiological measures could be collected during preference measurement tasks. For example, commercial software is already available to scan consumers' faces and extract information from their facial expressions (Teixeira et al., 2012). These additional measures could be incorporated in preference measurement models to further improve predictive performance and reduce the required length of questionnaires. Fourth, our bounded rationality framework may be used to shed new light on the impact of incentives in preference measurement. In our framework, consumers trade off the cognitive costs related to information processing with the benefits derived from their choices. Varying the incentives (e.g., how likely each choice is to be realized) would impact the expected benefits derived from each choice, which should then impact how much information (and possibly which information) gets processed by consumers during the task.

[8] All code used in this chapter is available upon request.

Essay 2: Attention, Information Processing and Choice in Incentive-Aligned Choice Experiments

Chapter 2

Attention, Information Processing and Choice in Incentive-Aligned Choice Experiments

2.1 Introduction

Choice experiments are used routinely in marketing, both by researchers and practitioners. A perennial issue associated with choice experiments is reduced attention (e.g., Ding et al. (2005), Liechty et al. (2005), and Netzer et al. (2008)). For example, Johnson (2008, p. 4) writes: "Although respondents do seem to use simplification strategies when filling out questionnaires, they probably work harder when making important real-life choices."

The most common approach for keeping subjects motivated during choice experiments and inducing them to behave more closely to how they would in real-life situation, is incentive alignment (e.g., Ding et al. (2005, 2009), Ding (2007), Dong et al. (2010), Park et al. (2008), and Toubia et al. (2012)). Incentive alignment induces truth-telling. Indeed, as long as respondents do not derive utility from lying, they have an incentive to tell the truth if their answers may have real consequences. For example, Ding (2007) showed that truth-telling is the Bayesian Nash equilibrium in his incentive-aligned procedure. However, truth-telling is sometimes confused with realism by researchers and practitioners. Incentive alignment is often viewed as the gold standard that ensures that a study reflects the real world.

In this chapter, we compare *hypothetical* choice tasks that have no probability of being realized, *probabilistically incentive-aligned* choice tasks in which the respondent's decision is realized with some probability strictly greater than 0 and strictly less than 1, and *deterministically incentive-aligned* choice tasks in which the decision is realized with probability 1. Incentive-aligned choice experiments typically use probabilistically incentive-aligned tasks while real-life choices are typically deterministically incentive-aligned (unless they involve gambles).

Based on the bounded rationality literature, we expect consumers to maximize not only the utility derived from the option they choose but also the utility derived from the process. The cognitive cost involved in processing choice-relevant information is the same irrespective of the incentives, but the expected benefits derived from the choice are smaller when incentives are probabilistic rather than deterministic. Therefore, if processing information is costly, subjects should process less choice-relevant information under probabilistic incentives than under deterministic incentives.

One stream of research (Arkes et al., 1986; Camerer and Hogarth, 1999; Gneezy and Rustichini, 2000; Heyman and Ariely, 2004; McGraw and McCullers, 1979) has suggested that modest incentives may actually diminish a consumer's intrinsic motivation, thereby *reducing* the amount of attention paid to the task. This would lead consumers to attend to hypothetical questions more carefully than they do to probabilistically incentive-aligned questions.

From a practical perspective, different levels of attention are likely to give rise to different behaviors and choices and therefore to lead researchers and practitioners to conclusions that may not hold in real-life situations.

We explore the link between incentives, attention, information processing, and behavior in choice experiments. We use an experimental design in which the probability ($prob$) that the consumer's choice will be realized varies from 0 to 1. We measure attention using both response times and eye-tracking data. We find a U-shaped relationship between the probability that the choice will be realized and the level of attention paid. Probability of 0 and 1 generate similar levels of attention. Probabilities between 0.01 and 0.99 generate levels of attention that are similar to each other but smaller than the extreme cases (0 or 1). Eye-tracking data allows us to shed light on information processing beyond attention. We find that respondents in conditions of probabilities of 0 and 1 allocate attention similarly across the choice-relevant information treatments. However, this does not imply that purely hypothetical questions ($prob = 0$) should be favored by researchers and practitioners. Indeed, consumers' decisions were different when presented with a purely hypothetical choice versus a choice that had a positive probability of being realized.

The rest of the chapter is organized as follows. We review relevant prior literature in Section 2.2 and describe our data and results in Section 2.3. In Section 2.4, we conclude and discuss implications for practice and research and possible remedies to the issues raised by our findings.

2.2 Prior Work

2.2.1 Incentive-aligned Choice Experiments

Incentive-aligned choice experiments are usually considered the gold standard in economics, psychology and marketing. Our experiment is conducted in the context of studies of preference measurement which increasingly rely on incentive-aligned choice experiments. Ding et al. (2005) proposed an incentive-aligned conjoint mechanism to offer additional motivation to respondents to provide truthful input and showed that it increased external validity in choice-based conjoint (CBC) experiments. Ding (2007) extended this method by allowing researchers to reward respondents from a limited set of products when some of the alternatives in all choice sets were not available as possible rewards. While Ding's (2007) model required estimates of price sensitivity, Dong et al. (2010) proposed an alternative approach based on an inferred rank order of the potential reward products that yielded similar predictive performance.

Other incentive-aligned preference measurements have been developed as well. Ding et al. (2009) proposed an online incentive-aligned method inspired by barter markets. Park et al. (2008) introduced a mechanism to elicit preferences using a web-based upgrading method (i.e., respondents would state their willingness to pay for an upgrade, and the transaction would be realized if the randomly generated price was smaller than stated willingness to pay). Toubia et al. (2012) developed and tested an incentive-aligned conjoint poker game to measure preferences.

All of those incentive-aligned preference measurement methods (like most other incentive-aligned choice experiments in the marketing literature) follow an approach known in economics as the random lottery mechanism (RLM). In an RLM, each choice has some probability of being realized and at most one choice is realized per subject. In other words, RLM implies probabilistically incentive-aligned tasks (i.e., the respondent's decision in each task is realized with some probability strictly greater than 0 and strictly less than 1). On the other hand, as previously mentioned, real-life choices are typically deterministically incentive-aligned (the decision is realized with probability 1).

2.2.2 Probabilistic Incentives vs. Deterministic Incentives

An implicit assumption typically made in incentive-aligned choice experiments is that respondents are fully rational. They therefore systematically process all the choice-relevant information (i.e., descriptions of the choice alternatives) and choose the alternative that provides the greatest utility. However, the literature on bounded rationality suggests that this assumption is not necessarily valid. Simon (1955) argued that the assumption that people process all the information is hardly satisfied given the complexity of a task environment and the decision-maker's limited computational capabilities. Payne et al. (1988, 1992, 1993) further argued that preferences are of a constructive nature, partly due to conflicting decision goals (e.g., maximizing choice accuracy vs. minimizing effort) and partly to decision complexity.

In the economics literature, Wilcox (1993) conceptualized the behavior of boundedly rational decision-makers in an RLM that uses a decision cost model according to which subjects trade off the expected utility related to their choices with the (cognitive) cost derived of the choice process itself. The cognitive cost is not a function of the probability that each choice will be realized. However, the expected utility from each choice consists of the (possibly weighted) probability that the choice will be realized multiplied by the utility of the chosen option. Therefore, the expected utility is smaller in probabilistically incentive-aligned choices than in deterministically incentive-aligned choices. Harrison (1994) and Smith and Walker (1993) argued that, in an RLM, the opportunity cost of deviating from the rational prediction in each choice is often negligible because it is weighted by the modest probability that the choice will be realized. Experimentally, Beattie and Loomes (1997) showed that the number of violations of rationality (arguably, a symptom of reduced effort) is smaller for real choices than for choices realized with some probability strictly less than 1. Therefore, we predict that consumers will process less information in probabilistically incentive-aligned choice questions than in deterministically incentive-aligned choice questions. This prediction is consistent with a large literature in economics that has shown that, when incentives are offered, the amount of effort tends to be monotonically increasing in the amount of the incentives (Camerer and Hogarth, 1999; Gneezy and Rustichini, 2000; Jenkins et al., 1998).

One limitation of extant research on the impact of incentives on effort is that effort is usually not measured directly. Economists often use deviations from rational behavior as an indicator of reduced effort (e.g., Beattie and Loomes (1997)). Some researchers have used response times as a proxy for effort (Wilcox, 1993). Other notable research has documented physical manifestations of effort induced by greater incentives by measuring pupil dilations (Kahneman and Peavler, 1969) or using Mouselab (Stone and Schkade, 1994). In this chapter, we focus on attention, information processing, and choice rather than on deviations from normative behavior. We use hypothetical, probabilistic, and deterministic incentives and measure attention and information processing using both response-time and eye-tracking data.

2.2.3 No Incentive vs. Probabilistic Incentives

As previously discussed, many studies have suggested that, *conditional* on offering incentives, greater incentives usually induce higher effort. However, several studies have compared the impact of offering modest incentives to that of offering no incentive at all and found that effort can actually be greater in the *absence* of incentives (Arkes et al., 1986; Camerer and Hogarth, 1999; Heyman and Ariely, 2004; McGraw and McCullers, 1979). For example, Gneezy and Rustichini (2000) found that subjects who were offered $0.10 per correct answer on an IQ test performed worse than subjects who were not rewarded for their performance. However, subjects who were offered greater incentives did perform better. Heyman and Ariely (2004) found that subjects exerted less effort in an online game when they were compensated at a low rate ($0.10) than when they received no compensation. (The authors found that performance was best when subjects were compensated at a high rate of $4.00.) One explanation for this phenomenon is that subjects are intrinsically motivated to exert effort when monetary rewards are absent and that the presence of incentives undermines intrinsic motivation (Camerer and Hogarth, 1999). A variant of this explanation is that two types of markets exist that determine the link between effort and payment: monetary and social (Heyman and Ariely, 2004). Offering incentives shifts the transaction between the researcher and the consumer from a social market (in which the consumer is doing the researcher a “favor”) to a monetary one (in which the consumer “works” for the researcher and responds to the amount of payment offered in compensation for effort). Therefore, it is not unreasonable to expect that consumers pay less attention in choice experiments that are probabilistically incentive-aligned than in choice experiments that are purely hypothetical. However it is an empirical question whether this latter effect would be found in the type of choice experiments used in marketing.

In summary, the literature suggests that consumers pay less attention to relevant information in probabilistically incentive-aligned choice experiments than in deterministically incentive-aligned ones. The literature also suggests that consumers may pay less attention when incentives are probabilistic than when no incentives are offered at all. However, to the best of our knowledge, no study has compared information processing in choice experiments involving no incentive, probabilistic incentives, and deterministic incentives. The experimental design described in the next section fills that gap; it uses both response-time and eye-tracking data as measures of attention and information processing in analyzing the impact of incentives on choices.

2.3 Experimental Design

We collected experimental data to explore the relationship between incentives, attention and information processing in choice experiments and the impact of those relationships on the choices made by consumers. We focused on the typical marketing context of preference measurement. Each respondent in our experiment chose one of eight products (snack boxes). We asked a single question per respondent because the choices had to be deterministically incentive-aligned for some respondents. Choices realized with probability of 1 will affect a consumer's response to any subsequent choice (e.g., the consumer may seek variety in subsequent choices). In the experiment, each respondent received 3 euros as compensation for participation. With some probability between 0 and 1, this money was automatically used to purchase the respondent's chosen product. The amount of attention spent by each respondent during the choice task was measured using response times and eye tracking.

This study compares hypothetical, probabilistic, and deterministic incentives in the context of choice experiments and is one of only a few studies that have been able to measure the impact of incentives on information processing directly using eye-tracking data and to link information processing to choice.

2.3.1 Methods

We tested our hypothesis in the context of online preference measurement and selected snack boxes as the product, which allowed us to offer customized products to respondents. Each snack box contained five components, giving rise to six attributes (including price) with four levels each (all of the products were available in the local market where the experiment was conducted): a drink (a can of Coke, a can of Fanta, a box of apple juice, or a box of Chocomel), a mini candy (a mini M&Ms, Snickers, Toblerone, or Mars), instant soup (tomato, chicken, goulash, or pea), sandwich spread (apple, honey, chocolate paste, or peanut butter), a healthy snack (Liga evergreen, Sultana, Mueslibar, or Fruitbar), and price (1.50, 2.00, 2.50, or 3.00 euros). Figure 2.1 provides a screenshot of the feature description provided to respondents.

Figure 2.1: Screenshot of the feature description provided to respondents.

Our experiment followed a between-subject design with five conditions that varied the probability that the respondent's choice would be realized from 0 to 1. In the *prob* = 0 condition, respondents received 3 euros for participation and did not received a snack box. In the *prob* = 0.01, 0.50, and 0.99 conditions, the respondent received his or her preferred alternative with probability 0.01, 0.50, and 0.99 respectively. Finally, in the *prob* = 1 condition, the respondent received the preferred alternative with probability 1. Whenever a respondent received a snack box, she or he also received the difference between 3 euros and the price of the snack box. Therefore, the total payment was held constant at 3 euros for all conditions.

Set-up

Respondents were recruited at a large European university. They participated in the survey in the university's behavioral lab while monitored by a free-standing, nonintrusive Tobii® 2150 eye tracker sampling infrared corneal reflections at 50Hz with a $0.35°$ spatial resolution and accuracy of $0.5°$. The stimuli were presented on a 21-inch LCD monitor with a display resolution of 1,600 by 1,200 pixels. The positions of the left and right eyes were recorded separately (van der Lans et al., 2011).

Procedure

All respondents who took the survey used the online platform developed by the author. The experiment consisted of the following steps. First, the online platform randomly assigned each respondent to one of the five conditions ($prob$ = 0, 0.01, 0.5, 0.99, and 1). Instructions were then provided and respondents completed a short quiz to ensure that they understood the instructions. A practice question followed involving one choice task with eight alternatives, followed by the main task, which consisted of another choice task with eight alternatives. The alternatives in each task were chosen randomly (once for all respondents; all respondents saw the same set of alternatives) subject to the constraint that each level of each attribute should be present in at least one of the alternatives. Figure 2.2 shows the interface in the main task. The descriptions of the eight alternatives were presented in a matrix format. We used text rather than pictures to eliminate any effect of longer fixation durations due to color contrast (Wedel and Pieters, 2008). We also made sure that the text in each cell of the matrix was vertically and horizontally distant enough from the text in the other cells so that each fixation only enabled respondents to identify one piece of information within the region of interest (ROI). Underwood and McConkie (1985) showed that the area from which words can be identified in a given fixation generally does not exceed seven or eight letter spaces to the right of the fixation. Pollatsek et al. (1993) conducted two experiments found no evidence that readers obtain semantic information from below the line or text. After completion of the main task, the reward received by each respondent was determined by the online system according to the incentive scheme in that condition and the respondent's choice. The reward was distributed to the respondent upon leaving the lab.

Dependent Variables

Response times for each respondent in each question were collected automatically in the database using time stamps and used as one measure of the amount of attention respondents paid to each question. Response time has long been studied as a variable associated with aspects of memory, attitudes, and decision-making (Aaker et al., 1980; Fazio and Olson, 2003; MacLachlan and Myers, 1983; MacLachlan et al., 1979; Tyebjee, 1979). Although response times can account for multiple latent factors such as deliberation, accessibility of memory, and conflict experience, it is one (noisy) way to measure attention.

Our second set of dependent variables came from the eye-tracking data. Eye tracking has become increasingly popular as a tool for directly measuring attention and involvement in various marketing contexts (Chandon et al., 2009; Pieters et al., 1999, 2002; Pieters and Warlop, 1999; Pieters and Wedel, 2004; Rosbergen et al., 1997; van der Lans et al., 2008a,b; Wedel and Pieters, 2000, 2008; Wedel et al., 2008). Eye tracking has also been used in the literature on preference measurement, albeit to address different questions (Musalem et al., 2013; Shi et al., 2013; Stüttgen et al., 2012; Toubia et al., 2012; Yang et al., forthcoming).

Eye-tracking data are composed of fixations and saccades based on eye-fixation duration (Wedel and Pieters, 2000). Fixations represent time periods when respondents fix their eyesight on a specific place; saccades represent eye movements between two fixations. In our study, fixations and saccades were differentiated by van der Lans et al.'s (2011) velocity-based algorithm. Data on *consecutive* fixations in the same cell were collapsed to one fixation since they were likely caused by respondents randomly moving their eyes in a very small range while blinking (nonconsecutive fixations in the same cell were recorded as distinct fixations).

We defined the ROI for each piece of information as the area within the boundary of a cell in a matrix of the number of attributes $\times$ the number of alternatives containing the choice-relevant information (see Figure 2.2).

Figure 2.2: Screenshot from main task.

MAIN QUESTION

Please indicate your favorite product from the set below.
You will receive 3 euros as compensation for participating in this study. With probability 1%, this money will be used to purchase your preferred product from the main question. In other words, with probability 99% you will receive 3 euros in cash, and with probability 1% you will receive your preferred product from the main question plus the difference between 3.00 euro and the price of this product.

	A	B	C	D	E	F	G	H
Drink	Apple juice	Chocomel	Coke	Coke	Coke	Chocomel	Apple juice	Fanta
Mini Candy	M&M's	Snickers	Mars	Snickers	M&M's	Toblerone	Mars	Snickers
Instant Soup	Pea	Pea	Pea	Goulash	Chicken	Pea	Tomato	Tomato
Sandwich Spread	Peanut butter	Honey	Apple spread	Chocolate paste	Apple spread	Chocolate paste	Chocolate paste	Peanut butter
Healthy Snack	Liga evergreen	Fruitbar	Sultana	Sultana	Sultana	Fruitbar	Sultana	Mueslibar
Price	2.00 euro	2.50 euro	1.50 euro	2.00 euro	1.50 euro	1.50 euro	3.00 euro	1.50 euro
	○	○	○	○	○	○	○	○

submit

Our basic measure of attention for the eye-tracking data was the proportion of information processed by each respondent, i.e., the proportion of cells for which at least one eye fixation was observed. Similar results were obtained when taking multiple visits to the same cell and/or the fixation duration into account (details are presented in Appendix A.3).

Finally, another dependent variable was the choice of snack box made by each respondent and the choice share of each alternative.

2.4 Results

We collected complete data on 86 respondents[1] across the five conditions. We now analyze our response-time and eye-tracking data and the choices made by these respondents.

2.4.1 Response Time

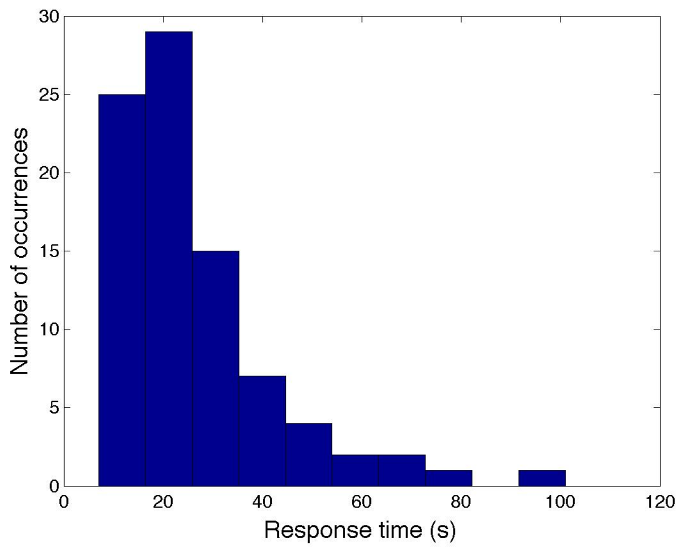
The average response time on the main choice task for all conditions was 26.87 seconds. Figure 2.3 shows the distribution of the response times. Figure 2.4 shows the average response time in each condition (as a function of the probability ($prob$) that the choice will be realized). We find a U-shaped relationship between this probability and response time; respondents tended to spend more time on questions with deterministic incentives ($prob$ = 1) and no incentives ($prob$ = 0) than on questions with probabilistic incentives ($prob$ = 0.01, 0.5, 0.99). Moreover, there appears to be little difference in the response times for the three probabilistic conditions ($prob$ = 0.01, 0.5 and 0.99). Since the distribution of response time was skewed, we compared response times across conditions using nonparametric ranksum tests. None of the pairwise differences between the probabilistic conditions ($prob$ = 0.01, 0.5 and 0.99) was statistically significant.[2] Therefore, we

[1] We recruited 120 respondents to participate in our study. We first looked at the raw eye-fixation data from each respondent during the study and detected time stamps without any affiliation of eye-fixation position as missing data. The respondents with missing data were also confirmed through a video of the choice experiment interface mapped with eye-tracking data during the study. We found that 34 of the 120 respondents the eye fixations were not recorded properly and excluded their eye tracking data. To have a consistent sample in our analysis for attention, information processing, and choices, we also excluded analysis of the response-time and choice data for those 34 respondents even though they were properly recorded.

[2] The median of the response time under $prob$ = 0.01, 0.5, and 0.99 is 23.5, 18.0, and 19.0 seconds respectively. The p-values were 0.32, 0.63, and 0.99 for ranksum tests for between $prob$ = 0.01 vs. $prob$ = 0.5, $prob$ = 0.01 vs.

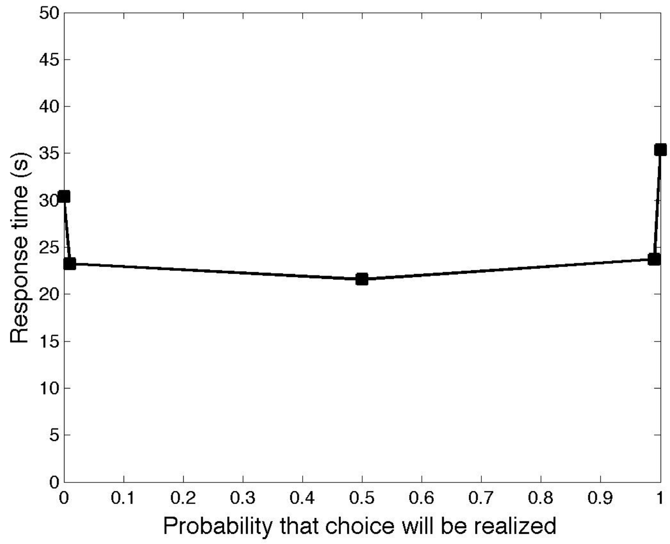
group these three conditions and compared the median response time across the conditions (19.0 seconds) to the median response times under $prob = 0$ (28.0 seconds) and $prob = 1$ (25.0 seconds). We find that respondents spent significantly less time answering a question when the incentives were probabilistic (compared to when there was no incentive ($p\text{-}value < 0.03$) and when the incentives were deterministic ($p\text{-}value < 0.04$)). Therefore, the analysis of response time confirms our prediction that choices associated with probabilistic incentives generate less attention than choices associated with deterministic incentives. Moreover, we find evidence that questions involving probabilistic incentives may generate less attention than purely hypothetical questions in the type of choice experiments used in marketing. The difference between the no-incentive condition and the deterministic incentive condition is not significant ($p\text{-}value = 0.64$).

Figure 2.3: Distribution of response times across all respondents.



$prob = 0.99$, and $prob = 0.5$ vs. $prob = 0.99$ respectively.

Figure 2.4: Average response time vs. probability that choice will be realized.



2.4.2 Eye Tracking

We now turn to eye-tracking data as a more direct measurement of attention. The average proportion of information visited across respondents was 53.08%. Figure 2.5 shows the distribution of the proportion of information visited for all respondents. Figure 2.6 shows the distribution of the number of visits (i.e., fixation) per piece of information (each piece of information is one cell that contains the level of one attribute for one alternative) across all pieces of information and respondents. That is, the bars in the histogram in Figure 2.6 add up to the number of respondents multiplied by the number of pieces of information. These two figures confirm earlier findings that a substantial proportion of choice-relevant information is not processed in preference measurement tasks and that ROIs with at least one fixation from a respondent are likely to have multiple fixations by the same respondent (Toubia et al., 2012; Yang et al., forthcoming).

Figure 2.5: Distribution of proportion of information visited across all respondents.

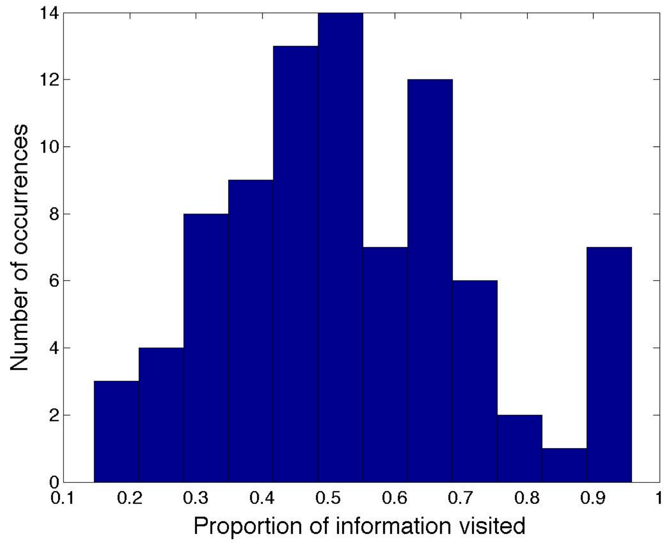
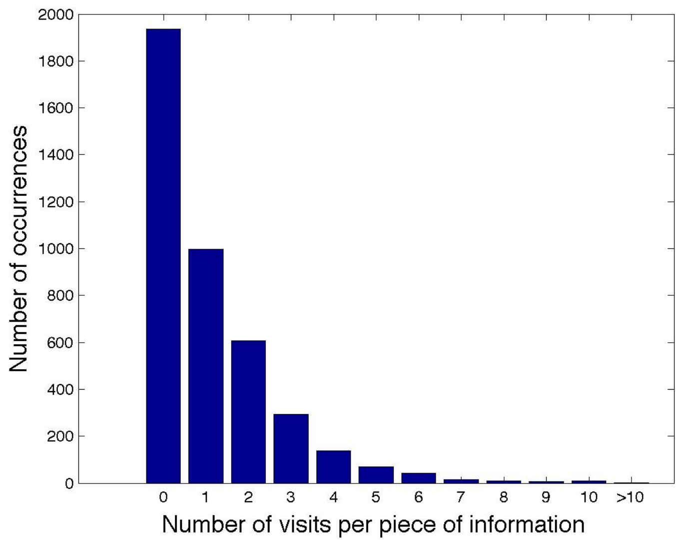
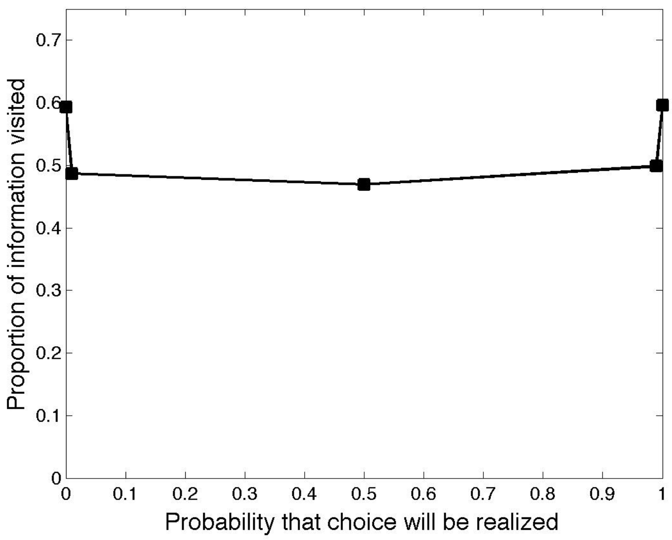


Figure 2.6: Distribution of the number of visits per piece of information.



In Figure 2.7 we plot the average proportion of information visited as a function of the probability (*prob*) that the choice will be realized. As with response time, the relationship between *prob* and the proportion of information visited is U-shaped. Respondents processed less information in probabilistic incentive-aligned questions than they did when there was no incentive and when the incentive was deterministic. We compared the proportion of information visited across conditions using t-tests (similar results were obtained with nonparametric tests but t-tests seemed appropriate because the distribution of the proportion of information visited is close to normal). We find no significant difference in the proportion of information visited between the three probabilistic conditions ($prob$ = 0.01, 0.5, and 0.99) and therefore group these conditions together.[3] The average proportion of information visited under no incentive, probabilistic incentives, and deterministic incentives was 0.59, 0.49, and 0.60 respectively. As with response time, the proportion of information visited is significantly lower in the probabilistic conditions than in the deterministic condition (p-$value$ < 0.03) and in the no-incentive condition (p-$value$ < 0.04). Again, there is no significant difference in the proportion of information visited between the no-incentive and the deterministic conditions (p-$value$ = 0.97). Similar results were obtained when using the total number of fixations per question as the measure of attention or the total fixation duration (see Appendix A.3).

Figure 2.7: Average proportion of information visited vs. probability that choice will be realized.



Therefore, the results using eye-tracking data confirm our results from the response-time data. Eye-tracking data allow us to take our analysis one step further by providing information on *which*

[3]The mean of proportion of information visited is 0.49, 0.47, and 0.50 when $prob$ = 0.01, 0.5, and 0.99 respectively. p-$value$ = 0.77, 0.85, and 0.61 for the comparison of $prob$ = 0.01 vs. $prob$ = 0.5, $prob$ = 0.01 vs. $prob$ = 0.99, and $prob$ = 0.5 vs. $prob$ = 0.99 respectively.

information was processed by respondents in different conditions. More precisely, we can study whether probabilistic incentives only reduce the overall level of attention or also distort how consumers process information. For example, we can study the impact of incentives on the number of attributes processed by subjects and on the number of alternatives processed.

Before analyzing the impact of incentives on the number of attributes and alternatives processed by respondents, we explore whether there is evidence for attribute-based and alternative-based processing in our data. Figure 2.8 plots the distribution of the number of attributes visited (at least one fixation) per alternative across all alternatives and respondents. Figure 2.9 plots the number of alternatives visited per attribute across all attributes and respondents. These figure suggests that consumers process the information presented to them using a mixture of attribute-based and alternative-based processing. Indeed, pure attribute-based searches would lead to some attributes not being visited at all, and pure alternative-based searches would lead to some alternatives not being visited at all. We find that an alternative (attribute) is completely ignored only 6.83% (1.16%) of the time.

Figure 2.8: Distribution of the number of attributes visited per alternative (across all alternatives and respondents).

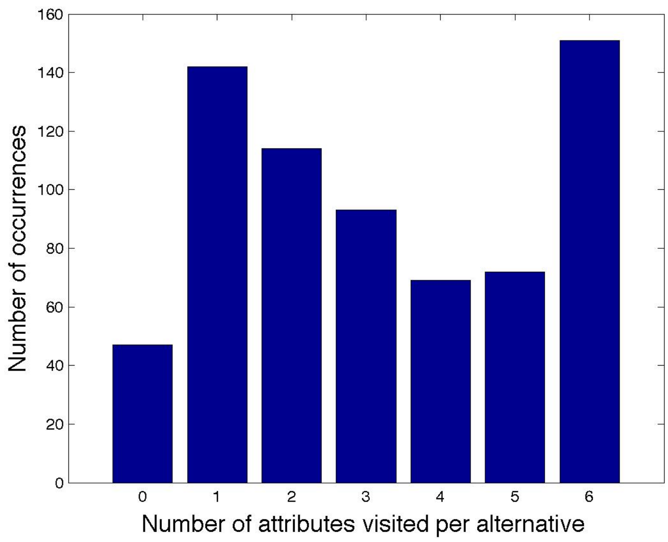
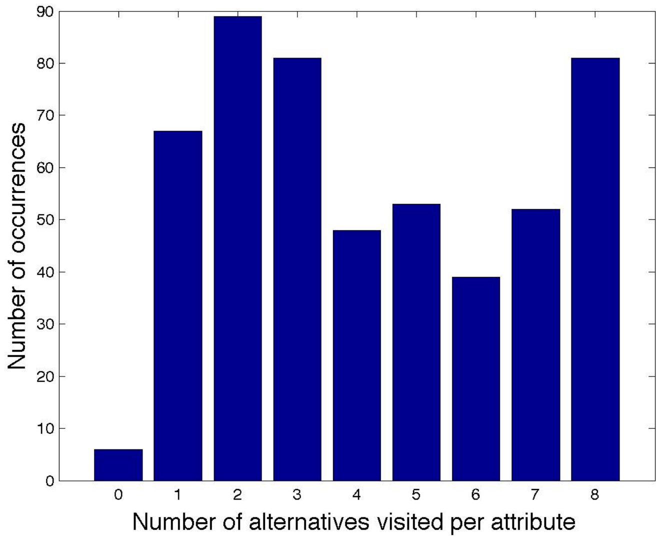


Figure 2.9: Distribution of the number of alternatives visited per attribute (across all attributes and respondents).



Based on the evidence just presented, we analyzed the impact of incentives on the number of attributes and on the number of alternatives processed by respondents using two multilevel mixed-effects ordered logistic regressions. In the first, the dependent variable was the number of alternatives visited in each attribute by each respondent (i.e., the number of observations in the regression was the number of respondents $\times$ the number of attributes). In the second, the dependent variable was the number of attributes visited in each alternative by each respondent (i.e., the number of observations in the regression was the number of respondents $\times$ the number of alternatives). In each case, the ordered logistic regression controlled for the probability conditions with two dummy variables associated with $0 < prob < 1$ (i.e., $prob$ = 0.01, 0.5, 0.99) and $prob = 1$ using $prob = 0$ as the baseline. We also added fixed effects for the six attributes in the first regression and the eight alternatives in the second. We took the individual effects as random and assumed that the individual effects follow identical independent normal distribution. The results are presented in Table 2.1.

Table 2.1: Number of alternatives (attributes) visited in an attribute (alternative) vs. probability that choice will be realized - mixed effects ordered logistic regression results.

(a) Number of alternatives visited in an attribute

Threshold 0-1	-9.01*
Threshold 1-2	-5.88*
Threshold 2-3	-4.35*
Threshold 3-4	-3.24*
Threshold 4-5	-2.59*
Threshold 5-6	-1.85*
Threshold 6-7	-1.24*
Threshold 7-8	-0.24
$0 < prob < 1$	-0.95*
$prob = 1$	0.01
Attribute 2	-1.81*
Attribute 3	-2.62*
Attribute 4	-3.51*
Attribute 5	-3.65*
Attribute 6	-3.16*
Random effects of individual (*stdev*)	1.52

(b) Number of attributes visited in an alternative

Threshold 0-1	-3.94*
Threshold 1-2	-1.97*
Threshold 2-3	-1.04*
Threshold 3-4	-0.35
Threshold 4-5	0.18
Threshold 5-6	0.82*
$0 < prob < 1$	-0.71*
$prob = 1$	0.05
Alternative 2	-0.39
Alternative 3	-0.37
Alternative 4	-0.22
Alternative 5	-0.32
Alternative 6	-0.22
Alternative 7	-0.91*
Alternative 8	-0.96*
Random effects of individual (*stdev*)	1.07

Note: * denotes a coefficient with *p-value* < 0.05.

We find that respondents processed significantly fewer alternatives per attribute and fewer attributes per alternative under probabilistic incentives than under no incentives (both *p-values* $<$ 0.05) and deterministic incentives (both *p-values* < 0.06). There is no significant difference in the number of alternatives visited per attribute or in the number of attributes visited per alternative for no incentive and deterministic incentives (both *p-values* > 0.90). This suggests that incentives influenced both the number of alternatives visited per attribute and the number of attributes visited per alternative in the same way that they influenced the total amount of information processed and response times.

We can also analyze the eye-tracking data at the level of each cell in the matrix containing the choice-relevant information (see Figure 2.2). Table 2.2 shows the share of fixations for each cell of the matrix across all conditions. For example, 0.03 in the first cell in Table 2.2 means that 3% of the fixations in the $prob = 0$ condition were made to the cell that contained the drink attribute for the first alternative. We compared the fixation shares for all conditions using chi-square tests of independence. We find that all pairwise comparisons are significant (p-$value$ $<$ 0.03) *except* the comparison between $prob = 0$ and $prob = 1$ (p-$value = 0.94$). This suggests that no incentive and deterministic-incentive alignments (the extreme conditions) not only raise the same level of attention but also lead to a similar spreading of attention across all of the choice-relevant information. Probabilistic incentives lead to different information processing patterns.[4]

[4]We also looked into the shares of fixation duration for the 48 pieces of information across different conditions and reached a similar conclusion. We applied a different statistical test due to the nature of the fixation-duration data. The results are available from the author upon request.

Table 2.2: Share of fixations in each piece of information across conditions.

		Probability that choice realized				
Alternatives	**Attributes**	0	0.01	0.5	0.99	1
Alternative 1	Drinks	0.03	0.04	0.05	0.04	0.04
	Candy	0.02	0.03	0.04	0.03	0.02
	Instant Soup	0.02	0.02	0.03	0.02	0.02
	Spread	0.01	0.02	0.02	0.02	0.02
	Healthy Snack	0.02	0.02	0.03	0.02	0.02
	Price	0.02	0.02	0.02	0.02	0.02
Alternative 2	Drinks	0.04	0.04	0.05	0.04	0.03
	Candy	0.02	0.02	0.03	0.02	0.02
	Instant Soup	0.01	0.02	0.01	0.02	0.01
	Spread	0.01	0.01	0.01	0.02	0.01
	Healthy Snack	0.01	0.00	0.02	0.02	0.01
	Price	0.01	0.01	0.01	0.02	0.01
Alternative 3	Drinks	0.04	0.05	0.05	0.03	0.04
	Candy	0.03	0.03	0.04	0.02	0.03
	Instant Soup	0.02	0.02	0.02	0.03	0.02
	Spread	0.02	0.02	0.00	0.01	0.02
	Healthy Snack	0.01	0.01	0.01	0.01	0.01
	Price	0.02	0.00	0.02	0.02	0.02
Alternative 4	Drinks	0.03	0.04	0.05	0.03	0.03
	Candy	0.03	0.02	0.04	0.03	0.03
	Instant Soup	0.02	0.03	0.03	0.02	0.01
	Spread	0.01	0.01	0.02	0.02	0.01
	Healthy Snack	0.02	0.01	0.01	0.01	0.01
	Price	0.03	0.01	0.02	0.03	0.02
Alternative 5	Drinks	0.04	0.04	0.04	0.02	0.04
	Candy	0.03	0.02	0.03	0.02	0.04
	Instant Soup	0.03	0.03	0.02	0.02	0.02
	Spread	0.02	0.02	0.02	0.01	0.02
	Healthy Snack	0.02	0.01	0.01	0.01	0.02
	Price	0.04	0.01	0.01	0.02	0.02
Alternative 6	Drinks	0.03	0.04	0.03	0.02	0.04
	Candy	0.02	0.02	0.02	0.03	0.03
	Instant Soup	0.02	0.02	0.02	0.03	0.02
	Spread	0.02	0.02	0.01	0.02	0.02
	Healthy Snack	0.01	0.02	0.01	0.02	0.03
	Price	0.02	0.02	0.01	0.03	0.03
Alternative 7	Drinks	0.03	0.03	0.03	0.02	0.02
	Candy	0.01	0.02	0.02	0.01	0.02
	Instant Soup	0.01	0.02	0.02	0.02	0.01
	Spread	0.01	0.02	0.02	0.01	0.01
	Healthy Snack	0.01	0.01	0.02	0.01	0.01
	Price	0.02	0.01	0.01	0.02	0.02
Alternative 8	Drinks	0.02	0.02	0.02	0.03	0.02
	Candy	0.01	0.03	0.01	0.02	0.02
	Instant Soup	0.02	0.03	0.01	0.02	0.02
	Spread	0.01	0.01	0.01	0.01	0.01
	Healthy Snack	0.01	0.02	0.01	0.01	0.01
	Price	0.02	0.01	0.01	0.02	0.01

2.4.3 Choice Shares

Our analysis of the response-time and eye-tracking data documents that respondents in our preference measurement task paid less attention under probabilistic incentives than under deterministic incentives and no incentive. Moreover, both the amount of attention and the allocation of attention across the choice-relevant information were similar for no incentive and deterministic incentives which begs a question. If purely hypothetical (i.e., not incentive-aligned) choice experiments and deterministically incentive-aligned choice experiments induce similar information processing, are hypothetical choice experiments best suited to predict real-life choices? To answer this question, we now turn to a comparison of choice shares across conditions.

Table 2.3 reports choice shares for the eight alternatives (presented in Figure 2.2) across all conditions. We compared the shares across conditions using Spearman's rank correlation, as reported in Table 2.4. We find that the shares in the *prob*=0 condition are not highly correlated with the shares in the *prob* = 1 condition ($\rho = 0.40$, *p-value* = 0.32). We also see that the shares in the probabilistic conditions are more highly correlated with those in the deterministic condition: *prob*=0.01 and *prob*=0.99 are both significantly correlated with *prob* = 1 ($\rho = 0.98$ and 0.79 respectively, *p-value* < 0.01 and 0.03 respectively) which *prob* = 0.5 has a lower correlation ($\rho = 0.57$, *p-value* = 0.14).

Table 2.3: Choice shares of different alternatives across conditions.

	Attribute						**Choice share when *prob*=**				
Alternative	Drink	Candy	Instant Soup	Spread	Healthy Snack	Price	0	0.01	0.5	0.99	1
1	Apple Juice	M&M'S	Pea	Peanut Butter	Liga Evergreen	2.00 Euro	0.14	0.19	0.18	0.18	0.20
2	Chocomel	Snickers	Pea	Honey	Fruitbar	2.50 Euro	0.05	0.00	0.06	0.00	0.00
3	Coke	Mars	Pea	Apple Spread	Sultana	1.50 Euro	0.14	0.00	0.06	0.06	0.00
4	Coke	Snickers	Goulash	Chocolate Paste	Sultana	2.00 Euro	0.05	0.06	0.12	0.18	0.07
5	Coke	M&M'S	Chicken	Apple Spread	Sultana	1.50 Euro	0.38	0.19	0.24	0.12	0.20
6	Chocomel	Toblerone	Pea	Chocolate Paste	Fruitbar	1.50 Euro	0.10	0.25	0.12	0.24	0.33
7	Apple Juice	Mars	Tomato	Chocolate Paste	Sultana	3.00 Euro	0.00	0.13	0.18	0.12	0.07
8	Fanta	Snickers	Tomato	Peanut Butter	Mueslibar	1.50 Euro	0.14	0.19	0.06	0.12	0.13

Table 2.4: Spearman's rank correlations between choice shares across conditions.

		Probability that choice will be realized				
		0	0.01	0.5	0.99	1
Probability that choice will be realized	0	1.00	0.36	0.15	-0.01	0.40
	0.01		1.00	0.47	0.73*	0.98*
	0.5			1.00	0.45	0.57
	0.99				1.00	0.79*
	1					1.00

Note: * denotes a correlation with *p-value* < 0.05.

Therefore, our analysis of the choice data suggests that hypothetical and deterministic incentive questions induce similar information processing but that consumers tend to choose differently *given the information they process* when questions are hypothetical. Therefore, we cannot conclude that hypothetical questions lead to the same choices as deterministically-aligned questions or that hypothetical questions should be used in practice. The source of the difference in the choice shares despite the similarities in information processing is a topic for further research. Memory encoding in hypothetical and incentive-aligned choices may be different, leading to different memory traces despite similar fixations (Wedel and Pieters, 2000). Also, the effort spent by consumers to compute their preferences and reach a decision given the information processed may vary as a function of incentives. According to the psychology literature, there should be variations in how choices are construed as a function of incentives. Indeed, hypothetical choices should be more distant psychologically and thus be construed at a higher level than incentive-aligned choices (Trope and Liberman, 2010). Intuitively, consumers may choose based on "abstract" preferences in hypothetical choices (e.g., "I like Coke") and use more concrete considerations when choices are incentive-aligned (e.g., "right now I feel like having a Fanta"). Finally, social desirability has been shown to bias hypothetical choices (Ding et al., 2005).

Regarding the comparison of choice shares for probabilistic and deterministic incentives, our evidence is mixed. Two of the three probabilistic conditions have a significant correlation with the deterministic condition. The null effect found in the other probabilistic condition may be a consequence of our sample size, and we hope that future research will provide additional comparisons of choice shares under probabilistic and deterministic incentives.

2.5 Conclusions

In this chapter, we explore the relationship between incentives, attention, information processing and choice in the context of choice experiments. Four features make our study unique: (*i*) we compare purely hypothetical, probabilistic, and deterministic incentives; (*ii*) we jointly study the impact of incentives on attention, information processing, and choice; (*iii*) we use eye-tracking data as a direct measure of information processing; and (*iv*) we use the typical marketing context of preference measurement.

We find a U-shaped relationship between the probability that a choice will be realized and the level of attention. Purely hypothetical choices ($prob = 0$) and deterministic choices ($prob = 1$) generate similar levels of attention and information processing. Choices that are to be realized with probabilities between 0.01 and 0.99 generate levels of attention that are similar to each other but smaller than the extreme cases. This result shows that incentive-alignment as it is typically implemented in choice experiments (using *probabilistic* incentives) does not motivate respondents to process as much information as they would in *deterministic* choices. While the utility related to processing choice-relevant information is the same irrespective of the probability that the choice will be realized, the expected utility derived from the chosen alternative is less in probabilistically incentive-aligned choices. As a result, boundedly rational consumers tend to process less information in probabilistically incentive-aligned choices than in deterministically incentive-aligned choices. In other words, although probabilistic incentives are enough to induce truth-telling, they do not appear to be enough to induce consumers to process choice-relevant information as carefully as they would if the choices were real. On the other hand, when incentives are absent, consumers appear to be intrinsically motivated to exert effort, leading to levels of attention that are similar to those in deterministic choices. We further provide evidence that attention was allocated similarly for no incentive and deterministic incentives. However, the choices that consumers made were different when there were no incentives.

Probabilistically incentive-aligned choice experiments are considered the gold standard by practitioners and researchers. Therefore, our findings have broad implications for theory and practice since they suggest that typical incentive-aligned choice experiments do not induce consumers to behave like they would when making real-life decisions. Hypothetical choice questions may increase the level of attention but give rise to different answers and thus they do not offer a viable alternative. Deterministically incentive-aligned choice questions induce biases when multiple questions of a similar type must be asked and are impractical and expensive. Thus hypothetical and deterministic choice questions may not be viable substitutes for probabilistic choice questions in practice. Nonetheless, our results raise serious questions about the accuracy of predictions based on probabilistic questions.

We argue that at least two types of solutions can be found in the recent literature and developed further in future research. The first would be to model the information acquisition process formally as a function of incentives, calibrate such a model using probabilistically incentive-aligned experiments, and predict real-life choices using counterfactual simulations. The model proposed by Yang et al. (forthcoming) may provide a building block for such an endeavor. However, our findings suggest that choices are not driven solely by information processing since we find that similar information processing can lead to different choices. Therefore, modeling information processing as a function of incentives may not be adequate to predict choices under various incentive schemes. A second potential approach is to develop choice experiments that increase respondents' intrinsic motivation to provide thoughtful information. This might be achieved through gamification. Some examples previously mentioned include Ding et al. (2009), Park et al. (2008), and Toubia et al. (2012).

Bibliography

Bibliography

Aaker, D. A., R. P. Bagozzi, J. M. Carman, J. M. MacLachlan. 1980. On using response latency to measure preference. *Journal of Marketing Research* **17**(2) 237–244.

Allenby, G. M., N. Arora, J. L. Ginter. 1995. Incorporating prior knowledge into the analysis of conjoint studies. *Journal of Marketing Research* **32**(2) 152–162.

Arkes, H. R., R. M. Dawes, C. Christensen. 1986. Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes* **37**(1) 93–110.

Assuncao, J. L., R. J. Meyer. 1993. The rational effect of price promotions on sales and consumption. *Management Science* **39**(5) 517–535.

Atchadé, Y. F., J. S. Rosenthal. 2005. On adaptive Markov chain Monte Carlo algorithms. *Bernoulli* **11**(5) 815–828.

Beattie, J., G. Loomes. 1997. The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty* **14**(2) 155–168.

Camerer, C. F., R. M. Hogarth. 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* **19**(1-3) 7–42.

Celeux, Gilles, Florence Forbes, Christian P Robert, D Michael Titterington, et al. 2006. Deviance information criteria for missing data models. *Bayesian Analysis* **1**(4) 651–673.

Chandon, P., J. Hutchinson, E. T. Bradlow, S. Young. 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. *Journal of Marketing* **73**(6) 1–17.

Ching, A. T., S. Imai, M. Ishihara, N. Jain. 2012. A practitioner's guide to Bayesian estimation of discrete choice dynamic programming models. *Quantitative Marketing and Economics* **10**(2) 1–46.

Chintagunta, Pradeep K, Ronald L Goettler, Minki Kim. 2012. New drug diffusion when forward-looking physicians learn from patient feedback and detailing. *Journal of Marketing Research* **49**(6) 807–821.

Ding, M. 2007. An incentive-aligned mechanism for conjoint analysis. *Journal of Marketing Research* **44**(2) 214–223.

Ding, M., R. Grewal, J. Liechty. 2005. Incentive-aligned conjoint analysis. *Journal of Marketing Research* **42**(1) 67–82.

Ding, M., Y. H. Park, E. T. Bradlow. 2009. Barter markets for conjoint analysis. *Management Science* **55**(6) 1003.

Dong, S., M. Ding, J. Huber. 2010. A simple mechanism to incentive-align conjoint experiments. *International Journal of Research in Marketing* **27**(1) 25–32.

Dube, J. P. H., G. J. Hitsch, P. Jindal. 2012. The joint identification of utility and discount functions from stated choice data: An application to durable goods adoption. *NBER working paper* .

Fazio, R. H., M. A. Olson. 2003. Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology* **54**(1) 297–327.

Gabaix, X., D. Laibson, G. Moloche, S. Weinberg. 2006. Costly information acquisition: Experimental analysis of a boundedly rational model. *The American Economic Review* **96**(4) 1043–1068.

Gelman, A., X. Meng, H. Stern. 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* **6**(4) 733–759.

Gilbride, T. J., G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science* **23**(3) 391–406.

Gilbride, T. J., G. M. Allenby. 2006. Estimating heterogeneous EBA and economic screening rule choice models. *Marketing Science* **25**(5) 494–509.

Gittins, J. C. 1979. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)* **41**(2) 148–177.

Gneezy, U., A. Rustichini. 2000. Pay enough or don't pay at all. *The Quarterly Journal of Economics* **115**(3) 791–810.

Hagerty, M. R., D. A. Aaker. 1984. A normative model of consumer information processing. *Marketing Science* **3**(3) 227–246.

Harrison, G. W. 1994. *Expected utility theory and the experimentalists*. Springer.

Hartmann, W. R., H. S. Nair. 2010. Retail competition and the dynamics of demand for tied goods. *Marketing Science* **29**(2) 366–386.

Hauser, J. R., G. L. Urban, B. D. Weinberg. 1993. How consumers allocate their time when searching for information. *Journal of Marketing Research* **30**(4) 452–466.

Heyman, J., D. Ariely. 2004. Effort for payment a tale of two markets. *Psychological Science* **15**(11) 787–793.

Huang, G., A. Khwaja, K. Sudhir. 2012. Short run needs and long term goals: A dynamic model of thirst management. *Working paper* .

Hutchinson, J. W., R. J. Meyer. 1994. Dynamic decision making: Optimal policies and actual behavior in sequential choice problems. *Marketing Letters* **5**(4) 369–382.

Jedidi, K., R. Kohli. 2005. Probabilistic subset-conjunctive models for heterogeneous consumers. *Journal of Marketing Research* **42**(4) 483–494.

Jedidi, K., R. Kohli. 2008. Inferring latent class lexicographic rules from choice data. *Journal of Mathematical Psychology* **52**(4) 241–249.

Jenkins, G. D., Jr, A. Mitra, N. Gupta, J. D. Shaw. 1998. Are financial incentives related to performance? a meta-analytic review of empirical research. *Journal of Applied Psychology* **83**(5) 777.

Johnson, E. J., R. J. Meyer, S. Ghose. 1989. When choice models fail: Compensatory models in negatively correlated environments. *Journal of Marketing Research* **26**(3) 255–270.

Johnson, R. 2008. A perspective on adaptive CBC (what can we expect from respondents?). *Sawtooth Software, Research Paper Series, Sequim, WA.* www.sawtoothsoftware.com/download/techpap/acbcperspective.pdf.

Kahneman, D., W. S. Peavler. 1969. Incentive effects and pupillary changes in association learning. *Journal of Experimental Psychology* **79**(2p1) 312.

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

Liechty, J. C., D. Fong, W. S. DeSarbo. 2005. Dynamic models incorporating individual heterogeneity: Utility evolution in conjoint analysis. *Marketing Science* **24**(2) 285–293.

MacLachlan, J., J. Czepiel, P. LaBarbera. 1979. Implementation of response latency measures. *Journal of Marketing Research* **16**(4) 573–577.

MacLachlan, J., J. G. Myers. 1983. Using response latency to identify commercials that motivate. *Journal of Advertising Research* **23**(5) 51.

McGraw, K. O., J. C. McCullers. 1979. Evidence of a detrimental effect of extrinsic incentives on breaking a mental set. *Journal of Experimental Social Psychology* **15**(3) 285–294.

Meyer, R. J. 1982. A descriptive model of consumer information search behavior. *Marketing Science* **1**(1) 93–121.

Misra, S., H. S. Nair. 2011. A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quantitative Marketing and Economics* **9**(3) 211–257.

Musalem, A., M. Meißner, J. Huber. 2013. Do motivated and incidental processing distort conjoint choices? *Working paper* .

Netzer, O., O. Toubia, E. T. Bradlow, E. Dahan, T. Evgeniou, F. M. Feinberg, E. M. Feit, S. K. Hui, J. Johnson, J. C. Liechty. 2008. Beyond conjoint analysis: Advances in preference measurement. *Marketing Letters* **19**(3) 337–354.

Park, Y. H., M. Ding, V. R. Rao. 2008. Eliciting preference for complex products: a web-based upgrading method. *Journal of Marketing Research* **45**(5) 562–574.

Payne, J. W., J. R. Bettman, E. J. Johnson. 1988. Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**(3) 534.

Payne, J. W., J. R. Bettman, E. J. Johnson. 1992. Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology* **43**(1) 87–131.

Payne, J. W., J. R. Bettman, E. J. Johnson. 1993. *The adaptive decision maker*. Cambridge University Press.

Pieters, R., E. Rosbergen, M. Wedel. 1999. Visual attention to repeated print advertising: A test of scanpath theory. *Journal of Marketing Research* **36**(4) 424–438.

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

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

Pieters, R., M. Wedel. 2004. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing* **68**(2) 36–50.

Pollatsek, A., G. E Raney, L. Lagasse, K. Rayner. 1993. The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* **47**(2) 179.

Rosbergen, E., R. Pieters, M. Wedel. 1997. Visual attention to advertising: A segment-level analysis. *Journal of Consumer Research* **24**(3) 305–314.

Rust, J. 1987. Optimal replacement of gmc bus engines: An empirical model of Harold Zurcher. *Econometrica* **55**(5) 999–1033.

Shi, S. W., M. Wedel, R. Pieters. 2013. Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Management Science* **59**(5) 1009–1026.

Simon, H. A. 1955. A behavioral model of rational choice. *The Quarterly Journal of Economics* **69**(1) 99–118.

Smith, V. L., J. M. Walker. 1993. Monetary rewards and decision cost in experimental economics. *Economic Inquiry* **31**(2) 245–261.

Stone, D. N., D. A. Schkade. 1994. Effects of attribute scales on process and performance in multiattribute choice. *Organizational Behavior and Human Decision Processes* **59**(2) 261–287.

Stüttgen, P., P. Boatwright, R. T. Monroe. 2012. A satisficing choice model. *Marketing Science* **31**(6) 878–899.

Teixeira, Thales, Michel Wedel, Rik Pieters. 2012. Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research* **49**(2) 144–159.

Toubia, O., M. G. de Jong, D. Stieger, J. Füller. 2012. Measuring consumer preferences using conjoint poker. *Marketing Science* **31**(1) 138–156.

Toubia, O., A. T. Stephen. 2013. Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter? *Marketing Science* **32**(3) 368–392.

Trope, Yaacov, Nira Liberman. 2010. Construal-level theory of psychological distance. *Psychological Review* **117**(2) 440.

Tversky, A., S. Sattath, P. Slovic. 1988. Contingent weighting in judgment and choice. *Psychological Review* **95**(3) 371.

Tyebjee, T. T. 1979. Response time, conflict, and involvement in brand choice. *Journal of Consumer Research* **6**(3) 295–304.

Underwood, N. R., G. W. McConkie. 1985. Perceptual span for letter distinctions during reading. *Reading Research Quarterly* **20**(2) 153–162.

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

van der Lans, R., R. Pieters, M. Wedel. 2008b. Eye-movement analysis of search effectiveness. *Journal of the American Statistical Association* **103**(482) 452–461.

van der Lans, R., M. Wedel, R. Pieters. 2011. Defining eye-fixation sequences across individuals and tasks: the binocular-individual threshold (bit) algorithm. *Behavior Research Methods* **43**(1) 239–257.

Wedel, M., R. Pieters. 2000. Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science* **19**(4) 297–312.

Wedel, M., R. Pieters. 2008. *Eye tracking for visual marketing*. Now Publishers Inc.

Wedel, M., R. Pieters, J. Liechty. 2008. Attention switching during scene perception: How goals influence the time course of eye movements across advertisements. *Journal of Experimental Psychology: Applied* **14**(2) 129.

Weitzman, M. L. 1979. Optimal search for the best alternative. *Econometrica* **47**(3) 641–654.

Wilcox, N. T. 1993. Lottery choice: Incentives, complexity and decision time. *The Economic Journal* **103**(421) 1397–1417.

Yang, C. L., O. Toubia, M. G. de Jong. forthcoming. A bounded rationality model of information search and choice in preference measurement . *Journal of Marketing Research* .

Yao, S., C. F. Mela. 2011. A dynamic model of sponsored search advertising. *Marketing Science* **30**(3) 447–468.

Yee, M., E. Dahan, J. R. Hauser, J. Orlin. 2007. Greedoid-based noncompensatory inference. *Marketing Science* **26**(4) 532–549.

ZDNet. 2012. Facebook acquires eye-tracking startup gazehawk URL `www.zdnet.com/blog/facebook/facebook-acquires-eye-tracking-startup-gazehawk/10218`.

Appendices

Appendix A

Appendices

A.1 Illustrative Example for Chapter 1

We illustrate our state variables and the computation of product-related utility, using a simple example. We assume one attribute (I=1) with three levels (L=3), and two alternatives per choice question (J=2). We assume that alternative 1 has attribute 1 at level 1 and alternative 2 has attribute 1 at level 2.

We have $I_1^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}$, and the partworths for the first attribute may be represented as: $\beta_1 = \begin{bmatrix} \beta_{11} \\ \beta_{12} \end{bmatrix}$ or $\widetilde{\beta}_1 = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ -\beta_{11} - \beta_{12} \end{bmatrix}$.

Before the first fixation the state variables have the following values:

-p=$\emptyset$

-$n_{1,1} = n_{1,2} = 0$

And we have the following:

-$w_{1,1} = w_{1,2} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$

-The expected product-related utility of alternative 1 is: $w_{1,1}\widetilde{\beta}_1 = 0$

-The expected product-related utility of alternative 2 is: $w_{1,2}\tilde{\beta}_1 = 0$

Suppose that the first fixation is to cell $(1,1)$ corresponding to attribute 1 of alternative 1. Then the state variables evolve to:

-p=$(1, 1)$

-$n_{1,1} = 1$

-$n_{1,2} = 0$

And we have the following:

-$w_{1,1} = [\frac{exp(\eta)}{2+exp(\eta)}, \frac{1}{2+exp(\eta)}, \frac{1}{2+exp(\eta)}]$

-$w_{1,2} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$

-The expected product-related utility of alternative 1 is now: $w_{1,1}\tilde{\beta_1} = \frac{exp(\eta)-1}{2+exp(\eta)}\beta_{11}$

-The expected product-related utility of alternative 2 is now: $w_{1,2}\tilde{\beta_1} = 0$

Suppose now that after $t = 15$, 10 fixations have been made to cell $(1, 1)$ and 5 fixations have been made to cell $(1, 2)$, and that the last fixation was on cell $(1, 2)$. Then we have:

-p=$(1, 2)$

-$n_{1,1} = 10$

-$n_{1,2} = 5$

And we have the following:

-$w_{1,1} = [\frac{exp(10\eta)}{2+exp(10\eta)}, \frac{1}{2+exp(10\eta)}, \frac{1}{2+exp(10\eta)}]$

-$w_{1,2} = [\frac{1}{2+exp(5\eta)}, \frac{exp(5\eta)}{2+exp(5\eta)}, \frac{1}{2+exp(5\eta)}]$

-The expected product-related utility of alternative 1 is now: $w_{1,1}\tilde{\beta_1} = \frac{exp(10\eta)-1}{2+exp(10\eta)}\beta_{11}$

-The expected product-related utility of alternative 2 is now: $w_{1,2}\tilde{\beta_1} = \frac{exp(5\eta)-1}{2+exp(5\eta)}\beta_{12}$

A.2 Simulation for Chapter 1

A.2.1 Data generation

We simulated a situation similar to our experiment. We used the same number of attributes and levels and the same experimental design. We simulated 70 participants completing the first 16 choice questions from the main task. For each participant and each choice question, we simulated the eye movements and the choice based on the learning parameter η, a set of individual-level search-related utility parameters $\theta_n = [\theta_{0n}, \theta_{1n}, \theta_{2n}, \theta_{3n}]$ defined as in Equation (4), and 18 individual-level partworths β_n .

The learning parameter was set to $\eta = 3$. All individual-level parameters were drawn from a multivariate normal distribution: $[\theta_n, \beta_n] \sim N([\theta_0, \beta_0], \begin{bmatrix} \Lambda_\theta & 0 \\ 0 & \Lambda_\beta \end{bmatrix})$, where Λ_θ was a diagonal matrix with $diag(\Lambda_\theta) = [0.1, 0.01, 0.1, 0.1]$, and Λ_β was the identity matrix. The average values of θ_n and β_n are reported in the table below.

A.2.2 Results

We calibrated our proposed model on the simulated dataset using the estimation procedure described in section 2.3. We performed a grid search on the parameter η by calibrating the model with $\eta = 0$ to 5 with increments of 1. The true value $\eta = 3$ was accurately selected based on the log-marginal density.

The average estimates of the relevant parameters are reported in the table below, together with 95% credible intervals. We see that all the search-related parameters as well as 16 out of 18 partworths are contained within the 95% credible intervals. The two partworths that fall outside of the 95% credible interval (1.9 GHz and 1 year Dell support) are still reasonably well recovered.

Table A.1: Simulation results

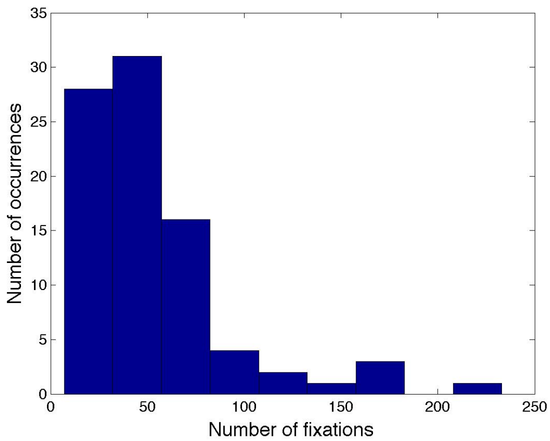
		True average value	**Estimated average value**	**95% credible interval**
Search-related parameters	θ_0	1.99	1.94	[1.79, 2.06]
	θ_1	-0.07	-0.07	[-0.09, -0.06]
	θ_2	-0.97	-0.99	[-1.02, -0.97]
	θ_3	0.70	0.74	[0.70, 0.79]
Processor speed	1.6 Ghz	-2.98	-3.28	[-3.87, -2.65]
	1.9 Ghz	-1.03	-0.63	[-0.99, -0.33]
	2.7 Ghz	1.08	0.80	[0.30, 1.29]
Screen size	26 cm	-3.05	-2.95	[-3.41, -2.63]
	35.6 cm	-1.14	-1.44	[-1.79, -1.04]
	40 cm	0.80	1.06	[0.65, 1.45]
Hard drive	160 GB	-2.83	-3.09	[-3.47, -2.78]
	320 GB	-1.02	-0.67	[-1.09, -0.32]
	500 GB	1.02	1.20	[0.90, 1.43]
Dell support	1 year	-2.85	-3.28	[-3.65, -2.94]
	2 years	-0.97	-1.03	[-1.23, -0.80]
	3 years	1.06	1.10	[0.79, 1.36]
Anti virus	30 days	-3.03	-3.05	[-3.44, -2.69]
	1 year	-1.12	-1.06	[-1.33, -0.73]
	2 years	1.08	1.01	[0.60, 1.57]
Price	350 euro	-3.10	-2.86	[-3.39, -2.32]
	500 euro	-0.96	-1.34	[-1.82, -0.86]
	650 euro	1.12	1.33	[1.07, 1.56]

A.3 Measuring attention using the total the number of fixations and fixation duration for Chapter 2.

The total number of eye fixations and fixation duration in a choice question offer another measure of attention. The average number of fixations and the average fixation duration in the main choice task was 53.76 and 17477.09 ms respectively across respondents in our data. Figure A.1 (a) and (b) shows the histogram of the number of fixations and fixation duration across all respondents. Figure A.2 (a) and (b) shows the average number of fixations and average fixation duration as a function of the probability *prob* that the choice will be realized.

Figure A.1: Distribution of number of fixations and fixation duration across all respondents.

(a) Number of fixations



(b) Fixation duration

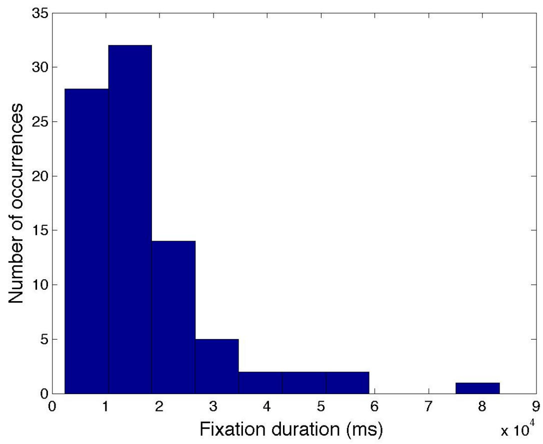
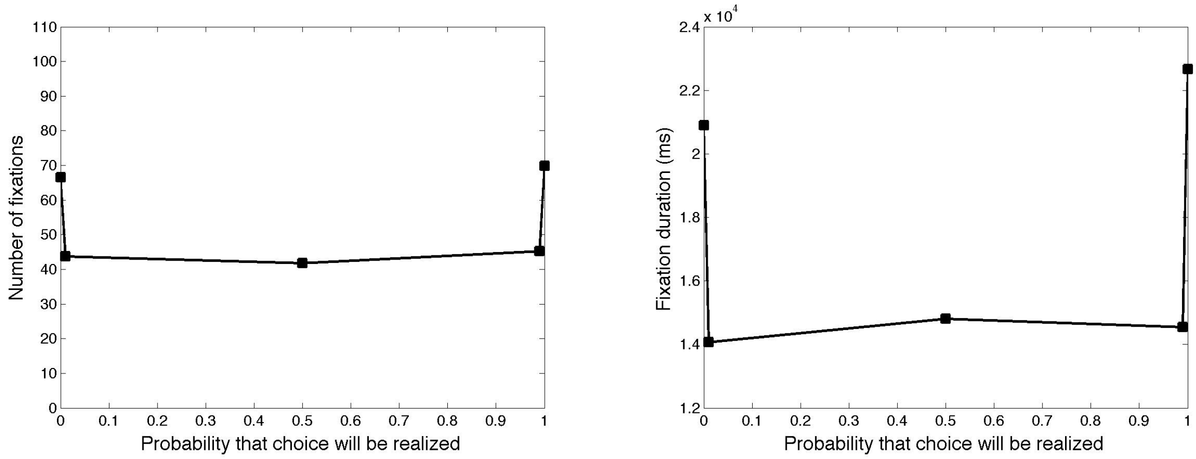


Figure A.2: Average number of fixations and fixation duration vs. probability that choice will be realized.

(a) Number of fixations

(b) Fixation duration



Like with response time and the average proportion of information visited, the relationship between *prob* and the number of fixations and fixation duration is U-shaped. We compare the number of fixations and fixation duration across conditions using non-parametric ranksum tests (similar results are obtained with t-tests; however non-parametric ranksum tests seem appropriate because the distribution of the number of fixations and fixation duration are skewed). We find no significant difference between the probabilistic conditions (*prob*=0.01, 0.5 and 0.99) for the number of fixations[1] as well as fixation duration,[2] and therefore group these three conditions. The median number of fixations under no incentives, probabilistic and deterministic incentives was respectively 56.00, 36.50 and 54.00; the median fixation duration under no incentives, probabilistic and deterministic incentives was respectively 17605.00, 11702.00 and 15725.00 ms. Like with response time and the proportion of information visited, the number of fixations and fixation duration in the probabilis-

[1]The median of number of fixations is 49, 35 and 31, when *prob*=0.01, 0.5 and 0.99 respectively. *p-value*=0.44, 0.76 and 0.78 for the comparison between *prob*=0.01 vs. *prob*=0.5, *prob*=0.01 vs. *prob*=0.99, and *prob*=0.5 vs. *prob*=0.99 respectively.

[2]The median of fixation duration is 13294.50, 11265.00 and 11087.00 ms, when *prob*=0.01, 0.5 and 0.99 respectively. *p-value*=0.79, 0.87 and 0.86 for the comparison between *prob*=0.01 vs. *prob*=0.5, *prob*=0.01 vs. *prob*=0.99, and *prob*=0.5 vs. *prob*=0.99 respectively.

tic conditions is significantly lower than that in the deterministic condition (*p-value*$<$0.04) and than that in the no-incentive condition (*p-value*$<$0.03). Again, there is no significant difference in the number of fixations and fixation duration between the no incentives and the deterministic conditions (*p-value*$>$0.89).