Cust. Need. and Solut.
DOI 10.1007/s40547-017-0072-0

RESEARCH ARTICLE

Advancing Non-compensatory Choice Models in Marketing

Anocha Aribarg[1] · Thomas Otter[2] · Daniel Zantedeschi[3] · Greg M. Allenby[3] · Taylor Bentley[4] · David J. Curry[5] · Marc Dotson[6] · Ty Henderson[4] · Elisabeth Honka[7] · Rajeev Kohli[8] · Kamel Jedidi[8] · Stephan Seiler[9] · Xin (Shane) Wang[10]

Abstract The extant choice literature has proposed different non-compensatory rules as a more realistic description of consumers' choice than a standard compensatory model. Some research has further suggested a two-stage sequential decision process of non-compensatory consideration and then compensatory choice, where the determinants of each stage may differ. Some aspects of non-compensatory choice modeling are under-studied. In this article, we hope to advance the understanding of non-compensatory choice models with the following aims: (a) providing an overview of existing representations for non-compensatory choice decisions, (b) discussing how such choice decisions can manifest from the economic search theoretical perspective, (c)

✉ Daniel Zantedeschi
zantedeschi.1@osu.edu

Anocha Aribarg
anocha@umich.edu

Thomas Otter
otter@marketing.uni-frankfurt.de

Greg M. Allenby
allenby.1@osu.edu

Taylor Bentley
Taylor.Bentley@mccombs.utexas.edu

David J. Curry
currydj@ucmail.uc.edu

Marc Dotson
marc_dotson@byu.edu

Ty Henderson
ty.henderson@mccombs.utexas.edu

Elisabeth Honka
elisabeth.honka@anderson.ucla.edu

Rajeev Kohli
rk35@gsb.columbia.edu

Kamel Jedidi
kj7@gsb.columbia.edu

Stephan Seiler
sseiler@stanford.edu

Xin (Shane) Wang
xwang@ivey.uwo.ca

1 Ross School of Business, University of Michigan, Ann Arbor, USA

2 Goethe University, Frankfurt am Main, Germany

3 Fisher College of Business, The Ohio State University, Columbus, USA

4 McCombs School of Business, The University of Texas at Austin, Austin, USA

5 Lindner College of Business, University of Cincinnati, Cincinnati, USA

6 Marriott School of Management, Brigham Young University, Provo, USA

7 Anderson School of Management, University of California Los Angeles, Los Angeles, USA

8 Graduate School of Business, Columbia University, New York, USA

9 Graduate School of Business, Stanford University, Stanford, USA

10 Ivey Business School, Western University, London, Canada

Published online: 22 June 2017

exploring the empirical identification of non-compensatory decisions using different data, and (d) presenting applications of non-compensatory choice models in novel domains.

Keywords Non-compensatory · Lexicographic · Elimination by aspects · Search · Identification · Error theory · Direct utility

1 Introduction

Choice models are an essential staple in the marketing research toolbox. The canonical choice model in marketing invokes a linear indirect utility function of the form $X'\beta + \beta p \cdot price + e$ where X encodes product attribute levels, β is a measure of preferences for these attribute levels, the product term of βp and price quantifies the disutility of price, and e summarizes influences that the analyst does not observe. This form of indirect utility function results from maximizing utility given a budgetary allotment [18]. Different distributional assumptions for e lead to different specifications of choice models. For finite attribute and preference values and a negative price coefficient βp, this model implies limited price differentials that exactly equate the indirect utility of two different alternatives, even conditional on error term realizations. A similar argument can be made for finding different attribute configurations that deliver the same indirect utility, depending on the scaling of variables in X. It is in this sense that the canonical choice model is fully "compensatory."

Given that this model is argued to embody rational choice behavior, non-compensatory models are often motivated by alluding to bounded rationality—the use of decision heuristics—by decision-makers. In this paper, we attempt to organize different motivations for departures from the canonical model just described that result in models that are widely regarded as non-compensatory. Non-compensatory models limit trade-offs among attributes, or between attributes and price that preserve a given preference ordering over alternatives. Also, we identify different degrees of departure from the fully compensatory model, with important implications for optimal actions derived from a particular model.

Psychology and consumer behavior research have consistently suggested that instead of a compensatory choice decision, consumers may also employ decision heuristics (e.g., a lexicographic rule) to minimize cognitive effort [14, 23, 29, 47]. In the behavioral economics literature, Simon [75] proposes that consumers are more likely to be *satisficers* than *maximizers*, and make their decisions based on bounded rationality (i.e., they stop searching once they find an alternative with satisfactory attribute values). However, even models of fully rational and informed consumers may give rise to non-compensatory decision-making once constraints faced by consumers become binding. Finally, rational, utility-maximizing behavior under limited information can motivate non-compensatory choice in a way that reconciles the traditional idea of utility maximizing behavior with decision strategies that are unlikely to identify the best alternative the decision-maker would pick in a hypothetical world of full information.

In this review paper, we propose a typology of and compare different non-compensatory choice models to establish linkages and understand distinctions among a collection of those suggested in the field of marketing and closely related areas. In this context, we explore the connections and differences between non-compensatory choice models that involve a two-stage process of consider-then-choose and search models in economics. We then discuss identification strategies for different forms of non-compensatory choice models given various types of data. We finally conclude with our perspectives on future research in the area of non-compensatory choice models.

2 Overview of Non-compensatory Choice Representations

2.1 Heuristic-Based Non-compensatory Models

Heuristic-based non-compensatory choice models [14] can be described as mathematical formulations constructed to describe a variety of rule-based strategies used by consumers as shortcuts to bypass an effortful compensatory decision. The lexicographic [29], disjunctive/conjunctive [23], and elimination-by-aspects models [9, 81] are well-known examples of this type of model. This class of model does not rely on utility theory as the basis of its conceptualization, although over time some researchers have attempted to bridge this gap (e.g., [44, 50, 51]). Many of these models originated from or are inspired by the mathematical psychology literature [23]. All models/rules in this class are operationalized at the attribute or aspect level.

It is useful to distinguish between fully *stochastic* and at least partially *deterministic* models derived from heuristics. While a stochastic structure makes it unlikely that some choices will occur, the deterministic structure makes it strictly impossible for them to happen. Early heuristic-based models are deterministic with the exception of the elimination-by-aspects model [81].

Deterministic In the lexicographic rule [29, 33], consumers focus on the attribute (e.g., mile per gallon) that is most important to them and simply choose the alternative (e.g., car) that is best for that particular attribute. If there is a tie, they compare the tied alternatives on the second most important attribute and so on until the tie is broken.[1] In the choice literature, researchers also use the term feature or attribute level to describe different values of a particular attribute (e.g., 25 vs.

[1] Some researchers use the term lexicographic to describe a one-shot choice heuristic where the alternative with the highest value of the most important attribute is chosen without considering a tie [14, 47], as compared to EBA, which is described as a sequential choice process that involves lexicographic rules.

31 miles per gallon). In the conjunctive and disjunctive rules [23, 25] consumers have threshold levels for all relevant attributes. While consumers only accept an alternative with all attributes passing their threshold levels in the conjunctive rule, they accept an alternative with at least one attribute passing its threshold level according to the disjunctive rule. Unlike the conjunctive rule where all attributes are considered simultaneously, a lexicographic rule relies on an ordered sequence of attributes based on their relative importance to eliminate alternatives until a choice is made. A deterministic version of the elimination-by-aspects (EBA) model [47, 58, 62] is equivalent to the lexicographic rule except that it is operationalized based on an aspect (i.e., an attribute level) instead of an attribute. When an attribute is binary, it can also be thought of as an aspect. An attribute with multiple levels can be considered a collection of aspects.[2] Yee et al. [87] also propose extended models of acceptance by aspect (ABA) and a mixture of EBA and ABA referred to as lexicographic by aspects (LBA).[3]

Stochastic Jedidi and Kohli [44] propose a subset-conjunctive rule with a probabilistic component to extend the conjunctive and disjunctive rules and subsume them as special cases. In a subset-conjunctive model, an alternative with n attributes is classified as being acceptable if it is satisfactory on at least k attributes, where 1 (disjunctive) $\leq k$ (subset-conjunctive) $\leq n$ (conjunctive). The original EBA model [81] is a stochastic rule because aspect selection follows a stochastic process instead of being deterministic (i.e., determined by the relative importance of aspects). Implementations of the original EBA model have been sparse in the marketing literature [35] due to its computational complexity [32, 68]. Most papers rely on carefully selected assumptions to circumvent problems, including the use of preference tree or hierarchical elimination structures [54, 79, 82] an MDS approach [52] and reliance on market share data instead of consumer choice [8] Kohli and Jedidi [50] specify an EBA model as a probabilistic version of a lexicographic rule, where each aspect has a random utility with an extreme value distribution. They argue that the probabilistic non-compensatory model obtains its deterministic counterpart as a limiting case, and includes preference trees and hierarchical elimination models [79] as its special cases.

Recent work by Kohli and Jedidi [51] further clarifies the relationship between EBA and the class of nested and cross-nested logit models. They propose an error theory for an extended EBA model in which (i) each alternative has a unique aspect[4] and (ii) the utility of a unique aspect is a function of attribute levels. In the simplest case, a unique aspect is defined by an attribute level only present in one alternative in a set. An "extended" aspect is defined as a composition of attribute levels. For example, an extended unique aspect could be given by the combination of the attribute levels "red" and "sedan" in a choice set of cars, if all other cars have different colors and are not sedans.

EBA captures a consideration-then-choice process, but without separately formulating a consideration stage and a choice stage. Choosing a "unique" aspect at any stage of elimination corresponds to choosing an alternative. Any preceding elimination stages represent a consideration phase in which alternatives are eliminated successively based on whether they exhibit momentarily focal aspects or not. If the momentarily focal aspect is shared by more than one alternative, all these alternatives are carried to the next stage and so on until a unique aspect becomes focal. The number of elimination stages in the consideration phase can differ from one choice occasion to another, as can the specific sequence of aspects used for eliminating alternatives. Nested logit and cross-nested logit models correspond to special cases of preference trees [79].

These results further clarify the formal relationship between compensatory and non-compensatory choice models. These models were previously known to have a common mathematical representation if we assume no error and consider discrete (or discretized) attributes. In this case, the simple linear ("part-worths") model is sufficient for representing lexicographic, conjunctive, disjunctive, and subset-conjunctive rules as special cases. In fact, for each of these rules, there are multiple parameterizations of a linear valuation model that can represent the same preferences (see [44, 45]). With error, however, the recent results by Kohli and Jedidi [51] show that the family of logit models can be represented as special cases of the extended EBA model.

Kohli and Jedidi [51] discuss identification issues, estimation, statistical testing, and data collection for EBA. Some of the multinomial, nested, and cross-nested logit models can be thought as special cases of the extended EBA model, and they suggest the use of likelihood ratio tests, Akaike information criterion (AIC), and Bayesian information criterion (BIC) to select among these models. Furthermore, Kohli and Jedidi (2016) illustrate this empirical approach using an application in which the utilities of alternatives are functions of covariates.

[2] Yee et al. [87] refer to an attribute as a feature and emphasize the difference between aspect and feature. Processing by feature implies that , for example, a brand was the feature chosen for evaluation first, then an ordering of all the aspects of that feature would take place before moving on to the next feature (e.g., price). Processing by aspect assumes that consumers split features into specific aspects and proceed through the decision process in a binary aspect-sorting manner. In this way, processing by feature is a special restricted case of processing by aspect.

[3] Lexicographic -by- aspects is a two-stage model with a non-compensatory decision in the first stage and compensatory decision in the second stage.

[4] Tversky [81] distinguishes unique and shared aspects in his original EBA paper. Moreover, McFadden [55] presents rigorous notions of the uniqueness of aspects defined over sets.

The application illustrates the process of building an extended preference tree, compares the choice processes for a preference tree and a nested logit model, and contrasts their marketing implications. Interestingly, in their empirical application related to transportation choices, the implied demand elasticities differ across these related models, suggesting a caution for researchers not to rely primarily on a standard logit model to derive marketing implications from different counterfactual scenarios.

2.2 Utility-Based Non-compensatory Models

Instead of completely abandoning the canonical fully compensatory model, some researchers posit a utility-based two-stage choice decision (e.g., [31, 66]) where consumers first rely on a non-compensatory rule to narrow the initial pool of alternatives to a consideration or choice set, and then engage in compensatory decision-making to select the best alternative from the reduced set. These models can be further classified into models that are entirely derived from a direct utility function that consumers maximize subject to an explicit set of constraints, and models where consumers employ simplifying heuristics before the constrained maximization that yields the optimal choice given the reduced set of considered alternatives.

The simplest type of economic constraint that can motivate a restricted set of considered brands is a monetary budget constraint that precludes consumers from choosing a brand priced above the available budget. For example, a consumer may not be able to afford a particular home and thus have a purchase probability of zero for homes beyond his budget, regardless of the direct utility provided by any particular home. As a consequence, effective compensation between the utility provided by a home and its price is only possible under the set of prices covered by the consumer's budget. This is an example of entirely rational, informed consumers making choices in a way not covered by the canonical fully compensatory model. However, the model implies preference-preserving trade-offs between (direct) utility and price upon relaxing the budget constraint.

It is again useful to distinguish between fully *stochastic* and at least partially *deterministic* models in the context of utility-based non-compensatory models. Partially deterministic models put zero probability on certain consideration sets, given model parameters, and thus may assign a probability of zero to particular choice outcomes. In contrast, fully stochastic models support with positive probability all possible choices from the universe of available alternatives.

Deterministic Early deterministic models propose the use of a deterministic EBA or EBA-like procedure based on a few attributes in the first stage to form a consideration set [3, 4, 28, 31]. More recently, Gilbride and Allenby [34] rely on conjunctive and disjunctive rules and propose a two-stage model where they specify screening rules and decision structures without additional error terms in the first stage and a compensatory choice structure in the second stage. These screening rules can be conceptualized to arise from constraints in a model of direct utility maximization, where in addition to the budget constraint other performance constraints exclude alternatives from being chosen. A digital camera, for example, may need to be waterproof in addition to costing less than a budgeted amount. Screening rules assume that some alternatives are not evaluated, and the associated error terms do not factor into the model likelihood. In this way, screening rules function as a hard or binding constraint in the decision process. Following the same deterministic structure in the screening stage, Gilbride and Allenby [35] propose an economic screening rule, in addition to an EBA model, as the result of a trade-off between cognitive effort and expected utility. Extending Gilbride and Allenby [34], Gilbride and Allenby [35] provide an economic justification (i.e., cognitive cost) for why a particular attribute is used to screen alternatives.

Finally, recent work [49] suggests that the aspects consumers screen on may be low-dimensional meta-attributes. Specifically, consumers may screen on the benefits that a product provides. For example, a consumer may screen on the cleaning performance of a vacuum and not its storage capacity and suction separately because these two attributes comprise the cleaning performance benefit for this consumer. Identifying both how attributes heterogeneously map to benefits and which benefits are screened on may require augmenting choice with auxiliary data on the benefits relevant to consumers.

Stochastic Given consumers' desire to limit evaluation costs in decision-making [73], Hauser and Wernerfelt [39] and Roberts and Lattin [66] propose a two-stage choice model where a consideration set is conceptualized as a result of an optimization problem where an alternative's expected consumption utility must exceed the mental cost of evaluating it in order to enter the consideration set. Unlike Gilbride and Allenby [35], there is an error term involved in their consideration utility. With the exception of Roberts and Lattin [66] that use a stated measure of consideration and Jedidi et al. [45] that employ a conjoint experiment, other two-stage models in this class are applied to scanner panel data (e.g., [5, 21, 74]). Consideration of an alternative—mostly a brand—is conceptualized to be driven by in-store promotions [2, 60, 74] and previous experience with the alternative [15]. While the early literature describes the consideration process as an evaluation cost threshold model (e.g., [39]), some later research takes a memory-based perspective [27, 59] or a probabilistic perspective—probabilities are constructed for all possible combinations of consideration sets [5, 21].

Figure 1 presents the typology of non-compensatory choice models based on our discussion in this section.

3 Search Vs. Non-compensatory Choice Process

The literature on non-compensatory choice and models of consumer search share many features, perhaps most saliently the fact that both describe a situation in which consumers do not evaluate all available options but rather make their purchase decision from a limited subset of products. Interestingly, to the best of our knowledge, there is a little linkage between the two kinds of literature and they are sometimes perceived to represent fundamentally different paradigms. We argue that consumer search and non-compensatory choice rules have much in common and that it is straightforward to view search models as a micro-foundation for the type of screening rules employed in non-compensatory choice models. Despite similarities, two apparent differences between the literature on non-compensatory models and the literature that derives consideration sets from search are that (1) search is typically modeled at the alternative level whereas non-compensatory screening rules are based on attributes or aspects and (2) search models derive search and purchase decisions from an inherently compensatory utility framework, although empirical search patterns can look much like non-compensatory screening rules.

Most models of consumer search assume that the consumer exhibits uncertainty with respect to one (or more) attribute(s) of a product (e.g., price or "fit") and engages in search to learn about these attributes (see, e.g., [16, 76, 84] for classic papers; [39, 42, 48, 56, 66, 70–72] for more recent work).[5] Search is then modeled at the alternative level in the sense that the consumer needs to pay a cost to search and, upon incurring this cost, the consumer learns about the previously unknown attribute(s) of the alternative. Under this set of assumptions, to save on search costs (which may include cognitive costs), consumers will not evaluate all alternatives and hence will make decisions with limited information. The consumer then makes his purchase decision by picking the product that gives him the highest utility among the alternatives in his consideration set (i.e., the set of alternatives searched by the consumer).

Models of search are characterized by a set of assumptions about the process by which the consumer searches: (1) information structure: what is unknown to the consumer initially but revealed after search and how consumers form expectations and (2) search process: how the consumer searches the different options available to him and how he optimally decides to stop searching. Typically, specific assumptions are imposed with regard to both (1) and (2), but a set of papers has made progress in identifying some of these components.[6]

With regard to the second point, two main methods of searching have been proposed in the literature, namely simultaneous and sequential search. Simultaneous search, which is sometimes also called fixed sample or non-sequential search—despite its name—does not imply that the consumer searches all products at the same time. Under simultaneous search, the consumer pre-commits to searching a specific set of alternatives and does not stop searching until he has collected the information for all products in his consideration set—even if he finds a very favorable option early on. The consumer searches the set of products that gives him the highest expected benefit net of his search costs among all possible search sets (see [42, 56, 76]). In contrast, under sequential search, the consumer searches one alternative at a time. After resolving his uncertainty about an alternative, the consumer decides whether he wants to continue searching or to stop searching and purchase a product among the products he has already searched. The consumer searches products in decreasing order of reservation utility— the utility that equates the marginal cost and the marginal benefit of a search (see [20, 48, 84]). While most empirical research has made an assumption on the search method consumers use, recent papers have investigated situations in which consumers' search methods are identified (e.g., [19, 43, 70]).

With regard to information structure, it is typically assumed that consumers are uncertain concerning just one product characteristic and that this characteristic is independently distributed across the available options. Independence is necessary for the optimal search rules developed by Weitzman [84] to be applicable in a sequential search model and lowers the computational burden in a simultaneous search model.[7] The assumption of independence, however, does not permit consumers to learn across evaluated alternatives.[8] The information structure regarding what the consumer is searching over is typically imposed based on the institutional features of the market. For instance, Honka [42] models search over the price for insurance contracts, whereas Kim et al. [48] model search over a "match value" or "fit" in the case of online camcorder purchases.

Turning to the relationship between search models and non-compensatory choice models, we note that search models are compensatory in nature; however, the searched set or consideration set might look much like one arising from a non-compensatory rule. For instance, suppose a researcher

[5] For detailed search literature overviews, see Baye et al. [10] and Ratchford [64].

[6] Our discussion focuses primarily on search models for differentiated goods rather than on search models for homogeneous goods.

[7] For a detailed discussion, see Honka and Chintagunta [43].

[8] Learning over the distribution of the unknown alternative has been modeled (see, e.g., [67]), but learning preferences more broadly (i.e., the consumer might try to figure out which characteristics he wants a digital camera to have during his search process) has not been tackled to the best of our knowledge.

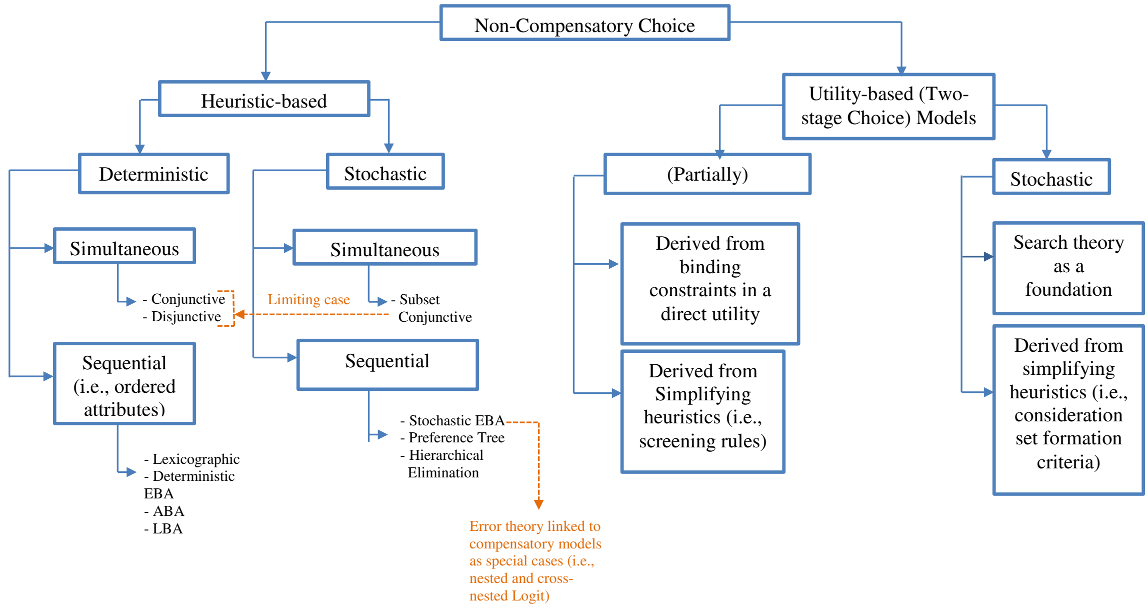

Fig. 1 Typology of non-compensatory choice models

observes that a consumer does not include any compact cars in his consideration set. The non-compensatory choice framework would describe such exclusion as the result of screening behavior (i.e., the consumer does not include any compact cars in his consideration set because these cars have the characteristic of being a compact car). The search framework would posit that such behavior is consistent with an underlying compensatory utility function. That is, the consumer does not like compact cars much, but might have considered and purchased a compact car at an extremely low price. However, (assuming that search is over price) the likelihood is scant of finding a price so low as to compensate for the undesired "compact" feature. Thus, the expected benefit falls short of the cost of search. Hence, products outside the consideration set will often include items with characteristics for which the consumer has little value. For such items, search costs are too high. Thus, search models can—and in fact are likely to—generate consideration sets that look much like the ones generated from a non-compensatory screening rule.

The second key difference is that the search literature has typically modeled search at the product rather than the product characteristic level [36, 38, 57, 86]. In other words, consumers resolve all uncertainty about a product after gathering information about a single product attribute (e.g., price or "fit").[9] While this is different from non-compensatory rules, which tend to be based on multiple attributes or aspects, the product-centric focus of search models is not a necessary feature of such models. Instead, to a large extent, the product focus is due to modeling challenges and the nature of most search data (often obtained from online browsing) that do not allow the researcher to observe which product attributes the consumer evaluated during search. However, on a conceptual level, it is possible to enrich search models by assuming that consumers have uncertainty over multiple attributes of each product and need to pay a search cost to learn about an attribute/product pair. A more granular model of search and information acquisition such as this could give rise to the development of optimal search rules that involve evaluating products on only one attribute first and thus would provide a micro-foundation for a process that looks like an attribute-based non-compensatory rule.

In summary, we argue that it is reasonable to view search models as a micro-foundation for screening rules that appear non-compensatory in nature. This begs (at least) two questions that have not yet been addressed. The first question is whether we can empirically distinguish (i.e., identify) non-compensatory choice rules from compensatory search behavior. And the second question is for what set of questions does this distinction matter? In other words, would elasticities and other (counterfactual) predictions differ much under one scenario vs. the other?

4 Identification of Non-compensatory Choice Processes

We organize the discussion on identification of non-compensatory models in different subsections. We first elaborate on general principles of identification. We then present a survey of empirical strategies related to the identification of specific non-compensatory models balancing more traditional research with recent research and venues for further investigation.

[9] Some progress in modeling search for multiple product attributes has been made under the assumption that the characteristics' distributions are independent [70].

4.1 General Principles

Identification of choice models represented with utility components above and beyond the linear additive formulation is a topic that has attracted a lot of interest in recent years [13]. There are several nuances in the concept of identification. The first is statistical, stating that a choice model is identifiable from the data if the mapping between parameters and the derived multinomial likelihood is one to one. However, these analytical conditions are not immune to some subtleties. For instance, two choice models can be separately identifiable but may produce predicted choice probabilities that are very close if not the same. This means that quantities relevant to marketers, such as market shares and welfare estimates, may achieve the same level of likelihood in the dataset but generate different counterfactuals and managerial implications.

Researchers are then motivated to explore the notion of identification beyond simple statistical exercises and consider, for instance, the economic identification of choice models. Economic identification strategies rely less on a mathematical process but instead use the variations and patterns in the data to argue for their consistency with a prescribed model. *Model-free evidence*, in particular, has become an important tool employed by researchers to justify the adoption of complex choice models based on empirically verifiable patterns of data and availability of alternatives compatible with data discontinuities. We refer to Reiss [65] for a thorough discussion of general strategies for model-free evidence and a survey of data discontinuities comprising exogenous shocks, statistical instruments, and natural experiments.[10]

The same principles also apply in the context of non-compensatory choice models. Researchers have relied on data discontinuities (i.e., situational factors affecting screening but not a choice) or the inclusion of auxiliary data, combining, for example, primary and secondary data sources, to improve the identification of choice probabilities.

Nonetheless, there are additional complications with both the statistical and economic identification of non-compensatory choice probabilities. First, the likelihood function of non-compensatory models is, by construction, *concentrated* in certain regions of the attributes and parameters space. In other words, there exist combinations of parameters and attribute configurations where choice probabilities are degenerate: some alternatives achieve a probability of zero (and log-likelihood of minus infinity, which causes numerical instability). In this case, it may become difficult to illustrate with model-free evidence the presence of zero-probability events, especially in large attribute spaces. Second, the concentrated likelihoods of non-compensatory models lack smoothness and maybe even lack (absolute) continuity (i.e., presence of point masses). This, in turn, may generate important economic implications such as a kinked demand function and the non-existence of demand elasticities. These aspects of non-compensatory choice models are not well documented in the literature. This may be in part due to the common practice of adding an error term to a consideration utility. Such practice restores smoothness and continuity in the likelihood but fundamentally violates the notion of a non-compensatory process. For instance, due to the presence of the additive error, it may be harder to identify consideration stages without relying on restrictive parametric assumptions. The lack of consideration becomes observables only for attributes levels (affecting the consideration evaluations) that completely overwhelm the influence of the error term.[11]

Moreover, while extreme preference coefficients may drive predicted choice probabilities to zero, the presence of error terms can ameliorate the identification process when dealing with large choice sets and mislead when looking for sources-of-volume calculations. Product categories can have upwards of hundreds of choice alternatives (e.g., beer, cereal, cars), and if choice among alternatives is based on models where each alternative is associated with an error term, the model will not fit the data well without some way to concentrate the likelihood. Price changes, for example, are watered down in models with an excessive number of error terms because certain realizations of errors will generate a positive probability of selection for every alternative regardless of price. Predictions from two-stage models that account for all model uncertainty in both consideration and choice stages, and not just the conditional demand model of the second stage, also suffer from this problem.

As mentioned earlier, the study of generalized error theories for non-compensatory choice models has advanced considerably in recent years due to the work of Jedidi and Kohli [50, 51], but there are several open avenues where identification of choice models should be explored. For example, models of non-compensatory discrete and continuous demand, as well as non-standard notions of complementarities and substitutability applicable, for instance, to the study of bundled goods.

The rigorous study of identification processes for non-compensatory choice models is not a sufficiently explored domain. Researchers are encouraged to adopt established tools such as model-free evidence or data fusion approaches to improve identification from data, but they ought to be careful when relying on standard functional forms for the error terms shaping the likelihood function of the choice process. A non-compensatory model can be thought of as "structural" [85] where restrictions imposed on the choice process need to be motivated by theoretical considerations and cannot be entirely justified by the underlying data.

[10] Note that the paradigms illustrated in Reiss [65] are intended for compensatory models with standard generalized extreme value errors.

[11] See Heckman and Vytlacil [41] for related notions of "identification at infinity."

We now turn to a selection of empirical strategies that can offer further insights about the identification process of non-compensatory rules.

4.2 Data Fusion Approach

Many utility-based non-compensatory models are implemented using conjoint experiment data [6, 34, 35]. Rather than accommodating non-compensatory attribute thresholds post hoc in the model, some have suggested delineating these thresholds in the initial stages of the choice experiment to optimize design efficiency [46, 78]. The initial stage of the Adaptive Conjoint Analysis (ACA) interview process asks respondents to rank attributes according to the personal importance and also includes an option asking respondents to identify attribute levels that are unacceptable. Although using self-explicated screening thresholds to prematurely restrict the choice experiment design space may be unwise, as individuals often violate their stated preferences when making actual/experimental choices [28], research suggests that stated measures of attribute thresholds can improve choice prediction in experimental contexts [1]. Rather than estimating the parameters governing such latent indicators using only stated choice data, more recently, researchers have leveraged auxiliary data such as self-stated consideration of attribute levels collected from surveys to improve estimation of the screening parameters and to explore relationships between screening and observable characteristics of the product stimuli [6, 7].

Dehmamy and Otter [26] investigate analytically how discrete-continuous choice data usefully constrain the choice between including observed covariates as drivers of consideration or drivers of utility given consideration in a two-stage model of choice. They also show that including drivers of consideration in the utility function may result in counter-intuitive sign reversals such as, for example, negative effects of the number of facings of a brand on a shelf.

An alternative route to improving identification for non-compensatory choice models is given by traditional research in the area of mathematical psychology [80]. The general idea from these developments is that individuals' evaluation of features is reflected in the response time necessary to make the decision: more complex decision rules may take longer on average than their heuristic counterparts.[12] In this respect, noteworthy and pioneering is the work by Busemeyer et al. [17]. They develop an empirical strategy based on attribute manipulations to test different versions of elimination models by looking at each model's implied response times. Unfortunately, this strategy, to the best of our knowledge, has never been implemented in an experimental or observational study despite its plausibility in the presence of emerging data structures that combine, for example, neuro-physiological or consumer path-to-purchase measurements with final choices [61].

There also exists a mathematical treatment of the joint structure of choice and response times: the "horse-race processes" where choice probabilities are mechanically linked to hazard rates determining response times [53]. These representations appear as an encouraging venue to identify non-compensatory choice models and test them against their compensatory counterparts, when the interactivity times (i.e., the time between the different steps of an EBA evaluation) are partially or entirely observed in addition to the ultimate choice. Regarding the earlier discussion on statistical identification, it is conceivable that two alternative non-compensatory choice models yield the same predicted choice probabilities from the data, but that these models can be distinguished by looking at their response time structures.

Two recent papers have empirically investigated the advantages of combining choice with response times. In a marketing research setting, Otter et al. [61] advance an integrated model improving identification of respondent preference for choice alternatives by accounting for heterogeneity in response times. Their findings provide consistent support for the endogenous nature of response times above and beyond the underlying choice probabilities. Clithero and Rangel [22], justified by neuroeconomic prescriptions, reinforce these findings by showing that out-of-sample forecasts of choice probabilities can be considerably improved by including response times in an indirect utility model. These two recent developments do not deal directly with non-compensatory decision processes, and this leaves room for further development and application of models of non-compensatory choices complemented by response times.

Finally, the advent of eye-tracking technology allows researchers to use eye movement data to identify a non-compensatory choice process [77, 86]. For example, Stüttgen et al. [77] propose a joint model of search and evaluation (i.e., a process to determine whether a product is satisfactory or not, leading up to choice using conjoint experiment data augmented with eye movement patterns). They show that consumers rely on a satisficing rule rather than on a utility-maximizing choice rule. That is, the consumer sequentially evaluates whether a given product has satisfactory attribute values and stops when she or he comes across the first satisfactory alternative. Nonetheless, the authors allow for consumers to continue searching for the purpose of verification [69].

4.3 Function Restrictions and Aggregate Data

Bentley and Seetharaman [12] argue that the full EBA model has never been applied to actual market demand data (e.g., scanner data) because it requires that consumers face different

[12] A complete treatment of the role of response times in identifying the underlying choice architecture is beyond the scope of this paper. We refer the interested reader to the recent survey by Gaissmaier et al. [30].

choice sets across choice occasions so that non-IIA switching patterns among alternatives can be explicitly observed and, therefore, exploited in the estimation of shared aspects. The only way to observe different choice sets in scanner data is when product stock outs occur and are observed. However, since stock outs are rare, they do not offer enough variation in choice sets across weeks for practical estimation purposes. The only recourse to estimating the EBA model was to use discrete choice experiments where choice set compositions were systematically varied via survey instruments [9, 11, 35, 68]. Further, the EBA model has difficulty incorporating continuous, time-varying covariates, such as price, because attributes in Tversky's original formulation are measured in discrete levels.

Bentley and Seetharaman develop (2016) an EBA model that allows for the simple and parsimonious inclusion of continuous, time-varying covariates while retaining the original EBA's ability to flexibly allow for unobserved similarities among alternatives. They show that temporal variation in marketing variables of alternatives (e.g., price, display, feature) across choice occasions provides choice set variation, and along with the corresponding substitution patterns, allows for the identification of the EBA model's shared aspects. Thus, these specific versions of the EBA model can be estimated on actual market demand data, rather than requiring the use of survey techniques.

5 Discussions of Applications and New Research Perspectives

In light of the different theoretical motivations (i.e., psychology, economics) for at least partially non-compensatory models of choice, it seems clear that the canonical fully compensatory choice model still popular in marketing applications can only provide a local approximation of consumers' choice behavior. Much of the current literature has focused on improving explanations of consumers' choices, including the measurement of economic quantities such as price elasticities and search costs. What we believe is still missing from the marketing literature is a wider investigation of the implications of different choice models for optimal marketing and competitive actions. In a field that is driven by applications and dependent on providing tools to support actual decisions, models that are more cumbersome to fit or harder to understand will only be accepted in mainstream applied marketing research if their implications for optimal actions are shown to be more sensible than those derived from the compensatory model. As an example, Sovinsky-Goree [37] investigates competitive implications of a two-stage choice model where the consideration set is determined by advertising, and advertising influences demand exclusively in the first stage.

While implications from the canonical compensatory choice model obviously differ from those obtained from various non-compensatory models at the individual consumer level, a largely unexplored question is the aggregation behavior of these models and the resulting implication for aggregate marketing actions. Much of the current empirical literature on demand for differentiated goods relies on the mixed logit model. A logical next step seems to investigate if and how adopting various non-compensatory model formulations result in different implications for aggregate pricing, product, and advertising decisions.

In addition to the application aspect, some researchers are attempting to conceptualize non-compensatory choice using alternative theoretical perspectives. For example, Wang [83] argues that even though they allow for preference heterogeneity, most choice models assume that consumers use the same decision process. The author proposes imputing belief networks directly from observed behavior to learn consumer decision-making processes without having to specify an explicit utility function. Furthermore, the attribute satiation assumption in choice modeling proposed by Kim et al. [49] implies that the utility function is monotone and subadditive. Using a neuroscience perspective, Curry and Wang [24] provide additional support for subadditivity using biologically oriented primitives based on the representational capacity of the brain that includes bounded output states and finite precision. They derive a subadditive utility function that closely matches the function derived by Kim et al. [49] using low-dimensional meta-attributes. Both functions exhibit an endogenous form of non-compensatory behavior largely unstudied in the extant literature. This form postulates that non-compensatory processing is a continuous but non-homothetic function of a choice option's overall utility. Hence, these approaches have the potential to link search models (with their holistic emphasis on options) and EBA-style models (with their emphasis on specific aspects of these options).

Another area that has been relatively under-explored is the evolution over time of choice heuristics or search behavior in dynamic choice settings. Typically, dynamic demand models for storable goods and durables analyze incentives to delay or accelerate purchases as a function of current and expected product prices (or other product/market characteristics such as new product entry). Clearly, consumer behavior in such settings depends crucially on the information processed over time and choice heuristics or search behavior conceivably change over time as a function of the relevant state variables. Concretely, in a storable good setting, a consumer might be less likely to pay attention to prices when he purchased recently and still has a high inventory of the product. Seiler [71] and Pires [63] model such dynamic changes in search behavior as a function of inventory holdings. Haviv [40] analyzes seasonal changes in search behavior in storable goods markets. To the best of our knowledge, the role of search in durable goods markets has not been explored in recent research.

Despite the extant literature on non-compensatory choice, more needs to be done to unify different specifications having been proposed, to better establish rigorous identification strategies from both statistical and economic perspectives, and to explore further implications of non-compensatory choice on optimal firm and public policies both at the individual choice and aggregate market share levels. We hope that this review will generate more interest in future work in this area.

References

1. Agarwal MK, Green PE (1991) Adaptive conjoint analysis versus selfexplicated models: some empirical results. In: *International Journal of Research in Marketing* 8(2):141–146
2. Allenby GM, Ginter JL (1995) Using extremes to design products and segment markets. In: *Journal of Marketing Research* 32(4): 392–403
3. Andrews R, Manrai AK (1998) Feature-Based Elimination: Model and Empirical Comparison. In: *European Journal of Operational Research* 111(2):248–267
4. Andrews RL, Manrai AK (1999) MDS maps for product attributes and market response: an application to scanner panel data. In: *Marketing Science* 18(4):584–604
5. Andrews RL, Srinivasan TC (1995) Studying consideration effects in empirical choice models using scanner panel data. In: *Journal of Marketing Research* 32(1):30–41
6. Aribarg A, Arora N, Henderson T, Kim Y (2014) Private label imitation of a national brand: implications for consumer choice and law. In: *Journal of Marketing Research* 51(6):657–675
7. Aribarg A, Burson K, Larrick RP (2017) Tipping the scale: discriminability. Effect on derived attribute importance. In: *Journal of Marketing Research* 54(2):279–292
8. Batsell RR, Polking JR (1985) A new class of market share models. In: *Marketing Science* 4(3):177–197
9. Batsell RR, Polking JC, Cramer RD, Miller CM (2003) Useful mathematical relationships embedded in Tversky's elimination by aspects model. In: *Journal of Mathematical Psychology* 47(5):538–544
10. Baye MR, Rupert Gattu J, Kattuman P, Morgan J (2009) Clicks, discontinuities, and firm demand online. In: *Journal of Economics & Management Strategy* 18(4):935–975
11. Ben-Akiva M, Boccara B (1995) Discrete choice models with latent choice sets. In: *International Journal of Research in Marketing* 12(1):9–24
12. Bentley T, Seetharaman PB (2017) Identifying unobserved similarity: estimating a fully flexible EBA model with standard marketing data. In: Working paper. University of, Texas at Austin
13. Berry ST, Haile PA (2009) Nonparametric identification of multinomial choice demand models with heterogeneous consumers. In: Working paper 15276. National Bureau of Economic Research
14. Bettman JR, Luce MF, Payne JW (1988) Constructive consumer choice processes. In: *Journal of Consumer Research* 25(3):187–217
15. Bronnenberg BJ, Vanhonacker WR (1996) Limited choice sets, local price response and implied measures of price competition. In: *Journal of Marketing Research* 33(2):163–173
16. Burdett K, Judd KL (1983) Equilibrium price dispersion. In: *Econometrica* 51(4):955–969
17. Busemeyer JR, Forsyth B, Nozawa G (1988) Comparisons of elimination by aspects and suppression of aspects choice models based on choice response time. In: *Journal of Mathematical Psychology* 32(3):341–349
18. Chandukala SR, Kim J, Otter T, Rossi PE, Allenby GM (2008) Choice models in marketing: economic assumptions, challenges and trends. In: *Foundations and Trends in Marketing* 2(2):97–184
19. Chen Y, Yang S (2007) Estimating disaggregate models using aggregate data through augmentation of individual choice. In: *Journal of Marketing Research* 44(4):613–621
20. Chen Y, Yao S (2016) Search with refinement. In: *Management Science* Article in Advance
21. Chiang J, Chib S, Narasimhan C (1998) Markov chain Monte Carlo and models of consideration set and parameter heterogeneity. In: *Journal of Econometrics* 89(1–2):223–248
22. Clithero JA, Rangel A (2015) Combining response times and choice data using a neuroeconomic model of the decision process improves out-of-sample predictions. In: *unpublished, California Institute of Technology* (2015).
23. Coombs CH (1951) Mathematical models in Psychological scaling. In: *Journal of the American Statistical Association* 46(256):480–489
24. Curry D, Wang X (2017) Commentary on "Benefit-based conjoint analysis" by Kim et al. 2016. In: Working paper University of Cincinnati. Lindner College of Business, Department of Marketing, pp 1–37
25. Dawes RM (1964) Social selection based on multidimensional criteria. In: *Journal of Abnormal and Social Psychology* 68:104–109
26. Dehmamy K, Otter T (2017) Consideration versus utility when choice is discrete continuous. In: Working paper Goethe University
27. Desai KK, Hoyer WD (2000) Descriptive characteristics of memory-based consideration sets: influence of usage occasion frequency and usage location familiarity. In: *Journal of Consumer Research* 27(3):309–323
28. Fader PS, McAlister L (1990) An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data. In: *Journal of Marketing Research* 27(3):322–332
29. Fishburn PC (1974) Exceptional paper-lexicographic orders, utilities and decision rules: a survey. In: *Management Science* 20(11): 1442–1471
30. Gaissmaier W, Fific M, Rieskamp J (2011) Analyzing response times to understand decision processes. In: A handbook of process tracing methods for decision research: a critical review and user's guide. Ed. by Schulte-Mecklenbeck, M., Küh- berger, A., and Ranyard, R. New York: Psychology Press, pp. 141–162.
31. Gensch DH (1987) A two-stage disaggregate attribute choice model. In: *Marketing Science* 6(3):223–239
32. Gensch DH, Ghose S (1992) Elimination by Dimensions. In: *Journal of Marketing Research* 29(4):417–429
33. Gigerenzer G, Goldstein DG (1996) Reasoning the fast and frugal way: models of bounded rationality. In: *Psychological review* 103(4):650–669
34. Gilbride TJ, Allenby GM (2004) A choice model with conjunctive, disjunctive, and compensatory screening rules. In: *Marketing Science* 23(3):391–406
35. Gilbride TJ, Allenby GM (2006) Estimating heterogeneous EBA and economic screening rule choice models. In: *Marketing Science* 25(5):494–509
36. Gilbride TJ, Currim IS, Mintz O, Siddarth S (2016) A model for inferring market preferences from online retail product information matrices. In: *Journal of Retailing* 92(4):470–485
37. Goeree MS (2008) Limited information and advertising in the U.S. personal computer industry. In: *Econometrica* 76(5):1017–1074
38. Hagerty MR, Aaker DA (1984) A normative model of consumer information processing, In: *Marketing Science* 84 (3):227–246
39. Hauser JR, Wernerfelt B (1990) An evaluation cost model of consideration sets. In: *Journal of Consumer Research* 16(4):393–408
40. Haviv A (2015) Does purchase without search explain counter-cyclic pricing? In: Simon School of Business working paper

41. Heckman JJ, Vytlacil E (2005) Structural equations, treatment effects, and econometric policy evaluation. In: *Econometrica* 73(3): 669–738
42. Honka E (2014) Quantifying search and switching costs in the US auto insurance industry. In: *The RAND Journal of Economics* 45(4): 847–884
43. Honka E, Chintagunta PK (2017) Simultaneous or sequential? Search strategies in the US auto insurance industry. In: *Marketing Science* 36(1)21-42
44. Jedidi K, Kohli R (2005) Probabilistic subset-conjunctive models for heterogeneous consumers. In: *Journal of Marketing Research* 42(4):483–494
45. Jedidi K, Kohli R, DeSarbo WS (1996) Consideration sets in conjoint analysis. In: *Journal of Marketing Research* 33(3):364–372
46. Johnson EJ, Payne JW, Bettman JR (1988) Information displays and preference reversals. In: *Organizational Behavior and Human Decision Processes* 42(1):1–21
47. Johnson EJ, Meyer RJ, Ghose S (1989) When choice models fail: compensatory models in negatively correlated environments. In: *Journal of Marketing Research* 26(3):255–270
48. Kim JB, Albuquerque P, Bronnenberg BJ (2010) Online demand under limited consumer search. In: *Marketing Science* 29(6):1001–1023
49. Kim DS, Bailey R, Hardt N, Allenby GM (2017) Benefit-based conjoint analysis. In: *Marketing Science* 36(1):54–69
50. Kohli R, Jedidi K (2015) Error theory for elimination by aspects. In: *Operations Research* 63(3):512–526
51. Kohli R, Jedidi K (2017) Relations between EBA and nested logit models. In: *Operations Research* Forthcoming
52. Manrai AK, Sinha P (1989) Elimination-By-Cutoffs. In: *Marketing Science* 8(2):133–152
53. Marley AAJ, Colonius H (1992) The "horse race" random utility model for choice probabilities and reaction times, and its compering risks interpretation. In: *Journal of Mathematical Psychology* 36(1):1–20
54. McFadden D (1973) University of California, Berkeley and Institute of Urban & Regional Development, Conditional logit analysis of qualitative choice behavior. Calif.: Institute of Urban and Regional Development, University of California, Berkeley
55. McFadden D (1981) Econometric models of probabilistic choice. In: Manski C, Mc Fadden D (eds) Structural analysis of discrete data with econometric applications. MIT Press, Cambridge, MA, pp. 198–272
56. Mehta N, Rajiv S, Srinivasan K (2003) Price uncertainty and consumer search: a structural model of consideration set formation. In: *Marketing Science* 22(1):58–84
57. Meyer RJ, Sathi A (1985) A multiattribute model of consumer choice during product learning. In: *Marketing Science* 4(1):41–61
58. Montgomery H, Svenson O (1976) On decision rules and information processing strategies for choices among multiattribute alternatives. In: *Scandinavian Journal of Psychology* 17(1):283–291
59. Nedungadi P (1990) Recall and consumer consideration sets: influencing choice without altering brand evaluations. In: *Journal of Consumer Research* 17(3):263–276
60. Nierop V, Erjen BB, Wedel M, Frances PH (2010) Retrieving unobserved consideration sets from household panel data. In: *Journal of Marketing Research* 47(1):63–74
61. Otter T, Allenby GM, van Zandt T (2008) An integrated model of discrete choice and response time. In: *Journal of Marketing Research* 45(5):593–607
62. Payne JW, Bettman JR, Johnson EJ (1988) Adaptive strategy selection in decision making. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14(3):534
63. Pires T (2016) Costly search and consideration sets in storable good markets. In: *Quantitative Marketing and Economics* 14(3):157–193
64. Ratchford BT (2009) Consumer search behavior and its effect on markets. Now Publishers Inc isbn: 1601982003
65. Reiss PC (2011) Structural workshop paper—descriptive, structural, and experimental empirical methods in marketing research. In: *Marketing Science* 30(6):950–964
66. Roberts JH, Lattin JM (1991) Development and testing of a model of consideration set composition. In: *Journal of Marketing Research*:429–440
67. Rothschild M (1974) Searching for the lowest price when the distribution of prices is unknown. In: *Journal of Political Economy* 82(4):689–711
68. Rotondo J (1986) Technical note—Price as an aspect of choice in EBA. In: *Marketing Science* 5(4):391–402
69. Russo JE, Leclerc F (1994) An eye-fixation analysis of choice processes for consumer nondurables. In: *Journal of Consumer Research* 21(2):274–290
70. De los Santos B, Hortacsu A, Wildenbeest MR (2012) Testing models of consumer search using data on web browsing and purchasing behavior. In: *American Economic Review* 102(6):2955–2980
71. Seiler S (2013) The impact of search costs on consumer behavior: a dynamic approach. In: *Quantitative Marketing and Economics* 11(2):155–203
72. Seiler S, Pinna F (2016) Consumer search: evidence from path-tracking data. In: *Marketing Science* Article in Advance
73. Shugan SM (1980) The cost of thinking. In: *Journal of Consumer Research* 7(2):99–111
74. Siddarth S, Bucklin RE, Morrison DG (1995) Making the cut: modeling and analyzing choice set restriction in scanner panel data. In: *Journal of Marketing Research* 32(3):255–266
75. Simon HA (1955) A behavioral model of rational choice. In: *The Quarterly Journal of Economics* 69(1):99–118
76. Stigler GJ (1961) The economics of information. In: *The Journal of Political Economy* 69(3):213–225
77. Stüttgen P, Boatwright P, Monroe RT (2012) A satisficing choice model. In: *Marketing Science* 31(6):878–899
78 Toubia O, Hauser JR, Simester DI (2004) Polyhedral methods for adaptive choice-based conjoint analysis. In: *Journal of Marketing Research* 41(1):116–131
79. Tversky A, Sattah S (1979) Preference trees. In: *Psychological Review* 86:542–573
80. Townsend JT (1990) Serial vs. parallel processing: sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. In: *Psychological Science* 1(1):46–54
81. Tversky A (1972) Choice by elimination. In: *Journal of mathematical psychology* 9(4):341–367
82. Urban GL, Johnson PL, Hauser JR (1984) Testing competitive market structures. In: *Marketing Science* 3(2):83–112
83. Wang X (2017) Uncovering unobserved heterogeneity using mixtures of neutral networks. In: Working paper. Ivey School of Business
84. Weitzman ML (1979) Optimal search for the best alternative. In: Econometrica 47(3):641–654
85. Wolpin KI (2013) The limits of inference without theory. MIT Press, Cambridge
86. Yang L, Toubia O, De Jong MG (2015) A bounded rationality model of information search and choice in preference measurement. In: *Journal of Marketing Research* 52(April):166–183
87. Yee M, Dahan E, Houser JR, Orlin J (2007) Greedoid-based noncompensatory inference. In: *Marketing Science* 26(4):532–549