Bounded Rationality in Choice Theory: A Survey*

Geoffroy de Clippel† and Kareen Rozen‡

Revised February 2023

Abstract

A vibrant literature incorporates elements of bounded rationality into choice theory. We survey this work, discussing five central ways in which the literature has modeled departures from the rational choice procedure. We discuss the variety of purposes axiomatic choice characterizations serve for these positive theories; explore the permissiveness of these theories; synthesize the welfare debate; and assess important future directions.

1 Introduction

The first order of business in a survey, especially one whose title has a slightly nebulous meaning, is to explain what it is about. We will thus begin by creating a common understanding, at least for our purposes here, of the terms *bounded rationality* and *choice theory*; and by extension, an understanding of what strands of the literature we will and will not discuss.

Evidence suggests that decision makers can have difficulty comprehending or paying attention to the available options; they may be fickle, changing their decisions based on seemingly irrelevant contextual information; they can be subject to a variety of judgment biases; and may be satisfied with outcomes that are simply good enough. Bounded rationality can capture such features, but is not an umbrella term for them. Nor is the term

*First version May 2022. We are grateful for valuable comments and suggestions from the editor, three anonymous referees, Victor Aguiar, Jose Apesteguia, Federico Echenique, Xavier Gabaix, Paola Manzini, Marco Mariotti, Yusufcan Masatlioglu, Hiroki Nishimura, Angelo Petralia, John Rehbeck, Ariel Rubinstein, Larry Samuelson, Yuval Salant, Rani Spiegler, Gerelt Tserenjigmid, and Christopher Tyson.

†Department of Economics. Email address: declippel@brown.edu.

‡Department of Economics. Email address: kareen_rozen@brown.edu.

a mere acknowledgement that humans' optimizing abilities fall short of super-computers. More than this, the term calls to mind structured and limited (*bounded*) departures from rationality. Boundedly-rational decision makers act 'coherently,' at least in some sense of that word. Our definition borrows from Rubinstein (1998, p. 2), who encouraged many strands of this literature: we have in mind decision makers who "make deliberate decisions by applying procedures that guide their reasoning about 'what' to do, and probably also about 'how' to decide." Complicating our paternal instincts to help them, such decision makers may or may not be aware of their 'irrationality', and even if aware, may or may not deem their choice procedures mistaken. In an elegant survey of evidence for bounded rationality, and why it should be incorporated into models, Conlisk (1996, p. 671) points to the importance of deliberation costs: "heuristics often provide an adequate solution cheaply whereas more elaborate approaches would be unduly expensive....[this tradeoff] tends to be pushed out of sight in economics by the emphasis on unbounded rationality."

Since Conlisk (1996)'s appeal, there has been a great deal of economic research incorporating bounded rationality. In this survey, we are concerned with the literature on individual decision making. We draw a slightly fuzzy line demarcating *choice theory* from its cousin, *decision theory.* Both are concerned with understanding the choices individuals make. Despite this common goal, they follow different approaches. The *modus operandi* in decision theory is to start with a preference ordering over some domain, thereby starting from the vantage point of rationality.[1] The aim is to impose axioms on the preference to characterize more specific functional forms; for example, expected utility for risk preferences, discounted utility for time preferences, preference for flexibility or commitment when facing menus of options.

Rather than taking the maximization of a preference ordering as a given, choice theory is rooted in the choices themselves. At its broadest level, choice theory is simply the study of choice patterns and the models that may generate them. Of course, since preference orderings can be summarized by the pairwise choices they generate, decision theory could be seen as a subfield of choice theory. However, the purpose of our survey is to focus on

[1]This does not mean that behavioral features are not of interest to decision theorists. Rather, part of the work is to construct the domain of the preference ordering in such a way that the behavior in question is expressible through preference maximization; this is what Lipman and Pesendorfer (2013) call the 'Krepsian' methodology, in recognition of that author's influence on the field. Spiegler (2013) discusses limitations of this approach: for instance, he is concerned with the potential tension between behavioral considerations and the assumption that the decision maker has the ability to formulate a complete and transitive preference over potentially complex objects, or the assumed ability to anticipate the psychological phenomenon in question when facing such objects (and without facing the behavioral bias at that point). In such cases, the work may provide foundations for behavioral effects, but not necessarily from the vantage point of bounded rationality.

choice patterns that need not be consistent with the maximization of a preference ordering (as such, the subset of choice theory that we study here is not intended to be normative in nature). Still, the demarcation line between choice theory and decision theory can be blurry: representation theorems using a weaker notion of preference can explain irrational choice patterns (e.g., incomplete preferences, intransitive indifferences). It should not be surprising, then, that progress in decision theory can inspire choice theories of bounded rationality, and vice versa.

More broadly, the choice-theory literature may draw on different sources of inspiration. An array of experimentally-observed anomalies from the economics, marketing and psychology literatures, such as status-quo bias, the attraction effect, and reference dependence, commonly inspires this work. Simple introspection about the confines of rationality (e.g., 'I am often indecisive'), or the influential suggestions of others (e.g., Herbert Simon), may play a role. And theories may be inspired by psychological evidence (e.g., imperfect perception). Indeed, there is some intersection with an interesting literature that builds decision-making models based on facts about human cognition, in the hope of simultaneously accounting for several biases; see, for instance, the surveys of Fehr and Rangel (2011), Rabin (2013) and Woodford (2020). This literature is not always interested in deriving the full testable implications of the models it develops, but we will see that it sometimes inspires choice-theoretic work that does. Such a choice-theoretic analysis may answer several questions of interest, such as (i) is the model falsifiable? (ii) can the testable implications be broken down into insightful principles or axioms? (iii) might the model have the same testable implications as other models whose motivations and structure seem entirely unrelated?

Our target literature is rich in several respects. Firstly, as may be inferred from above, bounded rationality may take many forms. Secondly, theories may be designed to explain different anomalies (and some theories may explain several). Thirdly, theories may differ in the type of 'choice data' they generate or require for testing, be it a single choice per menu, a set of choices per menu, or a probability distribution over choices per menu (i.e., stochastic choice functions, in which there has been a recent surge of interest). In the apt words of one referee, a goal of this survey is to be "helpful for general economists to wade their way through the maze of models." Towards this end, one must take a conceptual stand on how to taxonomize this literature. Though no taxonomy can be perfect, we see intellectual coherence in examining five central ways one may relax rationality's tenets. Rather than assuming the decision maker expertly maximizes a perfectly known, stable preference ordering in each menu she faces, works have considered situations where (i) she has an imperfect understanding of her preferences, (ii) she must somehow aggregate multiple, conflicting preferences, (iii) she may overlook some alternatives in the menu, (iv) she is subject to contextual effects, and (v) she is content making suboptimal, but not subpar, decisions. We group works at this conceptual level, rather than which anomalies they explain, or which choice data they presume.

These five central types of departures from the rational choice procedure serve not only as an organizing principle in this survey, but also clarify the forms of bounded rationality this survey will consider. With this understanding of the choice-theory literature in mind, which we develop in Sections 2 and 3, we then delve further into the questions this literature both studies and raises. We begin with the literature's *modus operandi*. The papers covered in this survey typically examine the axiomatic systems underlying the models proposed. While one might associate axioms with normative principles, bounded rationality theories are, clearly, not intended to be normative in nature. We discuss in Section 4 several important purposes such characterizations nonetheless serve. Next, we examine some difficult questions this literature raises. In particular, should we worry that these new choice theories are too permissive? And are they explaining the right patterns of choice? We explore these matters in Section 5. Finally, what normative implications can one draw from choices that are inconsistent with preference maximization? We synthesize in Section 6 a factious discourse on welfare inference. We conclude in Section 7 with our assessment of important future directions for research.

We now turn our attention to synthesizing our target literature. Knowing both its size and our own bounded rationality, we presume this survey has at least a handful of important and inadvertent omissions. For these we apologize.

2 Setting the stage

We now lay out the basic framework in preparation for the ideas to come. Rationality serves as our running example in this preliminary section, as its aspects are varyingly the point of departure or the inspiration for bounded rationality theories.

We consider a finite set X of conceivable options. A choice problem involves a nonempty subset $S \subseteq X$, which we call a menu. Denote by $\mathcal{P}(X)$ the set of all menus. Some strands of the literature in decision theory impose more structure on the options in X, such as when studying choices over lotteries or consumption streams. In those cases, it may be sensible to impose more structure on preferences and choices as well (e.g., expected utility maximization and departures from it). For the most part, this survey stays at a more abstract level, so that the main lessons remain valid whatever the structure imposed on X. Indeed, the basic choice-theoretic framework discussed here represents the common denominator of all choice situations. But some theories of bounded rationality naturally pertain to observable aspects or dimensions of the choice object and we will occasionally delve into those.

A *choice theory* provides a description of a process that the decision maker, henceforth DM, may use to make choices. For instance, rational choice theory, in its simplest incarnation, says that the DM has a strict preference ranking over X and always chooses her best available option. This defines a *choice function* $c : \mathcal{P}(X) \to X$, where $c(S) \in S$ represents the DM's choice from the menu S. Of course, we cannot observe the mental process described by rationality–or any theory, for that matter.[2] We can, however, understand the choice patterns a theory generates, even if (as will become clearer below) this does not provide irrefutable evidence of the DM's underlying choice process.

For example, suppose a rational DM is about to book a hotel for a vacation, after thoroughly reviewing three options: a hotel in New York, one in Las Vegas, and a ski resort in Colorado. The DM decides on New York, and is about to complete the purchase when she learns that the ski resort is now fully booked. If her preference ranking is stable, as rationality requires, then New York remains her best choice and she should pursue the reservation, as planned. This is an instance of the following well-known property:

Independence of Irrelevant Alternatives (IIA) For any menus R, S with $R \subseteq S$, if $c(S) \in R$ then $c(S) = c(R)$.

IIA *characterizes* all choice functions arising under rationality: a choice function can arise from the maximization of a strict preference ordering *if, and only if,* it satisfies IIA.[3]

In general, a property which characterizes the choice patterns of a theory may be considered insightful and meaningful without reference to the theory itself. A DM may find IIA normatively appealing (and follow it), without their choice process truly being rational (though their choice patterns will appear rational). Indeed, building on Rubinstein (1998), consider a DM who is partial towards a subset $F \subset X$ of food items in the cafeteria. The

[2]Some have argued, though, that observing neural activity can help predict choices; see, for instance, Smith et al. (2014) and references therein. At this point, choice data is still more readily available.

[3]Why is this? First suppose we have a rational choice function c. Since $c(S)$ is the most preferred alternative in S, then it is also most preferred in any subset R to which it belongs. Hence $c(S) = c(R)$, and IIA is satisfied. Conversely, suppose c is a choice function satisfying IIA. We may define a complete and transitive preference P as follows: $x_1 := c(X)$ is best, $x_2 := c(X \setminus \{x_1\}$ is second best, $x_3 := c(X \setminus \{x_2\}$ is third best, etc. Suppose, by contradiction, that the choice function c *does not* pick the P-maximal element in S for some set S. Denote by x_i that P-maximal element. By construction of P, i must the smallest index in S, so $c(S) = x_j$ for some $j > i$. But we know that $c(\{x_i, x_{i+1}, \ldots, x_{|X|}\}) = x_i$ while $c(S) \neq x_i$, and $x_i \in S \subset \{x_i, x_{i+1}, \ldots, x_{|X|}\}$. This contradicts c satisfying IIA.

cafeteria displays items in a fixed order, though the available options can vary (some racks may be empty). Moving along the cafeteria line, the DM chooses the first item that belongs to F, or the last item in the display if no element of F is available. Such a DM's choices will satisfy IIA, but their mental process is actually an instance of 'satisficing' (see Section 3.5). This points to the fact that different choice processes can have identical behavioral implications, *even when the resulting choices are rational.* We will encounter further such examples among bounded rationality theories. Ruminating further on choice processes could shed light on whether richer data would distinguish between those theories; for instance, we could compare choices before and after redesigning the cafeteria display.

As a step towards incorporating data into the analysis, it is helpful to differentiate between the set of choice functions that can potentially arise under a theory, and the choices that are actually observed. An *observed* choice function $c_{obs} : \mathcal{D} \to X$ describes the choice $c_{obs}(S)$ the DM has made from each menu S in the set of observed menus $\mathcal{D} \subseteq \mathcal{P}(X)$. If the data is consistent with a theory, then it should be possible to extend c_{obs} (i.e., fill in a choice for each unobserved problem) into a full choice function arising under that theory. This can be daunting, as $\mathcal{D}$ may be small relative to $\mathcal{P}(X)$. But testing consistency can be a straightforward process if using the right approach. This is easily demonstrable for rationality. As Samuelson (1938, 1948) pointed out in the context of consumer theory, the data $x = c_{obs}(S)$ reveals that the DM prefers x to every other element of S. A first observation is that IIA, which effectively requires asymmetry of the binary relation comprising all revealed preferences from nested sets in $\mathcal{D}$, paints an incomplete picture of rationality's empirical content unless $\mathcal{D}$ is sufficiently rich (e.g., contains all menus of sizes two and three).[4] The Strong Axiom of Revealed Preference (Houthakker 1950, Richter 1966), or SARP, characterizes the empirical content of rationality:[5]

Strong Axiom of Revealed Preference (SARP) Take any sequence $S_1, S_2, \ldots, S_n$ with $c_{obs}(S_{i+1}) \in S_i$ for all $i \in \{1, \ldots, n-1\}$. Then either $c_{obs}(S_1) \notin S_n$, or all the choices are the same ($c_{obs}(S_i) = c_{obs}(S_{i+1})$ for all i).

SARP boils down to ensuring the revealed preference is acyclic, and thus has a transitive

[4]For instance, suppose the observed choice function tells us $x = c_{obs}(\{x, y\})$, $y = c_{obs}(\{y, z\})$ and $z = c_{obs}(\{x, z\})$. The revealed preference is cyclic, in direct contradiction to rationality, but IIA is vacuously satisfied because $\{x, y, z\} \notin \mathcal{D}$. The Weak Axiom of Revealed Preference (WARP) provides an important generalization of IIA, ensuring that the revealed preference arising from any sets in the data is asymmetric. Formally, WARP says that for all R, S: $c(R), c(S) \in R \cap S \Rightarrow c(R) = c(S)$. But this still does not guarantee the ability to explain the data with a transitive preference.

[5]Houthakker (1960) phrased SARP for consumer demand. Richter (1966) generalizes to an abstract choice setting with potentially multi-valued choices, and refers to the property as Congruence.

completion to all of X.[6] If so, c_{obs} is easily extended into a rational choice function, by maximizing the completed preference out of sample. For bounded rationality theories, constructing revealed preferences may be more nuanced, but the essence of SARP – an acyclic revealed preference – can remain relevant for understanding testable implications.

Notice that SARP highlights the challenge of welfare inference. The fact that the revealed preference may be incomplete (at least without a rich-enough dataset) means that multiple preference orderings may explain the DM's choice data.[7] How do we know, then, what is welfare improving? Varian (1982) addresses this precise problem for rational but incomplete choice data. Taking a conservative approach, he asserts that x is revealed preferred to y only when *all* rationalizing preference orderings share that ranking. Complicating welfare analysis yet further, a similar multiplicity issue occurs for bounded rationality theories, even when all possible choices are observed.

So far, we have discussed only choice functions. More broadly, a theory may give rise to a *choice correspondence* $C : \mathcal{P}(X) \to \mathcal{P}(X)$, with $C(S) \subseteq S$ for each menu S. In its most general form, rationality allows for indifference, with $C(S)$ representing the set of preference maximizers in the menu S. In this case, rational choice correspondences are characterized by two properties, Sen's α and β (Sen, 1971). Condition α is a simple generalization of IIA to correspondences (it requires any choices in a menu S to also be selected in any smaller menu $R \subseteq S$ to which they belong); and, as we will see, it has found new life in recent bounded rationality theories as a property of alternatives *considered* rather than chosen. Condition β, sometimes called expansion consistency (it requires that if some choice in a subset is still chosen in a larger menu, then so are all the other choices from that subset), has played a role in some bounded rationality theories as well. We will also see that choice correspondences arise for multiple interesting reasons beyond indifference once one relaxes the assumption of rationality.

Regardless of the reason for multivalued choices, we must take a philosophical stand before analyzing an observed choice correspondence C_{obs}. Namely, should we interpret C_{obs} as representing *all* of the theoretically possible choices the DM could make (and thus test if there is a choice correspondence C under the theory such that $C_{obs}(S) = C(S)$ for each problem $S \in \mathcal{D}$), or should we worry whether other possible choices *could have been* recorded had we observed the DM for longer (and thus test if $C_{obs}(S) \subseteq C(S)$ for

[6] A relation $\succ$ is acyclic if one cannot find a sequence of options $x_1, x_2, \ldots, x_n$ comprising a $\succ$-cycle, $x_1 \succ x_2 \succ \ldots \succ x_n \succ x_1$. In general, a binary relation may be incomplete: not every pair of options need be comparable. A transitive completion of $\succ$ is an extension of $\succ$ into a complete relation that also satisfies transitivity. It is possible to find a transitive completion if, and only if, $\succ$ is acyclic.

[7] Of course, if we had a complete choice function at hand, then inferring the DM's entire preference ranking would be easy, but there would be nothing left to do, having observed all her choices already!

all $S \in \mathcal{D}$)? Under rationality, the former view yields crisp revealed preferences (all elements of $C_{obs}(S)$ are indifferent, and strictly better than those in $S \setminus C_{obs}(S)$), and the empirical content is given by the Generalized Axiom of Revealed Preference (GARP), which requires this revealed preference to be acyclic. The latter, cautious view of the data may be more realistic, but then observed choices can always be explained as arising from a rational DM who is indifferent over all of X. For this reason, one typically requires preferences to respect some natural criterion given the context (e.g., monotonicity for bundles, first-order stochastic dominance for lotteries) when taking the latter view.

It is important to keep in mind that choice correspondences describe the DM's possible selections from each menu, but ignore their relative frequencies. Letting $\Delta(\cdot)$ denote all lotteries over a finite set, such information can be captured by a *stochastic choice function* (SCF), $\sigma : \mathcal{P}(X) \to \Delta(X)$, which associates to each menu S a lottery $\sigma(S) \in \Delta(S)$. For a rational DM, who perfectly maximizes a stable preference, probabilities play no substantive role: only the support of $\sigma(S)$ matters. Probabilities become relevant when analyzing aggregate choices from rational DMs with heterogenous tastes, as suggested in the classic works Falmagne (1978), Barberà and Pattanaik (1986), and McFadden (1976). Notwithstanding, McFadden (1976, p. 371) points out that "the analyses may be on sounder behavioral ground in postulating random preference *within* each subject." Indeed, we will see that the bounded rationality literature has not only developed models of random utility but considered other sources of stochasticity as well.

Whether one considers deterministic or stochastic choice theories, observe that we have so far defined the objects of interest (c, C or σ) following the classic paradigm: they take choice menus as inputs, and output some selection(s) or lottery over the menu, as a description of the DM's decision. Yet theories could potentially enrich these objects in two different ways. One such enrichment regards the domain, affecting the very definition of a choice problem. Indeed, the menu may be paired with additional contextual information (e.g., an exogenously-known status quo, the known order in which alternatives are presented) from some possible set $\mathcal{I}_D$ that the modeler considers relevant, so that the domain is instead the product set $\mathcal{P}(X) \times \mathcal{I}_D$. The other manner of enrichment regards the range, encoding further information which is provided by the DM, whether consciously or not. The set $\mathcal{I}_R$ of potential decision-related information that may be collected could include, among other possibilities, response time data, eye-tracking information, or intermediate process data (i.e., prior to finalizing a decision). A theory may specify how such richer choice data may be used to sharpen our inferences. A conceptual difference between these two types of enrichments is that the latter is generated endogenously by the DM during the decision-making process, while the former is simply a feature of the choice setting (possibly exogenous, or possibly manipulable by the experimenter). Both enrichments may certainly coexist in a theory, though only a minority of the works in this literature invoke at least one. We come back to some of these ideas in Sections 3.4, 3.5, and 6.

3 Central concepts of bounded rationality

Rationality endows the DM with a single, stable and well-behaved (complete and transitive) preference, and assumes that she systematically identifies all available options and never fails to maximize her preference. More recent theories typically relax one or more of these assumptions. A boundedly rational DM may have difficulty distinguishing which is the better alternative. She may be conflicted between multiple objectives, and rely on rules of thumb to come to some compromise. She may not pay attention to all feasible options when determining which option is best. Perhaps her preference is not stable, varying with contextual information. Or she may simply be satisfied by fulfilling her preference to some degree, instead of maximizing it.

We follow this taxonomy as we survey recent theories of choice. This provides some structure which, we hope, helps organize one's understanding of the literature. Though useful, the taxonomy is by no means definitive. Some theories contribute to multiple categories, and we have had to make a judgment call on which category is most salient.

3.1 Imperfect Perception of Preferences

Psychophysics, a termed coined in the seminal work of Fechner (1860), is the branch of psychology studying the quantitative relationships that may exist between physical stimuli and the way we perceive them. For instance, how much louder must a sound be, or how much heavier must an object be, to be able to perceive the difference? The famous Weber-Fechner laws (Fechner 1860) hypothesize relationships between the intensity of increase in a physical stimulus and the (imperfect) ability to perceive it.

Luce (1956) applies these insights to decision making, observing that the utility of an option may have to be significantly larger than another for a DM to perceive the difference. This raises the possibility of intransitive indifference, as Luce illustrates using one's morning cup of coffee. Most would not notice the difference, and hence express indifference, between a cup of coffee with n grains of sugar and one with $n+1$ grains of sugar. Yet most people have some strict preference between getting one spoonful of sugar versus none, and certainly no one wants a pound bag of sugar poured into their coffee. Rational choice can be amended to accommodate such effects. Say the binary relation P has a *utility-threshold representation* if there exists a utility function $u : X \to \mathbb{R}$ and a threshold τ such that yPx if, and only if, $u(y) - u(x) > \tau$. The DM's possible choices from a menu S comprise those feasible options $x \in S$ for which one cannot find $y \in S$ such that yPx. Luce (1956) assumes τ is a non-negative constant, which amounts to P being a semi-order (Scott and Suppes, 1958).[8] The underlying idea is that utility differences may be too small to distinguish; but that if one can distinguish between two alternatives, then larger utility differences should be distinguishable too. Important feature of semi-orders are that they may be incomplete and have intransitive indifferences, though the term 'indifference' may be misleading in this context. A situation where a DM is truly indifferent between alternatives is conceptually different from one where she cannot distinguish between them, and Eliaz and Ok (2006) look for conditions on choice behavior identifying which is true.

A multitude of variants of Luce (1956)'s model have been studied, by varying properties of the threshold τ.[9] Moreover, several theories consider the sequential application of such thresholds or semi-orders. An early example is Tversky (1969), who envisions the alternatives as having multiple dimensions that matter to the DM. Options are assigned a utility along each dimension, and when comparing any pair of options, the DM lexicographically reviews the dimensions one by one. In this iterative process, she overlooks dimensions where the utility spread is less than a fixed threshold level, and finalizes her selection as soon as she finds a dimension where the spread exceeds that level. Developing the idea of similarity further, Rubinstein (1988) proposes a procedural model of decision-making over simple, binary lotteries; he suggests natural axioms of similarity for the probability and prize dimensions that yield a representation relating to semi-orders: one can find an increasing function of the prize (probability) such that the difference is noticed only if it exceeds a threshold. When comparing two lotteries, Rubinstein's DM can identify dominating lotteries, and otherwise goes for the lottery with the better prize (probability) when only the prizes (probabilities) are dissimilar. Manzini and Mariotti (2012b) build a full-fledged choice theory around the sequential application of semi-orders. The DM has a collection of semi-orders and a fixed sequence in which she applies them; she narrows her choice set down to a single element by further eliminating in round i

[8]Formally, a transitive and irreflexive relation P is a semi-order if both of the following hold for any x, x', y, y': (a) if xPy and $x'Py'$ then xPy' or $x'Py$, and (b) if xPy and yPx' then either $y'Px'$ or xPy'.

[9]Interval orders correspond to situations where τ is non-negative and can vary with y (see Mirkin (1974) and Fishburn (1985)). As easily checked, P is acyclic if, and only if, it admits a utility-threshold representation with a non-negative τ that can vary with both x and y. Authors also explored a variety of more specific functions τ; see the book of Aleskerov et al. (2007) for a thorough literature review.

options that are dominated according to the i^{th} semi-order.[10] Beyond their sequential application, Luce-type threshold functions serve as a building block for several other models discussed in this survey, most notably in relation to satisficing (Section 3.5).

Though Luce himself was a contributor to psychophysics, his 1956 assumption of deterministic perception (i.e., the DM always or never perceives the difference between two alternatives) departs from that field's typical presupposition of probabilistic perception.[11] One may indeed scrutinize deterministic perception in view of the accumulated evidence in both psychophysics and economics (Woodford, 2020). Consider an early experiment by Mosteller and Nogee (1951). Subjects face multiple questions that amount to choosing between, for instance, (a) 5 cents for sure, or (b) a lottery that pays $\$M$ with probability p. Classic choice theories boil down to specifying a threshold M^* such that option (a) is picked whenever $M < M^*$, while (b) is picked whenever $M > M^*$. Only for a single value of M, namely $M = M^*$, should a participant sometimes pick (a) and sometimes pick (b) when facing the same question on different occasions. Instead of a step function, Mosteller and Nogee find that the frequency with which a subject picks option (b) is typically an increasing, continuous S-shaped function of M.

In line with such evidence, Luce (1959)'s highly influential model forgoes his earlier approach: "A basic presupposition of this book is that choice behavior is best described as a probabilistic, not an algebraic, phenomenon" (p. 2). He instead takes the stance that perception is a root cause of stochasticity, observing that "The psychologist has been largely unwilling to make the economist's algebraic idealization, for in some measure the substance of his problem resides in the fact that people are unable to make consistent discriminations" (p. 21). In what is now known as the *Luce rule*, the DM has a utility function $u : X \to \mathbb{R}_{++}$ and selects an option x from the menu S with probability $\frac{u(x)}{\sum_{y \in S} u(y)}$. In subsequent work on random utility, one models the DM as having a utility function $v : X \to \mathbb{R}$ and maximizing a perceived utility function $v_k(x) = v(x) + \epsilon_k(x)$ in each instance k she makes a choice, where $\epsilon_k(x)$ is the realization of a random variable with a cumulative distribution function F that is independent of x (McFadden, 1981). Interestingly, the resulting probability of picking x is the Luce rule with $u(\cdot) = \exp(v(\cdot))$ when F is taken as a Gumbel distribution.[12] There have been many variants and generalizations of the Luce rule, some of which are discussed in the survey of Rieskamp, Busemeyer and

[10] This is the special case of the sequential shortlist methods studied by Manzini and Mariotti (2007) and Apesteguia and Ballester (2013) (see Section 3.3 of this survey), where each rationale used to construct the shortlists corresponds to a semi-order.

[11] For instance, a 'just-noticeable difference' is one that is perceived at least half the time.

[12] Often in econometrics, x is a vector of characteristics and v is linear, leading to the well-known multinomial logit model. If F were normally distributed, one would get the probit model instead.

Mellers (2006). We discuss here some recent contributions which address issues with, or connections between, models of perception; and we defer to the coming subsections those contributions which focus on other aspects of bounded rationality.

Contrary to semi-orders, which are specifically tailored to capture similarity, the Luce rule struggles in dealing with similar alternatives. Debreu (1960) first raised this critique, noting that contrary to Luce (1959)'s prediction, adding a second orchestra's recording of Beethoven's eighth symphony to a menu of Beethoven's eighth and Debussy's quartet should not overturn a DM's preference for the latter; a more popular incarnation of this example is that a DM selecting between bus and train travel should not be more likely to pick a bus when both red and blue buses are available. The nested logit model is a commonly-applied generalization that deals with similarity, by presuming the DM categorizes alternatives, draws a category using a Luce rule, and then draws from within the category using a Luce rule; see, for instance, McFadden (1978). Kovach and Tserenjigmid (2022a) provide the first axiomatic foundation for a non-parametric version of this model, pointing to a weakening of Luce's choice axiom that accommodates similarity.

The two different literatures spurred by Luce (1956, 1959) to address perception are merged in a hybrid model proposed by Horan (2021). In the model's first stage, the DM rules out alternatives dominated by a semi-order, and in the second stage, uses a Luce rule to choose from what remains. Importantly, the two stages are intricately tied. In the first stage, the DM only identifies an interval of utilities that may be assigned to each object, and the DM's semi-order is able to compare precisely those objects for which one utility interval is fully above the other. In the second-stage, the utility function the DM uses in the Luce rule must be *comonotone* with the utility intervals identified in the first stage: formally, if the utility of one object is higher than another, then each endpoint of the better object's interval must be greater than that for the worse object.

Horan's model accommodates an inability to distinguish between some alternatives in the first stage. Though not resolving Debreu (1960)'s critique, it addresses another quirk of Luce (1959)'s characterization theorem: requiring, for each menu, that the DM has a strictly positive probability of choosing any given alternative.[13] More permissive models, where the DM applies the Luce rule only to a set of considered alternatives, have also been proposed (see Section 3.3). An alternative way of addressing this, proposed in Apesteguia and Ballester (2020), is taking the closure of Luce's model. Connecting these different fixes, Doğan and Yildiz (2021) show that the only models at this boundary, beyond the

[13]As more fully discussed in Horan (2021), Luce (1959) also has a lesser known theorem with a partial characterization, that allows for some violations of the full-support assumption.

deterministic rational one, are 'preference-oriented' Luce Rules, whereby the DM applies a Luce rule, using some weights, to the maximizers of a weak preference ordering.

As noted above, Luce's model can be viewed as arising from a DM who maximizes a perceived utility function, where she receives a noisy signal on each option's value. One could have imagined that, instead of simply selecting the option with the highest signal realization, the DM uses those signals to update her prior over the options' values. In other words, there may be inference following perception. In that case, it would be sensible for the agent to pick the option with highest expected value given her updated beliefs. This is precisely the approach taken by Natenzon (2019), who introduces the *Bayesian probit* model. This model explains stochastic choice patterns consistent with the attraction and compromise effects, discussed in Section 3.2; and can also accommodate situations where choices might be nudged by options that are present, but cannot be chosen (e.g., unaffordable). Natenzon (2019)'s model may explain violations of regularity.[14] As discussed further in Section 4, regularity requires the probability of choosing a given alternative to weakly decrease when the menu enlarges. This property is satisfied by the Luce model and the random utility model (RUM) more broadly.

Finally, a vast literature on rational inattention,[15] surveyed by Mackowiak, Matějka and Wiederholt (2021), studies a DM's optimal choice of how much information to gather about alternatives, given information-processing costs. As this affects a DM's ability to distinguish between alternatives' characteristics and utilities, there would seem to be a close connection with perception à la Luce. Matějka and McKay (2015) formalize the connection within a discrete choice, rational-inattention framework, providing a microfoundation for a variant of Luce (1959)'s multinomial-logit model. Their incarnation incorporates both true payoffs of outcomes and the DM's beliefs about them. As such, it addresses the aforementioned critique of Debreu: believing the differently colored buses are essentially the same, the DM would treat buses as leading to the same outcome. On the other hand, their variant may violate regularity due to the endogeneity of (in)attention.

[14] Perception models not involving learning may explain regularity too; see Horan (2021), discussed above, and Echenique, Saito and Tserenjigmid (2018), who introduce a perception priority ordering.

[15] Beyond rational inattention, Gabaix (2019)'s survey places under a simple unifying framework a body of applied work interested in misperception of important objects such as utility, probability, and prices: the DM's perception of the object is a convex combination, parametrized by a weight m, of the truth with some assumed default value. In a multidimensional setting, Gabaix (2014) endogenizes the attention parameters for different dimensions of the problem as the outcome of sparsity-based optimization.

3.2 Resolving conflicting criteria

The beauty of the standard preference relation is that it succinctly encapsulates all the decision maker's principles and criteria. Yet aggregating rankings over different dimensions into a single, all-encompassing preference is perhaps easier said than done. Instances of bounded rationality can arise from the intricacies of reaching a compromise when contemplating conflicting objectives.

Such forces are often credited for causing a well-known violation of rationality known as the attraction (also known as decoy or asymmetric dominance) effect. Consider a DM in need of shampoo. It is plausible she has multiple conflicting criteria in mind, including, for instance, her budget and the inconvenience of running out of shampoo. Suppose shampoo bottles A and B are hard to compare from this perspective: A is larger, but B is cheaper. The DM ends up choosing B. To prevent this, the company selling A might try the following marketing trick: it adds a variant of A, a 'decoy' called A', that is clearly inferior to A, but hard to compare to B (Huber, Payne and Puto, 1982). For an extreme example, say A' has the same size as A, but is even more expensive. The goal isn't to make profits from selling A', but instead to increase the demand for A.[16] The underlying insight here, as Simonson (1989) and Shafir, Simonson and Tversky (1993) point out, is that in complex problems, DMs grasp for reasons: they tend to pick an option that can be justified in a simple way. In this case, the clear superiority of A over A' may nudge her to pick A from $\{A, A', B\}$, resulting in a violation of IIA.[17]

These effects are most often explained by explicitly modeling the DM as cardinally aggregating attributes. Some of this literature examines which dimensions DM weigh most, or focus on, when computing the aggregate utility of a choice, and how those weights can depend on the set at hand; other approaches invoke reference-dependent utility over attributes and are discussed further in Section 3.4. Bordalo, Shleifer and Gennaioli (2012, 2013) provide a theory of salience for multidimensional choice objects like lotteries and consumer bundles, where the DM focuses on an option's salient attributes when making her choice (e.g., price, quality, payoff in a given state). They specify formulations where the salience of an attribute increases the further it lies from the average level of that attribute in the menu, but with diminishing sensitivity: as all values of that attribute increase, the less salience matters. Kőszegi and Szeidl (2013) propose an alternative

[16]See Doyle, O'Connor, Reynolds, and Bottomley (1999) for a field experiment along these lines.

[17]Ok, Ortoleva and Riella (2015, footnote 3) list works documenting the attraction effect in varied choice domains, and Natenzon (2019, Footnote 3) provides references for examples in the animal world. Discussions on the meaning and replicability of the attraction effect can be found in Frederick, Lee, and Baskin (2014), Huber, Payne, and Puto (2014), Simonson (2014), and Yang and Lynn (2014).

formulation where the DM's aggregate utility from an option more heavily weights those attributes exhibiting a wide range of values in the menu.[18]

In cardinal models such as these, violations of rationality arise because the relative weight on different dimensions is affected by the attributes in the menu. For instance, a price tag of $50 for a bottle of wine may be very salient when the alternatives are cheaper, but less so when the alternatives are in the same price range. Some instances of the attraction effect can be explained this way, and Bordalo et al. (2013) identify when such preference reversals can arise (the specific parameters of the model will matter). An important feature of these papers is that the set of attributes is assumed to be fixed and exogenously given. If the modeler may freely choose the set of attributes, then even highly structured, context-dependent methods of aggregating utility from attributes are unlikely to impose any restriction on behavior (Ambrus and Rozen, 2015).

Other formulations of reason-based choice are ordinal. Kalai et al. (2002) posit that choice must be top-ranked according to one of the DM's preference orderings. Naturally, for this model to have testable implications, the DM's collection of preferences must be limited. de Clippel and Eliaz (2012) model the DM's choice when compromising conflictual criteria as the outcome of an (intra-personal) bargaining problem. They characterize an ordinal bargaining solution that triggers the attraction effect (as well as a compromise effect), under which the DM essentially applies a simple, successive-elimination heuristic when assessing the pros and cons of available options.

Though some of the above works explicitly consider multiattribute objects, it is important to note that a DM's conflicting objectives need not be tangible. Problems of temptation and self-control are oftentimes understood as resolving the tension between a long-term goal (e.g., healthy lifestyle) and short-term cravings (e.g., smoking).[19] de Clippel (2014) suggests a simple example of a choice procedure where the DM has limited willpower when choosing from menus. Her willpower is modeled as an integer k representing the number of alternatives she can overlook. The DM's choice maximizes her long-term preference over the set of options that are dominated by at most k elements according to her short-term craving. Masatlioglu, Nakajima and Ozdenoren (2020) study

[18] Tversky (1969), who models only binary choices, is a precursor to this approach.

[19] Many papers in decision theory, starting with the seminal paper by Gul and Pesendorfer (2001), study a DM's preferences for commitment as arising from the cost of exercising self-control. There are two reasons we consider this literature, nicely synthesized by Lipman and Pesendorfer (2013), outside the scope of the present survey. First, it focuses on rational preferences over menus. Second, while often including an (explicit or implicit) description of how the DM compromises conflicting objectives when choosing from a menu, such compromises are often modeled through standard utilitarian aggregation and thus generate a rational preference. In a sense, the cost of self control is sunk once the menu is chosen.

a model where the DM can overlook an option only if its temptation utility is close enough to that of the most-tempting option in the menu. In both cases, resolving the conflicting objectives means that choices from menus may violate rationality.

So far, we have only discussed manifestations of compromise in deterministic choice. However, some of the seminal papers in this literature were motivated by stochastic choice patterns. These influenced not only deterministic theories of choice in the face of conflicting objectives, but some more recent stochastic ones as well.

Observing (weakly) stochastically intransitive choices from binary menus,[20] Tversky (1969) was among the first to consider the aggregation of attributes. In his *additive difference model*, the DM sums a function of the attribute-by-attribute differences to determine which of two options is best. Formally, take an odd function $\Phi_i : \mathbb{R} \to \mathbb{R}$ and let $u_i : X \to R$ describe the level of attribute $i = 1, \ldots, n$ for each option. The DM weakly prefers x to y if $\sum_{i=1}^{n} \Phi_i\left(u_i(x) - u_i(y)\right) \geq 0$. In the spirit of Fechner (1860) and Luce (1959), Tversky extends to stochastic choices by positing that the probability x is chosen from the menu $\{x, y\}$ is given by $F(\sum_{i=1}^{n} \Phi_i\left(u_i(x) - u_i(y)\right)$ for some cumulative distribution function F. Though restricted to binary menus, his basic model relates to some later approaches. It is possible to extend Tversky (1969) to all menus, by defining the aggregate utility of an option $x \in S$ by $\sum_{i=1}^{n} \phi_i(\max_{b \in S} u_i(b) - \min_{a \in S} u_i(a)) u_i(x)$ for some functions ϕ_i (Ambrus and Rozen, 2015). In this case, Kőszegi and Szeidl (2013)'s focus-weighted model corresponds to ϕ_i being the identity for all i.

Tversky (1972)'s classic procedure of *elimination by aspects* is another example. In this model, the DM has several dichotomous criteria (i.e., an option either satisfies a criterion or not), and then sequentially eliminates items from the menu by applying these criteria in a random order. Such sequential elimination is ancestral to the shortlisting approach of Manzini and Mariotti (2007), where general criteria are applied in a deterministic order (discussed further in Section 3.3); as well as to Mandler, Manzini and Mariotti (2012)'s model of checklists, where dichotomous criteria are applied in a fixed order. Since the sequential application of rational choice correspondences still yields a rational correspondence, the latter describes a procedural model of choice that turns out to be indistinguishable from rationality. Similarly, Tversky (1972) retains a form of rationality, even though the random sequencing of criteria leads to stochastic choice behavior. Indeed, the resulting patterns are consistent with RUM, and any Luce (1959) rule can be captured

[20] That is, $p(x|\{x, y\}) \geq 1/2$, $p(y|\{y, z\}) \geq 1/2$, and yet $p(z|\{x, z\}) > 1/2$, where $p(\cdot|S)$ describes the probability of choosing an element from the menu S. We note that later work, mostly among mathematical psychologists, suggests that Tversky's data is not inconsistent with weak stochastic transitivity when using proper statistical testing; see the end of Section 5.

through elimination by aspects. But Tversky's theory also accommodates classic choice patterns contradicting Luce's model, addressing Debreu (1960)'s critique.

Generalizing Luce (1959) even further, Gul, Natenzon and Pesendorfer (2014) introduce *attribute rules*. An attribute is first selected at random by following a Luce rule, and then an option is selected within the feasible set by applying a Luce rule for the utilities pinned down by the selected attribute. Any stochastic choice function obtained through elimination by aspects can be obtained this way, and any attribute rule is consistent with RUM. In fact, any stochastic choice function obtained through random utility can be obtained as the limit of attribute rules. The procedure bears some superficial similarity to nested logit (and also addresses Debreu's critique), but there are some basic differences, including that attributes must be identified endogenously from choice frequencies.

When interpreted from an individual decision-making perspective, random utility models capture a fickle DM with multiple criteria or moods, who chooses according to the preference drawn at the time of decision. In general, RUM imposes no restriction on how many such criteria there may be. Several attempts have been made to be more parsimonious. Manzini and Mariotti (2018) explore a model of *dual random utility maximization*, which is more restrictive than RUM because it permits just two preference orderings (e.g., short-term and long-term preferences in problems of temptation, or selfish and altruistic preferences in resource allocation problems). Apesteguia, Ballester and Lu (2017) allow a DM to use only a set of preferences that single-cross under some ordering (as might occur if the DM's moodiness affects their coefficient of risk-aversion within a parametrized class of risk preferences), so that the resulting SCFs must still be within RUM. Filiz-Ozbay and Masatlioglu (2022) more broadly permit single-crossing, boundedly-rational behaviors (not just rational ones), and show this generalization allows for any behavior.

We conclude this subsection with some comments that go beyond individual decision making. First, some of the above models also have a population-based interpretation, which may be useful in contexts where the researcher only has access to aggregate data from individuals that have potentially conflicting objectives. For instance, RUM arises as a mixture of rational choice functions.[21] Second, a nascent literature considers the interesting possibility that individuals mutually influence each other, a feature generally absent from the models above. Cuhadaroglu (2017) examines when the choice correspondences of two DM's is consistent with each of them shortlisting a subset of alternatives using their own (potentially incomplete) preference, and then making a final selection based on

[21] Unlike rationality, some models are closed under heterogenous mixtures, which amounts to the existence of a representative agent. Apesteguia and Ballester (2016) show that Luce (1959) does not admit a representative agent, but RUM and Apesteguia, Ballester and Lu (2017)'s single-crossing variant do.

the preferences of the other. This is further explored in a stochastic-choice framework by Chambers, Cuhadaroglu and Masatlioglu (2023).

3.3 Overlooking some alternatives

Rationality requires the DM to give full consideration to all alternatives. Personal and empirical evidence calls this assumption into question. A class of bounded rationality theories examines the possibility that the DM still maximizes her preference ordering $\succ$, but only over the subset of alternatives $\Gamma(S)$ she actively considers ('pays attention to') from the potentially larger, feasible menu S.[22] Of course, without placing further restrictions on the considering-set mapping Γ, there is no hope for a meaningful theory: simply imagine that the DM considers just her choice $c(S)$ when facing the menu S, and any choice pattern can be explained! The literature analyzes heuristics and regularity properties for how DM narrows down her focus, whether consciously or not, to a subset of options. We begin with the first wave of theories, in which consideration sets are deterministic, before incorporating the possibility of random consideration that is studied in the second wave. Even within this first wave, two complementary approaches appear.[23]

One class of theories describes a mental process for how consideration sets arise. For instance, Manzini and Mariotti (2012a) model a DM who thinks of options as belonging to categories; in each menu, she focuses only on options belonging to undominated categories, according to some shading relation. Cherepanov, Feddersen and Sandroni (2013) posit that the DM restricts attention to options she can rationalize, that is, those she can justify as being maximal for at least one of her rationales (captured by binary relations). Manzini, Mariotti and Tyson (2013) model a DM who pays attention to any alternative whose value according to some criterion function exceeds the menu-dependent threshold, akin to a form of satisficing in the creation of the consideration set. Masatlioglu and Nakajima (2013) model consideration-set formation through an iterative search process.

Another class of theories imposes plausible conditions on consideration sets, without being too specific about how those sets form. An advantage is that these conditions may potentially encapsulate a variety of more detailed stories about consideration set formation. The conditions may be as simple as requiring the DM to consider a certain number of elements (possibly dependent on the size of the menu), as in Barberà, de Clippel, Neme

[22] A related structure arises when the DM may be better informed than the modeler about what options are actually unavailable (see, for instance, Brady and Rehbek (2016) or parts of Barseghyan et al. (2021)).

[23] Eliaz and Spiegler (2011) were among the first in economics to model some consideration-set formation (in a game-theoretic setting), but did not develop a full-fledged theory of attention formation.

and Rozen (2022).[24] But restrictions may also be placed on attention *across* menus. For instance, inspired by the marketing literature, Masatlioglu, Nakajima and Ozbay (2012) assume that a DM's consideration set does not change when overlooked alternatives become infeasible. Inspired by the notion of choice overload, Lleras, Masatlioglu, Nakajima and Ozbay (2017) instead posit that an option must be considered in a menu if it is considered in some larger menu. Naturally, the detail-free approach of this paragraph can encompass the detail-oriented one in the previous paragraph. For instance, the Lleras et al. (2017) property, which amounts to IIA on the consideration-set mapping, characterizes the consideration set mappings arising under Cherepanov et al. (2013).

Broadening the paradigm, one may generalize the 'preference' the DM maximizes over her consideration set to a relation that need not be an ordering. Manzini and Mariotti (2007)'s sequential rationality is in this vein. The DM effectively constructs a consideration set through successive shortlisting. In each stage, the DM shortlists alternatives which are undominated according to a 'shortlisting' relation that need only be asymmetric. When making a hiring decision, for instance, an employer may drop less-experienced applicants from consideration if experienced ones are available, and continue to sequentially trim the applicant pool using further criteria (e.g., education, skill). The DM has a final asymmetric relation which makes a decisive selection from the remaining options. There are several works delving further into this model or variations thereof. For instance, Au and Kawai (2011) study the restriction to transitive relations, while Dutta and Horan (2015) characterize what may be inferred about the underlying relations based on choices. Apesteguia and Ballester (2013) propose a related notion of sequential rationalizability which involves only acyclic relations.

The second wave of the literature, studying stochastic attention, presents an analogous modeling question: should we describe the probability a given attention set arises, or impose regularity properties on distributions of attention sets? Taking the former perspective, Manzini and Mariotti (2014) introduce an attention probability $\gamma(x)$ for each alternative x, and let $\prod_{b\in B}\gamma(b)\prod_{a\in A\setminus B}(1-\gamma(a))$ be the probability the DM pays attention to the subset B given a menu A. Two structural features are worth noting. First, the DM's attention set is empty with positive probability. In this case, the DM is assumed to pick a default alternative (e.g. not buying). A variant, studied in depth by Horan (2019), avoids resorting to default alternatives by renormalizing probabilities to

[24] Barseghyan, Coughlin, Molinari and Teitelbaum (2021) extend the classic random utility model, by similarly assuming agents' deterministic attention sets contain a minimal number of elements instead of actively considering all feasible alternatives. Geng and Ozbay (2021) also posit consideration sets of a fixed size (when the menu is large enough) but impose greater structure in how they are derived.

ensure the consideration set is nonempty (i.e., dividing by one minus the probability that the attention set from S is empty). A second feature is that attention probabilities are uncorrelated across alternatives, and independent of the set of feasible alternatives. This implies that the restriction of the model to deterministic attention is quite different from the deterministic models we discussed above: just a single option would be considered in all problems containing it (and the attention set is empty otherwise).

Several papers consider variations on Manzini and Mariotti (2014)'s probabilistic-attention process. Brady and Rehbeck (2016) fix a full-support probability distribution π over a collection of menus, and assume that the DM's preference is maximized over only a subset $B \subseteq A$ drawn with probability $\pi(B)/\sum_{C \subseteq A} \pi(C)$. Manzini and Mariotti (2014)'s model corresponds to the choice of distribution $\pi(A) = \prod_{b \in X \setminus A}(1-\gamma(b)) \prod_{a \in A} \gamma(a)$ for all menus A. Aguiar (2017) models a DM for whom a random subset (a 'mental category') of X becomes salient; the feasible options among those comprise her attention set, which may thus be empty. Demirkan and Kimya (2020) explore menu-dependent probabilities, and draw connections with other stochastic choice models.

Taking the detail-free perspective, Cattaneo, Ma, Masatlioglu and Suleymanov (2020) introduce a random-attention model characterized by a weak monotonicity condition on attention probabilities. This condition, discussed further in Section 4, captures a notion that alternatives compete for the DM's attention: the likelihood of a potential consideration set R improves when removing options outside of R, as they may attract the DM's attention away. This property nests multiple models of stochastic attention, including, for instance, randomizing over Masatlioglu et al. (2012)'s deterministic attention filters, Brady and Rehbek (2016)'s extension of Manzini and Mariotti (2014), and a model in the spirit of Tversky's (1972) elimination by aspects. Cattaneo, Cheung, Ma and Masatlioglu (2021) propose another detail-free model capturing attention overload; the probability an alternative is considered in a menu weakly decreases as the menu grows, a form of the standard regularity axiom but applied to consideration sets instead.

Other conditions on attention have been studied as well. For instance, Dardanoni, Manzini, Mariotti and Tyson (2020) capture cognitive constraints by a distribution over the number of elements in their attention sets, yielding a stochastic version of the model developed in Barberà et al. (2022). Echenique and Saito (2019) and Cerreia-Vioglio et al. (2021) generalize Luce (1959) to situations where the DM may overlook some alternatives by applying the Luce rule to only a subset of alternatives satisfying only basic properties.[25]

[25]Recall that Luce's characterization requires a positivity axiom on choice probabilities; see also the discussion in Cerreia-Vioglio et al. (2021).

For instance, as in Echenique and Saito (2019), the support of the randomization in each menu may be induced by undominated alternatives according to a strict partial order.

We wrap up this discussion with some broader comments. In the literature we survey, having limited 'attention' or 'consideration' is used nearly synonymously with the title of this subsection: overlooking some alternatives. By contrast, the aforementioned literature on rational inattention envisions attention as being closer to perception, in the sense of Section 3.1: it impacts the ability to distinguish the preference between alternatives. Caplin, Dean and Leahy (2019) bridge the worlds of 'perception' and 'overlooking alternatives' by considering a richer setting where, like in the rational attention literature, the DM has a prior about the alternatives, and contemplates which costly Blackwell experiment to take to sharpen their information before making a final decision. With costs modeled using Shannon entropy, they show that the DM's optimal information-gathering rule has a limited support. This means the DM essentially ignores some alternatives, thereby endogenizing the consideration set; but the preference and consideration set are not independent, which is a significant departure from the models above.

3.4 Reference points and other context dependence

Broadly speaking, context dependence means that choices are affected by features that would be deemed irrelevant under rationality.[26] Incorporating context dependence often involves enriching the dataset, and thus departing from the classic choice-theory paradigm. Of course, any choice function is explainable via contextual effects if no further structure is imposed. But a sizable and longstanding body of evidence points to rather specific forms of context dependence playing a role in choice behavior. This includes the attraction and compromise effects discussed in Section 3.2, but can more broadly include contextual information beyond the attributes of available options. For instance, the DM may be impacted by the size of the menu, or the order in which options are displayed. As extensively documented in Kahneman, Knetsch and Thaler (1991), DMs may also be impacted by highlighted alternative(s), such as in Thaler (1980)'s endowment effect or Samuelson and Zeckhauser (1988)'s status quo bias. The latter two phenomena are instances whereby a

[26] Determining what is irrelevant may be easier said than done. Suppose the DM maximizes a preference ordering which varies with the weather: she prefers ice cream over hot chocolate in warm weather, and vice versa in cold weather. On the one hand, one could argue that she does not maximize a single, stationary preference, and that her behavior is dependent on context (outside temperature). One could instead argue that ice cream in cold weather is actually a different consumption good than ice cream in warm weather, and that there is a single, stationary preference after all. The outcome of this judgement call might itself be context dependent: for instance, we may be more inclined to deem the DM irrational if her preferences instead varied with shelf placement at the supermarket.

reference point impacts the DM's evaluation of outcomes; but reference points can also interact, for instance, with the set of alternatives considered when attention is limited (as in Section 3.3). The possibility of reference dependence, and even the idea embodied in loss aversion that losses (relative to the reference point) loom larger than gains, are noted in experimental studies as early as Markowitz (1952). These concepts were developed further in Kahneman and Tversky (1979)'s seminal prospect theory for risky choices, and Tversky and Kahneman (1991)'s generalization to riskless ones.[27]

3.4.1 Exogenous context dependence and enriched data

We discuss here more recent developments, beginning with the umbrella framework of Salant and Rubinstein (2008). They introduce the notion of an extended choice problem (S, f), where S is a menu and f is a frame.[28] A frame captures observable information that the modeler believes may affect choice. This flexible notion can encapsulate, among other things, a status quo, the order in which options are presented, or how many times each alternative was advertised. The set of possible frames $\mathcal{F}$ thus enriches the domain of the choice correspondence. Salant and Rubinstein consider the (standard) choice correspondence that is induced by aggregating the choices from a menu across different possible frames. Depending on the underlying model of framing, they show it may be possible to explain this induced choice correspondence as the maximization of some preference relation; in other cases, it may be too coarse to offer any useful information. For instance, Rubinstein and Salant (2006) study a framework of choice from lists, where the choice function may depend on the order of presentation. They show list-based versions of the classic IIA and partition independence properties[29] are equivalent, and admit behaviors behaviors beyond rationality. Nonetheless, the choice correspondence induced from a list-based choice function (derived by aggregating choices across lists) satisfies WARP if, and only if, it that list-based choice function satisfies list-based IIA.

Other works also delve into structured manifestations of context dependence, but still share with Rubinstein and Salant (2006) and Salant and Rubinstein (2008) the simplifying feature that frames are specified as part of the data. We consider these next. Yet others, discussed much further below, address situations where the frame (typically a reference

[27]Barberis (2013) offers an excellent evaluation of prospect theory and its ensuing applications.

[28]Bernheim and Rangel (2009) introduce a similar framework to study welfare; see Section 6.

[29]List-based IIA means shortening a list by removing an unchosen alternative does not change choice. Partition independence generally means dividing a set into subsets, and restricting attention to the choices from those subsets, would not change choice; the list-based version means dividing a list into sublists and then selecting from the list of choices does not change choice.

point) is not specified in the data, but may potentially be inferred from choices themselves.

Masatlioglu and Ok (2005) enrich the domain of choice problems in a more specific way, by potentially designating a known status quo. In their framework, a choice problem (S, x) represents a situation where the DM faces a menu S with x as a status quo (whenever $x \in S$) or there is no status quo at all (when $x = \diamond$).[30] Allowing for choice correspondences over a compact metric space of alternatives, they propose axioms yielding a classic representation: when there is a status quo x, it receives an extra utility 'bump' of $\varphi(x)$ that another alternative has to overcome in order to be chosen. Imposing a weaker set of axioms over a finite set of alternatives yields another insightful representation. The DM has a vector-valued utility u and an increasing function f that aggregates these utilities, and without a status quo, the DM simply maximizes $f \circ u$ over the menu. A status quo, however, restricts which alternatives are considered: the DM maximizes $f \circ u$ over only those alternatives $y \in S$ where the vector of utilities $u(y)$ strictly dominates that of the status quo, and uses the status quo as the default choice in case no such y exists. Masatlioglu and Ok (2014) refine these ideas further, observing that choices may be seen as arising from utility maximization subject to 'psychological constraint sets' induced by the status quo. This may explain a variety of behaviors, ranging from the attraction effect to extreme status-quo bias (i.e., the status quo is always selected, because it is the only element in the constraint set), but it need not imply the endowment effect. An important feature of their models is that the DM acts rationally for any fixed status quo, but varying the status quo may alter the DM's selections even when the status quo is not chosen.

Under the above two models, the DM does not suffer from decision avoidance (the status quo is an option in the menu), nor does she suffer from choice overload (having a larger menu cannot be worse). Buturak and Evren (2017) demonstrate a channel through which status-quo bias may lead to both issues. They formalize a choice model with anticipated regret, where the DM has a fixed default option that does not belong to any menu (e.g., not choosing anything). The DM is unsure of her preference at the time of decision (she holds a probability distribution over her ultimate utility function), and experiences regret if she chooses an ex-post suboptimal option. Critically, the DM does not anticipate experiencing regret if she sticks with the status quo. Anticipated regret is thus naturally greater in larger choice sets, leading to a higher chance of picking the default.[31] Dean, Kibris and Masatlioglu (2017) put forth another explanation for choice overload which eschews decision avoidance. Generalizing Masatlioglu and Ok (2014), they posit

[30] Apesteguia and Ballester (2009) consider a related framework, and examine whether path independence properties can still arise when the choice becomes the next status quo in sequential choice problems.

[31] For another approach to choice overload, but which has some related axioms, see Gerasimou (2018).

that the DM is restricted by both a (status-quo induced) psychological constraint set and an attention correspondence satisfying IIA à la Lleras et al. (2017).[32] They axiomatically characterize this representation and test some of its features in a laboratory experiment.

Kovach and Suleymanov (2021) also intertwine attention and reference effects but, in contrast to Dean et al. (2017), they model attention as a random process. Their DM has a family of reference-dependent stochastic choice functions, indexed by the alternative serving as status quo; these arise through the maximization of a single preference relation over a stochastically-determined attention set that need only include the status quo. This is rather general, but has testable implications in the form of two simple axioms. Kovach and Suleymanov also consider imposing additional structure on the reference-dependent probability with which a given submenu is considered, such as the condition imposed by Cattaneo et al. (2020) or the more restrictive conditions of Manzini and Mariotti (2014), Brady and Rehbeck (2016), or Aguiar (2017). They examine which of these reference-enriched models of attention privilege the status quo alternative in choice behavior.

Notwithstanding these examples, models with an exogenously-specified reference point need not be models of status-quo bias. Rubinstein and Zhou (1999) axiomatize a model interpretable as aspirational choice: the DM selects an option which is closest (in Euclidean distance) to an exogenously given alternative. Kahneman and Tversky's seminal works let the reference point be the benchmark against which one assesses gains versus losses, and often simply define it as the DM's (known) current position. Nonetheless, they observe that the reference point "can also be influenced by aspirations, expectations, norms and social comparisons," and try not to spoil this challenge for others: "the question of the origin and the determinants of the reference state lies beyond the scope of the present article" (Tversky and Kahneman, 1991, pp. 1046-1047).

3.4.2 Endogenous context dependence

Ensuing theories have broached the matter of endogenous reference dependence from roughly three different angles, that we discuss in turn. A first approach is to have the theory, rather than the data, specify what the reference point must be as a function of the menu. This is more typically applied to multi-attribute settings, where the reference point can act as an aggregator. Kivetz, Netzer, and Srinivasan (2004) explain both the attraction and compromise effects using a 'contextual concavity' model that builds on Tversky and Kahneman (1991)'s model. Their DM's reference point is the (potentially infeasible) alternative constructed by taking for each attribute, the worst level in the menu.

[32] As well as the property that both alternatives are considered in binary menus.

The DM then values alternative $x \in S$ by $\sum_i (u_i(x_i) - \min_{y \in S} u_i(y_i))^{r_i}$, which sums for each attributes i, a concave CRRA function of the difference between the current alternative's utility and that of the reference point. Each utility u_i is assumed to be increasing, so the worst attribute level is readily identified as the lowest one. Kivetz, Netzer, and Srinivasan (2004)'s ability to explain both effects with their parameterization turns out to be no coincidence. In a deeper examination, Tserenjigmid (2019) considers a non-parametric generalization where the worst levels continue to serve as the reference point, and proves an equivalence between (i) satisfying Tversky and Kahneman (1991)'s diminishing sensitivity, (ii) displaying the compromise effect, and (iii) displaying the attraction effect.

A second approach, spearheaded by Köszegi and Rabin (2006) for consumption bundles and Köszegi and Rabin (2007b) for monetary lotteries, reimagines reference points through the lens of an equilibrium concept. They not only enrich Tversky and Kahneman (1991)'s functional form to include utility from consumption levels themselves, but also define the reference point against which 'gain-loss' utility is evaluated as the DM's rational expectations about her potentially stochastic outcome. Formally, the DM holds a probability distribution about what choice menus she may face, and what she would select out of each menu, that are indeed correct given her utility function. Hence the reference point is determined in a *personal equilibrium.* While this structure means only consumption utility is maximized in deterministic environments, gain-loss utility does matter whenever there is some uncertainty. Moreover, there may be multiple personal equilibria, corresponding to different possible reference points. Köszegi and Rabin propose focusing on the equilibrium leading to the highest utility–the DM's preferred personal equilibrium. Ensuing works have examined the testable implications of such an equilibrium view of the reference point. Freeman (2017, 2019) shows that the preferred personal equilibrium notion is indeed restrictive over riskless and risky choices, respectively. Masatlioglu and Raymond (2016) further characterize Köszegi and Rabin (2007b)'s preferences as the intersection of two classic non-expected utility theory models: quadratic utility (Machina 1982) and rank dependent preferences (Quiggin 1982).

A third approach is for the theory to specify how reference points affect choices, with neither the theory nor the data specifying what those reference points are. The crux of the matter is to test compatibility with the theory and potentially infer the unknowns. An early example of this approach is the 'triggered rationality' model proposed by Rubinstein and Salant (2006b) and further developed by Kibris, Masatlioglu, and Suleymanov (2021b).[33] The DM is envisioned as having a salience ordering over alternatives and a

[33]See also de Clippel and Rozen (2021) for an analysis in the presence of limited data.

family of preference relations, each indexed by an alternative $x \in X$. In each menu S, the DM's maximizes the reference-dependent preference associated with the most salient alternative in S. The structure of this model permits partial revelation of preferences and the salience ordering (and thus reference points): if x is most salient in a menu, then it is also most salient in any smaller menu in which it is contained, and the choices from those particular menus cannot mutually contradict rationality. Kibris, Masatlioglu, and Suleymanov (2021a) consider related ideas within a random-reference point framework: the DM may use a different reference point for the same menu at different times, and thereby apply a different reference-dependent preference. Each alternative has a salience weight and the reference point is drawn by applying a Luce rule using these weights.

Also in this vein is Ok, Ortoleva and Riella (2015), who envision reference points as impacting the comparative appeal of non-reference alternatives. Permitting choice correspondences, Ok et al. weaken WARP to accommodate such effects. To get a sense of how they define a 'revealed reference' in such a setting, suppose y is chosen from the menu $\{y, x\}$ but x is actually the choice from some triple $\{x, y, z\}$; in this case, z essentially 'helps' x to become chosen. They also introduce a weaker notion of a 'potential' reference point, which merely requires that z does not reduce the appeal of x. As opposed to a status quo bias, which could be captured by the models in the paragraph above, this model precludes the DM from choosing the reference point; it only serves to highlight other options. The DM has a set of possible utility functions and a reference mapping for each set. The reference mapping may be empty (in which case the DM behaves rationally), but if there is a reference point then the DM restricts attention to alternatives that are least as good as the reference point according to all her utility functions.

Beyond reference points, other forms of context-dependence may be inferred from choices as well. Without presuming the order in which the DM reviews options is known, Yildiz (2016) characterizes when potentially stochastic choices are explained by list-rationality (i.e., the DM has an asymmetric preference and iteratively compares the next item in the list to the best one so far). Kovach and Tserenjigmid (2022b) propose a variation of Luce's model that allows the DM to be affected by focal alternatives (e.g., some alternatives may be more prominently displayed). Each menu S is partitioned into focal alternatives $F(S)$ and non-focal ones (the complement), subsets which the modeler may potentially infer. The DM applies the Luce rule within each of these subsets, but the probability of choosing any focal alternative is biased up by a menu-dependent factor.[34]

Finally, there is also room for hybrid frameworks, as the above categories need not be

[34] Unlike Luce, their model does not fall within RUM: it permits regularity violations.

mutually exclusive. For instance, a theory can study endogenous context dependence with the aid of exogenous information. In Barbos (2010), the DM has a set of exogenously given categories of lotteries, and chooses the best lottery from some category; but the category selected is endogenously determined by a reference dependent utility that compares the best and worst lotteries in a category. Guney, Richter and Tsur (2018) enrich the standard choice domain by including exogenously-given information about potential alternatives that may be infeasible (e.g., your neighbor's luxury sports car) to characterize a model of aspiration-based choice.[35] They impose axioms over choice behavior that identify the DM's best alternative in the potential set Y as the aspiration level, and have the DM choose the alternatives in their menu $S \subseteq Y$ which is most similar to it, according to an endogenously determined metric. Also potentially allowing for exogenous and endogenous elements, Ellis and Masatlioglu (2022) propose a model of choice that contains several previously discussed theories. Their DM uses a reference point to group objects into categories, and has a monotone and additively separable utility function which depends on both category and reference point. However, the reference point in their model only determines an increasing transformation of utility conditional on category, which means it does not affect within-category choices. The DM's categories may be simple and exogenous (e.g., gains versus losses relative to some fixed point) but could also be endogenous (e.g., salience). They provide a behavioral foundation for Bordalo et al. (2013) and more broadly explore axioms that distinguish between known models nested in their framework.

3.5 Satisficing

A rational DM must identify all options in the menu, compare them to formulate a preference, and successfully maximize that preference. This is no simple feat, and likely requires time and effort. A more pragmatically-minded DM may instead decide to cut her search for options short, and thereby remain unaware of some alternatives. Assuming she does have a well-behaved preference ranking, she may still follow simplifying shortcuts in coming to a decision. These observations, first put forth by Simon (1955), motivate the multiple models discussed in this survey. Broadly speaking, Simon's idea of satisficing has the DM explore options in some order, and pick the first option she finds acceptable. Defining an acceptable set might entail having some reference utility level in mind. But there is an important difference with theories of reference dependence, which goes beyond whether or not the utility function itself is reference dependent: critically, satisficers seek

[35]As noted earlier, Natenzon (2019) also allows choices to be nudged by infeasible alternatives.

options which are merely good enough, instead of systematically striving for the best ones. There are many incarnations of this idea, depending on how one defines what 'acceptable' or 'good enough' means, and how the DM's search proceeds.

The typical, textbook introduction to satisficing goes as follows; see, for example, Rubinstein (2015). The DM has a utility function $u : X \to \mathbb{R}$, a threshold level $u^* \in \mathbb{R}$, and a search ordering $\succ$ of options in X. Facing a menu $S \subseteq X$, the DM reviews options according to $\succ$, and picks the first option x she encounters whose utility $u(x)$ is larger or equal to u^* (if none passes this test, then the DM picks the u-maximal option from S after having reviewed all of them). In addition to offering a first formalization of satisficing, this model is also commonly mentioned because it generates the same set of choice functions as rationality. Thus, though rationality and this form of satisficing encapsulate distinct choice processes, they are indistinguishable in terms of predicted choice behaviors.

More realistically, the threshold may vary with the menu. For instance, a low wage offer may be acceptable in an economic downturn, but not in an upswing. Aleskerov et al. (2007) formalize a menu-dependent model of satisficing: the DM has a utility threshold function $\theta : \mathcal{P}(X) \to \mathbb{R}$ and their choice correspondence consists, for each menu S, of *all* alternatives $x \in S$ with $u(x) \geq \theta(S)$. Tyson (2008) requires that if $S \subset T$, then $\theta(S) \geq \theta(T)$ whenever the best element in S would be satisfactory in T. His *expansive satisficing representation* is equivalent to one where the DM has an underlying weak order P, but maximizes only a menu-dependent weak order that is coarser than P (and gets coarser the larger the menu is).[36] Indeed, contrary to the possibly more common view that satisficing arises from a search process, Tyson takes the perspective that it arises from being content with only partially forming one's preference about the alternatives.

There is clearly an equivalent way to phrase satisficing representations of the kind Aleskerov et al. (2007) and Tyson (2008) discuss. Namely, one may define a departure function $\delta : \mathcal{P}(X) \to \mathbb{R}$ which indicates how far below the maximum utility level in a menu the DM is willing to depart. Her choice correspondence consists, for each menu S, of *all* alternatives $x \in S$ such that $u(x) \geq \max_{y \in S} u(y) - \delta(S)$. Capturing a DM whose satisficing is worsened by choice overload, Frick (2016) axiomatically characterizes *monotone* threshold representations where $\delta(S) \leq \delta(T)$ if $S \subseteq T$. Frick's model includes, as special cases, Tyson (2008)'s model as well as Luce (1956)'s choice by semiorders.

One may argue that endogenizing the threshold, as in the works above, requires the DM to have at least some understanding of the feasible set; else how could θ (or δ) be menu dependent? Depending on the assumed functional form, this may be as little information

[36] Tyson (2021) builds on this, specifying a parametric version amenable to game-theoretic applications.

as the number of alternatives in the menu. Menu-dependent thresholds may also naturally arise if DM is initially unaware of the menu she faces, but samples some available options, and then pick the first option surpassing the best one found during the sampling phase (a procedure reminiscent of the solution to the famous secretary problem). Still, in other search-based conceptions of satisficing, the DM may have more specific expectations about the feasible options in the menu, but may need to exert effort to actually identify them. This may be a reason, beyond other practical constraints, to avoid presuming that we can observe all options the DM considers satisfactory. For instance, an experienced recruiter seeking to fill a position may know the distribution of qualifications in a typical applicant pool, but still have to interview specific candidates to identify those with an acceptable combination of qualifications. 'Good enough' may then mean someone who is above average, or in the top quintile if one is more demanding. These variants of satisficing are discussed in Barberà, de Clippel, Neme and Rozen (2022).

More broadly, Salant (2011) provides an important theoretical foundation for viewing the DM's threshold as endogenously adjusting while she reviews the menu. Consistent with Simon (1955)'s view of satisficing as a phenomenon related to procedural and computational constraints, Salant explicitly models the DM as using only a finite-state automaton: the DM wishes to maximize expected utility subject to the choice rule requiring only few states.[37] He shows that satisficing is procedurally simple: a choice function has minimal state complexity if, and only if, it can be represented as a satisficing procedure. Moreover, history-dependent aspiration thresholds, and list-based effects (such as primacy and recency effects on choices) optimally emerge. Thus, viewed from the lens of a search process, a key aspect of satisficing pertains to the order in which options are reviewed.

This leads to several fundamental questions. Does the DM always follow the same order of contemplation? Does the modeler know the order in which the DM contemplated options each time she makes a choice? Can a third-party nudge this order? In some cases, the modeler may formulate a conjecture about the order in which the DM reviews options, and incorporate this into the theory to be tested. For instance, the order in which food is displayed in a buffet may be the relevant order for customers who walk through the cafeteria along the delineated path. Seeing that choices change when the buffet is rearranged proves the DM is not rational, but may still be compatible with satisficing. Hence additional data can disentangle models that may otherwise be behaviorally equivalent.

[37]Though Wilson (2014) studies a single decision problem rather than a choice theory, it also explicitly models the DM as having finitely many states and derives several biases in information processing. Important precursors to this strand of literature are Rubinstein (1986) and Abreu and Rubinstein (1988), who consider repeated-game players whose strategies are constrained to being finite-state automata.

One could imagine collecting richer data involving information about search, in addition to final choices. Caplin and Dean (2011) explore the theoretical analysis of such enriched data (see also Papi (2012)), noting that the search models they consider, though natural, do not have testable implications with standard choice data alone. Caplin, Dean and Martin (2011) implement a novel experimental design that partly reveals subjects' choice process (an instance of enriching the range of the choice object). In their experiment, each participant can update her selected option as time unfolds, knowing the selection will be finalized at a randomly selected time. Each option corresponds to a monetary payment, but effort is required to decipher the amount is (a sum of numbers written in plain english). Their choice process data is indicative of satisficing: option switches most often improve payoffs, and switches stop on average if, and only if, a certain payoff reservation level has been surpassed. Moreover, the data suggests the threshold is context-dependent, consistent with their model of search costs: the threshold decreases with the complexity of each option, and increases with the number of available options.

Reutskaja, Nagel, Camerer, and Rangel (2012) instead use eye-tracking technology to elicit information about search process in an experiment where participants select snacks under time pressure and choice overload. They propose three models to analyze the data. Common elements include a two-phase process (initial search, followed by choice); that a subject must 'fixate' on an item during the search phase to figure out its value; and that initial search is random in terms of the items' values. The models differ in the stopping rule DMs employ. Under optimal choice, the DM searches until all items have been fixated upon (or time runs out). Under satisficing, the DM stops when finding an item with value above a threshold. Under a more general hybrid model, the probability search stops increases in both the number of items already reviewed, and the value of the best item identified thus far. Given the preferences inferred from an initial, rating phase of the experiment, they find that their data is compatible with the more complex, hybrid model, but incompatible with the simpler versions of optimal choice and satisficing.

We are not aware of studies that directly test what characteristics of a choice problem determine the order of reviewing options. Certainly, the order in which options are listed matters in Caplin et al. (2011), and the items' screen locations impact choices in Reutskaja et al. (2012); hence in these studies, the domain of the choice object is also enriched. The search order may be randomly determined in some cases, or at least appear random to the modeler with limited information about what the DM observes. Aguiar, Boccardi and Dean (2016) and Aguiar and Kimya (2019) explore the testable implications of satisficing in that case. But stochasticity in a satisficer's choices may come from other sources as well.

Salant and Spenkuch (2022) consider a model where not only is search order randomly determined, but satisficing and random utility are married together: the DM selects the first item whose noisily-determined *estimated* value exceeds a threshold.

4 Axiomatic characterizations and their purposes

A choice theory typically tells a story about the DM's choice process. Yet without observing that process directly, a theory must be understood through the choice patterns it generates.[38] Under rationality, for instance, a chosen option remains selected even when other options turn out to be unavailable (IIA). A central part of many theoretical works in the literature is to develop behavioral characterizations of new choice theories, or new characterizations of old theories. Space constraints prevent an exhaustive discussion of these numerous results. Here we highlight a selection, to offer clear comparisons with benchmark characterizations and illustrate the range of purposes such results serve.

A first purpose is to improve our understanding of the theory. For instance, IIA is meaningful and understandable independently of Rationality: one would have no difficulty presenting the former before the latter to students in a course. Linking the two concepts is thus insightful, and one may be surprised when learning for the first time that choices satisfying IIA can always be viewed as maximizing an ordering. Secondly, a characterization can help us pinpoint the behavioral principle(s) that distinguish one theory from another, or perhaps lead us to discover that two theories with very different motivations turn out to be indistinguishable (at least based on standard choice data alone). Thirdly, characterization results may be valuable from a normative standpoint. For instance, some may find preference maximization more appealing knowing that it corresponds to the consistency in choices that is captured by IIA. Finally, a fourth purpose, absent direct observation of choice processes, is to provide choice-based tests for a theory. In some cases, a single axiom characterizes a theory; in other cases, the characterization may break the testable implications down into multiple, independent axioms. It is sometimes helpful to test axioms separately, to understand what underlying principles in the theory may hold more generally or are violated instead.

It is useful to keep in mind that there is no single 'characterization' for a choice theory. There may be multiple ways to express the underlying principles or testable implications, and different characterizations of a theory may advance different goals. For instance, both

[38] One could potentially rely on other observables, such as brain scans, time responses, etc. This approach is pursued in a few instances, though rarely so.

IIA and SARP characterize Rationality. The SARP-based characterization is preferable for testing, as it remains valid regardless of the collection of menus over which choices have been observed. It also forces the modeler to consider what parts of the DM's preference can be identified through her choices, which is relevant for welfare purposes and out-of-sample prediction. Yet SARP does not have all the advantages: IIA is simpler, and, as opposed to SARP, is a meaningful behavioral principle without regard to Rationality.

4.1 Deterministic theories

We now discuss some characterization results for bounded-rationality theory, keeping in mind the variety of purposes they may serve.

To begin, consider a classic characterization of correspondences arising from the selection of undominated alternatives according to an acyclic preference P.[39] As we have seen, choice correspondences can arise not only from indifference, but also under more general preference relations that relax completeness or transitivity. The particular model above captures the implications of a variety of theories. For instance, difficulty in distinguishing objective stimuli prompted Luce (1956) to rationalize choices using semi-orders. Generalizing Luce's threshold to depend on the alternatives compared has the same testable implications as selecting undominated alternatives according to an acyclic relation; so does Masatlioglu and Nakajima (2013)'s Markovian Choice by Iterative Search, when the starting point of the DM's search process is unobserved. Sen (1971) proves that a choice correspondence arises through this more permissive notion of rationality (i.e., there exists a P such that $C(S) = \{x \in S | \nexists y \in S : yPx\}$) if, and only if, it satisfies two conditions. The first is Condition α (mentioned in Section 2), and the other is a simple condition requiring any common choice from two sets to be chosen from their union:

Condition γ *If* $x \in C(R) \cap C(S)$, *then* $x \in C(R \cup S)$.

Left with multiple undominated options, it is natural to refine choices by applying one or more additional criteria. Manzini and Mariotti (2007)'s theory of choice from short-lists, discussed in Section 3.3, does precisely this. They show that the resulting choice functions are characterized by Condition γ and the following relaxation of IIA.[40]

[39] Once completeness is relaxed, maximization can mean two different things: choosing elements that dominate all others under the preference, or choosing those which are themselves undominated. The latter is weaker, and even then, acyclicity is the minimal condition ensuring a choice can be made. Moving further from the standard rationality, one could accommodate cyclic relations if one properly adjusts the notion of preference maximization (including notions, for instance, of top cycle and uncovered sets), see e.g. Ehlers and Sprumont (2007) and Lombardi (2008) for behavioral implications of such rationalizations.

[40] Rubinstein and Salant (2008) provide an alternate characterization based on revealed relations.

Weak WARP *Suppose* $\{x, y\} \subseteq R \subset S$. *If* $c(\{x, y\}) = c(S) = x$, *then* $c(R) \neq y$.

In essence, y cannot be chosen from a menu sandwiched in between two menus where x is chosen. Weak-WARP characterizes choice functions arising under two distinct theories, namely Manzini and Mariotti (2012a)'s Categorize-Then-Choose, and Cherepanov, Feddersen and Sandroni (2013)'s Basic Rationalization theory (see Section 3.3). This provides another instance of how different choice processes can be indistinguishable in terms of the choice patterns they generate. As we have already seen, such occurrences are not limited to bounded rationality theories: both a simple version of satisficing and Mandler et al. (2012)'s model of checklists yield rational behavior.

In line with the fact that they underlie several different theories, Weak WARP and Condition γ are arguably meaningful, stand-alone properties. Other axioms appearing in the literature are more closely tied to the defining features of the theory they characterize. Consider Masatlioglu et al. (2012)'s Limited Attention and Lleras et al. (2017)'s Choice Overload. Each of these works explain how to use choices to construct a revealed preference; they show its transitive closure is all that one can unambiguously identify about the DM's preference. Such characterizations are relevant for welfare assessments if the social planner's goal is to maximize that preference.[41] Each theory is fully characterized by one WARP-like axiom which is tantamount to the acyclicity of the revealed preference:

WARP for Limited Attention *For each nonempty menu* S, *one can find an element* x^* *such that the following condition holds for all menus* T *containing* x^*: *if* $c(T) \in S$ *and* $c(T) \neq c(T \setminus \{x^*\})$, *then* $c(T) = x^*$.

WARP for Choice Overload *For each nonempty menu* S, *one can find an element* x^* *such that the following condition holds for all menus* T *containing* x^*: *if* $c(T) \in S$ *and* $c(T') = x^*$ *for some* $T' \supset T$, *then* $c(T) = x^*$.

The common element is the ability to find a 'best' considered alternative x^*. The reason we know x^* is considered, though, depends on the defining structural feature of the theory: in Masatlioglu et al. (2012) choice only changes when removing considered alternatives, while in Lleras et al. (2017) alternatives chosen in supersets are considered in subsets.

Other axioms are meaningful only within the context of a specific bias, if not a specific theory; and certainly, not all axioms are derivates of standard rationality axioms. For instance, Sagi (2006) studies the implications of the following axiom: if an alternative x is preferred over y when y is the status quo, then x would be preferred over y *whatever* the

[41] If, instead of being paternalistic, the social planner's goal is to respect the DM's choices without trying to "correct" them, then knowing the entire choice function, namely the DM's choice from each menu, is information enough to make choices on her behalf whenever needed.

status quo. This is a natural behavioral postulate under status-quo bias, and is indeed satisfied, for instance, by Masatlioglu and Ok (2005); but Sagi (2006) shows the axiom is violated, perhaps surprisingly, by several well-known models of reference dependence, such as Tversky and Kahneman (1992)'s cumulative prospect theory. Such results can clarify the differences between theories and delineate which applications they fit best.

4.2 Limited Data

Many axiomatic results operate on choice functions, which describe choices from *all* menus. This is often a critical assumption (see, for instance, Manzini and Mariotti (2007), Masatlioglu et al. (2012) and Lleras et al. (2017) mentioned above). For the simplest illustration, a choice function is rational if, and only, if satisfies WARP; yet selecting a from $\{a, b\}$, b from $\{b, c\}$, and c from $\{a, c\}$ satisfies WARP and is clearly irrational. If testing is the goal, then this full-data assumption is hardly tenable. Choice menus may be beyond control in empirical settings, or not vary much. A controlled experiment may be challenging, as the number of menus grows exponentially in the number of options.[42]

de Clippel and Rozen (2021) explore the testing of bounded-rationality theories with limited data, pursuing a methodology comparable to the SARP test for Rationality. Observed choices are consistent with a choice theory if they can be 'completed' (by positing choices for unobserved menus) into a choice function compatible with the theory.[43] Many theories use an ordering (or simply an acyclic relation) to capture, for instance, preference, power, or salience. It may then be possible to capture the testable implications in terms of a condition in the spirit of SARP: namely, the ability to find an acyclic relation fulfilling all the 'restrictions' revealed by observed choices. Under Rationality, for instance, seeing x chosen when y is available leads to the simple restriction "x is better than y." For bounded rationality theories, restrictions might take more complex forms, such as "x is better than y or z." One aspect of SARP's appeal is the ability to tractably check it, such as by iteratively identifying a candidate worst element among surviving alternatives.[44] de Clippel and Rozen point out that this easy testing approach extends to a more generalized class of restrictions, including the one above. Indeed, even if we don't know whether it is y or z that is worse than x, we certainly know x cannot be the worst element in a set containing y and z. They show a SARP-like test applies to several models of bounded

[42] With 10 alternatives, subjects must answer over 1,000 questions to generate a full choice function.

[43] de Clippel and Rozen point out that some earlier attempts at discussing limited data, including for instance Manzini and Mariotti (2007, Corollary 1), used an inadequate definition of consistency that can result in false positives and issues with out-of-sample predictions.

[44] This is possible if, and only if, there is no cycle.

rationality, such as Choice Overload (Lleras et al. 2017) and Triggered Rationality (Rubinstein and Salant 2006b, Kibris et al. 2021). As can be seen from results in Hu et al. (2022) and Maniquet and Nosratabadi (2022), a SARP-like test also applies for several other models, including Masatlioglu and Ok (2005)'s status-quo biased choice. Still, de Clippel and Rozen show that testing can become NP-hard for theories generating more general classes of restrictions, such as Limited Attention (Masatlioglu et al. 2012) and Shortlisting (Manzini and Mariotti 2007). Simple testing à la SARP remains possible for these theories when the dataset is rich enough: for instance, containing all menus of size two and three for Shortlisting, or the intersection of any two menus that violate WARP in the case of Limited Attention. This points to a potential role for adaptive data collection.

4.3 Stochastic Choices

For practical testing purposes, it is important not only to consider the possibility of limited data, but also the possibility of randomness in the data that one does observe. Moving beyond deterministic choices, think now of a theory $\mathcal{T}$ as generating a set $\Sigma_{\mathcal{T}}$ of stochastic choice functions (with deterministic ones corresponding to the special case where a single option from each menu is selected with probability one).

Recognizing that individuals can make errors in empirical and experimental applications, a natural question arises: are deterministic theories, and their characterizations in terms of observables, valuable in practical settings? Depending on how one envisions errors, the answer to both questions may be yes. Note that any deterministic theory can be expanded to accommodate stochastic choices. While $\Sigma_{\mathcal{T}}$ defines a benchmark, the possibility of errors may draw the eye to SCFs that are 'close' to $\Sigma_{\mathcal{T}}$. For $0 \leq \varepsilon \leq 1$, say σ' is *ε-close* to σ if, for each menu S, there is $\varepsilon_S \in [0, \varepsilon)$ and a probability distribution $\delta \in \Delta(S)$ such that $\sigma'(\cdot|S) = (1 - \varepsilon_S)\sigma(\cdot|S) + \varepsilon_S\delta(\cdot)$. Let $\Sigma^{\varepsilon}_{\mathcal{T}} = \{\sigma' \in SCF \mid \sigma' \text{ is } \varepsilon\text{-close to some } \sigma \in \Sigma_{\mathcal{T}}\}$, where SCF is the set of all stochastic choice functions. Though we depart from Apesteguia and Ballester (2021)'s discussion, it is easy to see that σ belongs to $\Sigma^{\varepsilon}_{\mathcal{T}}$ if, and only if, ε is part of their 'maximal separation' of σ given the theory $\mathcal{T}$ (in their terminology, ε is the probability of trembles around the deterministic model). Here, we use this construction to point out that if $\varepsilon \leq 1/2$, or noise is at most moderate, then any consistency test developed for a deterministic theory $\mathcal{T}$ extends at once into a consistency test for the expanded $\Sigma^{\varepsilon}_{\mathcal{T}}$. To see why, first notice choices are incompatible with $\Sigma^{\varepsilon}_{\mathcal{T}}$ if in some menu, the modal choice has probability less than or equal to 1/2. Upon ruling that out, define a deterministic observed choice function c_{obs} by selecting from any observed menu S the majority choice (the option with probability strictly larger than 1/2). Compatibility with $\Sigma^{\varepsilon}_{\mathcal{T}}$ is then equivalent to c_{obs} being consistent with $\mathcal{T}$. Furthermore, any SCF in $\Sigma^{\varepsilon}_{\mathcal{T}}$ is ε-close to a *unique* choice function in $\mathcal{T}$.

In other words, we observe that one may still apply the deterministic-theory test over the choice function comprised of majority choices. Though simple, this observation is nonetheless powerful, illustrating the robustness of deterministic theories, as testing and identification remains possible even when accommodating sizeable noise (up to a 50% chance of *any* other stochastic behavior). Consider for instance $\Sigma^{1/2}_{RAT}$, the largest stochastic extension of Rationality that preserves identification. A large part of the stochastic-choice literature focuses on binary menus. Notice that consistency with $\Sigma^{1/2}_{RAT}$ over binary menus essentially amounts to Tversky's (1969) property of Weak Stochastic Transitivity (except for having strict instead of weak inequalities). Indeed, because the majority choice must be rational, $c(x|\{x, y\}) > 1/2$ and $c(y|\{y, z\}) > 1/2$ imply $c(x|\{x, z\}) > 1/2$, for all triplets x, y, z. The interpretation is simple: errors are modest enough that choice probabilities continue to reveal underlying preference comparisons, which must be acyclic.

Of course, one may be interested in a more discriminating theory that pins down choice probabilities. The most prominent example is Luce (1959)'s model: given a utility function u, the probability of picking x from a menu S is $u(x)/\sum_{y\in S} u(y)$. Luce introduces a new form of IIA, satisfied by his model, which requires that the likelihood ratio of picking x versus y does not depend on what other alternatives are available:

Luce IIA *The function* $\frac{\sigma(x|\cdot)}{\sigma(y|\cdot)}$ *is constant over menus containing* x *and* y*.*

As Luce (1959) points out while assuming positivity of SCFs, and Cerreia-Vioglio et al. (2021) show more generally, Luce IIA is equivalent to a simple postulate:

Luce Choice Axiom $\sigma(a|A) = \sigma(a|B)\sigma(B|A)$ *for all* $a \in B \subseteq A$*.*

This amounts to a law of conditional probability: the probability of choosing a is the probability of choosing a submenu containing a, and then choosing a from within that. Cerreia-Vioglio et al. (2021) prove that an SCF satisfies this choice axiom if and only if its support comprises a rational choice correspondence. Hence Luce's model can be seen as a particular tie-breaking rule among preference-maximal alternatives, and entails more rationality than might appear at first glance.

This result reinforces a longstanding impression that Luce's model is quite demanding. As surveyed by Rieskamp, Busemeyer and Mellers (2006), a large literature suggests, and oftentimes characterizes, more permissive alternatives. A prime example is RUM, which also happens to provide an interesting case study on the purposes of characterizations. RUM satisfies the following property discussed by Block and Marschak (1960):

Regularity *For all menus S containing both x and y, $\sigma(x|S \setminus \{y\}) \geq \sigma(x|S)$.*

Under RUM, $\sigma(x|S)$ represents the probability that x is top-ranked within S. That probability cannot decrease when removing y, as the property of being top-ranked is preserved. Regularity is a classic and intuitive property, and ensuing contributions have tended to be categorized by whether or not they satisfy it. But Regularity is only a necessary property of RUM, and does not describe its full testable implications. As shown by Falmagne (1978) and Bàrbera and Pattanaik (1986), σ is consistent with RUM (i.e., arises from a distribution over rational choice functions) if, and only if, it satisfies the Block and Marschak (1960) inequalities requiring

$$\sum_{T:S\subseteq T\subseteq X} (-1)^{|X\setminus T|}\sigma(x,T) \geq 0, \text{ for all menus } S \text{ and all } x \in S. \tag{1}$$

These inequalities are straightforward, but notoriously lacking in intuition. It took Gul and Pesendorfer (2006)'s more intuitive axiomatic characterization for random expected utility, which takes advantage of a richer domain of choice, to help revive hope in the literature that more insightful understanding of stochastic choice behaviors is within reach.

More recent contributions weakening Luce's model include Gul et al. (2014), who provide a characterization of their attribute rules via an elimination and weak-independence axiom; Apesteguia, Ballester and Lu (2017), whose single-crossing random utility model weakens Luce IIA to an easy-to-test Centrality axiom, which involves only triplets; and Kovach and Tserenjigmid's (2022a) non-parametric nested logit model, where Luce's IIA is weakened based on a natural definition of revealed similarities. Kovach and Tserenjigmid clarify the axiomatic differences between various generalizations.

Rather than maximizing a random preference, another avenue for reconciling stochastic choices and rationality has been explored: maximizing a stable preference ordering over a randomly determined consideration set. Yet even when new models have such significant conceptual innovations, earlier axiomatizations may still serve as useful springboards and benchmarks. Remember, for instance, from Section 3.3 the stochastic-attention model proposed by Manzini and Mariotti's (2014), where each option has a certain probability of being considered (a default option x^* is systematically available and considered). Behaviors under their model are characterized by the following two axioms:

i-Independence *The function $\frac{\sigma(x|\cdot\setminus\{y\})}{\sigma(x|\cdot)}$ is constant over menus containing both x and y. Similarly, $\frac{\sigma(x^*|\cdot\setminus\{y\})}{\sigma(x^*|\cdot)}$ is constant over menus containing y.*

i-Asymmetry *For any S and $x \neq y$, $\sigma(x|S\setminus\{y\}) \neq \sigma(x|S)$ implies $\sigma(y|S\setminus\{x\}) = \sigma(y|S)$.*

i-Independence is similar in structure to Luce IIA, in that it requires a probability ratio to be menu independent. But by contrast to Luce (1959), Manzini and Mariotti (2014) presume a stable preference observed without errors, and the ratio reflects the way attention is allocated across options in a feasible set. Indeed, the impact of dropping y on the probability of choosing x depends only on the preference ranking of these two options (and not the menu). Either x is preferred to y, and dropping y is irrelevant, or y is preferred to x, and the probability increase for picking x is exclusively determined by y's attention probability. As for i-Asymmetry, if dropping y affects the probability of choosing x, then it must be that y is preferred to x; but then it cannot be that x is preferred to y, so dropping x will not affect the probability of choosing y. Though clearly necessary, Manzini and Mariotti (2014) establish that these two axioms guarantee a stochastic-attention representation in their sense (provided choice probabilities are observed in a sufficiently rich set of menus). Furthermore, each stochastic choice function in the model is associated to a unique combination of attention probabilities and preference.

Manzini and Mariotti's model imposes specific identities on choice probabilities across menus. Another approach to characterizing the theory is knowing how to invert the map associating choice probabilities to the primitives. For instance, an option's choice probability when it is uniquely available (beyond the inferior default) reveals its attention probability. Aguiar, Boccardi, Kashaev and Kim (2023) show how the map from primitives to choice probabilities can be inverted with more general models as well (including Brady and Rehbeck (2016)'s logit-attention model, Aguiar (2017)'s variant of Tversky (1972)'s elimination by aspect model, and situations where underlying preferences are themselves stochastic). Their approach leverages information about choices from multiple menus. For some stochastic-attention theories, Dardanoni et al. (2020) show the distribution of cognitive characteristics can be inferred even from a single, sufficiently large menu, increasing the applicability of these theories in empirical work.

To revisit the contrast between detailed and detail-free approaches to stochastic attention, recall Cattaneo, Ma, Masatlioglu and Suleymanov (2020)'s model, in which attention need only be monotonic: the probability $S' \subseteq S$ is the consideration set does not decrease when removing $x \in S \setminus S'$. In their model, $\sigma(x|S) > \sigma(x|S \setminus \{y\})$ implies that, in menu S, there is a strictly positive probability of having an attention set A containing y and from which x is selected. It must then be that x is superior to y. The empirical content of their model is captured by the acyclicity of this revealed preference. Choice probabilities reveal information about the underlying preference, but only inequalities on choice probabilities, rather than specific identities, come into play. Even though both Manzini and Mariotti (2014) and Cattaneo et al. permit randomness exclusively through attention sets, only the former satisfies Regularity. Indeed, as can be seen from above, the testable implications in Cattaneo et al. only arise from violations of Regularity.

Like for deterministic theories, characterization results for stochastic theories come in a variety of formats, to fulfill different purposes. Results may deepen our insights into the model, or provide ways to identify underlying primitives for the purposes of welfare analysis. But new econometric issues emerge when randomness is expected. For instance, data is naturally finite, and can only approximate the probability distributions taken for granted in axioms and characterizations. To this end, Cattaneo et al. (2020, 2021) develop econometric testing methods and assess their finite-sample performance. Proper econometric tests still rely on characterization results as a stepping stone. Beyond issues of sampling variation, there are questions of goodness of fit, power, and identification. There is the challenge of unobserved heterogeneity in aggregate (as opposed to individual) data.[45] Recent work tackles various subsets of these issues; see, for instance, Cattaneo et al. (2020, 2021), Abaluck and Adams-Prassl (2021), Barseghyan, Coughlin, Molinari, and Teitelbaum (2021), Barseghyan, Molinari and Thirkettle (2021), Dardanoni et al. (2022), de Clippel and Rozen (2022), Filiz-Ozbay and Masatlioglu (2022), Kashaev and Aguiar (2022), and Aguiar et al. (2023). We further discuss some of these in the next section.

5 Judging bounded-rationality models

Rationality partially owes its status as the workhorse of economic decision-making to parsimony: it is a simple model that explains only a small and structured set of choice patterns. The purpose of relaxing rationality is, of course, to capture a wider range of behaviors (i.e., become more *permissive*). But, by doing so, the model loses some of its empirical content. In this section, we gauge how permissive recent bounded-rationality models are, and whether or not they expand rationality in the right direction, by accommodating new choice patterns that are empirically prevalent.

5.1 How permissive?

It's helpful to take a moment to reflect on what it means for a theory to be permissive, and what yardstick one should use to measure permissiveness. Here we take a fairly

[45]Heterogeneity may also occur in mixture choice data comprising the joint distribution of behavior for several agents; Dardanoni et al. (2022) introduce this richer framework and study its advantages.

simple, counting-based approach based on Selten's measure of a theory's area (Selten and Krischker 1983, Selten 1991), and reminiscent of Bronars (1987)'s benchmark for assessing power when testing consumer choices. Namely, when treating all choice functions as equally likely, what is the chance a choice function is explained by the theory?[46] Imposing such a uniform prior captures agnosticism about potential behaviors and their relative prevalence. But some may judge permissiveness against other benchmarks. Experts may categorize some choice functions as implausible or unlikely from the outset, particularly if additional structure is imposed on the problem.[47] For instance, in the context of choice over lotteries, the modeler may posit that DMs avoid stochastically-dominated lotteries. Or the modeler could use observed behavior in a population to form her prior. One may then want permissiveness to reflect those considerations. This discussion is reminiscent of issues encountered with Bayesian model comparison, where conclusions may be sensitive to the selected prior, and uniform priors tend to be imposed as a means of self restraint. Since we have here an abstract, unstructured framework of choices, and relatively little is understood thus far about how permissive or restrictive are the different theories proposed in the literature, we find a simple counting exercise to be worthwhile

As a first step towards this end, consider the permissiveness of the weak-WARP axiom, which underlies multiple characterizations of bounded-rationality theories (see Section 4). While Rationality accommodates a quarter of possible choice functions (6 out of 24) when $|X| = 3$, weak-WARP places no restriction at all (violations require three nested menus). The number of possible choice functions grows dramatically in $|X|$: a mere 24 when $|X| = 3$ becomes 768 when $|X| = 4$, 309 billion when $|X| = 5$, and over 10^{37} when $|X| = 6$.[48] The numbers become unfathomable for $|X| \geq 7$. By contrast, there are $n!$ rational choice functions when $|X| = n$, which grows more modestly. In particular, the fraction of rational choice functions quickly vanishes (e.g. 120 out of more than 309 billion when $|X| = 5$, down to less than one in 10^{34} when $|X| = 6$). The fraction of choice functions satisfying weak-WARP is of course larger. Does this fraction still vanishes as $|X|$ increases? And how much more permissive is weak-WARP compared to rationality?

As it turns out, it is extremely rare for a choice function to satisfy weak-WARP, but

[46]More generally, one could estimate the chance of being 'near' a theory, to allow for imperfect fit and have a richer understanding of the dispersion in possible behaviors; see de Clippel and Rozen (2022).

[47]Fudenberg, Gao and Liang (2021) define a theory's *restrictiveness* as the normalized average error of a theory across only those 'admissible' datasets satisfying some basic properties (e.g., respecting first-order stochastic dominance), so as to identify how much further structure the theory imposes.

[48]The number of menus in $|X|$ which contain n alternatives is given by the binomial coefficient $\binom{|X|}{n}$. For each option in each of these menus, one can construct a choice function that selects it, so the number of possible choice functions is $\prod_{n=1}^{|X|} n^{\binom{|X|}{n}}$. For instance, when $|X| = 3$, this is $3^1 \cdot 2^3 = 24$.

also extremely rare for a weak-WARP choice function to be rational when X is large. As we now show, these combined features are not specific to this axiom, but shared by many theories. For this, we introduce some rather mild regularity properties on how a theory applies as X varies. Let $\mathcal{X}$ be a countably infinite space of alternatives, with each conceivable set X a finite subset of $\mathcal{X}$. A theory $\mathcal{T}$ is a family $\{\mathcal{T}_X\}_{X \subseteq \mathcal{X}}$, with $\mathcal{T}_X$ defining all choice functions over X that are consistent with the theory given X. The first property captures an idea of neutrality, an implicit assumption of the theories we have discussed.

Property (i) Options' labels do not matter: given a bijection $\Psi : X \to X'$, c is consistent with the theory given X if, and only if, the translation of c using Ψ into a choice function over X' is consistent with the theory given X'.

The second property expresses a form of consistency in how X is treated by the theory, a bit in the spirit of reduced-game properties. This is satisfied by almost all the theories we have discussed, with the exception of theories where some aspirational alternative that is infeasible affects behavior (see, for instance, Section 3.4).

Property (ii) Take any $X \subset Y$, and a choice function c defined over Y. If c is consistent with $\mathcal{T}_Y$, then its restriction over X is consistent with $\mathcal{T}_X$.

The final property is also shared by many theories. It describes how a choice function consistent with $\mathcal{T}$ given X, and another choice function consistent with $\mathcal{T}$ given a disjoint set Y, can be naturally combined to obtain a choice function consistent with $\mathcal{T}$ on $X \cup Y$.

Property (iii) Suppose X and Y are disjoint, c is consistent with $\mathcal{T}_X$, and c' is consistent with $\mathcal{T}_Y$. The choice function selecting $c(S \cap X)$ from each menu $S \subseteq X \cup Y$ intersecting X, and $c'(S)$ otherwise, is consistent with $\mathcal{T}_{X \cup Y}$.

It is easy to check that the set of weak-WARP choice functions satisfies Property (iii). To demonstrate (iii) holds for rationality, note that one may extend to $X \cup Y$ by simply stacking the elements of X above Y in the preference. A similar idea extends to many other theories considered in this survey. For Triggered Rationality (Rubinstein and Salant 2006b; Kibris et al. 2021b) one may stack X above Y in both the preference and the salience ranking. For two-stage choice models with a consideration set Γ, stack the preference as above and define $\Gamma(S) = \Gamma^X(S \cap X)$ when $S \cap X \neq \emptyset$ and $\Gamma(S) = \Gamma^Y(S)$ otherwise; for both Masatlioglu et al. (2012) and Lleras et al. (2017), Γ inherits the defining property of the model from Γ^X and Γ^Y. For Shortlisting (Manzini and Mariotti 2007), simply assume elements of X eliminate all elements of Y in the extended shortlisting relation.

Of course, a theory can only have predictive power if it is *nontrivial*, in the sense that there exists a finite X for which only a strict subset of choice functions over X are consistent with the theory given X. It is *more permissive than rationality* if it contains all rational choice functions, and there is some finite X for which a non-rational choice function is consistent with the theory given X. The following observation sheds some light on the permissiveness of theories satisfying the properties above.

Observation 1. *Let $\mathcal{T}$ be a theory satisfying Property (i), and let $(X_n)_{n\geq 1}$ be a sequence of conceivable sets with strictly increasing cardinality.*

(A) *If $\mathcal{T}$ is nontrivial and satisfies Property (ii), then the fraction of choice functions consistent with $\mathcal{T}$ given X_n monotonically decreases to zero as $n \to \infty$.*

(B) *If $\mathcal{T}$ is more permissive than rationality and satisfies Property (iii), then the fraction of rational choice functions among those choice functions consistent with $\mathcal{T}$ given X_n converges to zero as $n \to \infty$.*

See the Appendix for the proof. Observation 1(A) does not mean that theories which violate Property (i) or (ii) necessarily explain large sets of behaviors. What the result does tell us is that virtually all theories defined until now are infinitely discerning when tested with enough conceivable alternatives. This might come as a (good) surprise about the restrictiveness of bounded rationality theories. Giarlotta, Petralia and Watson (2022) independently find a result similar to Observation 1(A) and also provide some numerical bounds for the fraction of choice functions permissible under some theories, as a function of the number of conceivable alternatives.[49] While Observation 1(B) might seem to contradict 1(A) at first glance, there is no inconsistency. Rather, Observation 1 points to the fact that bounded rationality theories may be much more permissive than rationality while still being highly discerning in which behaviors are allowed.

What about stochastic choice theories, which associate to each menu S a probability distribution? With $|X| = n$, let the binomial coefficient $\alpha(n, s) = n!/s!(n-s)!$ denote the number of size-s menus in X. Then the set of SCFs is a subset of $\mathbb{R}^{\sum_{s=1}^{n} \alpha(n,s)(s-1)}$. Each deterministic choice function defines a vector in that space, and the set of all possible SCFs is the convex hull of these (independent) vectors. By contrast, the set of SCFs arising under RUM is the convex hull of only those vectors arising from a rational choice function. For instance, the set of SCFs associated to five options is a convex polytope in $\mathbb{R}^{49}$ with over 309 billion extreme points, while RUM corresponds to a sub-polytope generated by only 120 of these vectors. Though the number of extreme points is dramatically smaller, there is nonetheless a positive measure of RUM SCFs.[50] Still, that measure converges to zero as n goes to infinity. A similar convergence result holds for most stochastic-choice

[49] Their notions of 'isomorphic choices' and 'hereditary property' play the same roles as (i) and (ii).

[50] To see this, observe that the SCF which is the uniform distribution over elements of each menu (aris-

theories, as Observation 1(A) above naturally generalizes beyond deterministic choices: if there is a set of options for which a theory is violated over a positive measure of SCFs, then analogs to Properties (i) and (ii) imply that the measure of permitted behaviors tends to zero as n goes to infinity (as consistency over X requires consistency over each atom of a partition where consistency has probability less than one).

5.2 How successful?

If we believe bounded rationality theories are not overly permissive, then we must naturally worry whether they are permissive enough: are the *right* patterns of choice being explained? Moreover, if we acknowledge that data is noisy and that theories may not be perfectly describe choices, how do we measure a theory's goodness of fit to the data?

To contemplate these questions, one must have an approach in mind for judging the success of theories. Selten's measure offers a fairly enduring and familiar method.[51] Wisely disadvantaging theories where almost anything goes, it subtracts the size of the theory's predicted behaviors (the theory's *area*, discussed in Section 5.1) from the frequency of correct predictions (the theory's *hit rate*). Though area generally shrinks to zero as the set of conceivable alternatives grows (see Section 5.1), it may well be positive and computable, or at least easily approximated by simulation, when the number of options is limited. Evaluating a theory's hit rate, however, presents an empirical challenge. Naturally, one must have data at hand and the econometric tools to properly evaluate it.

When it comes to developments in the last couple of decades, the picture is still incomplete. There are far fewer full-fledged tests of theories than there are theories themselves. We briefly discuss some of these empirically oriented works here. Of course, many works are motivated by previously observed behavioral patterns (as discussed in Section 3). Here, we focus on papers whose purpose is to provide proper tests using data.

Manzini and Mariotti (2006) elicit two complete choice functions from each participant in a laboratory experiment (each over a different grand set of four alternatives, the

ing from a uniform distribution over all preference orderings) *strictly* satisfies Block and Marshak (1960)'s inequalities (see Equation (1) of this survey). Letting $|X| = n$ and $|S| = m$, the inequality amounts to $\sum_{j=0}^{n-m}(-1)^j\binom{n-m}{j}\frac{1}{m+j} \geq 0$ for all $m \leq n$. The LHS equals $\sum_{j=0}^{n-m}(-1)^j\binom{n-m}{j}\int_0^1 x^{m+j-1}dx$. Rearranging, this equals $\int_0^1 x^{m-1}(\sum_{j=0}^{n-m}\binom{n-m}{j}(-x)^j)dx$. The term in parentheses is the binomial expansion of $(1-x)^{n-m}$, so the LHS equals $\int_0^1 x^{m-1}(1-x)^{n-m}dx$, which is strictly positive.

[51] Other approaches have been proposed as well. Most recently, Fudenberg, Kleinberg, Liang, and Mullainathan (2022)'s *completeness* evaluates the theory's reduction of out-of-sample prediction error beyond a naïve model, relative to the maximum possible reduction of error (estimated by machine learning). Fudenberg et al. (2021) consider the tradeoff between completeness and their notion of restrictiveness.

minimum number needed to potentially violate weak-WARP). Testing the plausibility of basic axioms, including those characterizing the models of Manzini and Mariotti (2007, 2012a), they show WARP is often violated, Expansion is violated at a slightly lower rate, and weak-WARP is violated at much lower rate.[52] The alternatives in this experiment are sequences of monetary payments in two or three installments; and Manzini et al. (2010)'s examination of the pairwise choices suggests that introducing an initial shortlisting stage, along the lines of Manzini and Mariotti (2007), outperforms discounted utility alone. Filiz-Ozbay and Masatlioglu (2022) also revisit Manzini and Mariotti (2006)'s dataset to illustrate how to identify behavioral types in their progressive random-choice model. For each grand set of alternatives, the three or four choice functions which comprise the vast majority of behaviors do have the progressive structure specified by their model.

Oprea (2020) and Banovetz and Oprea (2022) experimentally test the conceptualization of a DM as a finite-state automaton, such as in Salant (2011). Oprea (2020) establishes subjects are averse to using complex automata, by analyzing the willingness to pay to avoid using previously assigned rules in later problems. Banovetz and Oprea (2022) observe subjects endogenously forming their own automata to make repeated selections in a bandit problem, and note a preference for simple though suboptimal rules. Subjects choose more complex automata when a computer alleviates the mental cost of keeping track of states. Salant and Spenkuch (2022) test the impact of complexity in a large dataset of chess end-games; they examine the selection probability of a move as a function of its 'depth-to-mate', a proxy for the difficulty of evaluating the move's benefits. Consistent with their model of satisficing with noisy valuation, they find complexity has a negative (positive) effect on the choice probability of winning (losing) moves.

Moving to theories of stochastic choice with limited attention, Abaluck and Adams-Prassl (2021) experimentally manipulate subjects' attention and show they can empirically identify consideration probabilities under a class of models containing Manzini and Mariotti (2014). Barseghyan et al. (2021) analyze a version of RUM in a population of individuals, while presuming that choice sets are not observed and allowing for the possibility that DMs (households in their data) have limited attention. Requiring that consideration sets have at least two options, and accounting for heterogeneity in the data, they conclude that at least 75% of DMs consider strictly fewer alternatives than are ac-

[52] Manzini and Mariotti test the prevalence of the axioms directly. Another potential approach might instead borrow from Nielsen and Rehbeck (2022), who study whether violations of classic axioms of choice under risk are genuine or mere mistakes. Namely, they also ask DMs whether they would like to satisfy certain axioms (including the axioms' opposites to help avoid experimenter demand effects) and allow DMs to revise their choices if they ultimately contradict their axiomatic selections.

tually feasible. Aguiar et al. (2022) use a large online experiment to obtain aggregate stochastic choice data; they reject that RUM explains the data at all significance levels. They further consider two extensions of Manzini and Mariotti (2014)'s model: Brady and Rehbeck (2016) and Aguiar (2017). They reject that Aguiar (2017)'s model explains the data, but cannot reject Brady and Rehbeck (2016)'s logit attention model.

Certainly, one does not expect a model to perfectly explain choices, and some papers have explored different criteria to judge relative success. Natenzon (2019) uses stochastic-choice data from frog matings to estimate parameters of his Bayesian probit model. Using the Aikake information criterion (AIC), he observes that the Bayesian probit model fits the data better than the RUM, logit and probit models; his model also mostly outperforms these other ones in an out-of-sample prediction exercise using this data. Apesteguia and Ballester (2021) analyze a previously collected dataset where each subject makes a choice once per menu. Using the stochastic-choice dataset obtained by aggregating all subjects' choices from binary menus, they compare the performance of rationality, Luce (1959) and the single-crossing random-utility model of Apesteguia, Ballester and Lu (2017). Introducing a notion of 'maximal separation', they identify the (stochastic) choice function in a model that receives most weight under the observed behavior (with remaining weight placed on some unrestricted stochastic choice function capturing unmodeled error). They use the selected stochastic choice function to examine out-of-sample predictions, and find that the single-crossing random-utility model outperforms the others.

The above list is not exhaustive. But it is clear that steadfast conclusions about the performance of theories have not yet been reached, simply because more recent theories of bounded rationality have received relatively little empirical scrutiny (and sometimes none) in comparison to older theories such as rationality, some stochastic versions of transitivity, and RUM. Indeed, the importance of thorough testing is illustrated by Tversky (1969) and an ensuing literature in mathematical psychology. Tversky (1969) selected 8 out of 18 recruited subjects who, based on initial screening, exhibited a high likelihood of intransitivity. For those 8 subjects, Tversky collected stochastic choice data from all binary menus out of a grand set of five options, by asking 20 repeat questions per menu while implementing techniques to avoid memory effects. Tversky concludes there is substantial intransitivity in the data, based on counting violations of weak stochastic transitivity using the observed choice frequencies. Iverson and Falmagne (1985) point out, however, that a direct examination of choice frequencies does not constitute a proper statistical test; and upon taking sampling variation into account, they cannot reject transitivity at the 5% level for all but one of the 8 subjects Tversky used. Regenwetter et al. (2011)

replicate Tversky's experiment, adjusted to contemporaneous dollar values, over a full set of 18 subjects that are not pre-screened; they conclude that RUM is very successful in capturing the individual-level data. More recently, this literature has expanded beyond binary choice data. McCausland et al. (2020) collect for each subject 6 repeated choices from all possible menus over the same five lotteries studied in Tversky (again, adjusted to contemporaneous dollar values) and provides support for RUM.

McCausland et al. (2020) base their conclusions on Bayesian model comparison, a classic tool for model selection that accounts for sampling variation and a model's size (as measured by the prior probability of the model). de Clippel and Rozen (2022) suggest factoring in goodness of fit into Bayesian model comparison, as a theory is still valuable even when predicting behavior reasonably well. They suggest a two-step approach of fit-optimized Bayesian model comparison: find an optimal fit for each baseline model, and then compare across baseline models' by assessing the performance of the fit-optimized expansions. Measuring fit using an intuitive distance–an approach in the spirit of Apesteguia and Ballester (2021)'s maximal separation–they revisit Regenwetter et al. (2011)'s data to assess the performance of several models, both at the individual level and when explaining several subjects at once. Their analysis supports parsimonious specifications of rationality and the Luce rule.

6 Interpreting data: the welfare debate

Standard welfare economics is mainly concerned with interpersonal settings: since each person rationally maximizes their own utility, the main question is how to reconcile conflicting interests. Indeed, welfare is exceedingly simple when there is a single, rational DM; if she chooses x when y is available, then we automatically infer x is better than y. We need not worry, for instance, that she misperceived the utilities of x and y, did not pay attention to y, or liked x only because z was also feasible. But we must begin to worry about these things once bounded rationality enters the scene, even with a single DM. We begin to wonder, is the DM truly making the best choice? If she does not have a preference ordering, then how do we assess what is best? Can we differentiate between preferred choices and mistakes? Should we ever intervene, and how?

There have been many thought-provoking contributions, but no clear consensus on these questions. While all the approaches we discuss use choice behavior to make welfare inferences, they differ on whether the inference should be made in tandem with a model (and if so, which one). The underlying differences are philosophical: how much confidence should we place on choices as a reflection of preferences, and is it unnecessarily paternalistic to lack full confidence? We do not pretend to close the debate on these deep questions, but rather only survey the incarnations of different possible answers.

6.1 On model-based welfare inference

A first observation regarding model-based welfare inference is that the ability to draw clear conclusions can depend on the model in question. Consider, for instance, Manzini and Mariotti's Rational Shortlist Method: the model includes an asymmetric preference that the DM uses to choose from the shortlist, but the possibility that this preference may have cycles complicates welfare inference.[53] Alternatively, consider the theory of Triggered Rationality discussed earlier: even if we can infer all of the DM's reference-dependent preferences, which one is the *right* one to use for welfare analysis?

This brings us to the anti-paternalistic view of Sugden (2004), who observes modeling the individual as a collection of selves "comes at a cost: we are left with no preference-based concept of welfare that applies to the person as a continuing entity" (p. 1017).[54] He suggests viewing the agent as a "continuing locus of responsibility" (p. 1018), who might make sequential decisions based on changing objectives, and may end up in a worse place than where she began, but accepts that each decision was optimal at the time.

Though disagreeing on many other aspects, the works discussed here generally take the view that there is some room to help the DM make better choices. Besedeš et al. (2015)'s experiment on choice overload, for instance, suggest that such assistance would be helpful: they find individuals have difficulty identifying the choice architecture that most improves their decision quality. In view of the considerations above, though, some important conceptual ingredients are needed to make a convincing welfare inference. These include a single acyclic preference clearly designated for welfare purposes, and some confounding factor(s) that prevents its maximization. There is some disagreement on whether the confounding factors need to be modeled; but certainly, the large class of two-stage models, where the DM maximizes a preference ordering over a consideration set, conveniently fits this bill, as might models involving search costs or difficulty differentiating alternatives.

A second observation is that even if one fixes a model (with a single preference), differ-

[53]Such cycles are permitted whenever the shortlisting relation, which need not be welfare relevant, eliminates an element from the cycle and thereby permits the DM to make a choice.

[54]Can we infer that a multi-self model is redundant? Liang (2019) studies recovering the number of underlying preferences with high probability, by attempting to distinguish between error and heterogeneity in a multi-self model along the lines of Kalai et al. (2002), where each self may act as dictator in some information set. In the case that one preference is inferred, welfare evaluation is clearly possible.

ent preferences orderings may be 'plugged in' to explain the data. For instance, it may be possible to explain the same choices using two different preference orderings in Masatlioglu et al. (2012), by pairing them with different, appropriately chosen consideration set mappings. A milder issue actually arises with *rational* choices, as studied in Varian (1982): it may be impossible to infer a rational DM's entire preference ordering without having a rich-enough dataset (Varian considers budget data, which is naturally incomplete). As discussed in Section 2, Varian proceeds cautiously, deciding x is revealed preferred to y only if *all* rationalizing preference orderings share that ranking. The bounded rationality literature has generally followed Varian's lead, identifying a revealed preference only when all rationalizations under the model in question require that ranking to explain the data.

Neither Varian's assessments, nor those in the bounded-rationality literature, should be confused with a stronger statement one might desire: that *all models* which explain the data unambiguously require that preference ranking to hold. Unfortunately, there is no hope to achieve such consensus, even for basic choice patterns. Consider this data:

$$C(\{x,y,z\}) = \{y\},\ C(\{x,y\}) = \{x\},\ C(\{y,z\}) = \{z\},\ C(\{x,z\}) = \{x\}.$$

Masatlioglu et al. (2012) would conclude from these choices that y is preferred to z, while Cherepanov et al. (2013) would conclude the opposite. Similarly, while rational choice data is explainable by preference maximization, it is also explainable by preference minimization, leading to fully opposite welfare conclusions (Rubinstein and Salant, 2006b)![55] Ultimately, any inference of revealed preference à la Varian will be model dependent.

6.2 On model-free approaches

It would certainly be convenient to have a model-free approach, one that uses choices to infer a (possibly partial) welfare ranking over alternatives which is credible regardless of the DM's choice process. This requires us to assume, however, that the choices express the DM's true desires over the feasible set, given the context in which they made their choice. This is the philosophical basis of Bernheim and Rangel (2009), who partition choice data using ancillary conditions and offer ways to generalize the classic definition of revealed preference to nonstandard choice correspondences. They take a quite different view from model-based welfare analysis: "welfare is *defined* in terms of choice rather than

[55] As we know, rational choices also arise from satisficing with a fixed evaluation order, Mandler et al (2012)'s model of checklists, etc. Needing to make a normative judgment to pick a model isn't due to rationality's fairly weak requirements: Dillenberger et al. (2015) make a similar point for expected utility.

underlying objectives" and while "preferences and utility are useful *positive tools*....they play no direct role in normative analysis" (p. 52). They define a *generalized choice situation* as a pair (S, d) comprising a choice set S and an ancillary condition d.[56] The latter is some aspect of the situation, such as the presentation of information, time of day or presence of a 'default', which the social planner considers irrelevant to welfare but can nonetheless affect behavior. The objects the DM is willing to choose when facing (S, d) are denoted $C(S, d)$. The welfare-relevant domain $\mathcal{G}$ consists of those generalized choice situations that the social planner examines for normative guidance. Two assumptions are made. First, $C(S, d)$ is presumed nonempty for all $(S, d) \in \mathcal{G}$. Though this is a common technical assumption, those concerned about indecisiveness in behavioral settings may take issue with it. Second, it is assumed that every subset $S \subseteq X$ has some ancillary condition d with $(S, d) \in \mathcal{G}$. In other words, the choice from any set can be normatively relevant under some conditions, *and* the planner has such information for all sets.[57]

With these assumptions in mind, Bernheim and Rangel extend properties satisfied by standard revealed preference for *rational choices in the standard framework*, to define welfare relations for potentially irrational choices in generalized choice situations. For instance, when choices satisfy WARP, the standard strict revealed preference xPy derived from $x = C(\{x, y\})$ implies $y \notin C(\{S\})$ whenever $x, y \in S$. By analogy, Bernheim and Rangel's say that x is *unambiguously chosen over* y, denoted xP^*y, if, and only if, for all generalized choice situations in the welfare-relevant domain $\mathcal{G}$, $y \notin C(S, d)$ whenever $x, y \in S$. Bernheim and Rangel (2009, Theorem 1 and Corollary 1) provide assurance that P^* is a viable welfare relation: though potentially incomplete, it must be acyclic. This result hinges, though, on having data for all choice sets. Consider the following example of choices given the domain $X = \{x, y, z\}$ and ancillary conditions d, d', d'', d''':[58]

$$C(\{x, y\}, d) = \{x\},\ C(\{y, z\}, d') = \{y\},\ C(\{x, z\}, d'') = \{z\},\ C(\{x, y, z\}, d''') = \{x, y\}$$

[56] Salant and Rubinstein (2008) independently introduce the similar framework of *choice with frames*, but do not focus on welfare analysis.

[57] Whether the data is rich enough (in the sense above) depends on how the modeler differentiates between ancillary conditions and choice objects. Bernheim and Rangel distinguish between experience of choice and experience of consumption. Contrast the purchase of ice cream for consumption on a hot day, with the purchase of ice cream on a hot day for future consumption. A choice object is arguably tied to the experience of its consumption, and weather is reasonably a characteristic of the ice cream in the former case. In the latter case, weather must be taken as an ancillary condition for their framework to be helpful: if the choice environment is a characteristic of the choice object, then the planner's delegated decision (taken in potentially different weather) may not be equivalent to the DM's undelegated one.

[58] The ancillary conditions need not be distinct for this example. Also notice that the same example shows that P^* is not simply the asymmetric part of their *weakly* unambiguously preferred relation R', which says $xR'y$ if for all $(S, d) \in \mathcal{G}$, $x \in C(S, d)$ whenever $y \in C(S, d)$ and $x, y \in S$.

While these choices imply yP^*z, there would be a P^* cycle over X were the choice from $(\{x, y, z\}, d''')$ unobserved. The possibility P^* is cyclic when the data is incomplete may make the approach difficult to apply when the number of alternatives is large, as the number of choice problems grows exponentially. Taking this question to the data, Bouacida and Martin (2021) consider the grocery store purchases of 1,193 tracked individuals. All exhibit revealed-preference cycles, and 19% still have a cyclic welfare preference using the method above; they do, however, find a promising level of predictive power nonetheless.

To help address the matter of cyclicity, Nishimura (2018) develops the notion of the *transitive core*. He requires sufficiently rich data, in that the social planner must observe every binary choice from some space of alternatives $S \subseteq X$. Conceptually, Nishimura presumes the DM has a complete and transitive preference relation which, through some potentially unknown process, is imperfectly manifested in choices. The standard revealed preference $\succeq$ that is gleaned from those choices is complete but potentially intransitive, and is mapped by a *welfare evaluation rule* into a transitive but potentially incomplete binary relation. How does one know which welfare evaluation rule to use? Nishimura poses six properties of welfare evaluation rules he considers sensible, and shows these uniquely characterize, for each revealed relation $\succeq$, its transitive core $c(\succeq)$:

$$x \; c(\succeq) \; y \quad \text{if and only if} \quad \begin{cases} z \succeq x \text{ implies } z \succeq y \\ y \succeq z \text{ implies } x \succeq z \end{cases} \quad \text{for every } z \in S.$$

The assumption that the revealed preference is complete (that is, that we observe all binary choices) is critical for Nishimura's result: as he points out, "if we extend the domain of welfare evaluation rules by including incomplete preference relations, then there is no welfare evaluation rule that satisfies all the axioms" (p. 589). Two of Nishimura's axioms are basic technical conditions, such as invariance to the labeling of alternatives; another two are structural properties describing how the relation is preserved when shrinking or expanding the set of possible alternatives;[59] and the final two are simple consistency relationships with the standard revealed preference $\succeq$: the welfare relation must be a subset of $\succeq$, and must respect it for elements not involved in any cycles.

[59]'Upward consistency' says that if the welfare rule puts x above y in all spaces of three alternatives, then it puts x above y for all spaces; 'downward consistency' says that if the welfare rule puts x above y in a large space of alternatives, then it still does so in any smaller space containing x and y.

6.3 Illustrating the differences

Perhaps surprisingly, some of the divisions in the welfare debate are easily illustrated by Nishimura (2018)'s seemingly innocuous consistency properties. These imply that when the binary choice data is rational, then the welfare evaluation rule should coincide with the standard revealed preference. Consider the following possible choice data:

$$C(\{x_1, x_2, x_3, x_4\}) = x_4,\ C(\{x_2, x_3, x_4\}) = x_2,\ C(\{x_i, x_j\}) = x_{\min\{i,j\}} \text{ for } i, j \in \{1, 2, 3, 4\}$$

Nishimura's framework gleans information from binary sets and ignores the information from larger sets. His welfare evaluation rule would conclude x_i is better than x_j whenever $i < j$. Though it is also a 'model-free' approach, Bernheim and Rangel's P^* would not come to the same conclusion: for instance, it cannot compare x_4 with any other option (this must be true no matter how other choices are filled in to obtain a complete dataset).

Moreover, both of these 'model-free' approaches, which disagree with each other, also disagree with the welfare implications of some well-known models. For instance, using de Clippel and Rozen (2021, Proposition 6), it is easy to see that the above choices are consistent with Mastalioglu, Nakajima and Ozbay (2012)'s model of limited attention and that, for any rationalizing preference, Nishimura's ranking is nearly reversed: x_4 must be preferred to x_1, and it must also be preferred to at least one of x_2 and x_3. Having said this, some models of bounded rationality, such as Manzini and Mariotti (2007, 2012a) and Cherepanov, Feddersen and Sandroni (2013), do have the feature that when the model is consistent with the data, there exists a rationalizing preference that respects binary choices (even if other rationalizing preferences may contradict them). In those cases, Nishimura's welfare rule can be viewed as a selection among admissible preferences. Bernheim and Rangel (2008)'s approach, however, cannot, as the above-listed models infer x is preferred to y if $x = c(S)$, $y = c(T)$, and $x, y \in S \subset T$, in which case P^* is silent. On the other hand, Ok and Tserenjigmid (2022) find support for the Bernheim and Rangel welfare inference within a class of stochastic choice models (those induced by lack of strict preference).

6.4 Searching for common ground

All this leaves the literature on welfare in murky waters. In philosophical treaties on the subject, Rubinstein and Salant (2006b), Köszegi and Rabin (2007a) and Manzini and Mariotti (2014) argue that it is impossible to assess welfare without a model in mind. Köszegi and Rabin (2007a) make the point that 'mistakes' can be systematic, and meaningful welfare inferences are possible by incorporating them into models. Manzini and Mariotti (2014) emphasize that the model-free approach, which guarantees a transitive welfare preference by utilizing only a subset of observations, comes at a cost: the dropped observations may reveal information about the DM's underlying, boundedly-rational model.

What does seem to be clear is that there is no one-size-fits-all approach. Whether taking a 'model-free' or 'model-based' approach, some assumptions and restrictions are inevitably made. Several works that focus on welfare seem resigned to such a middle-ground perspective, taking advantage of restrictions while allowing for some flexibility. Chambers and Hayashi (2012) suggest that "instead of arguing that there is a 'right' or 'true' model ... the economist should recognize that the definition of welfare is a subjective notion" (p. 1819). Allowing for framing, they consider *situation-dependent* stochastic choice data from menus and draw on some standard axioms to suggest that the data should be aggregated linearly, using some weighting rule, to generate welfare inferences. Subjectivity enters through the modeler's ability to set those weights: for instance, the modeler might trust the data from smaller sets more and relatively overweight those choices. Apesteguia and Ballester (2015) also study welfare inference from stochastic choice data using an axiomatically-founded measure, though using a different construction. Fixing a preference, their *swaps* index averages (over all observations) the number of alternatives that rank above the choice; and their *swaps preference* minimizes this index. Though typically weighting observations by their relative frequency, they too accommodate weight adjustments based on contextual information. Though the above works specifically consider stochastic choice data, such approaches have also been studied for deterministic data. Indeed, Rubinstein and Salant (2012) occupy a similar space, though with more emphasis on the decision process. Combining the feature of a true underlying preference with frames (a.k.a. ancillary conditions), they envision a preference maximizer whose true, underlying preference becomes distorted by frames in specific, testable ways. A theory of the distortion formalize how closely choices reflect desires, and thus how seriously we should take them. The distortion might, for instance, only change the relative ranking of outcomes that are adjacent in the underlying preference (i.e., $x \succ y \succ z$ implies $x \succ' z$).

We close this section pointing to one direction in which there is mostly consensus in terms of its potential promise, and much fertile ground for development: the use of information beyond choices themselves to assess and improve welfare. For instance, Rubinstein and Salant (2012) take frames into account to test a conjectured theory of distortion and make welfare predictions. Benkert and Netzer (2018) delve deeply into this framework, studying properties of different distortion functions and when optimal ‘nudges’ (frames) to improve choices can be identified. Goldin and Reck (2020) build some additional structure into Salant and Rubinstein (2008)’s theory of choice with frames, and use pension-choice data with known defaults and indicators of decision quality (e.g., time pressure) to try to infer preferences in a population. Abaluck and Adams-Prassl (2021) estimate welfare improvements from default prescription drug plans under some stochastic consideration set models with defaults, including Manzini and Mariotti (2014).

Rather than being an exogenous part of the environment, like a frame, the additional data may be generated by the DM herself. Caplin et al. (2011)’s gathering of intermediate decision data, useful for the classification of DMs’ heterogenous search behaviors, may inform DM-specific choice architecture. Additional data may also help disentangle noise from intended behavior, getting to the heart of the welfare question: some irrational choices might simply be mistakes, and the DM may wish to revisit those decisions if confronted with them. Though they studied classic axioms for choice under risk, one could imagine taking the approach of Nielsen and Rehbeck (2022) to bounded rationality models by obtaining information on the DM’s preferences over relevant axioms of behavior. Response times, reflecting a more subconscious choice output of the DM, may be useful as well. Alós-Ferrer et al. (2021) enrich random-utility models with the classic chronometric function, which captures the notion that more difficult decisions take longer. In an elegant result pointing to the role of theory in deciphering such richer data, they show that response times can identify a DM’s preference while imposing strictly weaker assumptions on the noise structure than would be needed if using standard choice data alone.[60]

7 Looking Forward

A broader picture of the literature emerges from this survey. We see a voluminous and yet still vibrant literature, offering a rigorous and multifaceted treatment of boundedly-rational choice. The literature is neither adolescent nor anywhere close to retirement; rather, it is the perfect time for a mid-life assessment. There remains much room for growth in theoretical and applied directions, both of which may interact with the above-noted potential for enriched data.

Broadly speaking, there are two philosophical approaches in the literature. One focuses mostly on the axiomatic analysis of choice procedures. Another focuses on modeling the DM’s procedural and computational constraints and studying the implications, following

[60]Rubinstein (2007) spurred interest in response times; see also Rubinstein (2013), Woodford (2014), and Fudenberg et al. (2018), among others.

Simon (1955)'s lead and answering Conlisk (1996)'s appeal. There can be an admittedly fine line between these approaches. For instance, assuming the DM pays attention to only the first n alternatives in a list (corresponding to an extreme type of attention-cost function) may be categorized under the latter approach, while assuming a condition on attention correspondences that generalizes the first-n-alternatives constraint may be categorized under the former. Similarly, though it is not explicitly couched in terms of a cost, sequential application of criteria to eliminate alternatives may be seen as arising from a cognitive constraint that the DM is unable to aggregate criteria, and can mentally process only one criterion at a time. Still, there is an arguably greater focus on the axiomatic approach in the choice theory literature. Despite several important contributions to the modeling of procedural and computational constraints, we believe there remains much potential for further developments in this vein. Such advances would build additional bridges to a literature on psychologically-founded models of decision making and a literature in game theory that incorporates computational and/or procedural constraints.

Another relatively underdeveloped area, despite some important contributions, is in the testing and comparative evaluation of theories (and their underlying axioms). The resulting challenges blaze open important and wide avenues for research. Shedding greater light on these questions will require investment by experimental economists in the creation of rich data sets. But it also requires theoretical advances. Indeed, this survey lays bare the recent surge of research in the area of stochastic choice. The near-impossibility of having fine-grained, individual-level stochastic-choice data, and the confounds arising in aggregate-level stochastic-choice data, makes the intersection of econometric theory and choice-theoretic treatments of bounded rationality fertile ground for future research.

Appendix

Proof of Observation 1. By Property (i), it is without loss of generality to consider $\mathcal{X} = \{x_1, x_2, \dots\}$ and the sequence $X_n = \{x_1, \dots, x_n\}$.

(A) By nontriviality, there is m such that there is a fraction $0 < p_m < 1$ of choice patterns consistent with $\mathcal{T}$ given X_m. Define a sequence of blocks of m elements each, with $B_1 = \{x_1, \dots, x_m\}$, $B_2 = \{x_{m+1}, \dots, x_{2m}\}$, etc. By Property (i), for each block B_j, the fraction of choice functions over B_j that are consistent with $\mathcal{T}$ given B_j is p_m. The fraction of all choice functions over X_n that are consistent with $\mathcal{T}$ given X_n is bounded above by the fraction of all choice functions over X_n satisfying the weaker condition, in view of Property (ii), that for each block $B_j \subseteq X_n$, the restriction to B_j is consistent with

$\mathcal{T}$ given B_j. Since the blocks are mutually disjoint, the latter fraction is $p_m^{(n \bmod m)}$. Hence the fraction of interest is bounded above by a number that tends to zero.

To show the fraction itself monotonically decreases in n, consider two finite sets $X \subset Y$. For $Z \in \{X, Y\}$, let $CF(Z)$ denote the set of all choice functions defined on Z, and let $W(Z)$ denote the subset of those choice functions consistent with $\mathcal{T}$. We have

$$|W(Y)| = \sum_{c \in W(X)} |\{c' \in W(Y) \text{ s.t. } c'|_X = c\}| \leq \sum_{c \in W(X)} |\{c' \in CF(Y) \text{ s.t. } c'|_X = c\}|,$$

where the equality follows from (ii) and the inequality follows from $W(Y) \subseteq CF(Y)$. But $|CF(Y)| = \sum_{c \in CF(X)} |\{c' \in CF(Y) \text{ s.t. } c'|_X = c\}|$, and $|\{c' \in CF(Y) \text{ s.t. } c'|_X = c\}|$ is constant in $c \in CF(X)$. Let K denote this constant. Hence $\frac{|W(Y)|}{|CF(Y)|} \leq \frac{K|W(X)|}{K|CF(X)|} = \frac{|W(X)|}{|CF(X)|}$. (B) We know there is m and a non-rational choice function c consistent with $\mathcal{T}$ given X_m. Define a sequence of blocks of m elements: $B_1 = \{x_1, \ldots, x_m\} = X_m$, $B_2 = \{x_{m+1}, \ldots, x_{2m}\}$, etc. Suppose $n > m$ and X_n contains $B_1, \ldots, B_J$. Construct J choice functions by:[61] $c^j(S) = c(S \cap B_{k(j)})$ if $S \cap (\cup_{k \leq j} B_k) \neq \emptyset$ where $k(j)$ is the smallest k such that $S \cap B_k \neq \emptyset$; and $c^j(S)$ is the lowest-indexed element of S if $S \cap (\cup_{k \leq j} B_k) = \emptyset$.

Each of the c^j's is consistent with $\mathcal{T}$ by Property (iii). To see this, applying Property (iii) with c over $X = B_1$ and the choice function c' defined on $Y = \{x_{m+1}, x_{m+2} \ldots\}$ which picks the lowest index in each set (being rational, this is consistent with $\mathcal{T}$) shows c^1 is consistent with $\mathcal{T}$. Suppose by induction c^{j-1} is consistent with $\mathcal{T}$. Applying Property (iii) with c over $X = B_j$ and the choice function c' defined on $Y = \{x_{(j+1)m+1}, x_{(j+1)m+2}, \ldots\}$ that picks the lowest index in each set, yields a choice function $\tilde{c}$ on $\{x_{jm+1}, x_{jm+2}, \ldots\}$ which is consistent with $\mathcal{T}$. Apply Property (iii) again with c^{j-1} restricted to $X = \cup_{k \leq j-1} B_k$ and $\tilde{c}$ on $Y = \{x_{jm+1}, x_{jm+2}, \ldots\}$ to show c^j is also consistent with $\mathcal{T}$.

Consider distinct relabelings $\Psi : X_n \to X_n$ and $\Psi' : X_n \to X_n$. The choice function $\hat{c}^j$ derived from c^j by applying Ψ is distinct from the choice function $\bar{c}^j$ derived from c^j by applying Ψ': the choice from X_n under c^j, call it x, is $\Psi(x)$ under $\hat{c}^j$ and $\Psi'(x)$ under $\bar{c}^j$. We are done if $\Psi(x) \neq \Psi'(x)$. Otherwise, let $y \neq x$ be the choice from $X_n \setminus \{x\}$ under c^j. The same menu arises when applying Ψ or Ψ', and the selected option is $\Psi(y)$ under $\hat{c}^j$, or $\Psi'(y)$ under $\bar{c}^j$. We are done showing $\hat{c}^j \neq \bar{c}^j$ if $\Psi(y) \neq \Psi'(y)$. If not, iterate the reasoning. We must reach a step where $\hat{c}^j \neq \bar{c}^j$; else we'd establish $\Psi = \Psi'$.

Certainly, c^j and c^ℓ are distinct choice functions for $j \neq \ell$; and for any common relabeling Ψ, $\hat{c}^j$ and $\hat{c}^\ell$ are also distinct. We now show $\hat{c}^j$ is distinct from $\bar{c}^\ell$, where $\bar{c}^\ell$ is

[61] Formally, c is defined over B_1. With a slight abuse of notation, we use the same letter c to denote its natural 'translation' to B_j (simply shifting all indices up by $(j-1)m$).

the choice function derived from c^ℓ when applying a different relabeling Ψ'. Without loss, suppose $j < \ell$. As c^j coincides with c^ℓ for each menu S that intersects $\cup_{k\leq j} B_k$, there is no loss in assuming $\Psi = \Psi'$ on $\cup_{k\leq j} B_k$ (otherwise the reasoning from the previous paragraph would show $\hat{c}^j$ is indeed distinct from $\bar{c}^\ell$). Hence the image (R) of the remaining options is the same under Ψ and Ψ'. The conclusion follows, as $\bar{c}^\ell$ is not rational over R but $\hat{c}^j$ is.

Hence we can construct at least $J(n!)$ choice functions consistent with $\mathcal{T}$ given X_n, where $J = (n \bmod m)$. As there are $n!$ rational choice function, the fraction of choice functions consistent with $\mathcal{T}$ given X_n that are rational goes to zero as $n \to \infty$. *Q.E.D.*

References

Abaluck, Jason and Abi Adams-Prassl (2021), What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses, *The Quarterly Journal of Economics*, 136(3), 1611–1663.

Abreu, Dilip and Ariel Rubinstein (1988), The Structure of Nash Equilibrium in Repeated Games with Finite Automata, *Econometrica*, 56(6), 1259–1281.

Aguiar, Victor (2017), Random Categorization and Bounded Rationality, *Economics Letters*, 159, 46–52.

Aguiar, Victor, Maria Jose Boccardi and Mark Dean (2016), Satisficing and Stochastic Choice, *Journal of Economic Theory*, 166, 445–482.

Aguiar, Victor, Maria Jose Boccardi, Nail Kashaev and Jeongbin Kim (2023), Random Utility and Limited Consideration, *Quantitative Economics*, 14, 71–116.

Aguiar, Victor, and Mert Kimya (2019), Adaptive Stochastic Search, *Journal of Mathematical Economics*, 81, 74–83.

Aleskerov, F., D. Bouyssou, and B. Monjardet (2007), *Utility Maximization, Choice and Preference*, Springer Verlag.

Alós-Ferrer, Carlos, Ernst Fehr and Nick Netzer (2021), Time Will Tell: Recovering Preferences When Choices are Noisy, *Journal of Political Economy*, 129(6), 1828–1877.

Ambrus, Attila and Kareen Rozen (2015), Rationalizing Choice with Multiself Models, *The Economic Journal* , 125(585), 1136–1156.

Apesteguia, Jose and Miguel Ballester (2009), A Theory of Reference-Dependent Behavior, *Economic Theory*, 40, 427–455.

Apesteguia, Jose and Miguel Ballester (2013), Choice by Sequential Procedures, *Games and Economic Behavior*, 77, 90–99.

Apesteguia, Jose and Miguel Ballester (2015), A Measure of Rationality and Welfare, *Journal of Political Economy*, 123(6), 1278–1310.

Apesteguia, Jose and Miguel Ballester (2016), Stochastic Representative Agent, *Unpublished manuscript*.

Apesteguia, Jose and Miguel Ballester (2021), Separating Predicted Randomness from Residual Behavior, *Journal of the European Economic Association*, 19(2), 1041–1076.

Apesteguia, Jose, Miguel Ballester, and Jay Lu (2017), Single-Crossing Random Utility Models, *Econometrica*, 85(2), 661–674.

Au, Pak Hung and Keiichi Kawai (2011), Sequentially Rationalizable Choice with Transitive Rationales, *Games and Economic Behavior*, 73(2), 608–614.

Banovetz, James and Ryan Oprea (2022), Complexity and Procedural Choice, *mimeo*.

Barberà, Salvador, Geoffroy de Clippel, Alejandro Neme, and Kareen Rozen (2022), Order-k Rationality, *Economic Theory*, 732, 1135–1153.

Barberà, Salvador and Prasanta K. Pattanaik (1986), Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings, *Econometrica*, 54(3), 707–715.

Barberis, Nicholas (2013), Thirty Years of Prospect Theory in Economics: A Review and Assessment, *Journal of Economic Perspectives*, 27(1), 173–196.

Barbos, Andrei (2010), Context effects: A Representation of Choices from Categories, *Journal of Economic Theory*, 145, 1224–1243.

Barseghyan, Levon, Maura Coughlin, Francesca Molinari, and Joshua C. Teitelbaum (2021), Heterogeneous Choice sets and Preferences, *Econometrica*, 89(5), 2015–2048.

Barseghyan, Levon, Francesca Molinari, and Matthew Thirkettle (2021), Discrete Choice under Risk with Limited Consideration, *American Economic Review*, 111(6), 1972–2006.

Benkert, Jean-Michel and Nick Netzer (2018), Informational Requirements of Nudging, *Journal of Political Economy*, 126(6), 2323–2355.

Bernheim, Douglas and Antonio Rangel (2009), Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics, *The Quarterly Jour-*

nal of Economics, 124(1), 51–104.

Besedeš, Tibor, Cary Deck, Sudipta Sarangi and Mikhael Shor (2015), Reducing Choice Overload without Reducing Choices, *The Review of Economics and Statistics*, 97(4), 793–802.

Block, H.D. and J. Marschak (1960), Random Orderings and Stochastic Theories of Response, in *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Eds. Olkin et al., Stanford, CA: Stanford University Press, 97–132.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012), Salience Theory of Choice Under Risk, *The Quarterly Journal of Economics*, 127(3), 1243–1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013), Salience and Consumer Choice, *Journal of Political Economy*, 121(5), 803–843.

Brady, Richard L. and John Rehbeck (2016), Menu-Dependent Stochastic Feasibility, *Econometrica*, 84(3), 1203–1223.

Bouacida, Elias and Daniel Martin (2021), Predictive Power in Behavioral Welfare Economics, *Journal of the European Economic Association*, 19(3), 1556–1591.

Buturak, Gökhan and Özgür Evren (2017), Choice Overload and Asymmetric Regret, *Theoretical Economics*, 12(3), 1029–1056.

Caplin, Andrew and Mark Dean (2011), Search, Choice and Revealed Preference, *Theoretical Economics*, 6, 19–48.

Caplin, Andrew, Mark Dean and John Leahy (2019), Rational Inattention, Optimal Consideration Sets and Stochastic Choice, *Review of Economic Studies*, 86(3), 1061–1094.

Caplin, Andrew, Mark Dean and Daniel Martin (2011), Search and Satisficing, *American Economic Review*, 101(7), 2899–2922.

Cattaneo, Matias, Xinwei Ma, Yusufcan Masatlioglu, and Elchin Suleymanov (2020), A Random Attention Model, *Journal of Political Economy*, 128(7), 2796–2836.

Cattaneo, Matias, Paul Cheung, Xinwei Ma and Yusufcan Masatlioglu (2021), Attention Overload, *mimeo.*

Cerreia-Vioglio, Simone, Per Olov Lindberg, Fabio Maccheroni, Massimo Marinacci and Aldo Rustichini (2021), A Canon of Probabilistic Rationality, *Journal of Economic Theory*, 196, 105289.

Chambers, Christopher, Tugce Cuhadaroglu and Yusufcan Masatlioglu

(2023), Behavioral Influence, *Journal of the European Economic Association*, 21(1), 135–166.

Chambers, Christopher and Takashi Hayashi (2012), Choice and Individual Welfare, *Journal of Economic Theory*, 147, 1818–1849.

Cherepanov, Vadim, Tim Feddersen and Alvaro Sandroni (2013), Rationalization, *Theoretical Economics*, 8, 775–800.

Conlisk, John (1996), Why Bounded Rationality?, *Journal of Economic Literature*, 34, 669–700.

Cuhadaroglu, Tugce (2017), Choosing on Influence, *Theoretical Economics*, 12, 477–492.

Dardanoni, Valentino, Paola Manzini, Marco Mariotti Henrik Petri and Christopher J. Tyson (2022), Mixture Choice Data: Inferring Preferences and Cognition, *Journal of Political Economy*, forthcoming.

Dardanoni, Valentino, Paola Manzini, Marco Mariotti and Christopher J. Tyson (2020), Inferring Cognitive Heterogeneity from Aggregate Choices, *Econometrica*, 88(3), 1269–1296.

de Clippel, Geoffroy (2014), Behavioral Implementation, *American Economic Review*, 104(10), 2975–3002.

de Clippel, Geoffroy and Kfir Eliaz (2012), Reason-Based Choice: A Bargaining Rationale for the Attraction and Compromise Effects, *Theoretical Economics*, 7, 125–162.

de Clippel, Geoffroy and Kareen Rozen (2021), Bounded Rationality and Limited Datasets, *Theoretical Economics*, 16(2), 359–380.

de Clippel, Geoffroy and Kareen Rozen (2022), Which Performs Best? Comparing Discrete-Choice Models, *working paper*.

Dean, Mark, Özgür Kibris and Yusufcan Masatlioglu (2017), Limited Attention and Status Quo Bias, *Journal of Economic Theory*, 169, 93–127.

Debreu, Gérard (1960), Review of R. D. Luce, Individual Choice Behavior: A Theoretical Analysis, *American Economic Review*, 50, 186–188.

Demirkan, Yusufcan and Mert Kimya (2020), Hazard Rate, Stochastic Choice and Consideration Sets, *Journal of Mathematical Economics*, 87, 142–150.

Doğan, Serhat and Kemil Yildiz (2021), Odds Supermodularity and the Luce Rule, *Games and Economic Behavior*, 126, 443–452.

Doyle, John R., David J. O'Connor, Gareth M. Reynolds, and Paul A. Bottomley (1999), The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases, *Psychology and Marketing*, 16(3), 225–243.

Dutta, Rohan and Sean Horan (2014), Inferring Rationales from Choice: Identification for Rational Shortlist Methods, *American Economic Journal: Microeconomics*, 7(4), 179–201.

Echenique, Federico and Kota Saito (2019), General Luce Model, *Economic Theory*, 68(4), 811–826.

Echenique, Federico, Kota Saito and Gerelt Tserenjigmid (2018), The Perception-Adjusted Luce Model, *Mathematical Social Sciences*, 93, May 67–76.

Ehlers, Lars, and Yves Sprumont (2008), Weakened WARP and Top-Cycle Choice Rules, *Journal of Mathematical Economics*, 44(1), 87–94.

Eliaz, Kfir and Efe Ok (2006), Indifference or Indecisiveness? Choice Theoretic Foundations of Incomplete Preferences, *Games and Economic Behavior*, 56, 61–86.

Eliaz, Kfir and Ran Spiegler (2011), Consideration Sets and Competitive Marketing, *Review of Economic Studies*, 78(1), 235–262.

Ellis, Andrew and Yusufcan Masatlioglu (2022), Choice with Endogenous Categorization, *Review of Economic Studies*, 89(1), 240–278.

Falmagne, J.C. (1978), A Representation Theorem for Finite Random Scale Systems, *Journal of Mathematical Psychology*, 18, 52–72.

Fechner, Gustav (1860), Elemente der Psychophysik, Leipzig: Breitkopf und Härtel.

Fehr, Ernst and Antonio Rangel (2011), Neuroeconomic Foundations of Economic Choice–Recent Advances, *Journal of Economic Perspectives*, 25(4), 3–30.

Filiz-Ozbay, Emel and Yusufcan Masatlioglu (2022), Progressive Random Choice, *Journal of Political Economy*, forthcoming

Fishburn, Peter (1985), *Interval Orders and Interval Graphs:A Study of Partially Ordered Sets*, John Wiley, New York.

Frederick, Shane, Leonard Lee and Ernest Baskin (2014), The Limits of Attraction, *Journal of Marketing Research*, 51(4), 487–507.

Freeman, David (2017), Preferred Personal Equilibrium and Simple Choices, *Journal of Economic Behavior and Organization*, 143, 165–172.

Freeman, David (2019), Expectations-Based Reference-Dependence and Choice Under Risk, *Economic Journal*, 129(622), 2424–2458.

Frick, Mira (2016), Monotone Threshold Representations, *Theoretical Economics*, 11, 757–772.

Fudenberg, Drew, Wayne Gao and Annie Liang (2021), How Flexible is that Functional Form? Quantifying the Restrictiveness of Theories, *Unpublished manuscript.*

Fudenberg, Drew, Jon Kleinberg, Annie Liang, and Sendhil Mullainathan (2022), Measuring the Completeness of Economic Models, *Journal of Political Economy*, 130(4), 956–990.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2018), Speed, Accuracy and the Optimal Timing of Choices, *American Economic Review*, 108(12), 3651–3684.

Gabaix, Xavier (2014), A Sparsity-Based Model of Bounded Rationality, *The Quarterly Journal of Economics*, 129(4), 1661–1710.

Gabaix, Xavier (2019), "Behavioral Inattention" in Handbook of Behavioral Economics, Vol. 2. Eds. Douglas Bernheim, Stefano DellaVigna and David Laibson. Elsevier, 261–343.

Geng, Sen and Erkut Ozbay (2021), Shortlisting procedure with a limited capacity, *Journal of Mathematical Economics,* 94.

Gerasimou, Georgios (2018), Indecisiveness, Undesirability and Overload Revealed Through Rational Choice Deferral, *The Economic Journal,* 128(614), 2450–2479.

Giarlotta, Alfio, Angelo Petralia and Stephen Watson (2022), Bounded Rationality is Asymptotically Rare, *Journal of Economic Theory*, 204, 105509.

Goldin, Jacob and Daniel Reck (2020), Revealed Preference Analysis with Framing Effects, *Journal of Political Economy*, 128(7), 2759–2795.

Gul, Faruk, Paulo Natenzon and Wolfgang Pesendorfer (2014), Random Choice as Behavioral Optimization, *Econometrica*, 82(5), 1873–1912.

Gul, Faruk and Wolfgang Pesendorfer (2001), Temptation and Self Control, *Econometrica*, 69(6), 1403–1435.

Gul, Faruk and Wolfgang Pesendorfer (2006), Random Expected Utility, *Econometrica*, 74(1), 121–146.

Guney, Begum, Michael Richter and Matan Tsur (2018), Aspiration-Based Choice, *Journal of Economic Theory*, 176, 935–956

Horan, Sean (2019), Random Consideration and Choice: A Case Study of 'Default' options, *Mathematical Social Sciences*, 102, 73–84.

Horan, Sean (2021), Stochastic Semi-Orders, *Journal of Economic Theory*, 192, 105171.

Houthakker, Hendrik (1950), Revealed Preferences and the Utility Function, *Economica*, 17, 159–74.

Hu, Gaoji, Jiangtao Li, John K-H Quah and Rui Tang (2022), A Theory of Revealed Indirect Preference, *mimeo*.

Huber, Joel, John W. Payne, Christopher Puto (1982), Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis., *Journal of Consumer Research*, 9, 90–98.

Hung Au, Pak and Keiichi Kawai (2011), Sequentially Rationalizable Choice with Transitive Rationales, *Games and Economic Behavior*, 73(2), 608–614.

Iverson, G. and J.C. Falmagne (1985), Statistical Issues in Measurement, *Mathematical Social Sciences*, 10, 131–153.

Kahneman, Daniel, Jack Knetsch and Richard Thaler (1991), Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias, *The Journal of Economic Perspectives*, 5(1), 193–206.

Kahneman, Daniel and Amos Tversky (1979), Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, 47(2), 263–292.

Kalai, Gil, Ariel Rubinstein, and Ran Spiegler (2002), Rationalizing Choice Functions by Multiple Rationales, *Econometrica*, 70(6), 2481–2488.

Kashaev, Nail and Victor Aguiar (2022), A Random Attention and Utility Model, *Journal of Economic Theory*, 204, 105487.

Kibris, Özgür, Yusufcan Masatlioglu, and Elchin Suleymanov (2021a), A Random Reference Model, *Unpublished manuscript*.

Kibris, Özgür, Yusufcan Masatlioglu, and Elchin Suleymanov (2021b), A Theory of Reference Point Formation, *Economic Theory*, forthcoming.

Kivetz, Ran, Netzer, Oded, Srinivasan, V. (2004), Alternative Models for Capturing the Compromise Effect, *Journal of Marketing Research*, 41(3), 237–257.

Kószegi, Botond and Matthew Rabin (2006), A Model of Reference-Dependent Preferences, *The Quarterly Journal of Economics*, 121(4), 1133–1165.

Kószegi, Botond and Matthew Rabin (2007a), Mistakes in Choice-Based Welfare Analysis, *American Economic Review*, 97(2), 477–481.

Kőszegi, Botond and Matthew Rabin (2007b), Reference-Dependent Risk Attitudes, *American Economic Review*, 97(4), 1047–1073.

Kőszegi, Botond and Adam Szeidl (2013), A Model of Focusing in Economic Choice, *The Quarterly Journal of Economics*, 53–104.

Kovach, Matthew and Elchin Suleymanov (2021), Reference Dependence and Random Attention, *Unpublished manuscript.*

Kovach, Matthew and Gerelt Tserenjigmid (2022a), Behavioral Foundations of Nested Stochastic Choice and Nested Logit, *Journal of Political Economy*, 130(9), 2411–2461.

Kovach, Matthew and Gerelt Tserenjigmid (2022b), The Focal Luce Model, *American Economic Journal: Microeconomics*, 14(3): 378–413.

Liang, Annie (2019), Inference of Preference Heterogeneity from Choice Data, *Journal of Economic Theory*, 179, 275–311.

Lipman, Bart and Wolfgang Pesendorfer (2013), "Temptation." In: Acemoglu, Daron, Manuel Arellano and Eddie Dekel (eds.), *Advances in Economics and Econometrics*, Cambridge University Press. 243–288.

Lleras, Juan, Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Ozbay (2017), When More is Less: Choice by Limited Consideration, *Journal of Economic Theory*, 170, 70–85.

Lombardi, Michele (2008), Uncovered Set Choice Rules, *Social Choice and Welfare*, 31(2), 271–279.

Luce, R. Duncan (1956), Semiorders and a Theory of Utility Discrimination, *Econometrica*, 24, 178–191.

Luce, R. Duncan (1959), *Individual Choice Behavior: A Theoretical Analysis.* John Wiley, New York.

Machina, Mark (1982), 'Expected Utility' Analysis without the Independence Axiom, *Econometrica*, 50(2), 277–323.

Mackowiak, Bartosz, Filip Matějka and Mirko Wiederholt (2021), Rational Inattention: A Review, *Journal of Economic Literature*, forthcoming.

Mandler, Michael, Paola Manzini, and Marco Mariotti (2012), A Million Answers to Twenty Questions: Choosing by Checklist, *Journal of Economic Theory*, 147, 71–92.

Maniquet, Francois and Hassan Nosratabadi (2022), Welfare Analysis When

Choice is Status-Quo Biased, *Journal of Mathematical Economics*, 102, 102718.

Manzini, Paola, and Marco Mariotti (2006), Revealed Preferences and Boundedly Rational Choice Procedures: an Experiment, *IZA Discussion Paper No. 2341*.

Manzini, Paola, and Marco Mariotti (2007), Sequentially Rationalizable Choice, *American Economic Review*, 97(5), 1824–1839.

Manzini, Paola, and Marco Mariotti (2012a), Categorize Then Choose: Boundedly Rational Choice and Welfare, *Journal of the European Economics Association*, 10(5), 1141–1165.

Manzini, Paola, and Marco Mariotti (2012b), Choice by Lexicographic Semiorders, *Theoretical Economics*, 7(1), 1–23.

Manzini, Paola, and Marco Mariotti (2014), Stochastic Choice and Consideration Sets, *Econometrica*, 3, 1153–1176.

Manzini, Paola, and Marco Mariotti (2018), Dual Random Utility Maximisation, *Journal of Economic Theory*, 177, 162–182.

Manzini, Paola, Marco Mariotti, and Luigi Mittone (2010), Choosing Monetary Sequences: Theory and Experimental Evidence, *Theory and Decision*, 69, 327–354.

Manzini, Paola, Marco Mariotti, and Christopher Tyson (2013), Two-Stage Threshold Representations, *Theoretical Economics*, 8, 875–882.

Markowitz, Harry (1952), The Utility of Wealth, *Journal of Political Economy*, 60, 150–158.

Masatlioglu, Yusufcan and Daisuke Nakajima (2013), Choice by Iterative Search, *Theoretical Economics*, 8, 701–728.

Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Ozbay (2012), Revealed Attention, *American Economic Review*, 102(5), 2183–2205.

Masatlioglu, Yusufcan, Daisuke Nakajima and Emre Ozdenoren (2020), Willpower and Compromise Effect, *Theoretical Economics*, 15(1), 279–317.

Masatlioglu, Yusufcan and Efe Ok (2005), Rational Choice with Status Quo Bias, *Journal of Economic Theory*, 121, 1–29.

Masatlioglu, Yusufcan and Efe Ok (2014), A Canonical Model of Choice with Initial Endowments, *The Review of Economic Studies*, 81(2), 851–883.

Masatlioglu, Yusufcan and Collin Raymond (2016), A Behavioral Analysis of Stochastic Reference Dependence, *American Economic Review*, 106(9), 2760–2782.

Matějka, Filip and Alisdair McKay (2015), Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Mode, *American Economic Review*, 105(1), 272–290.

McCausland, William, Clintin Davis-Stober, A.A.J. Marley, Sanghyuk Park and Nicholas Brown (2020), Testing the Random Utility Hypothesis Directly, *The Economic Journal*, 130, 183–207.

McFadden, Daniel (1976), *Quantal Choice Analysis: A Survey*, Annals of Economic and Social Measurement, 5, 363–390.

McFadden, Daniel (1978), "Modeling the Choice of Residential Location." In: A. Karlqvist, J. W. Weibull, L. Lundqvist, and F. Snickars (eds.), *Spatial Interaction Theory and Planning Models*, North-Holland, 72–77.

Mirkin, Boris (1974), *Group Choice*, Winston, New York.

Mosteller, F. and Nogee, P. (1951), An Experimental Measurement of Utility, *Journal of Political Economy*, 59(5), 371–404.

Natenzon, Paulo (2019), Random Choice and Learning, *Journal of Political Economy*, 127(1), 419–457.

Nielsen, Kirby and John Rehbeck (2022), When Choices are Mistakes, *American Economic Review*, 112(7), 2237–2268.

Nishimura, Hiroki (2018), The Transitive Core: Inference of Welfare from Nontransitive Preference Relations, *Theoretical Economics*, 13, 579–606.

Ok, Efe, Pietro Ortoleva and Gil Riella (2015), Incomplete Preferences Under Uncertainty: Indecisiveness in Beliefs vs. Tastes, *Econometrica*, 80(4), 1791–1808.

Ok, Efe, Pietro Ortoleva and Gil Riella (2015), Revealed (P)Reference Theory, *American Economic Review*, 105(1), 299–321.

Ok, Efe and Gerelt Tserenjigmid (2022), Indifference, Indecisiveness, Experimentation and Stochastic Choice, *Theoretical Economics*, 17(2), 651–686.

Oprea, Ryan (2020), What Makes a Rule Complex, *American Economic Review*, 110(12), 3913–3951.

Papi, Mauro (2012), Satisficing Choice Procedures, *Journal of Economic Behavior and Organization*, 84(1), 451–462.

Quiggin, John (1982), A Theory of Anticipated Utility, *Journal of Economic Behavior and Organization*, 3(4), 323–343.

Rabin, Matthew (2013), Incorporating Limited Rationality into Economics, *Journal of Economic Literature*, 51(2), 528–543.

Regenwetter, Michel, Jason Dana and Clintin Davis-Stober (2011), Transitivity of Preferences, *Psychological Review*, 118(1), 42–56.

Reutskaja, Elena, Rosemarie Nagel, Colin Camerer and Antonio Rangel (2011), Search Dynamics in Consumer Choice Under Time Pressure: An Eye-Tracking Study, *American Economic Review*, 101(2), 900–926.

Rieskamp, J., J. R. Busemeyer and B.A. Mellers (2006), Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice, *Journal of Economic Literature*, 44(3), 631–661.

Rubinstein, Ariel (1986), *Finite Automata Play the Repeated Prisoner's Dilemma*, *Journal of Economic Theory*, 39, 83–96.Rubinstein, Ariel (1988), *Similarity and Decision-Making Under Risk*, *Journal of Economic Theory*, 46, 145–153.

Rubinstein, Ariel (1998), *Modeling Bounded Rationality.* MIT Press, Cambridge, Massachusetts.

Rubinstein, Ariel (2007), Instinctive and Cognitive Reasoning: A Study of Response Times, *The Economic Journal*, 117(523), 1243–1259.

Rubinstein, Ariel (2013), Response Time and Decision Making: An Experimental Study, *Judgment and Decision Making*, 8(5), 540–551.

Rubinstein, Ariel (2015), *Lecture Notes in Microeconomic Theory: The Economic Agent*.

Rubinstein, Ariel and Yuval Salant (2006a), A Model of Choice from Lists, *Theoretical Economics*, 1, 3–17.

Rubinstein, Ariel and Yuval Salant (2006b), Two Comments on the Principle of Revealed Preference, *mimeo.*

Rubinstein, Ariel and Yuval Salant (2008), Some Thoughts on the Principle of Revealed Preference. In *Handbooks of Economic Methodologies* (Andrew Caplin and Andrew Schotter, eds.), New York: Oxford University Press, 2008, 115–124.

Rubinstein, Ariel and Yuval Salant (2012), Eliciting Welfare Preferences from Behavioral Data Sets, *Review of Economic Studies*, 79, 375–387.

Rubinstein, Ariel and Lin Zhou (1999), Choice Problems with a 'Reference' Point, *Mathematical Social Sciences*, 37, 205–209.

Sagi, Jacob (2006), Anchored Preference Relations, *Journal of Economic Theory*, 130, 283–295.

Salant, Yuval (2011), Procedural Analysis of Choice Rules with Applications to Bounded Rationality, *American Economic Review*, 101, 724–748.

Salant, Yuval and Ariel Rubinstein (2008), (A, f): Choice with Frames, *Review of Economics Studies*, 75, 1287–1296.

Salant, Yuval and Jörg Spenkuch (2022), Complexity and Choice, *Working paper.*

Samuelson, Paul (1948), Consumption Theory in Terms of Revealed Preference, *Economica*, 15, 243–253.

Samuelson, William and Richard Zeckhauser (1988), Status Quo Bias in Decision Making, *Journal of Risk and Uncertainty*, 1, 7–59.

Selten, Richard (1991), Properties of a Predictive Measure of Success, *Mathematical Social Sciences*, 21, 153–167.

Selten, Richard and S. Krischker (1983), Comparison of Two Theories for Characteristic Function Experiments. *Aspiration Levels in Bargaining and Economic Decision Making* (edited by R. Tietz), Springer, Berlin and Heidelberg, 259–264.

Sen, Amartya (1971), Choice Functions and Revealed Preference, *Review of Economic Studies*, 38, 307–317.

Sen, Amartya (1993), Maximization and the Act of Choice, *Econometrica*, 65, 745–779.

Shafir, Eldar, Itamar Simonson and Amos Tversky (1993), Reason-Based Choice, *Cognition*, 49, 11–36.

Simon, Herbert (1955), A Behavioral Model of Rational Choice, *Quarterly Journal of Economics*, 69, 99–118.

Simonson, Itamar (1989), Choice Based on Reasons: The Case for the Attraction and Compromise Effects, *Journal of Consumer Research*, 16, 158–174.

Simonson, Itamar (2014), Vices and Virtues of Misguided Replications: The Case of Asymmetric Dominance, *Journal of Marketing Research*, 51(4), 514–519.

Smith, Alec, B. Douglas Bernheim, Colin F. Camerer, and Antonio Rangel (2014), Neural Activity Reveals Preferences without Choices, *American Economic Journal: Microeconomics*, 6(2), 1–36.

Spiegler, Ran (2013), “Comments on ‘Behavioral’ Decision Theory.” In: Acemoglu, Daron, Manuel Arellano and Eddie Dekel (eds.), *Advances in Economics and Econometrics*, Cambridge University Press, 289–302.

Thaler, Richard (1980), Toward a Positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization*, 1, 39–60.

Tserenjigmid, Gerelt (2019), Choosing with the Worst in Mind: A Reference-Dependent Model, *Journal of Economic Behavior and Organization*, 157, 631–652.

Tversky, Amos (1969), "Intransitivity of Preferences," *Psychological Review*, 76(1), 31–48.

Tversky, Amos (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79, 281–299.

Tversky, Amos and Daniel Kahneman (1991), Loss Aversion in Riskless Choice: a Reference Dependent Model, *The Quarterly Journal of Economics*, 106, 1039–1061.

Tversky, Amos and Daniel Kahneman (1992), Advances in Prospect theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 5(4), 297–323.

Tyson, Christopher (2008), "Cognitive Constraints, Contraction Consistency, and the Satisficing Criterion", *Journal of Economic Theory*, 138(1), 51–70.

Tyson, Christopher (2021), Exponential Satisficing, *American Economic Journal: Microeconomics*, 13, 439–467.

Varian, Hal (1982), The Nonparametric Approach to Demand Analysis, *Econometrica*, 50(4), 945–973.

Wilson, Andrea (2014), Bounded Memory and Biases in Information Processing, *Econometrica*, 82(6), 2257–2294.

Woodford, Michael (2014), Stochastic Choice: An Optimizing Neuroeconomic Model, *American Economic Review*, 104(5), 495–500.

Woodford, Michael (2020), Modeling Imprecision in Perception, Valuation, and Choice, *Annual Review of Economics*, 12, 579–601.

Yang, Sybil, and Michael Lynn (2014), More Evidence Challenging the Robustness and Usefulness of the Attraction Effect, *Journal of Marketing Research*, 51, 508–513.

Yildiz, Kemal (2016), List-Rationalizable Choice, *Theoretical Economics*, 11, 587–599.