Theoretical Economics 11 (2016), 757–772 1555-7561/20160757

Monotone threshold representations

Mira Frick
Cowles Foundation for Research in Economics, Yale University

Motivated by the literature on "choice overload," we study a boundedly rational agent whose choice behavior admits a *monotone threshold representation*: There is an underlying rational benchmark, corresponding to maximization of a utility function v, from which the agent's choices depart in a menu-dependent manner. The severity of the departure is quantified by a threshold map δ, which is monotone with respect to set inclusion. We derive an axiomatic characterization of the model, extending familiar characterizations of rational choice. We classify monotone threshold representations as a special case of Simon's theory of "satisficing," but as strictly more general than both Tyson's (2008) "expansive satisficing" model as well as Fishburn (1975) and Luce's (1956) model of choice behavior generated by a semiorder. We axiomatically characterize the difference, providing novel foundations for these models.

Keywords. Bounded rationality, threshold representations, satisficing, choice, revealed preference.

JEL classification. D01, D11, D80, D81.

1. Introduction

The classical model of rational choice is based on two fundamental postulates: When choosing from a menu A, an agent considers acceptable precisely those alternatives in A that are optimal according to some underlying preference ranking; this preference ranking is assumed (I) to be independent of the particular menu at hand and (II) to satisfy the axioms of a weak order.

These two postulates are jointly called into question by a growing literature (spanning psychology, marketing, and behavioral economics) on the phenomenon of "choice overload."[1] This literature has sought to corroborate the intuition that individuals have limited cognitive resources, which are put under greater strain by larger menus of alternatives. In line with this intuition, but contrary to (I), experimental studies suggest that

Mira Frick: mira.frick@gmail.com
This paper is a revised and condensed version of Chapter 3 of my Ph.D. dissertation at Harvard University. I am indebted to Drew Fudenberg and Tomasz Strzalecki for generous advice and encouragement, and to the co-editor and two anonymous referees for detailed and insightful comments that greatly improved the paper. For helpful feedback, I also thank seminar audiences at Harvard and at the 2011 Econometric Society Summer Meetings in St. Louis, as well as Nageeb Ali, Gabriel Carroll, Jerry Green, Ryota Iijima, Sangram Kadam, Yuichiro Kamada, Jacob Leshno, Bart Lipman, Morgan McClellon, Efe Ok, Pietro Ortoleva, Drazen Prelec, Assaf Romm, Todd Sarver, and Chris Tyson.

[1]For a detailed survey, including additional references, see Broniarczyk (2008).

DOI: 10.3982/TE1547

agents' choice procedures vary with the menu, with consumers faced with larger menus resorting to greater use of simplifying choice heuristics (e.g., Payne 1976, Payne et al. 1993) and achieving lower levels of choice accuracy relative to their "ideal" benchmark (e.g., Jacoby et al. 1974, Malhotra 1982).[2] Contrary to (II), such choice heuristics typically do not involve the maximization of an underlying weak order.[3,4]

This paper proposes a parsimonious extension of the classical model that accommodates these findings. As is well known, the classical model is equivalent to the Weak Axiom and, when the domain X of alternatives is finite, to the existence of a utility function $v: X \to \mathbb{R}$ such that the set $c(A)$ of acceptable alternatives in menu $A \subseteq X$ can be represented as

$$c(A) = \Big\{x \in A : \max_{y \in A} v(y) - v(x) = 0\Big\}. \tag{1}$$

We generalize the classical utility maximizing representation in (1) to a *monotone threshold representation* of the form

$$c(A) = \Big\{x \in A : \max_{y \in A} v(y) - v(x) \leq \delta(A)\Big\}, \tag{2}$$

where δ is a threshold function mapping menus of alternatives to nonnegative real numbers and we assume that δ is weakly increasing with respect to set inclusion.

The fully rational agent of the classical model ranks alternatives in any menu according to the weak order represented by v. The monotone threshold model captures a boundedly rational agent who departs from the maximization of v in a menu-dependent way. To model this departure in a parsimonious manner, the agent is assumed to maximize a menu-dependent *semiorder*, according to which y is preferred to x in A if and only if $v(y)$ exceeds $v(x)$ by more than $\delta(A)$. The semiorder is consistent with the underlying rational benchmark v in the sense that y is preferred to x only if $v(y) > v(x)$; but it is less discriminating, with the threshold $\delta(A)$ quantifying the menu-dependent extent of the departure from v.

Going back to Luce (1956), menu-*independent* semiorders have been used to model cognitively constrained agents who either deliberately resort to simplifying heuristics that ignore small differences in some decision-relevant criteria (e.g., price differences of a few cents) and/or are simply unable to discriminate between some alternatives (e.g.,

[2]In these studies, subjects are asked to choose from sets of hypothetical alternatives (e.g., houses in Malhotra 1982), each of which is described in terms of a range of attributes. The ideal benchmark is obtained by first eliciting consumers' most preferred levels for each attribute on which the study provides information.

[3]For example, in Payne (1976), subjects facing menus of six or more alternatives frequently reported using choice procedures reminiscent of Simon's (1997) "satisficing" model and/or Tversky's (1972) "elimination-by-aspects" model.

[4]Another well documented manifestation of "choice overload" is that consumers are more likely to walk away from larger menus without making any choice (e.g., Iyengar and Lepper 2000). This is not the focus of the present paper. However, if "walking away" is modeled as an outside option x^* that is available in every menu A and has value $v(x^*)$, then the monotone threshold model in (2) could accommodate agents who never choose x^* from some binary menu $\{x^*, y_1\}$ where $v(y_1) > v(x^*)$, but sometimes choose x^* from the larger menu $\{x^*, y_1, y_2, \dots, y_n\}$.

similar varieties of toothpaste). Moving beyond this, the threshold δ of the agent's menu-*dependent* semiorder in our model is monotonic with respect to set inclusion, making the departure from the rational benchmark more severe the larger the menu. Thus, Lucean "limited discrimination" is paired with an "overload effect," enabling us to capture the finding that larger menus exacerbate agents' cognitive limitations, decreasing choice accuracy and increasing their use of simplifying heuristics.

Relying on observable choice data alone, how can an external observer test whether an agent's behavior is consistent with the monotone threshold model? Remarkably, Section 2 shows that all observable implications of the monotone threshold model are fully encapsulated by the acyclicity of a simple relation S^c derived from the agent's choice data c (Theorem 1); S^c encodes two intuitive ways in which c reveals one alternative to be superior to another according to *any* rational benchmark from which the agent could conceivably be departing. This characterization answers a question left open in Aleskerov et al. (2007)[5] and can equivalently be stated in terms of a relaxation of the Weak Axiom, which we call Occasional Optimality. Our proof provides a fully constructive procedure for obtaining a monotone threshold representation $\langle v, \delta \rangle$ for given choice data c.

In Section 3, we relate the monotone threshold model to other threshold models in the literature and build on the characterization result of Section 2 to provide new foundations for these models. We show that adding the well known Contraction axiom (Sen's α) to Occasional Optimality yields Luce's (1956) model of choice generated by a menu-independent semiorder, which differs from the monotone threshold model in that it is consistent with postulate (I) of the classical model and does not accommodate the overload effect. By contrast, adding the Strong Expansion axiom (Sen's β) to Occasional Optimality yields Tyson's (2008) "expansive satisficing" model; this is consistent with postulate (II) of the classical model, in that the agent's preference over the options in any given menu is assumed to be a *weak order* rather than a semiorder, and hence cannot accommodate agents who in the face of large menus resort to heuristics that ignore small differences between alternatives. Finally, the intersection of Luce's and Tyson's models is precisely the classical model, and all aforementioned models are special cases of Simon's theory of "satisficing" as axiomatized by Aleskerov et al. (2007).

Other related literature

In addition to the aforementioned papers, our model relates to Ortoleva's (2013) representation of an agent who dislikes large menus because of the greater "cost of thinking" involved in choosing from them. However, at a methodological level, Ortoleva's primitive is a preference over lotteries of menus, whereas our primitive is the agent's

[5]In Chapter 5 of Aleskerov et al. (2007), they introduce the concept of a monotone threshold representation, which they refer to as "utility maximization with an isotone threshold." However, while Aleskerov et al. provide necessary and sufficient conditions for a choice function to admit a *general* threshold representation, where the threshold map δ is not required to be nondecreasing with respect to set inclusion (we discuss this model in Section 3), their study of monotone threshold representations (pp. 190–193) only establishes a (straightforward) necessary condition for this more restrictive type of representation to exist.

choice *from* menus, which is arguably more readily observable. At the conceptual level, Ortoleva's agent anticipates always choosing an *optimal* element from any menu, but might sometimes dislike larger menus because identifying their optimal elements entails a greater cost of thinking; our agent, by contrast, does not always choose an optimal (according to the underlying rational benchmark) element from every menu, precisely because his cognitive limitations prevent him from doing so.

More broadly, our approach in this paper fits into an emerging literature in decision theory that seeks to characterize bounded rationality in terms of axioms on observable choice behavior. Some deviations from the fully rational paradigm that have been studied include status quo bias (Masatlioglu and Ok 2005), framing effects (Salant and Rubinstein 2008), sequential elimination of options (Mandler et al. 2012, Manzini and Mariotti 2007, Manzini and Mariotti 2012), limited attention (Masatlioglu et al. 2012, Ellis 2016), and sequential consideration of options (Caplin and Dean 2011, Masatlioglu and Nakajima 2013).

Among these, our paper relates most closely to Masatlioglu et al. (2012) and Manzini and Mariotti (2012). However, Masatlioglu et al. (2012) study an agent who maximizes a stable, menu-independent weak order in any given menu, but departs from the classical paradigm in that this maximization is carried out only over a limited subset of alternatives from the menu (his "consideration set"). This departure can be viewed as a reaction to choice overload, which is in some sense the opposite of our model: Our agent always considers *all* items in any given menu, but as a result of the taxing nature of this process perceives coarser preferences in larger menus. Manzini and Mariotti's (2012) agent also employs the heuristic of ignoring small differences in some decision-relevant criteria, but their model prescribes the successive application of *multiple* semiorders (each representing a different decision-relevant dimension) that, in contrast to our overload effect, are again assumed to be menu-independent. Another distinction with Masatlioglu et al. (2012) and Manzini and Mariotti (2012) (and many of the other papers cited in the preceding paragraph) is that our primitive is a choice *correspondence* rather than a single-valued choice function. This is key in enabling us to capture an agent who perceives coarser preferences when faced with larger menus and hence might choose different alternatives from the same menu on different occasions. Since this is the only departure from rational choice that we seek to model, an agent satisfying the monotone threshold axioms but whose choice correspondence is single-valued, is in fact fully rational.

2. Monotone threshold representations

Throughout this paper, let $X \neq \varnothing$ denote a finite set of *alternatives* and let $\mathcal{A} := \{A \subseteq X : A \neq \varnothing\}$ denote the set of *menus* (which are assumed to be nonempty). Letters A and B always denote menus, and x, y, and z denote alternatives. A *choice correspondence* on X is a map $c : \mathcal{A} \to \mathcal{A}$ such that $c(A) \subseteq A$ for all A; by definition, we only consider nonempty choice correspondences. We study the class of choice correspondences that admit a monotone threshold representation.

Definition 1. A choice correspondence c on X admits a *threshold representation* if there exist functions $v: X \to \mathbb{R}$ and $\delta: \mathcal{A} \to \mathbb{R}_+$ such that for every A,

$$c(A) = \Big\{x \in A : \max_{y \in A} v(y) - v(x) \leq \delta(A)\Big\}.$$

We call v the *fully rational benchmark* and call δ the *departure threshold* of the representation. The threshold representation $\langle v, \delta \rangle$ is called a *monotone threshold representation* (MTR) if δ is nondecreasing with respect to set inclusion, that is, $\delta(A) \leq \delta(B)$ whenever $A \subseteq B$.

As motivated in the Introduction, the monotone threshold model captures two anomalies: The agent has "limited discrimination" between the alternatives in any menu A, represented by the fact that his choices from A do not maximize v, but rather the *semiorder*[6] according to which y is preferred to x if and only if $v(y) - v(x) > \delta(A)$. Moreover, there is an "overload effect," captured by the assumption that the extent $\delta(A)$ of the departure in A from the rational benchmark v is more severe the larger is A. We interpret $c(A)$ as the set of alternatives that the agent considers acceptable after contemplating the entire menu A; these are the alternatives that we might observe him select on different occasions. To keep the departure from the classical model as parsimonious as possible, we are not concerned with predicting the relative *frequency* with which any particular alternative is chosen.[7] For the same reason, we do not seek to model other menu-dependent departures from fully rational choice, such as the "attraction effect,"[8] which we view as orthogonal to the choice overload phenomenon.

2.1 *Characterization*

Our main result, Theorem 1, identifies testable conditions on an agent's choice behavior c that are equivalent to c admitting a monotone threshold representation. We first define the following revealed preference relations.

Definition 2. Given a choice correspondence c on X, the induced relations R^c, Q^c, and S^c are defined as follows[9] for all $x, y \in X$:

(i) We have $x \, R^c \, y$ if and only if there exists A such that $x \in c(A)$ and $y \in A \setminus c(A)$.

[6] A relation K on X is a *semiorder* if it is *irreflexive* ($\forall x \in X \; \neg x \, K \, x$), *semitransitive* (for all $w, x, y, z \in X$ such that $w \, K \, x$ and $x \, K \, y$, we have $w \, K \, z$ or $z \, K \, y$), and satisfies the *interval order condition* (for all $w, x, y, z \in X$ such that $w \, K \, x$ and $y \, K \, z$, we have wKz or yKx).

[7] By contrast, taking as primitive an agent's stochastic choice rule, Fudenberg et al. (2015) study so-called additive perturbed utility representations. As a special case, they consider an agent who has limited discrimination, in the sense that his choices are "more uniform" the larger is the menu he faces. In contrast with the monotone threshold model, their representation assumes that for any menu A, *all* alternatives from A are chosen with strictly positive probability.

[8] See Huber et al. (1982) for the original formulation of this effect and Ok et al. (2014) for an axiomatic model.

[9] Relation R^c is familiar to the literature; see, for example, Aleskerov et al. (2007). The definitions of relations Q^c and S^c appear to be new.

(ii) We have $x\, Q^c\, y$ if and only if there exists A such that $y \in A$ and $c(A) \nsubseteq c(A \cup \{x\})$.

(iii) We have $x\, S^c\, y$ if and only if $x\, R^c\, y$ or $x\, Q^c\, y$.

In the classical model, where c is generated by a weak order W, it is easy to see that $W = R^c = Q^c$, so that W is fully revealed by either R^c or Q^c. By contrast, if c admits a monotone threshold representation $\langle v, \delta \rangle$, then the relations R^c and Q^c need not coincide and the agent's choice behavior does not in general fully reveal his rational benchmark.[10] Nevertheless, R^c and Q^c each partially reveal v:

If $x\, R^c\, y$, then there is a menu A containing both x and y such that $x \in c(A)$ but $y \notin c(A)$. Thus, $\max_A v - v(x) \leq \delta(A) < \max_A v - v(y)$, implying $v(x) > v(y)$. Intuitively, according to the rational benchmark the best element in A is preferred to y by enough of a margin to overcome the agent's limited discrimination in A, but it is not sufficiently preferred to x. Thus, x must be better than y.

If $x\, Q^c\, y$, then for some menu B containing y, there exists $w \in c(B) \setminus c(B \cup \{x\})$. Hence, $\max_B v - v(w) \leq \delta(B) \leq \delta(B \cup \{x\}) < \max_{B \cup \{x\}} v - v(w)$, which implies $v(x) = \max_{B \cup \{x\}} v > \max_B v \geq v(y)$. Intuitively, due to the overload effect, the agent is less discriminating when faced with the menu $B \cup \{x\}$ than when faced with its subset B. So if x is not better (according to the rational benchmark) than the best alternative in B, then everything considered acceptable in B would also be acceptable in $B \cup \{x\}$, along with possibly some additional elements. Since this is not the case, x must be better than everything in B, including y.

The above discussion shows that if c admits a monotone threshold representation, then *any* rational benchmark v from which the agent might conceivably be departing must extend R^c and Q^c. Hence, even if the choice data do not allow us to fully identify the rational benchmark v, we can at the very least conclude that the relation S^c, which subsumes both R^c and Q^c, must be *acyclic*: That is, for all $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n \in X$ such that $x_i\, S^c\, x_{i+1}$ for $i = 1, \ldots, n-1$, we must have $\neg x_n\, S^c\, x_1$.

Remarkably, we will see in Theorem 1 that acyclicity of S^c is not only necessary for c to admit a monotone threshold representation, but in fact completely encapsulates *all* the observable implications of the model. To fully bring out the connection with the classical model, we note that acyclicity of S^c can also be stated in terms of a relaxation of the Weak Axiom. Consider the following two equivalent formulations of the Weak Axiom:

EQUIVALENT FORMULATIONS OF THE WEAK AXIOM.

- *Formulation 1.* For all A, for all $x \in c(A)$, and for any B containing x, if $c(B) \cap A \neq \varnothing$, then $x \in c(B)$.

[10] For example, consider c on $X = \{w, x, y, z\}$ given by $c(\{w, x\}) = \{w\}$, $c(\{y, z\}) = \{z\}$, $c(\{w, y, z\}) = \{w, z\}$, and $c(A) = A$ for all other menus A. Then $R^c = \{(w, x), (z, y), (w, y)\} \neq Q^c = \{(w, x), (z, y), (z, w)\}$. Also, c admits the distinct MTRs $\langle v_1, \delta \rangle$ and $\langle v_2, \delta \rangle$, where $\delta(\{w, x\}) = \delta(\{y, z\}) = 0 = \delta(\{a\})$ for all a, $\delta(\{w, y\}) = \delta(\{w, z\}) = \delta(\{w, y, z\}) = 2$, $\delta(B) = 4$ for all other menus B, and $v_1(z) = v_2(z) = 5$, $v_1(w) = v_2(w) = 3$, $v_1(x) = v_2(y) = 2 > 1 = v_1(y) = v_2(x)$.

- *Formulation 2.* For all A, for all $x \in c(A)$, and for any B containing x, if $y \in A$, then $c(B) \subseteq c(B \cup \{y\})$.

In both formulations, a particular kind of optimality requirement is imposed on *all* elements of $c(A)$. In the language of Definition 2, Formulation 1 requires all elements in $c(A)$ to be R^c-maximal. It is easy to see that this is equivalent to Formulation 2, which imposes Q^c-maximality on all elements in $c(A)$.

At the same time, a relation K on a finite set X is acyclic if and only if every nonempty subset $A \subseteq X$ has a K-maximal element.[11] Therefore, acyclicity of $S^c = R^c \cup Q^c$ is equivalent to every menu A containing *at least one* distinguished element x_A that is *both* R^c-maximal *and* Q^c-maximal; moreover, by R^c-maximality we must have $x_A \in c(A)$. This is the content of the following relaxation of the Weak Axiom, where the changes are highlighted in boldface:

Condition 1 (Occasional optimality). For all A, **there exists** $x_A \in c(A)$ such that for any B containing x_A, the following statements hold:

(i) If $c(B) \cap A \neq \varnothing$, then $x_A \in c(B)$. **And**

(ii) If $y \in A$, then $c(B) \subseteq c(B \cup \{y\})$.

While the Weak Axiom requires that *any* alternative we might observe the agent choose from A will be S^c-optimal, Occasional Optimality requires only that *at least some* of the agent's potential choices from A are optimal, thus justifying the name of the axiom.[12]

We are now ready to state the representation theorem:

Theorem 1. *Suppose c is a choice correspondence on X. The following are equivalent:*

(i) Choice correspondence c admits a monotone threshold representation.

(ii) The relation S^c is acyclic.

(iii) Choice correspondence c satisfies Occasional Optimality.

The proof is given in Appendix A. The argument that (ii) implies (i) is fully constructive: If S^c is acyclic, then we can construct a weak order W^c on X that extends S^c. The utility v is then constructed in such a way as to represent W^c while also satisfying the following "increasing differences" property:

$$\text{If } v(x) > v(y) \text{ and } v(x) > v(w)\text{, then } v(x) - v(w) > v(y) - v(z) \text{ for all } z. \tag{3}$$

[11] Compare, for example, Kreps (1988, Propositions 2.7 and 2.8 (pp. 12–13)).

[12] The move from the universal imposition of a certain property in the Weak Axiom to the requirement that it hold only on a distinguished subset of elements in the Occasional Optimality axiom is in the same spirit as other relaxations of the Weak Axiom in the recent literature on boundedly rational choice, for example, the Reducibility axiom in Manzini and Mariotti (2012) or the WARP with Limited Attention axiom in Masatlioglu et al. (2012). Note that despite the existential formulation of these axioms, they are (at least in principle) testable by observation, because the global domain X of alternatives is assumed to be finite.

This property allows us to inductively define the threshold δ as follows:

- Set $\delta(\{x\}) := 0$ for all $x \in X$.
- If $|A| \geq 2$ and $c(A) = \operatorname{argmax}_A v$, set $\delta(A) := \max_{B \subsetneq A} \delta(B)$.
- If $|A| \geq 2$ and $c(A) \supsetneq \operatorname{argmax}_A v$, set $\delta(A) := \max_A v - \min_{c(A)} v$.

By an inductive argument involving several cases, we verify that $\langle v, \delta \rangle$ thus constructed is indeed an MTR of c.

3. Related threshold models

In this section, we contrast the monotone threshold model with related models in the literature.

Taking their cue from Simon's theory of satisficing, according to which agents choose "alternative[s] that meet or exceed specified criteria," but that are "not guaranteed to be either unique or in any sense the best,"[13] Aleskerov et al. (2007) study choice correspondences c that admit a *satisficing representation* (SR): There exist functions $u : X \to \mathbb{R}$ and $\theta : \mathcal{A} \to \mathbb{R}$ such that for every A, we have that $c(A) = \{x \in A : u(x) \geq \theta(A)\}$. Setting $u = v$ and $\delta(A) = \max_{y \in A} v(y) - \theta(A)$, it is clear that any satisficing representation can be converted into a *general* threshold representation and vice versa; but Example 1 in Appendix B exhibits a choice correspondence with a satisficing representation that does *not* admit a *monotone* threshold representation. Thus, the satisficing model incorporates the intuition of an agent who departs from maximization of a rational benchmark v in a menu-dependent manner, but unlike the monotone threshold model, it is too general to capture the fact that this departure is due to choice overload caused by larger menus. In the previous section, our argument that Q^c reveals the agent's preference relied on the overload effect, but the argument for R^c did not. Hence, it comes as no surprise that the satisficing model is fully characterized by acyclicity of R^c, as is shown in Aleskerov et al.[14]

At the same time, the monotone threshold model is strictly more general than the following two special cases of the satisficing model: Luce's (1956) model of choice generated by a menu-*independent* semiorder[15] is equivalent to a *constant threshold representation* (CTR) with menu-independent threshold $\delta \in \mathbb{R}_+$. Tyson (2008) studies choice correspondences c that admit an *expansive satisficing representation* (ESR): There exists a satisficing representation $\langle u, \theta \rangle$ of c with the property that whenever $A \subseteq B$ and $\max_{y \in A} u(y) \geq \theta(B)$, then $\theta(A) \geq \theta(B)$.

Like the monotone threshold model, the constant threshold model satisfies limited discrimination, violating postulate (II) of the classical model (as formulated in the Introduction). But since the departure threshold is independent of menu size, it is consistent

[13] Simon (1997, p. 125).

[14] Compare Aleskerov et al. (2007, Corollary 5.4 (p. 167)). Aleskerov et al. refer to acyclicity of R^c as the Strong Axiom of Revealed Strict Preference and refer to satisficing representations as "overvalue" choice rules.

[15] The model was later axiomatized by Jamison and Lau (1973) and Fishburn (1975).

with postulate (I) and does not capture an additional overload effect. By contrast, the expansive satisficing model can accommodate an overload effect, whereby larger menus make agents less discriminating, but unlike the monotone threshold model, it is consistent with postulate (II) of the classical model, because Tyson's agent can be seen as "locally rational": Tyson[16] shows that the ESR model is equivalent to the agent maximizing a menu-dependent preference relation P_A in each A, where these preference relations are coarser than an underlying weak order W,[17] and more so the larger the menu,[18] but additionally, *each menu-dependent preference P_A is itself a weak order*.

To axiomatically elucidate the gap between the monotone threshold and CTR and ESR models, the following lemma builds on Theorem 1 to obtain novel foundations for the latter two representations. Recall the following two well known conditions, which are jointly equivalent to the Weak Axiom (cf. Sen 1969).

Condition 2 (Contraction/Sen's α). For all A, B such that $A \subseteq B$, we have $c(B) \cap A \subseteq c(A)$.

Condition 3 (Strong expansion/Sen's β). For all A, B such that $A \subseteq B$ and $c(B) \cap A \neq \varnothing$, we have $c(A) \subseteq c(B)$.

Lemma 1. *Suppose c is a choice correspondence on X. Then the following statements hold:*

(i) Choice correspondence c admits a CTR if and only if c satisfies Occasional Optimality and Contraction.

(ii) Choice correspondence c admits an ESR if and only if c satisfies Occasional Optimality and Strong Expansion.

See Appendix A for the proof.

Since Contraction and Strong Expansion are jointly equivalent to the Weak Axiom, the class of choice correspondences admitting *both* a CTR and an ESR is precisely the class of choice correspondences with a classical utility maximizing representation. By contrast, Example 2 (respectively, Example 3) in Appendix B exhibits a choice correspondence that admits a CTR but not an ESR (respectively, an ESR but not a CTR), showing that Strong Expansion ("local rationality") and Contraction ("menu independence") are independent in the presence of Occasional Optimality. Finally, Example 4 exhibits a choice correspondence with an MTR that admits *neither* a CTR nor an ESR, showing that the monotone threshold model can simultaneously accommodate failures of postulates (I) and (II) of the classical model. Figure 1 summarizes the relationships between the various threshold models.

We conclude with a brief discussion of Example 4, showing that the monotone threshold model is strictly more general than the ESR model. The violation of Strong

[16] Compare Tyson's (2008) Theorem 3 (p. 58) and Theorem 5B (p. 59).

[17] In the sense that $\bigcup_{A \in \mathcal{A}} P_A \subseteq W$.

[18] In the sense that if $x, y \in A \subseteq B$ and $x \, P_B \, y$, then $x \, P_A \, y$ (i.e., the relations are "nested").

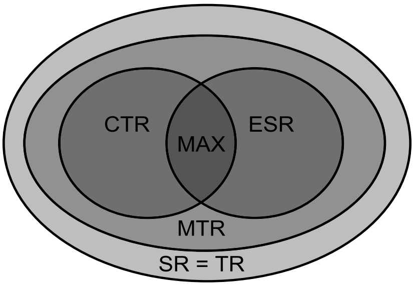

FIGURE 1. Relationship between classes of choice functions. SR, TR, MTR, ESR, CTR, and MAX denote the class of choice functions admitting a satisficing, threshold, monotone threshold, expansive satisficing, constant threshold, and utility maximizing representation, respectively.

Expansion in this example takes the following form: For some menu A and alternatives $x, y \in A$ and $z \notin A$ with $v(z) > v(y) > v(x)$, we have that $x, y \in c(A)$, $y, z \in c(A \cup \{z\})$, and $x \notin c(A \cup \{z\})$. Concretely, such a situation might arise if a consumer employs a "rule of thumb" of ignoring price differences of less than 10 cents when faced with menus of A's size or larger, and if x, y, z are (otherwise equally attractive) candy bars priced at \$0.99, \$0.95, and \$0.87, respectively. The example relies crucially on the consumer's limited discrimination: Adding z to A *indirectly* helps him choose between x and y, even though his heuristic does not *directly* discriminate between the two—a situation that cannot arise if a consumer is locally rational in the sense of Tyson's model.[19]

This example appears to be consistent with findings from consumer psychology. As discussed in the Introduction, consumers' choice accuracy from menus that exceed a certain size is in general far from perfect,[20] but various studies also suggest that the *addition* of alternatives to a menu can have an ambiguous effect on choice accuracy: summarizing these findings, Broniarczyk (2008) writes that "the addition of product alternatives to a choice set initially increases a consumer's choice accuracy, but the continued addition of product options results in a decrease in a consumer's choice accuracy."[21] While the measures of choice accuracy employed by this literature vary, one indicator is a consumer's *worst-possible* choice from a menu. Contrary to these findings, the ESR model implies that either $c(A \cup \{z\}) = \{z\}$ (so that choice accuracy is perfect) or $c(A) \subseteq c(A \cup \{z\})$ (so that the worst-possible choice from $c(A \cup \{z\})$ is at least as bad as from $c(A)$). By contrast, the monotone threshold model allows that $c(A) \nsubseteq c(A \cup \{z\}) \neq \{z\}$, and hence can accommodate improved, but still imperfect, choice accuracy as a result of adding z to A.

[19]More precisely, in Tyson's model, if $\neg z\, P_{A\cup\{z\}}\, y$ and $z\, P_{A\cup\{z\}}\, x$, then because $P_{A\cup\{z\}}$ is a weak order, we must in fact have that $y\, P_{A\cup\{z\}}\, x$ (i.e., the consumer *directly* perceives a preference for y over x in $A \cup \{z\}$). But by nestedness (see footnote 18), this implies $y\, P_A\, x$, contradicting $x \in c(A)$.

[20]For example, in Malhotra (1982), consumers' average probability of choosing the alternative from a menu A that is closest (in terms of Euclidean distance) to their ideal benchmark is 0.34 if $|A| = 15$ and 0.23 if $|A| = 25$. Recall footnote 2 for an explanation of the ideal benchmark.

[21]Broniarczyk (2008, p. 12).

Appendix A: Proofs

Proof of Theorem 1. We will prove the equivalence of (i) and (ii). The equivalence of (ii) and (iii) is immediate from the discussion in the main text.

(i) $\Rightarrow$ (ii). Suppose c admits an MTR $\langle v, \delta\rangle$. We prove that for any $x, y \in X$, $x\, S^c\, y$ implies $v(x) > v(y)$: acyclicity of S^c then follows from acylicity of $>$ on $\mathbb{R}$. Suppose that $x\, S^c\, y$, so either $x\, R^c\, y$ or $x\, Q^c\, y$. If $x\, R^c\, y$, then there is some A with $x \in c(A)$ and $y \in A \setminus c(A)$. Then $\max_A v - v(x) \leq \delta(A) < \max_A v - v(y)$, so $v(x) > v(y)$. If $x\, Q^c\, y$, then there exists A and $z \in A$ such that $y \in A$, $z \in c(A)$, and $z \notin c(A \cup \{x\})$. Then $\max_A v - v(z) \leq \delta(A) \leq \delta(A \cup \{x\}) < \max_{A \cup \{x\}} v - v(z)$, so we must have $v(x) = \max_{A\cup\{x\}} v > \max_A v \geq v(y)$.

(ii) $\Rightarrow$ (i). Suppose S^c is acyclic. By finiteness of X, we can extend S^c to a weak order[22] (in fact even a strict total order)[23] T^c on X by an inductive argument on the cardinality of $|X|$.[24] If $|X| = 1$, there is nothing to prove. If $|X| = n > 1$, pick any S^c-minimal element $a \in X$ (such an a exists, because S^c is acyclic and X is finite). Supposing that S^c has been extended to a strict total order T^c on $X \setminus \{a\}$, setting $x\, T^c\, a$ for all $x \in X \setminus \{a\}$ gives a strict total order on X that extends S^c. Thus, the set of weak orders on X extending S^c is nonempty, finite (since X is finite), and partially ordered by set inclusion, so we can pick a minimal element W^c.[25] Let E^c denote the associated equivalence relation, that is, $x\, E^c\, y$ if and only if $\neg x\, W^c\, y$ and $\neg y\, W^c\, x$.

To construct the utility $v: X \to \mathbb{R}$, pick a single element from each equivalence class of E^c and enumerate these elements in W^c-increasing order by $x_0, x_1, \ldots, x_{n-1}$ (where n is the number of equivalence classes). Then set $v(y) := 2^i$ for any $y \in X$ with $y\, E^c\, x_i$. Clearly v represents W^c, which extends S^c, so for all $x, y \in X$ such that $x\, S^c\, y$, we have $v(x) > v(y)$. Note also that v satisfies the following "increasing differences" condition: For all $w, x, y, z \in X$,

$$\text{if } v(x) > v(y) \text{ and } v(x) > v(w), \text{ then } v(x) - v(w) > v(y) - v(z). \tag{4}$$

For every A, note that the set $\operatorname{argmax}_A v$ of v-maximal elements in A is contained in $c(A)$. Indeed, if not, then $x\, R^c\, y$ for some $x \in c(A)$ and $y \in \operatorname{argmax}_A v$, implying $v(x) > v(y) = \max_A v$, a contradiction. So we can inductively define the threshold map $\delta : \mathcal{A} \to \mathbb{R}_+$ as follows:

- Set $\delta(\{x\}) := 0$ for all $x \in X$.
- If $|A| \geq 2$ and $c(A) = \operatorname{argmax}_A v$, set $\delta(A) := \max_{B \subsetneq A} \delta(B)$.

[22] Relation K on X is a *weak order* if it is *asymmetric* ($\nexists x, y \in X$ such that $x\, K\, y$ and $y\, K\, x$) and *negatively transitive* ($\forall x, y, z \in X$ such that $\neg x\, K\, y$ and $\neg y\, K\, z$, we have $\neg x\, K\, z$).

[23] Relation K on X is a *strict total order* if it is *trichotomous* ($\forall x, y \in X$ exactly one of the following holds: $x = y$, $x\, K\, y$, or $y\, K\, x$) and *transitive* ($\forall x, y, z \in X$ such that $x\, K\, y$ and $y\, K\, z$, we have $x\, K\, z$).

[24] Alternatively, one can invoke Szpilrajn's (1930) Embedding Theorem.

[25] The proof does not rely on W^c being minimal. The reason we choose a minimal weak order extension of S^c is so as not to attribute departures from rationality to the agent "unless absolutely necessary": specifically, if c satisfies the Weak Axiom, then the construction in the proof ensures that $W^c = S^c$ and the departure threshold δ is identically 0.

- If $|A| \geq 2$ and $c(A) \supsetneq \text{argmax}_A v$, set $\delta(A) := \max_{x,y \in c(A)}(v(x) - v(y)) = \max_A v - \min_{c(A)} v$.

To prove that v and δ constitute an MTR of c, we show by induction on the cardinality of A that the following statements hold:

(a) We have $c(A) = \{x \in A : \max_A v - v(x) \leq \delta(A)\}$.

(b) If $B \subseteq A$, then $\delta(B) \leq \delta(A)$.

(c) We have $\delta(A) \leq \max_{x,y \in A}(v(x) - v(y))$, and if $|A| \geq 2$ and $c(A) = \text{argmax}_A v$, then $\delta(A) \leq \max_{x,y \in A \setminus \text{argmax}_A v}((v(x) - v(y))$.

If $|A| = 1$, then (a), (b), and (c) obviously hold. If $|A| \geq 2$ and (a), (b), and (c) hold for sets of cardinality less than $|A|$, then we consider separately the cases where $c(A) = \text{argmax}_A v$ or $c(A) \supsetneq \text{argmax}_A v$.

Case 1: First, suppose that $c(A) = \text{argmax}_A v$. Then $\delta(A) := \max_{B \subsetneq A} \delta(B) = \delta(B_0)$, say. Thus, property (b) is immediate. If $|B_0| = 1$, then $\delta(A) = \delta(B_0) = 0$ and properties (a) and (c) are equally obvious. So suppose that $|B_0| \geq 2$. Since $\delta(B_0) \geq 0$ by inductive hypothesis, we certainly have $c(A) = \text{argmax}_A v \subseteq \{x \in A : \max_A v - v(x) \leq \delta(B_0)\}$. To prove (a), we must also show that if $w \in A \setminus \text{argmax}_A v$, then $\max_A v - v(w) > \delta(B_0)$. There are two cases to consider:

Suppose first that $\max_A v = \max_{B_0} v$. By (b) applied to sets of cardinality less than $|A|$, we can assume that $|B_0| = |A| - 1$, so $B_0 \cup \{z\} = A$ for some $z \in A$. Then we must have $c(B_0) = \text{argmax}_{B_0} v$: Indeed, if not, we have $c(B_0) \nsubseteq c(B_0 \cup \{z\}) = c(A) = \text{argmax}_A v$. But then $z\ Q^c\ y$ for all $y \in B_0$, so $v(z) > \max_{B_0} v$, contradicting $\max_A v = \max_{B_0} v$. Thus, by inductive hypothesis and since $|B_0| \geq 2$, the second part of (c) applies to B_0 and yields $\delta(B_0) \leq \max_{x,y \in B_0 \setminus \text{argmax}_{B_0} v}(v(x) - v(y)) = \max_{x,y \in B_0 \setminus \text{argmax}_A v}(v(x) - v(y))$.

Suppose now that $\max_A v > \max_{B_0} v$. Then by (c) applied to B_0, we again get $\delta(B_0) \leq \max_{x,y \in B_0}(v(x) - v(y)) = \max_{x,y \in B_0 \setminus \text{argmax}_A v}(v(x) - v(y))$.

So in either case, $\delta(B_0) \leq \max_{x,y \in B_0 \setminus \text{argmax}_A v}(v(x) - v(y))$. But note that the increasing differences property (4) implies that $\max_{x,y \in B_0 \setminus \text{argmax}_A v}(v(x) - v(y)) < \max_A v - v(w)$ for all $w \in A \setminus \text{argmax}_A v$. Thus, for all $w \in A \setminus \text{argmax}_A v$, $\max_A v - v(w) > \delta(B_0)$, establishing (a). Finally, property (c) holds for A, because by the above we have $\delta(A) = \delta(B_0) \leq \max_{x,y \in B_0 \setminus \text{argmax}_A v}(v(x) - v(y)) \leq \max_{x,y \in A \setminus \text{argmax}_A v}(v(x) - v(y))$.

Case 2: Now suppose that $c(A) \supsetneq \text{argmax}_A v$. Then $\delta(A) := \max_A v - \min_{c(A)} v > 0$. Thus, $c(A) \subseteq \{x \in A : \max_A v - v(x) \leq \delta(A)\}$ is immediate. Conversely, if $z \in A \setminus c(A)$, then $y\ R^c\ z$ for any $y \in c(A)$, so $v(z) < \min_{c(A)} v$, whence $\max_A v - v(z) > \delta(A)$. This proves (a). Property (c) is immediate by definition of $\delta(A)$. Finally, to prove (b), consider $B \subseteq A$. If $|B| = 1$, then $\delta(B) = 0 < \delta(A)$. So suppose $|B| \geq 2$. Again there are two cases:

First suppose that $\max_B v = \max_A v$. By inductive hypothesis, we can assume $|B| = |A| - 1$, say $B \cup \{z\} = A$ for some $z \in A$. Then $\max_B v = \max_A v \geq v(z)$, so $\neg z\ Q^c\ y$ for some $y \in B$, whence $c(B) \subseteq c(B \cup \{z\}) = c(A)$. Thus, $\min_{c(B)} v \geq \min_{c(A)} v$. So if $B \supsetneq \text{argmax}_B v$, then $\delta(B) := \max_B v - \min_{c(B)} v \leq \max_A v - \min_{c(A)} v = \delta(A)$, as required. At the same time, if $c(B) = \text{argmax}_B v$, then the second part of (c) applied to B yields $\delta(B) \leq \max_{x,y \in B \setminus \text{argmax}_B v}(v(x) - v(y)) < \max_A v - \min_{c(A)} v$ by the increasing differences property (4). So again $\delta(B) \leq \delta(A)$.

Alternatively, suppose $\max_B v < \max_A v$. Then part (c) applied to B yields $\delta(B) \leq \max_{x,y \in B}(v(x) - v(y)) < \max_A v - \min_{c(A)} v$ by the increasing differences property (4). So again $\delta(B) \leq \delta(A)$, completing the proof. □

Proof of Lemma 1. (i) Suppose first that c admits a CTR $\langle v, \delta \rangle$. Then Occasional Optimality holds by Theorem 1. To prove Contraction, suppose that $x \in c(B) \cap A$ with $A \subseteq B$. Then $\max_B v - v(x) \leq \delta$ and $\max_B v \geq \max_A v$, so $\max_A v - v(x) \leq \delta$, whence $x \in c(A)$.

For the converse, suppose that c satisfies Occasional Optimality and Contraction. By Theorem 1, c admits an MTR, $\langle v, \delta \rangle$, say. By Luce (1956), c admits a CTR if and only if c is generated by a semiorder[26] (a simple proof can be found in Aleskerov et al. 2007).[27] Define the pairwise revealed preference relation P^c on X by $x \, P^c \, y$ if and only if $x \neq y$ and $c(\{x, y\}) = \{x\}$. Then it is sufficient to prove that c is generated by P^c and that P^c is a semiorder. To see that P^c generates c, fix B. If $x \in c(B)$ and $y \in B$, then applying Contraction with $A = \{x, y\}$ yields $x \in c(\{x, y\})$, so $\neg y \, P^c \, x$. Conversely, if $x \in B \setminus c(B)$, then there exists $y \in B$ such that $v(y) - v(x) > \delta(B) \geq \delta(\{x, y\})$. So $c(\{x, y\}) = \{y\}$, whence $y \, P^c \, x$. Hence, $c(B) = \{x \in B : \forall y \in B \neg y \, P^c \, x\}$, that is, P^c generates c. We now show that P^c is a semiorder: Irreflexivity is clear. If $x \, P^c \, y$ and $y \, P^c \, z$, then since P^c generates c and c is nonempty, we must have $c(\{x, y, z\}) = \{x\}$. Thus, $v(x) - v(z) > \delta(\{x, y, z\}) \geq \delta(\{x, z\}) \geq 0$, so that $x \neq z$ and $c(\{x, z\}) = \{x\}$, that is, $x \, P^c \, z$. This shows that P^c is transitive. To prove that P^c is semitransitive, suppose to the contrary that $x \, P^c \, y$ and $y \, P^c \, z$, but $\neg x \, P^c \, w$ and $\neg w \, P^c \, z$. The relationships $x \, P^c \, y$ and $\neg x \, P^c \, w$ together with transitivity of P^c imply that $\neg y \, P^c \, w$, so $w \in c(\{x, y, w\})$ and $y \notin c(\{x, y, w\})$, whence $w \, S^c \, y$. But $y \, P^c \, z$ and $\neg w \, P^c \, z$ imply $z \in c(\{w, z\})$ and $z \notin c(\{w, z, y\})$, so $y \, Q^c \, w$, whence $y \, S^c \, w$. This contradicts acyclicity of S^c (which holds by Theorem 1). Finally, to prove that P^c satisfies the interval order condition, suppose to the contrary that $x \, P^c \, y$ and $w \, P^c \, z$, but $\neg x \, P^c \, z$ and $\neg w \, P^c \, y$. Then by transitivity of P^c, we also have $\neg y \, P^c \, z$ and $\neg z \, P^c \, y$. Hence, $z \in c(\{x, y, z\})$ and $y \notin c(\{x, y, z\})$, so $z \, S^c \, y$; also, $y \in c(\{w, y, z\})$ and $z \notin c(\{w, y, z\})$, so $y \, S^c \, z$. This again contradicts acyclicity of S^c.

(ii) Tyson (2008) proves that c admits an ESR if and only if c satisfies Strong Expansion and P^c is acyclic,[28] where P^c is the pairwise revealed preference relation defined in part (i) above. By Theorem 1, Occasional Optimality is equivalent to acyclicity of S^c, which implies acyclicity of P^c since $P^c \subseteq S^c$. So it suffices to prove that if c satisfies Strong Expansion, then $S^c \subseteq P^c$. Suppose that $x \, S^c \, y$, so either $x \, R^c \, y$ or $x \, Q^c \, y$. If $x \, R^c \, y$, then there is B such that $x \in c(B)$ and $y \in B \setminus c(B)$. Applying Strong Expansion with $A = \{x, y\}$ gives $c(\{x, y\}) = \{x\}$, so $x \, P^c \, y$. If $x \, Q^c \, y$, then there is A such that $y \in A$ and $c(A) \nsubseteq c(A \cup \{x\})$. Applying Strong Expansion with $B = A \cup \{x\}$ yields $c(A \cup \{x\}) = \{x\}$. Thus $x \, R^c \, y$, and we are back in the previous case. □

[26]Recall the definition of a semiorder in footnote 6.
[27]Compare Aleskerov et al. (2007) Theorem 3.2 (p. 66).
[28]Compare Tyson (2008) Theorem 5B (p. 59). Note that Tyson refers to acyclicity of P^c as base acyclicity.

Appendix B: Separating examples for Section 3

Example 1 (SR $\not\subseteq$ MTR). Let c be the choice correspondence on $X = \{x, y, z\}$ with satisficing representation $\langle u, \theta \rangle$, where $u(x) = 1$, $u(y) = 2$, $u(z) = 3$, $\theta(\{a\}) = \theta(\{a, b\}) = 1$ for all a, b, and $\theta(\{x, y, z\}) = 3$. Then $c(\{x, z\}) = \{x, z\}$ and $c(\{x, y, z\}) = \{z\}$, so $y\, Q^c\, z$ and $z\, R^c\, y$, producing the cycle $y\, S^c\, z\, S^c\, y$. Hence, by Theorem 1, c does not admit an MTR. ♢

Example 2 (CTR $\not\subseteq$ ESR). Let c be the choice correspondence on $X = \{x, y, z\}$ induced by the CTR $\langle v, \delta \rangle$, where $v(z) = 3$, $v(y) = 2$, $v(x) = 1$, $\delta = 1$. Then $c(\{x, y\}) = \{x, y\}$ and $c(\{x, y, z\}) = \{y, z\}$. So setting $A = \{x, y\}$ and $B = \{x, y, z\}$ yields a violation of Strong Expansion. Thus, by Lemma 1, c does not admit an ESR. ♢

Example 3 (ESR $\not\subseteq$ CTR). Let c on $X = \{x, y, z\}$ be given by $c(\{x, y\}) = \{x, y\}$, $c(\{x, z\}) = \{z\} = c(\{y, z\})$, and $c(X) = X$. Setting $A = \{x, z\}$, $B = X$ yields a violation of Contraction, so by Lemma 1, c does not admit a CTR. But $\langle u, \theta \rangle$ with $u(z) = 3$, $u(y) = 2 = u(x)$, $\theta(X) = 1$, $\theta(\{y, z\}) = \theta(\{x, z\}) = \theta(\{z\}) = 3$, and $\theta(A) = 2$ for all other A is an ESR of c. ♢

Example 4 (MTR $\not\subseteq$ ESR $\cup$ CTR). Let c be the choice correspondence on $X = \{x, y, z\}$ induced by the MTR $\langle v, \delta \rangle$, where $v(z) = 4$, $v(y) = 2$, $v(x) = 1$, $\delta(\{a\})$, $\delta(\{a, b\}) = 1$ for all $a, b \in X$ with $a \neq b$, and $\delta(X) = 2$. Then $c(\{y, z\}) = \{z\}$, $c(\{x, y\}) = \{x, y\}$, and $c(\{x, y, z\}) = \{y, z\}$. So setting $A = \{y, z\}$ and $B = \{x, y, z\}$ yields a violation of Contraction, whence c does not admit a CTR. And setting $A = \{x, y\}$ and $B = \{x, y, z\}$ yields a violation of Strong Expansion, whence c does not admit an ESR. ♢

References

Aleskerov, Fuad, Denis Bouyssou, and Bernard Monjardet (2007), *Utility Maximization, Choice and Preference*, second edition. Springer, Berlin. [759, 761, 764, 769]

Broniarczyk, Susan M. (2008), "Product assortment." In *Handbook of Consumer Psychology* (Curtis P. Haugtvedt, Paul M. Herr, and Frank R. Kardes, eds.), 755–779, Psychology Press, New York. [757, 766]

Caplin, Andrew and Mark Dean (2011), "Search, choice, and revealed preference." *Theoretical Economics*, 6, 19–48. [760]

Ellis, Andrew (2016), "Foundations for optimal inattention." Working paper, London School of Economics. [760]

Fishburn, Peter C. (1975), "Semiorders and choice functions." *Econometrica*, 43, 975–977. [757, 764]

Fudenberg, Drew, Ryota Iijima, and Tomasz Strzalecki (2015), "Stochastic choice and revealed perturbed utility." *Econometrica*, 83, 2371–2409. [761]

Huber, Joel, John W. Payne, and Christopher Puto (1982), "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis." *Journal of Consumer Research*, 9, 90–98. [761]

Iyengar, Sheena S. and Mark R. Lepper (2000), "When choice is demotivating: Can one desire too much of a good thing?" *Journal of Personality and Social Psychology*, 79, 995–1006. [758]

Jacoby, Jacob, Donald E. Speller, and Carol A. Kohn (1974), "Brand choice behavior as a function of information load." *Journal of Marketing Research*, 11, 63–69. [758]

Jamison, Dean T. and Lawrence J. Lau (1973), "Semiorders and the theory of choice." *Econometrica*, 41, 901–912. [764]

Kreps, David M. (1988), *Notes on the Theory of Choice*. Westview Press, Boulder. [763]

Luce, R. Duncan (1956), "Semiorders and a theory of utility discrimination." *Econometrica*, 24, 178–191. [757, 758, 759, 764, 769]

Malhotra, Naresh K. (1982), "Information load and consumer decision making." *Journal of Consumer Research*, 8, 419–430. [758, 766]

Mandler, Michael, Paola Manzini, and Marco Mariotti (2012), "A million answers to twenty questions: Choosing by checklist." *Journal of Economic Theory*, 147, 71–92. [760]

Manzini, Paola and Marco Mariotti (2007), "Sequentially rationalizable choice." *American Economic Review*, 97, 1824–1839. [760]

Manzini, Paola and Marco Mariotti (2012), "Choice by lexicographic semiorders." *Theoretical Economics*, 7, 1–23. [760, 763]

Masatlioglu, Yusufcan and Daisuke Nakajima (2013), "Choice by iterative search." *Theoretical Economics*, 8, 701–728. [760]

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay (2012), "Revealed attention." *American Economic Review*, 102, 2183–2205. [760, 763]

Masatlioglu, Yusufcan and Efe A. Ok (2005), "Rational choice with status quo bias." *Journal of Economic Theory*, 121, 1–29. [760]

Ok, Efe A., Pietro Ortoleva, and Gil Riella (2014), "Revealed (p)reference theory." *American Economic Review*, 105, 299–321. [761]

Ortoleva, Pietro (2013), "The price of flexibility: Towards a theory of thinking aversion." *Journal of Economic Theory*, 148, 903–934. [759]

Payne, John W. (1976), "Task complexity and contingent processing in decision making: An information search and protocol analysis." *Organizational Behavior and Human Performance*, 16, 366–387. [758]

Payne, John W., James R. Bettman, and Eric J. Johnson (1993), *The Adaptive Decision Maker*. Cambridge University Press, New York. [758]

Salant, Yuval and Ariel Rubinstein (2008), "(A, f): Choice with frames." *Review of Economic Studies*, 75, 1287–1296. [760]

Sen, Amartya K. (1969), "Quasi-transitivity, rational choice and collective decisions." *Review of Economic Studies*, 36, 381–393. [765]

Simon, Herbert A. (1997), *Models of Bounded Rationality*, Vol. 3. MIT Press, Cambridge, Massachusetts. [758, 764]

Szpilrajn, Edward (1930), "Sur l'extension de l'ordre partiel." *Fundamenta Mathematicae*, 16, 386–389. [767]

Tversky, Amos (1972), "Elimination by aspects: A theory of choice." *Psychological Review*, 79, 281–299. [758]

Tyson, Christopher J. (2008), "Cognitive constraints, contraction consistency, and the satisficing criterion." *Journal of Economic Theory*, 138, 51–70. [757, 759, 764, 765, 769]

Co-editor Faruk Gul handled this manuscript.

Submitted 2013-5-28. Final version accepted 2015-8-7. Available online 2015-8-12.