

Choice Overload and Asymmetric Regret*

Gökhan Buturak[†] and Özgür Evren[‡]

December 3, 2016

Abstract

We propose a model of “choice overload” which refers to a stronger tendency to select the default option in larger choice problems. Our main finding is a behavioral characterization of an asymmetric regret representation that depicts a decision maker who does not consider the possibility of experiencing regret for choosing the default option. By contrast, the value of ordinary alternatives is subject to regret. The calculus of regret for ordinary alternatives is identical to that in Sarver’s (2008) anticipated regret model, despite the fact that the primitives of the two theories are different. Our model can also be applied to choice problems with the option to defer the decision.

JEL Classification: D11, D81

Keywords: Choice overload, anticipated regret, subjective states, choice deferral

*Earlier versions of this paper circulated under the titles “Rational Choice Deferral” and “A Theory of Choice When ‘No Choice’ is an Option.” We owe special thanks to the Editor Faruk Gul and anonymous referees for their thoughtful suggestions which improved the paper significantly. We also thank Mark Dean, Kfir Eliaz, Tore Ellingsen, Guillaume Fréchet, Georgios Gerasimou, Itzhak Gilboa, Jens Josephson, Asen Kochov, Bart Lipman, Marco Mariotti, Yusufcan Masatlioglu, Efe Ok, Pietro Ortoleva, David Pearce, Debraj Ray, Ariel Rubinstein, Avital Shamir, Ennio Stacchetti, Mark Voorneveld and the seminar participants at Caltech, NES, NHH, NYU, SSE, UAB, UPF and University of Stavanger. The usual disclaimer applies.

[†]Freelance Researcher. PhD, Stockholm School of Economics. Email: buturak@gmail.com

[‡]New Economic School, 100 Novaya Street, Skolkovo, Moscow 143025, Russia. Email: oevren@nes.ru

1 Introduction

Choice overload, also known as “overchoice,” refers to a stronger tendency to stick to the default option in choice problems that contain many alternatives, where the default option is the alternative that obtains if the decision maker (DM) does not actively select any other alternative. Once we depart from the “rational agent” paradigm, one can think of several reasons for choice overload. In particular, some researchers suggest that, absent a well-defined ranking of alternatives, the DM may regret choosing any given alternative upon learning more about her tastes (or alternatives), and that the likelihood of experiencing regret may increase with the size of the choice set (Iyengar and Lepper, 2000; Anderson, 2003; Inbar, Botti and Hanks, 2011).

In practice, regret, or anticipation of it, seems to affect people’s behavior asymmetrically, with a bias towards the default option, leading to choice overload. For example, in a field study, Iyengar and Lepper (2000) find that a small tasting booth in a grocery store can generate much more sales than a larger one. In the same study, customers report greater subsequent satisfaction with their selections when the set of options is limited. In a laboratory experiment with economic incentives, Dean (2008) confirms that larger choice sets may reinforce subjects’ tendency to select the default option.¹

In this paper, we propose a model of choice overload driven by anticipated regret. Our main finding is a behavioral characterization of an asymmetric regret representation. The DM (behaves as if she) is uncertain of her tastes at the time of choice. She anticipates experiencing regret if her choice turns out to be inferior *ex post*, upon resolution of the uncertainty. Thus, an ordinary alternative is evaluated with its expected utility minus a regret term. On the other hand, when evaluating the default option, the DM does not consider the possibility of experiencing regret, leading to a bias towards the default option. Moreover, this bias is stronger in larger choice sets because the regret term for ordinary alternatives increases when additional alternatives become available.

In the remainder of this section, we take a closer look at our representation, followed by a literature review. We introduce the formal setup in Section 2, while Section 3 is devoted to our axioms and representation theorem. In Section 4, we formalize the notion of choice overload and present some comparative statics exercises. Section 5 relates our model to Sarver’s (2008) theory of anticipated regret. Finally, in Section 6, we discuss a dynamic setup where the default option acts as a means of deferring choice. The appendix contains the proofs and some further supplementary material.

¹Also, a field study by Redelmeier and Shafir (1995) shows that the presence of similar medications (instead of a single one) might lead physicians to avoid prescribing any medication if their effectiveness is doubtful.

1.1 Overview of the Representation and Axioms

We model the DM's subjective uncertainty with a probability measure μ over a set \mathcal{U} of ex post utility functions. Each element of \mathcal{U} , referred to as a *state*, is an expected utility function over a space of lotteries, Δ . We think of these lotteries as ordinary alternatives. The utility of ordinary alternatives is context dependent and includes a negative regret term. Specifically, $E_\mu(u(p) - K(\max_{q \in x} u(q) - u(p)))$ gives the *net expected utility* of selecting an alternative p from a set $x \subseteq \Delta$, where E_μ stands for the expectation operator over $u \in \mathcal{U}$ with respect to the probability measure μ . We view the term $K(\max_{q \in x} u(q) - u(p))$ as the *ex post regret* in state u that the DM anticipates experiencing upon selecting p from x . Thus, the ex post regret is proportional to the maximum utility that the DM could have attained if she were not to select p , while the parameter K measures the strength of regret. So, the net expected utility of selecting p from x is the expectation of utility minus regret, $u(p) - K(\max_{q \in x} u(q) - u(p))$.

The ex post regret upon selection of a given ordinary alternative p increases with the size of the choice set that the DM faces. That is, $x \subseteq y$ implies $K(\max_{q \in x} u(q) - u(p)) \leq K(\max_{q \in y} u(q) - u(p))$ at any state u . Consequently, the net expected utility of a given ordinary alternative decreases with the size of the choice set. By contrast, the utility of the default option is a context independent number, a . Our interpretation of this pattern is that, when selecting the default option, the DM does not take into account the possibility of experiencing regret.²

To summarize, our representation describes a choice correspondence such that, given a set x of ordinary alternatives:

(i) the DM selects an element p' of x if and only if

$$E_\mu \left(u(p') - K(\max_{q \in x} u(q) - u(p')) \right) = \max_{p \in x} E_\mu \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) \geq a;$$

(ii) she selects the default option if and only if $\max_{p \in x} E_\mu \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) \leq a$.

Let us now illustrate how this representation can generate choice overload.

Example 1. A grocery store will introduce one or more exotic, herbal jams to its product line. Their supplier provides two options: A rose jam (r) and a hibiscus jam (h). The store

²The default option is an object that does not belong to Δ . A particular implication of this assumption is that the default option does not enter the calculus of regret for ordinary alternatives. This seems reasonable because if the DM were to take into account the possibility of experiencing regret for choosing an ordinary alternative over the default option, presumably, she would also be able to take into account the opposite scenario, i.e., the possibility of experiencing regret for choosing the default option.

manager will base his decision on the projected behavior of a generic shopper, who is our DM. The DM is not familiar with either type of jam, and she is uncertain of her tastes. She has two equally likely ex post utility functions, u^1 and u^2 , defined as follows:

	u^1	u^2
r	5	1
h	1	5

The DM’s regret parameter is $K = 2$, and her default option is to not buy a herbal jam, which yields the utility level 0.

When there is only one ordinary alternative, the DM does not experience regret according to our representation. Thus, if the store offers only one type of a jam, the DM’s expected utility from that jam will be $5/2 + 1/2 = 3$. As $3 > 0$, the DM will purchase the offered product in this case.

On the other hand, if the store offers both jams, then purchasing either will induce an ex post regret of $2(5 - 1) = 8$ with probability $1/2$. Thus, in this case, the net expected utility of a jam will be $3 - 8/2 = -1 < 0$. Consequently, the simultaneous presence of two jams will cause the DM to refrain from purchasing any. \square

Behavioral characterization of our representation demands two substantive axioms. The first one is a general version of the weak axiom of revealed preferences (WARP) that is confined to instances in which the DM does not select the default option. This axiom enables our model to accommodate a context dependent attitude towards the default option, while disciplining the choices among ordinary alternatives. We call this property “exclusive WARP.”

The second axiom, called “asymmetric alpha,” ensures that the context dependence embodied in the model works in the same direction as the findings on choice overload. This axiom asserts that if an ordinary alternative p is selected from a given set, then it should also be selected from any subset that contains p . Since the default option is present in any choice problem, it follows that the DM has a stronger tendency to select the default option when she faces a larger choice set.

Apart from these two axioms, we also impose a non-triviality condition and some independence and continuity properties.

1.2 Related Literature

Our definition of regret follows Sarver’s (2008) anticipated regret model, which takes as primitive a preference relation over menus, i.e., choice sets. Aside from different primitives, the main novelty of the present approach is the asymmetry embodied in our representation.

Specifically, in our theory, only ordinary alternatives are subject to anticipated regret, and this is precisely how we accommodate the findings on choice overload. By contrast, in Sarver’s theory, anticipated regret influences the value of *all* alternatives uniformly, holding fixed the menu that the DM faces. Consequently, the corresponding choice behavior is compatible with WARP. Despite these differences, Sarver’s representation theorem plays a key role in the proof of our main result. A more detailed discussion of the connections between the two theories can be found in Section 5.

The classical regret theory, due to Bell (1982), Loomes and Sugden (1982, 1987) and Sugden (1993), envisions a DM endowed with a general regret/rejoice functional that can lead to cyclical choices among any set of alternatives. The predictions of our theory is more disciplined thanks to exclusive WARP, which rules out cycles among ordinary alternatives. Indeed, in our theory, the net expected utility of selecting p from a given choice set exceeds that of selecting q if and only if $E_\mu(u(p)) \geq E_\mu(u(q))$, which means that the DM’s choices among ordinary alternatives can also be represented with the (gross) expected utility function $p \rightarrow E_\mu(u(p))$.

Minimax regret models (e.g., Hayashi, 2008; Stoye, 2011) portray a DM who selects an alternative that minimizes the maximum expected regret, where the maximum is taken over a set of priors on exogenously given states. In these models, the value of any alternative, be it a default option or not, includes a regret term, in contrast to the asymmetry embodied in our model. Moreover, violations of WARP are solely driven by ambiguity, as opposed to risk, and disappear completely unless the DM holds multiple priors. On a related note, in our representation, “preference uncertainty” is subjective, as opposed to the Savagean approach with exogenous states adopted in minimax regret models.

Apart from his experimental findings, Dean (2008) proposes a theoretical model of choice overload that focuses on incomplete preferences. His most closely related representation depicts, roughly, a DM who selects an ordinary alternative if and only if that alternative is ranked above any other option according to an incomplete preference relation.

Gerasimou (2016) provides axiomatic foundations for a choice rule that resembles the one proposed by Dean (2008). While neither of these models admits an anticipated regret interpretation, Gerasimou’s axioms are closely related to ours.³ In fact, except our independence axioms, which have no place in Gerasimou’s ordinal setup, all of our substantive axioms do hold in the latter model. In particular, the contraction consistency axiom of Gerasimou is a direct analogue of our asymmetric alpha, the only difference being that the empty set, which represents deferral in Gerasimou’s model, takes the role of the default option in our model.⁴ Gerasimou also assumes a variant of WARP that is stronger than

³Needless to say, we formulated our axioms independently.

⁴One of the main differences between Gerasimou (2016) and Dean (2008) is the same: In the former

our exclusive WARP. Thus, our findings imply that in a cardinal setup with suitable independence properties, the incomplete preference relation envisioned by Gerasimou (and Dean) can actually be replaced by an expected utility function $p \rightarrow E_\mu(u(p))$, as far as the ranking of ordinary alternatives is concerned. However, this does not mean that our model is more general because an independence axiom does, indeed, play a role in our derivation of a complete ranking of ordinary alternatives. (We elaborate on this in Section 3.)

Dean, Kıbrıs and Masatlıoğlu (in press) relate choice overload to limited attention. A key feature of their model is that if an ordinary alternative p attracts the DM’s attention in a large set, then it also does so in any subset that contains p . However, the converse does not hold in general, leading to potential violations of exclusive WARP. Specifically, an ordinary alternative p may be selected over another ordinary alternative q in a given set, and yet, the DM may switch to q in a larger set if p happens to slip her attention.

By holding the default option fixed, in this paper we abstract from the traditional status-quo bias, which refers to an enhanced preference towards an alternative when that alternative is designated as the status-quo. To accommodate this phenomenon, a variety of reference dependent choice models were proposed, pioneered by Kahneman and Tversky’s (1979) theory of loss aversion. Typically, the models in this strand of literature satisfy WARP for a fixed status-quo option.⁵ To the best of our knowledge, the only exceptions that also accommodate choice overload are the aforementioned papers by Dean (2008) and Dean et al. (in press).

2 The Model

B stands for a finite set of riskless prizes, while Δ denotes the set of all lotteries on B . We equip Δ with the Euclidean norm $\|\cdot\|$ and the usual algebraic operations. An *ordinary alternative*, denoted as p, p', q, r etc., refers to a generic element of Δ . By a *choice set* we mean a nonempty closed subset of Δ , and denote the choice sets as x, y, z etc. \mathcal{X} stands for the collection of all choice sets equipped with the Hausdorff metric d_H .⁶

We assume that, in addition to ordinary alternatives, there exists a fixed default option (or, a status-quo alternative) that is available in every choice problem. \ominus denotes this default option, which is an object that does not belong to Δ . Accordingly, a *choice correspondence* c is defined as a nonempty valued correspondence from \mathcal{X} into $\Delta \cup \{\ominus\}$

model, the option to defer the decision replaces the default option. In addition, Gerasimou drops a secondary decision criterion considered by Dean, and thereby, formulates more compactly the idea of “incomplete preference maximization.”

⁵Recent contributions of this sort include Masatlıoğlu and Ok (2005, 2014), Sagi (2006), Apesteguía and Ballester (2009), Ortoleva (2010), Riella and Teper (2014), and Ok, Ortoleva and Riella (2015).

⁶ $d_H(x, y) := \max\{\max_{p \in x} \min_{q \in y} \|p - q\|, \max_{q \in y} \min_{p \in x} \|p - q\|\}$ for every $x, y \in \mathcal{X}$.

such that, for every $x \in \mathcal{X}$,

$$c(x) \subseteq x \cup \{\ominus\}.$$

Following the standard interpretation in choice theory, if an object belongs to $c(x)$ we understand that the DM in question may select that object in the choice problem $x \cup \{\ominus\}$.

Our representation suggests that the DM is uncertain of her tastes at the time of choice. We model the DM's tastes with expected utility functions on Δ . We utilize the same notation for an expected utility function and the associated utility vector (or index). That is, $u(p) = \sum_{b \in B} u_b p_b = u \cdot p$.

Set

$$\mathbb{R}_0^B := \left\{ u \in \mathbb{R}^B : \sum_{b \in B} u_b = 0 \right\} \quad \text{and} \quad \mathcal{U} := \{ u \in \mathbb{R}_0^B : \|u\| = 1 \}.$$

We view \mathcal{U} as a canonical state space because any non-constant von Neumann-Morgenstern preference on Δ can be represented with a function in \mathcal{U} .⁷ Finally, we write $E_\mu(f(u))$ in place of $\int_{\mathcal{U}} f(u) \mu(du)$, for a continuous function $f : \mathcal{U} \rightarrow \mathbb{R}$ and a (countably additive, Borel) probability measure μ on \mathcal{U} .

The next definition formalizes our representation notion.

Definition 1. An *asymmetric regret representation* (henceforth, *AR representation*) for a choice correspondence c consists of a probability measure μ on \mathcal{U} , and a pair of numbers K and a , with $K \geq 0$, such that the following two statements hold for every $x \in \mathcal{X}$ and $p' \in x$:

(i) $p' \in c(x)$ if and only if

$$E_\mu \left(u(p') - K(\max_{q \in x} u(q) - u(p')) \right) = \max_{p \in x} E_\mu \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) \geq a;$$

(ii) $\ominus \in c(x)$ if and only if $\max_{p \in x} E_\mu \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) \leq a$.

In what follows, (μ, K, a) stands for a generic AR representation.

As we discussed in Section 1.1, the parameter a represents the utility of \ominus , which is a context independent number, while $E_\mu(u(p) - K(\max_{q \in x} u(q) - u(p)))$ is the expected utility of selecting p from x , net of the regret term $K(\max_{q \in x} u(q) - u(p))$. It should also be noted that

$$\arg \max_{p \in x} E_\mu(u(p)) = \arg \max_{p \in x} E_\mu \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) \quad \forall x \in \mathcal{X}.$$

⁷In that respect, we follow Dekel et al. (2001) and Sarver (2008), among many others.

Thus, if it is nonempty, the set of ordinary alternatives that the DM may select from a given choice set x coincides with the maximizers of the gross expected utility function $p \rightarrow E_\mu(u(p))$ over x .

3 Representation Theorem

We now turn to behavioral characterization of AR representations. Our first axiom is a general version of WARP.

A1: Exclusive WARP. If $x \subseteq y$ and $c(x), c(y) \subseteq \Delta$, then $c(y) \cap x \neq \emptyset$ implies $c(y) \cap x = c(x)$.

Observe that the scope of this axiom is limited to choice sets x and y such that \ominus does not belong to $c(x)$ or $c(y)$. Thus, exclusive WARP does not impose any restriction on DM's decisions to select \ominus , leading to a (possibly) context dependent attitude towards the default option. By contrast, Arrow's (1959) classical formulation of WARP applies to any pair of choice sets x, y with $x \subseteq y$. This is the only difference between exclusive WARP and Arrow's formulation.

We complement exclusive WARP with an asymmetric version of Sen's (1971) property alpha.⁸

A2: Asymmetric Alpha. If $x \subseteq y$, then $p \in c(y) \cap x$ implies $p \in c(x)$.

Unlike in exclusive WARP, the sets $c(y)$ and $c(x)$ in the statement of asymmetric alpha may also contain \ominus . In particular, it follows that if an ordinary alternative $p \in x$ is selected over the default option from a set y that contains x , then p should also be selected from the small set x . However, asymmetrically, we do not demand the same from the default option. Thus, it remains possible to have $c(y) = \{\ominus\}$ and $\ominus \notin c(x)$ for some $x, y \in \mathcal{X}$ with $x \subseteq y$. Indeed, this is precisely the pattern observed in the findings on choice overload. By contrast, the classical version of property alpha does not make such a distinction between the available alternatives.

Our independence axiom consists of three parts, each focusing on a different scenario about the contents of $c(x)$ and $c(y)$, given a pair of sets x and y that will be mixed with each other. By a mixture of x and y , we mean the set $\alpha x + (1-\alpha)y := \{\alpha p + (1-\alpha)q : p \in x, q \in y\}$ for some $\alpha \in [0, 1]$.

A3: Independence. (i) If $c(x) \cap \Delta \neq \emptyset$ and $c(y) \cap \Delta \neq \emptyset$, then for every $p \in x, q \in y$

⁸“Property alpha” is the term introduced by Sen (1971) to refer to Chernoff's (1954) Postulate 4. As shown by Sen, this property and a dual property beta are jointly equivalent to WARP.

and $\alpha \in (0, 1)$,

$$p \in c(x) \text{ and } q \in c(y) \quad \Leftrightarrow \quad \alpha p + (1 - \alpha)q \in c(\alpha x + (1 - \alpha)y).$$

(ii) If $c(x) = \{\ominus\}$ and $\ominus \in c(y)$, then $c(\alpha x + (1 - \alpha)y) = \{\ominus\}$ for every $\alpha \in (0, 1)$.

(iii) For every $p, q, r \in \Delta$ and $\alpha \in [0, 1]$,

$$p \in c(\{p, q\}) \text{ and } c(\{\alpha q + (1 - \alpha)r\}) \cap \Delta \neq \emptyset \quad \Rightarrow \quad c(\{\alpha p + (1 - \alpha)r\}) \cap \Delta \neq \emptyset.$$

Part (i) of this axiom is a fairly standard independence property that is satisfied in the classical model of choice under risk. One notable implication of this part of the axiom is that if $c(x)$ and $c(y)$ both contain ordinary alternatives, then $c(\alpha x + (1 - \alpha)y)$ must also contain some ordinary alternatives. Part (ii) is a dual property which says that if the DM does not select an ordinary alternative from x , and if she also selects \ominus given y , then she must select \ominus uniquely when she faces $\alpha x + (1 - \alpha)y$ for any $\alpha \in (0, 1)$. As for part (iii), suppose $c(\{p\}) = \{p\}$ while $c(\{r\}) = \{\ominus\}$. Then, given any $\alpha \in (0, 1)$, we may well have $c(\{\alpha p + (1 - \alpha)r\}) = \{\ominus\}$. However, following the logic of the classical independence axiom, this possibility can be ruled out if p is revealed preferred to some q such that $\alpha q + (1 - \alpha)r$ is revealed preferred to \ominus . This is the content of part (iii).

Our next axiom is a standard topological continuity property.

A4: Continuity. Let (x_n) be a sequence in \mathcal{X} that converges to x .

(i) If $p_n \in c(x_n) \cap \Delta$ for every n and $p_n \rightarrow p$, then $p \in c(x)$.

(ii) If $\ominus \in c(x_n)$ for every n , then $\ominus \in c(x)$.

We also require a Lipschitz continuity property, which takes the role of the corresponding axiom of Sarver (2008). This property can be interpreted along the lines of Dekel, Lipman, Rustichini and Sarver (2007).

A5: L-Continuity. There exist $y^*, y_* \in \mathcal{X}$ and a number $m > 0$ such that for every $x, y \in \mathcal{X}$ and $\alpha \in (0, 1)$ with $d_H(x, y) \leq \alpha/m$,

$$\ominus \in c(\alpha y^* + (1 - \alpha)y) \quad \Rightarrow \quad \ominus \in c(\alpha y_* + (1 - \alpha)x).$$

Our final axiom is a non-triviality condition.

A6: Non-Triviality. There exist $p^*, p_* \in \Delta$ such that $c(\{p^*\}) = \{p^*\}$ and $c(\{p_*\}) = \{\ominus\}$.

This axiom rules out the cases in which the default option is the best or worst alternative.

In terms of an AR representation (μ, K, a) , (A6) means that

$$E_\mu(u(p^*)) > a > E_\mu(u(p_*)) \quad \text{for some } p^*, p_* \in \Delta. \quad (1)$$

Throughout the paper, we say that an AR representation is non-trivial if it satisfies (1).

Our main representation theorem reads as follows.

Theorem 1. *A choice correspondence c on \mathcal{X} satisfies the axioms (A1)-(A6) if and only if it admits a non-trivial AR representation.*

Towards the proof of Theorem 1, in Appendix B we first establish an auxiliary representation (Theorem 0) that dispenses with asymmetric alpha as well as part (iii) of the independence axiom. Essentially, this auxiliary representation delivers a von Neumann-Morgenstern preference \succsim on Δ and a continuous, affine⁹ function $\Psi : \mathcal{X} \rightarrow \mathbb{R}$ such that, for every $x \in \mathcal{X}$,

$$\begin{aligned} c(x) \cap \Delta \neq \emptyset &\Rightarrow c(x) \cap \Delta = \{p \in x : p \succsim q \forall q \in x\}, \\ c(x) \cap \Delta \neq \emptyset &\Leftrightarrow \Psi(x) \geq 0, \quad \text{and} \quad \ominus \in c(x) \Leftrightarrow \Psi(x) \leq 0. \end{aligned} \quad (2)$$

The first part of this expression means that as far as the ordinary alternatives are concerned, the DM is a standard preference maximizer. In particular, the relation \succsim represents the DM's ranking of ordinary alternatives. On the other hand, the ranking of the default option is context dependent, as depicted in the second part of (2). Specifically, if $\Psi(x) \geq 0$, the best ordinary alternatives in x are selected over \ominus , whereas the opposite behavior obtains when $\Psi(x) \leq 0$.

We elicit the DM's ranking of ordinary alternatives from local choice data, focusing on a small neighborhood of an ordinary alternative p^* with $c(\{p^*\}) = \{p^*\}$. The role of exclusive WARP is to ensure that c can be "rationalized" by a preference relation \succsim in this neighborhood. From part (i) of the independence axiom, it follows that \succsim is a von Neumann-Morgenstern preference. The very same axiom also implies that \succsim can be extended to the entire space Δ (uniquely), in such a way that the first implication in (2) holds true. In turn, part (ii) of the independence axiom has a significant role in the derivation of an affine function Ψ that satisfies the second line in expression (2).

The remainder of the proof of Theorem 1 builds upon asymmetric alpha and part (iii) of the independence axiom. Claim 6 in Appendix C shows that part (iii) of the independence

⁹A function $\Psi : \mathcal{X} \rightarrow \mathbb{R}$ is *affine* if $\Psi(\lambda x + (1 - \lambda)y) = \lambda\Psi(x) + (1 - \lambda)\Psi(y)$ for every $x, y \in \mathcal{X}$ and $\lambda \in (0, 1)$. An affine function on Δ is defined analogously.

axiom implies, for any $p, q \in \Delta$,

$$p \succsim q \quad \Leftrightarrow \quad \Psi(\{p\}) \geq \Psi(\{q\}). \quad (3)$$

So, the function $p \rightarrow \Psi(\{p\})$ represents the DM's ranking of ordinary alternatives. Finally, asymmetric alpha helps us show that Ψ can be written as a positive affine transformation of the maximum values that a net expected utility function attains over choice sets.¹⁰ That is, there exist a probability measure μ on \mathcal{U} and three numbers K, α, γ with $K \geq 0$ and $\alpha > 0$, such that for every $x \in \mathcal{X}$,

$$\Psi(x) = \alpha \max_{p \in x} E_{\mu} \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) + \gamma. \quad (4)$$

From (2), (3) and (4), it easily follows that the parameters μ, K and $a := -\gamma/\alpha$ constitute an AR representation for the choice correspondence c .

4 Comparative Statics

As we mentioned earlier, choice overload refers to a stronger tendency to select the default option in larger choice problems. The following definition formalizes this phenomenon.

Definition 2. A choice correspondence c *exhibits choice overload* at $x \in \mathcal{X}$ if $c(x) = \{\ominus\}$ and there exists a $p \in x$ such that $p \in c(\{p\})$. We say that c *exhibits choice overload* if there exists such an x .

In our model, $p \in c(\{p\})$ means that p is revealed preferred to \ominus . According to the standard choice theory, if $\ominus \in c(x)$ for a choice set x , and if x contains an alternative p that is revealed preferred to \ominus , then p should also belong to $c(x)$. Thus, the pattern in Definition 2 can be viewed as a boundedly rational mode of behavior. It is also clear that this pattern corresponds to a particular form of choice overload in which the presence of many ordinary alternatives, as opposed to a single one, triggers the choice of the default option.

In fact, our theory attributes such instances to anticipated regret. To see this point, let c be a choice correspondence that admits an AR representation (μ, K, a) , and set

$$\phi(p) := E_{\mu}(u(p)) \quad \forall p \in \Delta. \quad (5)$$

Observe that if there is only one ordinary alternative, selecting that alternative does not

¹⁰More specifically, this step of the proof follows from Sarver's (2008) representation theorem, while asymmetric alpha establishes the main link between the two theories. (More on this in Section 5 below.)

inflict regret. That is, with $y = \{p\}$, we have $\max_{q \in y} u(q) - u(p) = 0$ for every $u \in \mathcal{U}$. Hence, the net expected utility of selecting p from $\{p\}$ is equal to $\phi(p)$, which implies

$$p \in c(\{p\}) \iff \phi(p) \geq a. \quad (6)$$

So, given a choice set x that contains an alternative p with $p \in c(\{p\})$, putting aside the expected regret terms, the alternative that maximizes ϕ over x would surely yield an expected utility that exceeds a . It follows that we can have $c(x) = \{\ominus\}$ only because of the negative impact of anticipated regret.

Henceforth, the term ‘‘choice overload’’ refers to Definition 2.

Proposition 1. *Let c be a choice correspondence that admits a non-trivial AR representation (μ, K, a) . Then, c exhibits choice overload if and only if $K > 0$ and the support of μ contains at least two distinct points.*

Intuitively, Proposition 1 means that the DM exhibits choice overload if and only if she faces a subjective uncertainty and $K > 0$ so that this uncertainty leads to instances of regret. For further insight, suppose $\mu = \delta_{\hat{u}}$ for some $\hat{u} \in \mathcal{U}$.¹¹ Then, the expected regret term $KE_{\mu}(\max_{q \in x} u(q) - u(p))$ is equal to ex post regret at the state \hat{u} , given by $K(\max_{q \in x} \hat{u}(q) - \hat{u}(p))$. Moreover, by definition of ϕ , $\mu = \delta_{\hat{u}}$ implies $\phi(p) = \hat{u}(p)$ for every $p \in \Delta$. Finally, recall that if it is nonempty, the set $c(x) \cap \Delta$ equals $\arg \max_{p \in x} \phi(p)$. It follows that if the support of μ contains only one point, then the expected regret term equals 0 for every $x \in \mathcal{X}$ and any ordinary alternative that the DM may choose from x . In this case, c admits a standard utility representation, which does not allow choice overload. Specifically, we have $c(x) = \arg \max_{t \in x \cup \{\ominus\}} g(t)$ for every $x \in \mathcal{X}$, where

$$g(t) := \begin{cases} \phi(t) & \text{for } t \in \Delta, \\ a & \text{for } t = \ominus. \end{cases} \quad (7)$$

Similarly, an AR representation with $K = 0$ reduces to the standard model above.

Conversely, if $K > 0$ and the support of μ contains two distinct points, then for any $p \in \Delta$ with $\phi(p) = a$, there exists a choice set x containing p such that c exhibits choice overload at x . In fact, any neighborhood of p contains such an x . We refer to Lemma 1 in Appendix D for the details of this construction, which completes the proof of Proposition 1.

Motivated by Proposition 1, we shall say that a non-trivial AR representation (μ, K, a) is *strictly non-trivial* if $K > 0$ and the support of μ contains at least two distinct points.

The following definition proposes a comparative measure of choice overload.

¹¹Throughout the paper, δ_u denotes the degenerate probability measure supported at u .

Definition 3. Let c and c' be a pair of choice correspondences. c' is *more choice overload prone than* c if for any $x \in \mathcal{X}$, whenever c exhibits choice overload at x , so does c' .

Clearly, if c does not exhibit choice overload, then any other choice correspondence is more choice overload prone than c . Hence, our focus will be on choice correspondences that exhibit choice overload, i.e., on strictly non-trivial AR representations.

Proposition 2. Let (μ, K, a) and (μ', K', a') be strictly non-trivial AR representations for c and c' , respectively. Assume further that $\mu = \mu'$. Then, c' is more choice overload prone than c if and only if $K' \geq K$ and $a' = a$.

This result shows that holding fixed the belief μ , the DM's tendency to exhibit choice overload can be strengthened by increasing the regret parameter K . Moreover, the utility of the default option, a , should be kept constant to make sure that the DM's behavior does not change in choice problems that contain only one ordinary alternative.

Roughly, Definitions 2 and 3 suggest that if c' is more choice overload prone than c , we must have

$$p \in c(\{p\}) \quad \Rightarrow \quad p \in c'(\{p\}). \quad (8)$$

Indeed, c can exhibit choice overload at x only if $p \in c(\{p\})$ for some $p \in x$, and similarly for c' . Expression (8) is equivalent to saying that $\phi(p) \geq a$ implies $\phi'(p) \geq a'$, because expression (6) also applies to c' , ϕ' and a' . Moreover, $\phi = \phi'$ assuming $\mu = \mu'$. So, it follows that if c' is more choice overload prone than c , we must have $a \geq a'$.

On the other hand, the first part of the definition of choice overload, i.e., the condition $c(x) = \{\ominus\}$, pushes both a and K in the opposite direction. Following the logic of expression (8), c' is more choice overload prone than c only if $c(x) = \{\ominus\}$ implies $c'(x) = \{\ominus\}$. That is, c' must exhibit a stronger preference for \ominus (relative to ordinary alternatives) than c does. In turn, this effect can be decomposed into two parts. First, the representation of c' must attach a larger utility to \ominus , so that $a' \geq a$. Second, the net expected utility of ordinary alternatives should be smaller according to c' due to a larger expected regret functional, which means $K' \geq K$.

At first sight, one might think that, in the statement of Proposition 2, the assumption $\mu = \mu'$ can be replaced with the weaker condition $\phi = \phi'$. However, this contention is not correct, because the behavior of the expected regret term $KE_\mu(\max_{q \in x} u(q) - u(p))$ tightly depends on the probability measure μ . Put differently, even if $K' \geq K$, at least for some choice sets, the representation (μ', K', a) may induce smaller expected regret terms than (μ, K, a) does unless μ and μ' satisfy certain conditions beyond the assumption $\phi = \phi'$. We provide an overview of these conditions in Appendix A.

Our last result highlights the role of the parameter a , in line with the related remarks on Proposition 2. Holding fixed the net expected utility of ordinary alternatives, an increase

in a corresponds to a stronger tendency to select the default option, irrespective of the choice set that the DM faces:

Proposition 3. *Let (μ, K, a) and (μ', K', a') be non-trivial AR representations for c and c' , respectively. Assume further that $\mu = \mu'$ and $K = K'$. Then, $a \leq a'$ if and only if for every $x \in \mathcal{X}$,*

$$c(x) = \{\ominus\} \quad \Rightarrow \quad c'(x) = \{\ominus\}.$$

5 Relation to Sarver's Menu-Choice Model

The primitive of Sarver's (2008) theory is a preference relation \succsim^* on the collection of choice sets, \mathcal{X} . His main result delivers a probability measure μ on \mathcal{U} and a number $K \geq 0$ such that the maximum net expected utility function, $x \rightarrow \max_{p \in x} E_\mu(u(p) - K(\max_{q \in x} u(q) - u(p)))$, represents \succsim^* . In the present context, $x \succsim^* y$ should be interpreted as saying that the best ordinary alternative in x leads to a higher net expected utility than the best ordinary alternative in y .

In the proof of Theorem 1, we define a binary relation \succsim^* on \mathcal{X} as $x \succsim^* y$ if and only if $\Psi(x) \geq \Psi(y)$, where Ψ is the function in expression (2). The main behavioral property demanded by Sarver's representation theorem is the *dominance axiom* which asserts that

$$\{p\} \succsim^* \{q\} \text{ and } p \in x \quad \Rightarrow \quad x \succsim^* x \cup \{q\}.$$

Intuitively, this axiom means that the presence of an ordinary alternative q can only make the DM worse off unless q is strictly better than any other ordinary alternative that is available. Our asymmetric alpha has a similar flavor. Letting $y := x \cup \{q\}$, this axiom can be interpreted as saying that if, given the choice set y , the DM prefers to select an ordinary alternative p over \ominus despite the negative effect of q , then she should also select p upon removal of q .

Building upon asymmetric alpha, Claim 7 in Appendix C shows that the relation \succsim^* induced by the function Ψ satisfies the dominance axiom, in addition to all other axioms of Sarver. Then, we apply Sarver's representation theorem to deduce Theorem 1 from our auxiliary representation.

To relate the comparative statics of the two models, let \succsim^* and \succsim'^* stand for a pair of preference relations on \mathcal{X} . As a comparative measure of "regret aversion," Sarver proposes the following definition: For every $p \in \Delta$ and $x \in \mathcal{X}$,

$$\{p\} \succ^* x \quad \Rightarrow \quad \{p\} \succ'^* x. \tag{9}$$

This means that, compared to \succsim'^* , the relation \succsim^* is less averse towards choice sets with multiple elements, which pose the danger of regret. In a sense, property (9) is stronger than our comparative measure of choice overload because the former applies to any $(p, x) \in \Delta \times \mathcal{X}$, whereas our definition focuses on instances with $c(x) = \{\ominus\}$ and $p \in c(\{p\})$ for some $p \in x$. It is this difference that allows us to conclude that any choice correspondence is more choice overload prone than another one that does not exhibit choice overload. By contrast, any pair of preference relations on \mathcal{X} must have a non-trivial relationship whenever they are ranked according to Sarver’s regret aversion.¹² Remarkably, however, for choice correspondences that exhibit choice overload, the parametric characterization of our comparative measure is equivalent to that of Sarver, putting aside the additional parameter a in our model.

Similarly, for choice correspondences that exhibit choice overload, the uniqueness properties of our representation are identical to those of Sarver’s, aside from straightforward adjustments necessitated by the presence of the parameter a . It should be noted, however, that in both models the parameters μ and K can be identified only jointly, but not separately. In other words, without altering the associated choice correspondence, one can change K by manipulating μ , and vice versa. While Appendix A contains some related remarks on the comparative statics of our model, a detailed discussion of the uniqueness issue can be found in an earlier version of the present paper, Buturak and Evren (2015).

6 Choice Deferral: A “Rational” Form of Choice Overload

In many choice problems, the default option acts as a flexible alternative that allows the DM to defer the decision temporarily. For example, a person who has a certain budget to buy a new TV set may decide to stick to her old TV for a while in order to reflect on her tastes or the available alternatives. Experimental studies on such choice problems document the same pattern as in the notion of choice overload: Larger choice sets reinforce subjects’ tendency to select the default option (Tversky and Shafir, 1992; Dhar, 1997; White and Hoffrage, 2009).

Under suitable assumptions, our theory can also be applied to such dynamic problems. The dynamic setting, however, requires a “preference for flexibility” interpretation along the lines of the menu-choice literature pioneered by Kreps (1979) and Dekel et al. (2001). Specifically, suppose that the DM faces a choice set $x \subseteq \Delta$ at a given point of time, stage

¹²Indeed, if (9) holds, then $\{p\} \succ^* \{q\}$ implies $\{p\} \succ'^* \{q\}$. So, \succsim^* and \succsim'^* must agree on the ranking of singletons.

1. She has to select an alternative from x , but she can also postpone this decision to a later point, stage 2, by selecting the default option \ominus at stage 1. On the other hand, if she selects an ordinary alternative at stage 1, she has to consume it perpetually. Moreover, at stage 1, she faces subjective uncertainty about her tastes which will be resolved by the beginning of stage 2.

If we model the DM's subjective uncertainty with a probability measure μ on \mathcal{U} , the expected lifetime utility of selecting \ominus at stage 1 given a choice set x can be formulated as $a + KE_\mu(\max_{q \in x} u(q))$. Here, a represents the utility of consuming \ominus at stage 1, whereas $K \geq 0$ measures the importance of future consumption relative to instantaneous consumption, which may depend on the DM's time preferences/discount factor as well as the relative duration of the two stages. The term $\max_{q \in x} u(q)$ is the ex post utility level that the DM will attain at state u upon deferring choice at stage 1. By the same logic, if we rule out potential effects of anticipated regret, the expected lifetime utility of selecting an ordinary alternative p can be expressed as $(1 + K)E_\mu(u(p))$.

These specifications lead to the following choice rule: For every $x \in \mathcal{X}$ and $p' \in x$,

(i) $p' \in c(x)$ if and only if

$$(1 + K)E_\mu(u(p')) = (1 + K) \max_{p \in x} E_\mu(u(p)) \geq a + KE_\mu\left(\max_{q \in x} u(q)\right);$$

(ii) $\ominus \in c(x)$ if and only if $(1 + K) \max_{p \in x} E_\mu(u(p)) \leq a + KE_\mu\left(\max_{q \in x} u(q)\right)$.

It can easily be verified that the statements (i) and (ii) are equivalent to the corresponding statements in the definition of an AR representation. Thus, axioms (A1)-(A6) provide a behavioral foundation also for the dynamic representation above. However, this does not mean that the static and dynamic versions of our theory are conceptually equivalent. In particular, if the default option acts as a means of deferring choice, there seems to be no reason to interpret the choice overload pattern in Definition 2 as a violation of WARP. After all, the DM selects the default option not to consume it perpetually, but to keep her options open temporarily, just as in the aforementioned menu-choice models on preference for flexibility. In this sense, the pattern in Definition 2 can be viewed as a rational form of choice overload in dynamic problems with the option to defer choice.

Appendix

Throughout the appendix, we often write $\max_x u$ and $\arg \max_x u$ in place of $\max_{p \in x} u(p)$ and $\arg \max_{p \in x} u(p)$, respectively.

A. On the Role of Beliefs in Comparative Statics

How can we obtain a more choice overload prone AR representation by modifying the DM's belief? An earlier version of this paper, Buturak and Evren (2015), provides formal results that answer this question.¹³ In this appendix, we summarize the content of these results, which are quite involved. At the outset, it should be noted that any change in the DM's belief also necessitates changes in other parameters in order to obtain a new representation that is more choice overload prone.

Consider a choice correspondence c that admits a non-trivial AR representation (μ, K, a) . The gross expected utility function ϕ , defined in equation (5), can equivalently be thought of as a vector in \mathbb{R}_0^B . In fact, as a vector, ϕ is equal to the expectation of the identity function $u \rightarrow u$ with respect to μ . That is, $\phi = (E_\mu(u_b))_{b \in B}$. Hence, we shall refer to ϕ as the *mean of μ* .

Set $u_\phi := \phi / \|\phi\|$. Given the non-triviality assumption, ϕ is non-zero, and u_ϕ is a well-defined element of \mathcal{U} . Moreover, since u_ϕ and ϕ are collinear, $K(\max_{q \in x} u_\phi(q) - u_\phi(p)) = 0$ for any $x \in \mathcal{X}$ and $p \in c(x)$. In this sense, u_ϕ is a *regret-free state*.

To clarify the main idea, suppose, for the moment, that we expand our state space so that every point in \mathbb{R}_0^B qualifies as a state. As is well-known, for any $x \in \mathcal{X}$, the support function $u \rightarrow \max_x u$ is convex in $u \in \mathbb{R}_0^B$ (see, e.g., Schneider, 1993, Section 1.7). Thus, replacing a probability measure on \mathbb{R}_0^B with a mean-preserving spread of that measure induces larger expected regret terms. Intuitively, this corresponds to an increase in subjective uncertainty, which decreases the net expected utility of ordinary alternatives, just as in the case of a risk-averse individual who does not like mean-preserving spreads of monetary lotteries.

Let us now consider an example where the original belief μ is supported over \mathcal{U} . Pick any \bar{u} in the support of μ that is distinct from the regret-free state u_ϕ and suppose $\mu(\{\bar{u}\}) > 0$. Let $\{v^1, \dots, v^n\} \subseteq \mathbb{R}_0^B$ be a finite set that contains \bar{u} in its convex hull. That is, let $\bar{u} = \sum_{i=1}^n \alpha^i v^i$ for some $\{\alpha^1, \dots, \alpha^n\} \subseteq [0, 1]$ with $\sum_{i=1}^n \alpha^i = 1$. Then, we can construct a new probability measure μ' on \mathbb{R}_0^B by transferring the mass $\mu(\{\bar{u}\})$ to the points v^1, \dots, v^n so that $\mu'(\{v^i\}) = \alpha^i \mu(\{\bar{u}\})$ for every i . By construction, ϕ' and ϕ , i.e., the means of μ' and μ , are equal to each other. In fact, μ' is a mean-preserving spread of μ because the former probability measure is obtained from the latter by replacing \bar{u} with multiple points, v^1, \dots, v^n . Since the support functions are convex, from Jensen's inequality it then follows that $E_{\mu'}(\max_x u) \geq E_\mu(\max_x u)$ for every $x \in \mathcal{X}$.¹⁴ Moreover, with $\phi' = \phi$, for any $x \in \mathcal{X}$ and

¹³In Buturak and Evren (2015), the comparative measure of choice overload is defined in a slightly different way, but that definition is equivalent to the present one. We thank a referee for suggesting the present version of Definition 3.

¹⁴Indeed, $E_{\mu'}(\max_x u) - E_\mu(\max_x u) = \mu(\{\bar{u}\})(E_\eta(\max_x u) - \max_x \bar{u})$, where η is the probability measure

$p \in x$ this implies

$$\begin{aligned} E_{\mu'} \left(u(p) - K(\max_x u - u(p)) \right) &= \phi'(p) - K \left(E_{\mu'}(\max_x u) - \phi'(p) \right) \\ &\leq \phi(p) - K \left(E_{\mu}(\max_x u) - \phi(p) \right) = E_{\mu} \left(u(p) - K(\max_x u - u(p)) \right). \end{aligned}$$

So, replacing μ with the mean-preserving spread μ' decreases the net expected utility of ordinary alternatives. Consequently, the choice correspondence represented by (μ', K, a) is more choice overload prone than that represented by (μ, K, a) .

Adapting this method to the state space \mathcal{U} requires further work that also includes a shift in the regret parameter K . Since the unit ball in \mathbb{R}_0^B is a strictly convex set, the given point \bar{u} cannot be expressed as a convex combination of other states in \mathcal{U} . Yet, we can find some states $v^1, \dots, v^n \in \mathcal{U}$ and weights $\alpha^1, \dots, \alpha^n \in [0, 1]$ such that the convex combination $v := \sum_{i=1}^n \alpha^i v^i$ is collinear with \bar{u} . Then, the difference between v and \bar{u} can be compensated with a larger regret parameter K' . Specifically, we can select a K' such that $K' \|v\| \geq K \|\bar{u}\|$. As a further difficulty, if we only replace \bar{u} as described above, then ϕ' , i.e., the mean of the new probability measure, will not be collinear with ϕ . However, depending on the structure of the support of μ , we can restore the equality of ϕ' and ϕ by repeating the replacement process for other points in the support of μ , in addition to the given point \bar{u} . Following these steps, we can obtain a new representation (μ', K', a) that is more choice overload prone than the original representation.

Finally, if $\mu(\{u_\phi\}) > 0$, it is possible to obtain a more choice overload prone representation also by transferring some mass from the regret-free state u_ϕ to other states in \mathcal{U} . This process is less demanding because, unlike all other states, we do not have to worry about the possibility of decreasing the ex post regret at the state u_ϕ . Hence, this method does not necessitate to increase the parameter K in order to obtain a more choice overload prone representation. Moreover, unlike the previous method, a mass transfer from u_ϕ to suitably selected states in \mathcal{U} would induce a ϕ' that is collinear with the original mean ϕ , even if we do not reduce the mass of any other point in the support of μ . On the other hand, with this method we cannot retain the condition $\phi' = \phi$. Thus, the parameter a should be replaced with $a' := \alpha a$, where $\alpha \in (0, 1)$ is the number with $\phi' = \alpha \phi$, so that the new representation displays the same behavior as the original representation whenever there is only one ordinary alternative.¹⁵

For further details and examples on the role of beliefs in comparative statics, we refer the reader to Buturak and Evren (2015).

on \mathbb{R}_0^B that attaches the mass α^i to the point v^i for $i = 1, \dots, n$. Furthermore, the mean of η equals \bar{u} by construction, and hence, Jensen's inequality implies $E_\eta(\max_x u) \geq \max_x \bar{u}$.

¹⁵If one wishes, it is possible to restore the condition $\phi' = \phi$ by suitably adjusting K .

B. An Auxiliary Representation

In this appendix, we prove the following auxiliary representation that will act as our main tool in the proof of Theorem 1.

Theorem 0. *A choice correspondence c on \mathcal{X} satisfies the axioms (A1), (A3i), (A3ii), (A4) and (A6) if and only if there exist continuous and affine functions $\varphi : \Delta \rightarrow \mathbb{R}$ and $W : \mathcal{X} \rightarrow \mathbb{R}$ such that:*

(i) *For every $x \in \mathcal{X}$ and $p' \in x$,*

$$\begin{aligned} p' \in c(x) &\Leftrightarrow \varphi(p') = \max_{p \in x} \varphi(p) \geq W(x), \\ \ominus \in c(x) &\Leftrightarrow \max_{p \in x} \varphi(p) \leq W(x). \end{aligned} \tag{10}$$

(ii) *$\varphi(p^*) > W(\{p^*\})$ and $\varphi(p_*) < W(\{p_*\})$ for some $p^*, p_* \in \Delta$.*

In this representation, the function φ is a standard expected utility function that represents the DM's ranking of ordinary alternatives. The term $W(x)$ is a threshold level that varies with the choice set x , and allows the representation to accommodate a context dependent attitude towards the default option. The key feature of this representation is that, given a choice set x , the DM opts for an ordinary alternative as opposed to the default option if and only if $\max_{p \in x} \varphi(p)$ exceeds $W(x)$.

To relate Theorem 0 to Theorem 1, set $\Psi(x) := \max_x \varphi - W(x)$ for every $x \in \mathcal{X}$, and denote by \succsim the preference relation on Δ represented by φ . Then, expression (10) implies (2), while the latter expression plays a key role in the proof of Theorem 1, as we noted in Section 3.

On the other hand, Theorem 0 might also prove useful in alternative models that depict different forms of context dependence because it dispenses with asymmetric alpha and part (iii) of the independence axiom.

Proof of Theorem 0. We omit the “if” part of the proof, which is a routine exercise. For the “only if” part, let c be a choice correspondence on \mathcal{X} that satisfies the axioms (A1), (A3i), (A3ii), (A4) and (A6).

Fix a pair of ordinary alternatives p^*, p_* such that $c(\{p^*\}) = \{p^*\}$ and $c(\{p_*\}) = \{\ominus\}$, as in the non-triviality axiom. Put $\mathcal{X}^* := \{x \in \mathcal{X} : c(x) \subseteq \Delta\}$ and $\mathcal{X}_* := \{x \in \mathcal{X} : c(x) = \{\ominus\}\}$.

Claim 1. *\mathcal{X}^* and \mathcal{X}_* are relatively open subsets of \mathcal{X} .*

Proof. Part (ii) of the continuity axiom implies that $\{x \in \mathcal{X} : \ominus \in c(x)\}$ is a closed subset of \mathcal{X} . Hence, $\{x \in \mathcal{X} : c(x) \subseteq \Delta\}$ is open. Using compactness of Δ and part (i) of the

continuity axiom, it can easily be verified that $\{x \in \mathcal{X} : c(x) \cap \Delta \neq \emptyset\}$ is also closed, which implies that $\{x \in \mathcal{X} : c(x) = \{\ominus\}\}$ is open. \square

Since \mathcal{X}^* is an open subset of \mathcal{X} , clearly, there exists a number $\alpha^* \in (0, 1)$ such that $\alpha^*x + (1 - \alpha^*)\{p^*\} \in \mathcal{X}^*$ for every $x \in \mathcal{X}$. Define a binary relation \succsim on Δ as, for every $p, q \in \Delta$,

$$p \succsim q \quad \Leftrightarrow \quad \alpha^*p + (1 - \alpha^*)p^* \in c(\alpha^*\{p, q\} + (1 - \alpha^*)\{p^*\}).$$

Note that the relation \succsim is complete by definitions. We shall now show that \succsim is transitive. Take any $p, q, r \in \Delta$ with $p \succsim q$ and $q \succsim r$. Then,

$$\alpha^*p + (1 - \alpha^*)p^* \in c(\alpha^*\{p, q\} + (1 - \alpha^*)\{p^*\}), \quad (11)$$

$$\alpha^*q + (1 - \alpha^*)p^* \in c(\alpha^*\{q, r\} + (1 - \alpha^*)\{p^*\}). \quad (12)$$

Put $\tilde{x} := \alpha^*\{p, q, r\} + (1 - \alpha^*)\{p^*\}$ and $\tilde{z} := \{\eta \in \{p, q, r\} : \alpha^*\eta + (1 - \alpha^*)p^* \in c(\tilde{x})\}$. Observe that if $p \in \tilde{z}$, which means $\alpha^*p + (1 - \alpha^*)p^* \in c(\tilde{x})$, then $c(\tilde{x}) \cap (\alpha^*\{p, r\} + (1 - \alpha^*)\{p^*\}) \neq \emptyset$, and exclusive WARP implies $\alpha^*p + (1 - \alpha^*)p^* \in c(\alpha^*\{p, r\} + (1 - \alpha^*)\{p^*\})$. Similarly, if $q \in \tilde{z}$, then $c(\tilde{x}) \cap (\alpha^*\{p, q\} + (1 - \alpha^*)\{p^*\}) \neq \emptyset$. Hence, in this case, from (11) and exclusive WARP it follows that $p \in \tilde{z}$. Analogously, $r \in \tilde{z}$ implies $q \in \tilde{z}$, as a result of (12). Moreover, \tilde{z} is nonempty by construction. It follows that $p \in \tilde{z}$ in all contingencies, and hence, $p \succsim r$.

Let us now show that \succsim satisfies the classical independence axiom. Pick any $p, q, r \in \Delta$ and $\gamma \in (0, 1)$. Suppose $p \succsim q$, meaning that $p' := \alpha^*p + (1 - \alpha^*)p^*$ belongs to $c(x')$ where $x' := \alpha^*\{p, q\} + (1 - \alpha^*)\{p^*\}$. Set $r' := \alpha^*r + (1 - \alpha^*)p^*$. Observe that

$$\gamma x' + (1 - \gamma)\{r'\} = \alpha^*\{\gamma p + (1 - \gamma)r, \gamma q + (1 - \gamma)r\} + (1 - \alpha^*)\{p^*\}. \quad (13)$$

Similarly,

$$\gamma p' + (1 - \gamma)r' = \alpha^*(\gamma p + (1 - \gamma)r) + (1 - \alpha^*)p^*. \quad (14)$$

Moreover, independence(i) implies $\gamma p' + (1 - \gamma)r' \in c(\gamma x' + (1 - \gamma)\{r'\})$. By (13) and (14), this simply means $\gamma p + (1 - \gamma)r \succsim \gamma q + (1 - \gamma)r$, as we sought.

To verify continuity of \succsim , let $(p_n), (q_n)$ be convergent sequences in Δ such that $\alpha^*p_n + (1 - \alpha^*)p^* \in c(\alpha^*\{p_n, q_n\} + (1 - \alpha^*)\{p^*\})$ for every n . Then, from the continuity axiom, it readily follows that $\alpha^*\lim p_n + (1 - \alpha^*)p^* \in c(\alpha^*\{\lim p_n, \lim q_n\} + (1 - \alpha^*)\{p^*\})$. Hence, $\lim p_n \succsim \lim q_n$ if $p_n \succsim q_n$ for every n . This proves that \succsim is also continuous.

By the properties of \succsim that we have established, there exists an expected utility function $\varphi : \Delta \rightarrow \mathbb{R}$ that represents \succsim . Next, we prove that φ also represents the restriction of c to Δ .

Claim 2. *If $c(x) \cap \Delta \neq \emptyset$, then $c(x) \cap \Delta = \arg \max_x \varphi$.*

Proof. Let $x \in \mathcal{X}$ and $p' \in c(x) \cap \Delta$. Then, independence(i) yields $\alpha^* p' + (1 - \alpha^*) p^* \in c(\alpha^* x + (1 - \alpha^*) \{p^*\})$. From the definition of α^* and exclusive WARP, it follows that $\alpha^* p' + (1 - \alpha^*) p^* \in c(\alpha^* \{p', q\} + (1 - \alpha^*) \{p^*\})$ for every $q \in x$. That is, $p' \in \arg \max_x \varphi$. Thus, $c(x) \cap \Delta \subseteq \arg \max_x \varphi$ for every $x \in \mathcal{X}$.

To establish the converse inclusion, pick any $x \in \mathcal{X}$ and $p \in \arg \max_x \varphi$. Recall that $c(\alpha^* x + (1 - \alpha^*) \{p^*\}) \subseteq \Delta$. Pick any $q \in x$ with $\alpha^* q + (1 - \alpha^*) p^* \in c(\alpha^* x + (1 - \alpha^*) \{p^*\})$. Then, $c(\alpha^* x + (1 - \alpha^*) \{p^*\}) \cap (\alpha^* \{p, q\} + (1 - \alpha^*) \{p^*\}) \neq \emptyset$. Moreover, $\alpha^* p + (1 - \alpha^*) p^* \in c(\alpha^* \{p, q\} + (1 - \alpha^*) \{p^*\})$ by definitions of p and φ . Hence, exclusive WARP implies $\alpha^* p + (1 - \alpha^*) p^* \in c(\alpha^* x + (1 - \alpha^*) \{p^*\})$. Finally, from independence(i) it follows that $p \in c(x)$ provided that $c(x) \cap \Delta \neq \emptyset$. \square

The next claim will be useful in the derivation of the function W .

Claim 3. (i) *If $c(x) \cap \Delta \neq \emptyset$ and $c(\{p\}) = \{p\}$, then $c(\alpha x + (1 - \alpha) \{p\}) \subseteq \Delta$ for every $\alpha \in (0, 1)$.*

(ii) $\ominus \in c(x) \cap c(y)$ *implies $\ominus \in c(\alpha x + (1 - \alpha)y)$ for every $\alpha \in (0, 1)$.*

Proof. We start with the proof of (i). Pick any $x \in \mathcal{X}$ and $p \in \Delta$ such that $c(x) \cap \Delta \neq \emptyset$ and $c(\{p\}) = \{p\}$. Suppose by contradiction that there exists an $\alpha \in (0, 1)$ such that $\ominus \in c(\alpha x + (1 - \alpha) \{p\})$. Set $z := \alpha x + (1 - \alpha) \{p\}$. Observe that $\gamma z + (1 - \gamma) \{p_*\} = \gamma \alpha x + (1 - \gamma \alpha) \{p_\gamma\}$ for any $\gamma \in (0, 1)$, where

$$p_\gamma := \frac{\gamma(1 - \alpha)}{1 - \gamma\alpha} p + \frac{1 - \gamma}{1 - \gamma\alpha} p_*.$$

It is also clear that $\lim_{\gamma \rightarrow 1} p_\gamma = p$. Thus, by Claim 1, $c(p_\gamma) = \{p_\gamma\}$ for all sufficiently large $\gamma \in (0, 1)$. From independence(i), it follows that $c(\gamma \alpha x + (1 - \gamma \alpha) \{p_\gamma\}) \cap \Delta \neq \emptyset$ for any such γ . On the other hand, independence(ii) implies $c(\gamma z + (1 - \gamma) \{p_*\}) = \{\ominus\}$ for any $\gamma \in (0, 1)$, which is a contradiction.

To prove (ii), let $x, y \in \mathcal{X}$ be such that $\ominus \in c(x) \cap c(y)$. Fix any $\alpha, \gamma \in (0, 1)$ and put $x_\gamma := \gamma x + (1 - \gamma) \{p_*\}$. Then, independence(ii) implies $c(x_\gamma) = \{\ominus\}$. Thus, by applying independence(ii) to the sets x_γ and y , we also see that $c(\alpha x_\gamma + (1 - \alpha)y) = \{\ominus\}$. Since γ is an arbitrary number in $(0, 1)$, from the continuity axiom it follows that $\ominus \in c(\lim_{\gamma \rightarrow 1} \alpha x_\gamma + (1 - \alpha)y) = c(\alpha x + (1 - \alpha)y)$, as we sought. \square

Claim 4. *There exists a continuous and affine function $W : \mathcal{X} \rightarrow R$ such that for every $x \in \mathcal{X}$,*

$$\begin{aligned} c(x) \cap \Delta \neq \emptyset &\Leftrightarrow \max_x \varphi \geq W(x), \\ \ominus \in c(x) &\Leftrightarrow \max_x \varphi \leq W(x). \end{aligned} \tag{15}$$

Proof. Fix an $x \in \mathcal{X}_*$. By Claim 1, the sets $\{\lambda \in [0, 1] : c(\lambda x + (1 - \lambda)\{p^*\}) = \{\ominus\}\}$ and $\{\lambda \in [0, 1] : c(\lambda x + (1 - \lambda)\{p^*\}) \subseteq \Delta\}$ are relatively open in $[0, 1]$. Since both of the former sets are also disjoint and nonempty, their union cannot be equal to the connected set $[0, 1]$. That is, there exists a $\lambda \in (0, 1)$ such that $\ominus \in c(\lambda x + (1 - \lambda)\{p^*\}) \neq \{\ominus\}$. In fact, this number, which we denote by $\lambda^*(x)$, is the unique number in $[0, 1]$ that satisfies the latter two properties. Indeed, for any $\lambda > \lambda^*(x)$ the set $\lambda^*(x)x + (1 - \lambda^*(x))\{p^*\}$ can be expressed as a convex combination of $\lambda x + (1 - \lambda)\{p^*\}$ and $\{p^*\}$. Thus, if $c(\lambda x + (1 - \lambda)\{p^*\}) \cap \Delta$ were nonempty for some $\lambda \in (\lambda^*(x), 1]$, Claim 3(i) would imply $c(\lambda^*(x)x + (1 - \lambda^*(x))\{p^*\}) \subseteq \Delta$, which contradicts the definition of $\lambda^*(x)$. Hence, $c(\lambda x + (1 - \lambda)\{p^*\}) = \{\ominus\}$ for $\lambda > \lambda^*(x)$. Similarly, $\lambda < \lambda^*(x)$ implies $c(\lambda x + (1 - \lambda)\{p^*\}) \subseteq \Delta$ by Claim 3(i).

Let us now show that $\lambda^*(\cdot)$ is continuous on \mathcal{X}_* . Pick a sequence (x_n) in \mathcal{X}_* that converges to some $x \in \mathcal{X}_*$. It suffices to find a subsequence (x_{n_k}) such that $\lim_k \lambda^*(x_{n_k}) = \lambda^*(x)$. For each n , put $\lambda_n^* := \lambda^*(x_n)$ and $z_n := \lambda_n^* x_n + (1 - \lambda_n^*)\{p^*\}$. Then, $\ominus \in c(z_n) \neq \{\ominus\}$ for every n . In particular, we can pick a sequence of ordinary alternatives (q_n) such that $q_n \in c(z_n)$ for every n . Since $[0, 1] \times \Delta$ is compact, there exists a subsequence $(\lambda_{n_k}^*, q_{n_k})$ that converges to some $(\lambda, q) \in [0, 1] \times \Delta$, which also implies $\lim_k z_{n_k} = \lambda x + (1 - \lambda)\{p^*\}$. From the continuity axiom, it then follows that q and \ominus both belong to $c(\lambda x + (1 - \lambda)\{p^*\})$. Hence, λ satisfies the defining properties of the unique number $\lambda^*(x)$, implying that $\lambda = \lambda^*(x)$.

Since $\lambda^*(x) \in (0, 1)$ for every $x \in \mathcal{X}_*$, we can define a function $h : \mathcal{X}_* \rightarrow \mathbb{R}$ as $h(x) := 1/\lambda^*(x)$. The next step is to show that h is affine on \mathcal{X}_* . Let $x, y \in \mathcal{X}_*$ and $\gamma \in (0, 1)$. Note that $\gamma x + (1 - \gamma)y$ also belongs to \mathcal{X}_* by independence(ii). Put $\tau := (1 + \lambda^*(x)(1 - \gamma)/\lambda^*(y)\gamma)^{-1}$ so that

$$\gamma = \frac{\tau \lambda^*(x)}{\tau \lambda^*(x) + (1 - \tau) \lambda^*(y)} \quad \text{and} \quad 1 - \gamma = \frac{(1 - \tau) \lambda^*(y)}{\tau \lambda^*(x) + (1 - \tau) \lambda^*(y)}.$$

Then,

$$\begin{aligned} & \tau (\lambda^*(x)x + (1 - \lambda^*(x))\{p^*\}) + (1 - \tau) (\lambda^*(y)y + (1 - \lambda^*(y))\{p^*\}) \\ &= (\tau \lambda^*(x) + (1 - \tau) \lambda^*(y)) \left(\frac{\tau \lambda^*(x)}{\tau \lambda^*(x) + (1 - \tau) \lambda^*(y)} x + \frac{(1 - \tau) \lambda^*(y)}{\tau \lambda^*(x) + (1 - \tau) \lambda^*(y)} y \right) \\ & \quad + (1 - (\tau \lambda^*(x) + (1 - \tau) \lambda^*(y))) \{p^*\} \\ &= (\tau \lambda^*(x) + (1 - \tau) \lambda^*(y)) (\gamma x + (1 - \gamma)y) + (1 - (\tau \lambda^*(x) + (1 - \tau) \lambda^*(y))) \{p^*\}. \end{aligned}$$

Moreover, Claim 3(ii) and independence(i) jointly imply

$$\ominus \in c(\tau (\lambda^*(x)x + (1 - \lambda^*(x))\{p^*\}) + (1 - \tau) (\lambda^*(y)y + (1 - \lambda^*(y))\{p^*\})) \neq \{\ominus\}.$$

It follows that $\lambda^*(\gamma x + (1 - \gamma)y) = \tau\lambda^*(x) + (1 - \tau)\lambda^*(y)$, i.e., $h(\gamma x + (1 - \gamma)y) = (\tau\lambda^*(x) + (1 - \tau)\lambda^*(y))^{-1}$. Since the latter number equals both $\theta := \gamma/(\tau\lambda^*(x))$ and $\theta' := (1 - \gamma)/((1 - \tau)\lambda^*(y))$, we see that $h(\gamma x + (1 - \gamma)y) = \tau\theta + (1 - \tau)\theta' = \gamma/\lambda^*(x) + (1 - \gamma)/\lambda^*(y) = \gamma h(x) + (1 - \gamma)h(y)$, as we sought.

In order to extend the function h to \mathcal{X} , pick a number $\gamma_* \in (0, 1)$ such that $\gamma_*x + (1 - \gamma_*)\{p_*\} \in \mathcal{X}_*$ for every $x \in \mathcal{X}$. (The existence of such a γ_* is guaranteed by Claim 1.) For every $x \in \mathcal{X}$, set

$$h_1(x) := \frac{h(\gamma_*x + (1 - \gamma_*)\{p_*\}) - (1 - \gamma_*)h(\{p_*\})}{\gamma_*}.$$

Observe that by affinity of h , we have $h_1(x) = h(x)$ for every $x \in \mathcal{X}_*$. Moreover, the function h_1 is continuous and affine on \mathcal{X} by the corresponding properties of the maps h and $x \rightarrow \gamma_*x + (1 - \gamma_*)\{p_*\}$ on \mathcal{X}_* and \mathcal{X} , respectively.

Now, fix a number $\beta > 0$ and set, for each $x \in \mathcal{X}$,

$$W(x) := \max_x \varphi + \beta(h_1(x) - 1).$$

Then, $W(x)$ is a continuous and affine function of $x \in \mathcal{X}$ by the corresponding properties of $\max_x \varphi$ and $h_1(x)$.

Pick any $x \in \mathcal{X}$. To verify (15), it suffices to establish the following three properties:

(i) $h_1(x) > 1$ if $x \in \mathcal{X}_*$; (ii) $h_1(x) = 1$ if $\ominus \in c(x) \neq \{\ominus\}$; and (iii) $h_1(x) < 1$ if $x \in \mathcal{X}^*$.

Since $h_1 = h$ on \mathcal{X}_* , property (i) immediately follows from the definitions. To prove (ii), suppose $\ominus \in c(x) \neq \{\ominus\}$. Let (γ_n) be a sequence in $(0, 1)$ that converges to 1 and put $x_n := \gamma_n x + (1 - \gamma_n)\{p_*\}$ for each n . Observe that by independence(ii), $x_n \in \mathcal{X}_*$ for every n . We claim

$$\lim_n \lambda^*(x_n) = 1. \tag{16}$$

Otherwise, there exist an $\varepsilon \in (0, 1)$ and a subsequence (n_k) such that $\lambda_{n_k}^* := \lambda^*(x_{n_k}) \leq 1 - \varepsilon$ for each k . By passing to a further subsequence if necessary, we can assume that $(\lambda_{n_k}^*)$ converges to some $\lambda \in [0, 1 - \varepsilon]$. Then, the continuity axiom and the definition of $\lambda^*(\cdot)$ imply $\ominus \in c(\lim_k \lambda_{n_k}^* x_{n_k} + (1 - \lambda_{n_k}^*)\{p_*\}) = c(\lambda x + (1 - \lambda)\{p_*\})$. This contradicts Claim 3(i) and proves (16).

By definitions and affinity of h_1 , we then see that $h_1(x) = \lim_n \gamma_n h_1(x) + (1 - \gamma_n)h_1(\{p_*\}) = \lim_n h_1(x_n) = \lim_n h(x_n) = \lim_n 1/\lambda^*(x_n) = 1$. This proves (ii).

Finally, to prove (iii), suppose that $c(x) \subseteq \Delta$. Following an argument that we used when defining the function $\lambda^*(\cdot)$, there exists a $\gamma \in (0, 1)$ such that $\ominus \in c(\gamma\{p_*\} + (1 - \gamma)x) \neq \{\ominus\}$. Observe that $h_1(\{p_*\}) > 1$ while $h_1(\gamma\{p_*\} + (1 - \gamma)x) = 1$ by properties (i) and (ii), respectively. Since h_1 is an affine function, it follows that $1 = \gamma h_1(\{p_*\}) + (1 - \gamma)h_1(x) >$

$\gamma + (1 - \gamma)h_1(x)$, implying that $1 > h_1(x)$. \square

It is easy to check that Claims 2 and 4 complete the proof of Theorem 0.

C. Proof of Theorem 1

For the “if” part of Theorem 1, it suffices to show that a choice correspondence that admits an AR representation must satisfy L-continuity—the necessity of all other axioms is fairly obvious. Given an AR representation (μ, K, a) , we denote by Φ the corresponding net expected utility function. That is, for every $x \in \mathcal{X}$ and $p \in x$,

$$\Phi(p, x) := \int_{\mathcal{U}} \left(u(p) - K(\max_{q \in x} u(q) - u(p)) \right) \mu(du). \quad (17)$$

Furthermore, whenever ϕ is non-zero as an element of \mathbb{R}_0^B , we set $u_\phi := \phi / \|\phi\|$ as in Appendix A.

Claim 5. *Let c be a choice correspondence on \mathcal{X} that admits an AR representation (μ, K, a) . Then, c satisfies the L-continuity axiom.*

Proof. Recall that an affine function on \mathcal{X} is Lipschitz continuous iff it is a positive affine transformation of a function of the form $x \rightarrow \int_{\mathcal{U}} \max_x u \eta(du)$ for a (finite) signed measure η on \mathcal{U} (Dekel et al., 2007, Supplementary Material).

Set $\Psi(x) := \max_{p \in x} \Phi(p, x) - a$ for $x \in \mathcal{X}$, so that $\ominus \in c(x)$ iff $\Psi(x) \leq 0$. Note that $\max_{p \in x} \Phi(p, x) = \int_{\mathcal{U}} \max_x u \eta(du)$, where $\eta := (1 + K) \|\phi\| \delta_{u_\phi} - K\mu$ if $\phi \neq 0$, and $\eta := -K\mu$ if $\phi = 0$. Thus, Ψ is Lipschitz continuous on \mathcal{X} . That is, there exists a number $\bar{m} > 0$ such that $\Psi(x) - \Psi(y) \leq \bar{m}d_H(x, y)$ for every $x, y \in \mathcal{X}$.

If Ψ is constant over \mathcal{X} , then either $\ominus \in c(x)$ for every $x \in \mathcal{X}$ or $\ominus \notin c(x)$ for every $x \in \mathcal{X}$. Hence, in this case, the conclusion of L-continuity holds trivially, for any $y^*, y_* \in \mathcal{X}$.

Suppose now that Ψ is not constant, so that $\Psi(y^*) > \Psi(y_*)$ for some $y^*, y_* \in \mathcal{X}$. Set $m := \bar{m} / (\Psi(y^*) - \Psi(y_*))$. Then, $\Psi(x) - \Psi(y) \leq m(\Psi(y^*) - \Psi(y_*))d_H(x, y)$ for every $x, y \in \mathcal{X}$. In particular, $d_H(x, y) \leq \alpha/m$ implies $\Psi(x) - \Psi(y) \leq (\Psi(y^*) - \Psi(y_*))\alpha / (1 - \alpha)$ for any $\alpha \in (0, 1)$ and $x, y \in \mathcal{X}$. Since Ψ is an affine function, the latter inequality simply means $\Psi(\alpha y_* + (1 - \alpha)x) \leq \Psi(\alpha y^* + (1 - \alpha)y)$. It follows that $\ominus \in c(\alpha y^* + (1 - \alpha)y)$ implies $\ominus \in c(\alpha y_* + (1 - \alpha)x)$ for any $\alpha \in (0, 1)$ and $x, y \in \mathcal{X}$ with $d_H(x, y) \leq \alpha/m$, as demanded by L-continuity. \square

To prove the “only if” part of Theorem 1, let c be a choice correspondence on \mathcal{X} that satisfies the axioms (A1)-(A6). Define the points p^*, p_* and the functions φ, W as in the proof of Theorem 0.

For every $x \in \mathcal{X}$, set

$$\Psi(x) := \max_x \varphi - W(x).$$

By Claim 4, for any $x \in \mathcal{X}$, we have $\ominus \in c(x)$ iff $\Psi(x) \leq 0$, while $c(x) \cap \Delta \neq \emptyset$ iff $\Psi(x) \geq 0$. In particular, $\Psi(\{p^*\}) > 0 > \Psi(\{p_*\})$, and the function $p \rightarrow \Psi(\{p\})$ is not constant over Δ .

We shall next show that the functions $\varphi(p)$ and $\Psi(\{p\})$ induce the same preference relation over Δ .

Claim 6. *For any $p, q \in \Delta$, we have $\varphi(p) \geq \varphi(q)$ if and only if $\Psi(\{p\}) \geq \Psi(\{q\})$.*

Proof. Let us write $\Psi(p)$ and $c(p)$ in place of $\Psi(\{p\})$ and $c(\{p\})$, respectively. As noted earlier, the function $\Psi(p)$ is not constant over Δ . Thus, clearly, it suffices to show that $\varphi(p) \geq \varphi(q)$ implies $\Psi(p) \geq \Psi(q)$.

Pick any $p, q \in \Delta$ with $\varphi(p) \geq \varphi(q)$. Set $p' := \alpha^*p + (1 - \alpha^*)p^*$ and $q' := \alpha^*q + (1 - \alpha^*)p^*$. Note that $\varphi(p) \geq \varphi(q)$ simply means $p' \in c(\{p', q'\})$. Moreover, we have $\Psi(q') > 0$ by definition of α^* . Let $\gamma \in (0, 1)$ be the number such that $\Psi(\gamma q' + (1 - \gamma)p_*) = 0$, so that $c(\gamma q' + (1 - \gamma)p_*) \cap \Delta \neq \emptyset$. Then, independence(iii) implies $c(\gamma p' + (1 - \gamma)p_*) \cap \Delta \neq \emptyset$, i.e., $\Psi(\gamma p' + (1 - \gamma)p_*) \geq 0$. It follows that $\Psi(\gamma p' + (1 - \gamma)p_*) \geq \Psi(\gamma q' + (1 - \gamma)p_*)$. Since Ψ is affine and $\alpha^*, \gamma > 0$, we conclude that $\Psi(p) \geq \Psi(q)$, as we sought. \square

Now, let us define a binary relation \succ^* on \mathcal{X} as, for every $x, y \in \mathcal{X}$,

$$x \succ^* y \iff \Psi(x) \geq \Psi(y).$$

The next claim shows that \succ^* satisfies all axioms demanded by Sarver's (2008) representation theorem.

Claim 7. *\succ^* is a complete and transitive binary relation that also satisfies the following properties:*

- (i) *If $x \succ^* y \succ^* z$, then there exist $\alpha, \alpha' \in (0, 1)$ such that $\alpha x + (1 - \alpha)z \succ^* y \succ^* \alpha'x + (1 - \alpha')z$.*
- (ii) *If $x \succ^* y$, then $\alpha x + (1 - \alpha)z \succ^* \alpha y + (1 - \alpha)z$ for every $z \in \mathcal{X}$ and $\alpha \in (0, 1]$.*
- (iii) *If $\{p\} \succ^* \{q\}$ and $p \in x$, then $x \succ^* x \cup \{q\}$.*
- (iv) *There exist $y^*, y_* \in \mathcal{X}$ and a number $m > 0$ such that $\alpha y^* + (1 - \alpha)y \succ^* \alpha y_* + (1 - \alpha)x$ for every $\alpha \in (0, 1)$ and $x, y \in \mathcal{X}$ with $d_H(x, y) \leq \alpha/m$.*

Proof. We shall only prove (iii) and (iv), for the remaining claims are well-known implications of the fact that Ψ is a real valued, affine function over \mathcal{X} .

To prove (iii), pick any $p, q \in \Delta$ and $x \in \mathcal{X}$ such that $\{p\} \succ^* \{q\}$ and $p \in x$. Then, Claim 6 and the definition of \succ^* imply $\varphi(p) \geq \varphi(q)$. Pick $r \in \{p^*, p_*\}$ and $\alpha \in (0, 1]$

such that $\Psi(\alpha(x \cup \{q\}) + (1 - \alpha)\{r\}) = 0$. Put $y := \alpha x + (1 - \alpha)\{r\}$ and observe that $y \cup \{\alpha q + (1 - \alpha)r\} = \alpha(x \cup \{q\}) + (1 - \alpha)\{r\}$. Hence, $\Psi(y \cup \{\alpha q + (1 - \alpha)r\}) = 0$, which implies $c(y \cup \{\alpha q + (1 - \alpha)r\}) \cap \Delta \neq \emptyset$. Moreover, since $\varphi(p) \geq \varphi(q)$ and $\alpha p + (1 - \alpha)r \in y$, there exists a $p' \in y$ such that $\varphi(p') \geq \varphi(q')$ for every $q' \in y \cup \{\alpha q + (1 - \alpha)r\}$. Then, $p' \in c(y \cup \{\alpha q + (1 - \alpha)r\})$ by Claim 2, while asymmetric alpha implies $p' \in c(y)$. In particular, $c(y) \cap \Delta \neq \emptyset$, and hence, $\Psi(y) := \Psi(\alpha x + (1 - \alpha)\{r\}) \geq 0 = \Psi(\alpha(x \cup \{q\}) + (1 - \alpha)\{r\})$. As in the last argument of Claim 6, this simply means $\Psi(x) \geq \Psi(x \cup \{q\})$, i.e., $x \succ^* x \cup \{q\}$. This proves (iii).

For the proof of (iv), let $y^*, y_* \in \mathcal{X}$ and $m > 0$ be as posited by the L-continuity axiom. Pick any $x, y \in \mathcal{X}$ and $\alpha \in (0, 1)$ such that $d_H(x, y) \leq \alpha/m$. It remains to show that $\Psi(\alpha y^* + (1 - \alpha)y) \geq \Psi(\alpha y_* + (1 - \alpha)x)$.

Put $z^* := \alpha y^* + (1 - \alpha)y$ and $z_* := \alpha y_* + (1 - \alpha)x$. Let $r \in \{p^*, p_*\}$ and $\gamma \in (0, 1]$ be such that $\Psi(\gamma z^* + (1 - \gamma)\{r\}) = 0$. Set

$$\alpha' := \gamma\alpha, \quad y' := \frac{\gamma(1 - \alpha)}{1 - \gamma\alpha}y + \frac{1 - \gamma}{1 - \gamma\alpha}\{r\}, \quad \text{and} \quad x' := \frac{\gamma(1 - \alpha)}{1 - \gamma\alpha}x + \frac{1 - \gamma}{1 - \gamma\alpha}\{r\}.$$

Observe that $\gamma z^* + (1 - \gamma)\{r\} = \alpha' y^* + (1 - \alpha')y'$, while $\gamma z_* + (1 - \gamma)\{r\} = \alpha' y_* + (1 - \alpha')x'$. It is also easy to check that $d_H(x', y') = \gamma(1 - \alpha)(1 - \gamma\alpha)^{-1}d_H(x, y) \leq \gamma(1 - \alpha)(1 - \gamma\alpha)^{-1}\alpha/m \leq \alpha'/m$. Moreover, by construction, \ominus belongs to $c(\alpha' y^* + (1 - \alpha')y')$. Thus, the L-continuity axiom implies that \ominus also belongs to $c(\alpha' y_* + (1 - \alpha')x')$. In turn, this is equivalent to saying that $0 \geq \Psi(\gamma z_* + (1 - \gamma)\{r\})$. Thus, $\Psi(z^*) \geq \Psi(z_*)$ by a usual argument. \square

From Claim 7 it follows that Sarver's (2008) representation theorem applies to the binary relation \succ^* . That is, there exists a probability measure μ on \mathcal{U} and a number $K \geq 0$ such that for every $x, y \in \mathcal{X}$,

$$x \succ^* y \quad \Leftrightarrow \quad \max_{p \in x} \Phi(p, x) \geq \max_{p \in y} \Phi(p, y),$$

where $\Phi(p, x)$ is defined as in equation (17).

Since the function $x \rightarrow \max_{p \in x} \Phi(p, x)$ is continuous on \mathcal{X} , the binary relation \succ^* is continuous in a standard, topological sense. Proposition 2 of Dekel et al. (2001) shows that an affine functional that represents a continuous binary relation on \mathcal{X} is unique up to positive affine transformations. Thus, there exist $\alpha > 0$ and $\gamma \in \mathbb{R}$ such that $\Psi(x) = \alpha \max_{p \in x} \Phi(p, x) + \gamma$ for every $x \in \mathcal{X}$. In particular, $\Psi(x) \geq (\leq) 0$ iff $\max_{p \in x} \Phi(p, x) \geq (\leq) -\gamma/\alpha$.

Hence, if we let $a := -\gamma/\alpha$, it follows that for every $x \in \mathcal{X}$,

$$\begin{aligned} c(x) \cap \Delta \neq \emptyset &\Leftrightarrow \max_{p \in x} \Phi(p, x) \geq a, \\ \ominus \in c(x) &\Leftrightarrow \max_{p \in x} \Phi(p, x) \leq a. \end{aligned} \tag{18}$$

We shall next show that

$$c(x) \cap \Delta \neq \emptyset \quad \Rightarrow \quad c(x) \cap \Delta = \arg \max_{p \in x} \Phi(p, x). \tag{19}$$

First observe that by Claim 6 and the definitions, the function $p \rightarrow \Phi(p, \{p\})$ represents the same preference relation over Δ as the function φ . Moreover, it is easily verified that $\Phi(p, x) = (1 + K)\Phi(p, \{p\}) - K \int_{\mathcal{U}} \max_x u \mu(du)$ for every $p \in \Delta$ and $x \in \mathcal{X}$. By combining these two observations, we see that $\arg \max_{p \in x} \Phi(p, x) = \arg \max_x \varphi$ for every $x \in \mathcal{X}$. Thus, (19) follows from Claim 2.

Finally, note that (18) and (19) are jointly equivalent to the statements (i) and (ii) in Definition 1. Thus, we have shown that (μ, K, a) is an AR representation for c , which completes the proof of Theorem 1.

D. Proofs of Propositions 1-3

Note that for any AR representation (μ, K, a) and any $x \in \mathcal{X}$, we have

$$\begin{aligned} \max_{p \in x} \Phi(p, x) &= \max_{p \in x} \int_{\mathcal{U}} \left(u(p) - K(\max_x u - u(p)) \right) \mu(du) \\ &= \max_{p \in x} \int_{\mathcal{U}} u(p) \mu(du) - K \left(\int_{\mathcal{U}} \max_x u \mu(du) - \max_{p \in x} \int_{\mathcal{U}} u(p) \mu(du) \right) \\ &= \max_x \phi - K \left(\int_{\mathcal{U}} \max_x u \mu(du) - \max_x \phi \right). \end{aligned}$$

Here, the second equality is a consequence of the additivity of the expectation operator, whereas the first and third equalities follow from the definitions of Φ and ϕ , respectively. To simplify our notation, let us set $\mathcal{R}(x) := K \left(\int_{\mathcal{U}} \max_x u \mu(du) - \max_x \phi \right)$, which also equals $K \int_{\mathcal{U}} (\max_x u - u(\bar{p})) \mu(du)$ for any $\bar{p} \in \arg \max_x \phi$ and $x \in \mathcal{X}$. It then follows that

$$\max_{p \in x} \Phi(p, x) = \max_x \phi - \mathcal{R}(x) \quad \forall x \in \mathcal{X}.$$

It is also worth noting that $\mathcal{R}(\{p\}) = 0$ for every $p \in \Delta$. Thus, the choice correspondence

c represented by (μ, K, a) exhibits choice overload at some $x \in \mathcal{X}$ iff

$$\max_{p \in x} \Phi(p, x) < a \leq \max_x \phi. \quad (20)$$

In what follows, Δ° stands for the interior of Δ . That is, $\Delta^\circ := \{p \in \Delta : p_b > 0 \forall b \in B\}$.

Lemma 1. *Let (μ, K, a) be a non-trivial AR representation for c . Assume further that $K > 0$ and the support of μ contains at least two distinct points. Then, there exists a $\bar{p} \in \Delta^\circ$ with $\phi(\bar{p}) = a$. Moreover, any neighborhood of $\{\bar{p}\}$ contains a set $x \in \mathcal{X}$ such that c exhibits choice overload at x .*

Proof. Let p_* and p^* be as in the non-triviality axiom so that $\phi(p_*) < a < \phi(p^*)$. Since Δ° is a dense subset of Δ , without loss of generality we can assume $p_* \in \Delta^\circ$ by the first part of the continuity axiom. Clearly, there exists an $\alpha \in (0, 1)$ with $\phi(\alpha p_* + (1 - \alpha)p^*) = a$. Since $\alpha > 0$, it is also clear that $\bar{p} := \alpha p_* + (1 - \alpha)p^*$ belongs to Δ° as well.

By assumption, the support of μ contains a point \bar{u} that is distinct from u_ϕ . Since \bar{u} and u_ϕ represent distinct preferences on Δ , there exist $q, q' \in \Delta$ such that

$$\bar{u}(q - q') > 0 \geq u_\phi(q - q').$$

As \bar{p} belongs to Δ° , there exists an $\varepsilon > 0$ such that $x^\beta := \{\bar{p}, \bar{p} + \beta(q - q')\} \subseteq \Delta$ for every $\beta \in (0, \varepsilon]$. Observe that $\phi(\bar{p}) \geq \phi(\bar{p} + \beta(q - q'))$ because ϕ and u_ϕ represent the same preference over Δ . Thus, $\max_{r \in x^\beta} \Phi(r, x^\beta) = \phi(\bar{p}) - \mathcal{R}(x^\beta)$, while $\mathcal{R}(x^\beta) = K \int_{\mathcal{U}} (\max_{x^\beta} u - u(\bar{p})) \mu(du)$. Furthermore, $\bar{u}(\bar{p} + \beta(q - q')) > \bar{u}(\bar{p})$, implying that the function $u \rightarrow \max_{x^\beta} u - u(\bar{p})$ attains a strictly positive value at \bar{u} . Since \bar{u} belongs to the support of μ , and $u \rightarrow \max_{x^\beta} u - u(\bar{p})$ is continuous and nonnegative, it follows that $\int_{\mathcal{U}} (\max_{x^\beta} u - u(\bar{p})) \mu(du) > 0$. Then, $\mathcal{R}(x^\beta) > 0$ because $K > 0$. Moreover, $\mathcal{R}(x^\beta) > 0$ implies $\phi(\bar{p}) > \max_{r \in x^\beta} \Phi(r, x^\beta)$ by definitions. As $\phi(\bar{p}) = a$, we can then conclude that c exhibits choice overload at x^β . This completes the proof because β is an arbitrary number in $(0, \varepsilon]$, and $\lim_{\beta \rightarrow 0} x^\beta = \{\bar{p}\}$. \square

Proof of Proposition 1. The “if” part of the proposition follows from Lemma 1 immediately. For the “only if” part, note that if $K = 0$ or the support of μ consists of a single point, then $\mathcal{R}(x) = 0$ for every $x \in \mathcal{X}$. This, in turn, implies $\max_{p \in x} \Phi(p, x) = \max_x \phi$, which rules out instances of the form (20). \square

Proof of Proposition 2. Let (μ, K, a) and (μ', K', a') be strictly non-trivial AR representations for c and c' , respectively, and assume $\mu = \mu'$.

Suppose c' is more choice overload prone than c . Let \bar{p} , ε and x^β be as in the proof of

Lemma 1, so that c exhibits choice overload at x^β for every $\beta \in (0, \varepsilon]$. Then, c' must also do so, which means that for every $\beta \in (0, \varepsilon]$,

$$\max_{r \in x^\beta} \Phi'(r, x^\beta) < a' \leq \max_{x^\beta} \phi'. \quad (21)$$

Observe that $\max_{r \in x^\beta} \Phi'(r, x^\beta)$ converges to $\Phi'(\bar{p}, \{\bar{p}\}) = \phi'(\bar{p})$ as $\beta \rightarrow 0$ because x^β converges to $\{\bar{p}\}$. Similarly, $\max_{x^\beta} \phi'$ also converges to $\phi'(\bar{p})$. Hence, in (21) passing to limit as $\beta \rightarrow 0$ yields $\phi'(\bar{p}) = a'$. Moreover, $\mu = \mu'$ implies $\phi(\bar{p}) = \phi'(\bar{p})$, while $\phi(\bar{p}) = a$ by definition of \bar{p} . So, $a = a'$, as we sought.

It remains to show that $K \leq K'$. Fix any $\beta \in (0, \varepsilon]$, and recall that $\mathcal{R}(x^\beta) > 0$. To simplify our notation, let us write x in place of x^β . By contradiction, suppose $K > K'$. With $\mu = \mu'$, this implies

$$\mathcal{R}(x) = K \int_{\mathcal{U}} \left(\max_x u - u(\bar{p}) \right) \mu(du) > K' \int_{\mathcal{U}} \left(\max_x u - u(\bar{p}) \right) \mu(du) = \mathcal{R}'(x).$$

Following usual arguments, let $\ominus \in c'(\gamma x + (1 - \gamma)\{\tilde{p}\}) \neq \{\ominus\}$ for some $\gamma \in (0, 1]$ and $\tilde{p} \in \Delta$. Set $x_\gamma := \gamma x + (1 - \gamma)\{\tilde{p}\}$. Observe that $\mathcal{R}(x_\gamma) = \gamma \mathcal{R}(x)$ because \mathcal{R} is an affine function and $\mathcal{R}(\{\tilde{p}\}) = 0$. Similarly, $\mathcal{R}'(x_\gamma) = \gamma \mathcal{R}'(x)$. Thus, $\mathcal{R}(x) > \mathcal{R}'(x)$ implies $\mathcal{R}(x_\gamma) > \mathcal{R}'(x_\gamma)$. Also, $\max_{x_\gamma} \phi' = \mathcal{R}'(x_\gamma) + a'$ by definition of x_γ . Hence, with $a = a'$ and $\phi = \phi'$, we see that

$$\mathcal{R}(x_\gamma) + a > \mathcal{R}'(x_\gamma) + a' = \max_{x_\gamma} \phi. \quad (22)$$

Moreover, since $\mathcal{R}'(x_\gamma) \geq 0$, the equality in (22) also entails $\max_{x_\gamma} \phi \geq a' = a$. It follows that $\max_{x_\gamma} \phi \geq a > \max_{x_\gamma} \phi - \mathcal{R}(x_\gamma) = \max_{r \in x_\gamma} \Phi(r, x)$, which means that c exhibits choice overload at x_γ . On the other hand, c' does not exhibit choice overload at x_γ because $c'(x_\gamma) \neq \{\ominus\}$ by definitions. This proves the “only if” part of the proposition.

For the “if” part, note that $K' \geq K$ and $\mu' = \mu$ imply $\max_{p \in x} \Phi'(p, x) \leq \max_{p \in x} \Phi(p, x)$ for every $x \in \mathcal{X}$. If, in addition, a' is equal to a , from the characterization (20) it easily follows that c' is more choice overload prone than c because ϕ' is equal to ϕ as well. \square

Proof of Proposition 3. Let (μ, K, a) and (μ', K', a') be non-trivial AR representations for c and c' , respectively. Assume further that $\mu = \mu'$ and $K = K'$. Then, $\max_{p \in x} \Phi(p, x) = \max_{p \in x} \Phi'(p, x)$ for every $x \in \mathcal{X}$. Recall that $c(x) = \{\ominus\}$ iff $\max_{p \in x} \Phi(p, x) < a$, and similarly, $c'(x) = \{\ominus\}$ iff $\max_{p \in x} \Phi'(p, x) < a'$. Hence, clearly, if $a \leq a'$, then $c(x) = \{\ominus\}$ implies $c'(x) = \{\ominus\}$ for every $x \in \mathcal{X}$.

Conversely, suppose $c(x) = \{\ominus\}$ implies $c'(x) = \{\ominus\}$. Since c satisfies the non-triviality axiom, we can find a convergent sequence (p_n) in Δ such that $c(\{p_n\}) = \{\ominus\}$ for every n ,

and $\ominus \in c(\{\lim p_n\}) \neq \{\ominus\}$. By the latter condition, we have $\phi(p) = a$, where $p := \lim p_n$. Moreover, $c(\{p_n\}) = \{\ominus\}$ implies $c'(\{p_n\}) = \{\ominus\}$, which means $\phi'(p_n) < a'$. Thus, $\lim \phi'(p_n) \leq a'$, whereas $\lim \phi'(p_n) = \lim \phi(p_n) = \phi(p) = a$. So, $a \leq a'$. \square

References

- Anderson, Christopher J., The psychology of doing nothing: forms of decision avoidance result from reason and emotion, *Psychological Bulletin* **129** (2003), 139–167.
- Apesteguia, Jose and Miguel A. Ballester, A theory of reference-dependent behavior, *Economic Theory* **40** (2009), 427–455.
- Arrow, Kenneth J., Rational choice functions and orderings, *Economica* **26** (1959), 121–127.
- Bell, David E., Regret in decision making under uncertainty, *Operations Research* **30** (1982), 961–981.
- Buturak, Gökhan and Özgür Evren, Choice overload and asymmetric regret, New Economic School (2015), mimeo. (Available at SSRN: <http://ssrn.com/abstract=2347128>)
- Chernoff, Herman, Rational selection decision functions, *Econometrica* **22** (1954), 422–443.
- Dean, Mark, Status quo bias in large and small choice sets, New York University (2008), mimeo.
- Dean, Mark, Özgür Kıbrıs and Yusufcan Masatlioglu, Limited attention and status quo bias, *Journal of Economic Theory*, in press.
- Dekel, Eddie, Barton L. Lipman and Aldo Rustichini, Representing preferences with a unique subjective state space, *Econometrica* **69** (2001), 891–934.
- Dekel, Eddie, Barton L. Lipman, Aldo Rustichini and Todd Sarver, Representing preferences with a unique subjective state space: a corrigendum, *Econometrica* **75** (2007), 591–600.
- Dhar, Ravi, Consumer preference for a no-choice option, *Journal of Consumer Research* **24** (1997), 215–231.
- Dubra, Juan, Fabio Maccheroni and Efe A. Ok, Expected utility theory without the completeness axiom, *Journal of Economic Theory* **115** (2004), 118–133.
- Gerasimou, Georgios, Indecisiveness, undesirability and overload revealed through rational choice deferral, University of St Andrews (2016), mimeo.
- Hayashi, Takashi, Regret aversion and opportunity-dependence, *Journal of Economic Theory* **139** (2008), 242–268.
- Inbar, Yoel, Simona Botti and Karlene Hanko, Decision speed and choice regret: When haste feels like waste, *Journal of Experimental Social Psychology* **47** (2011), 533–540.
- Iyengar, Sheena S. and Mark R. Lepper, When choice is demotivating: can one desire too much of a good thing?, *Journal of Personality and Social Psychology* **79** (2000), 995–1006.
- Kahneman, Daniel and Amos Tversky, Prospect theory: An analysis of decision under risk, *Econometrica* **47** (1979), 263–291.

- Kreps, David M., A representation theorem for ‘preference for flexibility’, *Econometrica* **47** (1979), 565–577.
- Loomes, Graham and Robert Sugden, Regret theory: An alternative theory of rational choice under uncertainty, *Economic Journal* **92** (1982), 805–824.
- Loomes, Graham and Robert Sugden, Some implications of a more general form of regret theory, *Journal of Economic Theory* **41** (1987), 270–287.
- Masatlioglu, Yusufcan and Efe A. Ok, Rational choice with status quo bias, *Journal of Economic Theory* **121** (2005), 1–29.
- Masatlioglu, Yusufcan and Efe A. Ok, A canonical model of choice with initial endowments, *Review of Economic Studies* **81** (2014), 851–883.
- Ok, Efe A., Pietro Ortoleva and Gil Riella, Revealed (p)reference theory, *American Economic Review* **105** (2015), 299–321.
- Ortoleva, Pietro, Status quo bias, multiple priors and uncertainty aversion, *Games and Economic Behavior* **69** (2010), 411–424.
- Redelmeier, Donald A. and Eldar Shafir, Medical decision making in situations that offer multiple alternatives, *Journal of the American Medical Association* **273** (1995), 302–305.
- Riella, Gil and Roe Teper, Probabilistic dominance and status quo bias, *Games and Economic Behavior* **87** (2014), 288–304.
- Sagi, Jacob S., Anchored preference relations, *Journal of Economic Theory* **130** (2006), 283–295.
- Sarver, Todd, Anticipating regret: Why fewer options may be better, *Econometrica* **76** (2008), 263–305.
- Schneider, Rolf, *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge, Cambridge University Press, 1993.
- Sen, Amartya K., Choice functions and revealed preference, *Review of Economic Studies* **38** (1971), 307–317.
- Stoye, Jörg, Axioms for minimax regret choice correspondences, *Journal of Economic Theory* **146** (2011), 2226–2251.
- Sugden, Robert, An axiomatic foundation for regret theory, *Journal of Economic Theory* **60** (1993), 159–180.
- Tversky, Amos and Eldar Shafir, Choice under conflict: The dynamics of deferred decision, *Psychological Science* **3** (1992), 358–361.
- White, Chris M. and Ulrich Hoffrage, Testing the tyranny of too much choice against the allure of more choice, *Psychology & Marketing* **26** (2009), 280–298.