



## Management Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### A Foundation for Cue-Triggered Behavior

Daniele Pennesi

To cite this article:

Daniele Pennesi (2020) A Foundation for Cue-Triggered Behavior. Management Science

Published online in Articles in Advance 04 Jun 2020

. <https://doi.org/10.1287/mnsc.2020.3632>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2020, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# A Foundation for Cue-Triggered Behavior

Daniele Pennesi<sup>a</sup>

<sup>a</sup>Department of Economic and Social Sciences, Mathematics and Statistics, School of Management and Economics, University of Torino, 10134 Torino, Italy

Contact: [daniele.pennesi@unito.it](mailto:daniele.pennesi@unito.it),  <http://orcid.org/0000-0001-8505-9666> (DP)

Received: January 28, 2019

Revised: January 9, 2020; February 25, 2020

Accepted: March 7, 2020

Published Online in Articles in Advance:  
June 4, 2020

<https://doi.org/10.1287/mnsc.2020.3632>

Copyright: © 2020 INFORMS

**Abstract.** Exposure to environmental cues triggers sudden preference reversals in several choice contexts, including consumption and intertemporal, social, and risky choices. This paper introduces a dual-self model of cue-triggered behavior that (1) is based on a general mechanism that makes it applicable to many choice contexts, (2) allows a sharp comparative analysis of the responsiveness to cues, (3) can explain a wide range of behavioral anomalies, from a cue-triggered present bias to high-frequency variations in social and risk preferences, and (4) can inform the design of managerial interventions and advertising strategies employing environmental cues. Testable restrictions combining choice and nonchoice data fully characterize the model.

**History:** Accepted by Manel Baucells, decision analysis.

**Keywords:** environmental cues • dual-self • time preferences • choice under risk • advertising

## 1. Introduction

Environmental cues are apparently irrelevant aspects of the decision environment that generate sudden variations in individual behavior. Cues have a central role in explaining important dynamic consumption patterns, such as habit formation and addiction (Laibson 2001, Bernheim and Rangel 2009). Environmental cues are also widely used by policy makers to nudge citizens (Pollay et al. 1996) and by managers to nudge workers and customers (Ashforth and Mael 1989, Krishna 2012). The relevance of cues as determinants of individual choices has spurred the development of formal models of cue-triggered behavior (Laibson 2001, Bernheim and Rangel 2009). A limitation common to existing models is that they are devised to describe cue-triggered *consumption*. In fact, exposure to cues generates preference reversals in several contexts beyond consumption, including intertemporal choices (Kim and Zauberman 2013), social choices (Bennett 2009, Andreoni et al. 2017), and choices under risk (Guiso et al. 2018).

The aim of the present work is to develop a unified model of behavior that can describe cue-triggered preference reversals observed across choice domains. In so doing, we also provide policy makers and managers with a flexible model that can inform the design of public policies, marketing strategies, and managerial interventions employing cues. We develop our model within the dual-self paradigm, where a “hot” and a “cold” self, with potentially conflicting preferences, interact to decide. Exposure to a cue arouses the hot self, thus giving it (more) decision weight at the moment of choice. For instance, a picture of a starving

child triggers the desire to make a donation by arousing the (more) “altruistic self.” In our preferred interpretation, the selves represent distinct regions of the brain differentially activated by the cue.<sup>1</sup> Although each self is “rational,” their interaction—mediated by the cue—generates “irrational” behaviors.

In the domain of intertemporal choice, our model is consistent with a cue-triggered present bias (e.g., Laibson 1997). It arises when the selves have different time preferences, even if neither of the former is present biased. In the domain of social choice, our model can explain high-frequency variations in prosocial behavior that are otherwise difficult to rationalize by fluctuations of individual preferences (e.g., Blanco et al. 2011). Our model can also explain behavior triggered by apparently unrelated cues, such as impatience or risk seeking triggered by exposure to sexual cues (see, e.g., Knutson et al. 2008, Van den Bergh et al. 2008, and Kim and Zauberman 2013) and risk aversion triggered by exposure to graphic images (Guiso et al. 2018). The model also sheds light on the long-standing debate over the welfare effect of (noninformative) advertising. From a welfare viewpoint, we offer two possibilities as for what utility may be deemed as prescriptive: the utility of the cold self or the combined total utility. In the former case, advertisement that stirs the individual toward the hot state is always detrimental. In the latter case, it may not be so.

The structure of the model makes it possible to introduce a comparative notion of responsiveness to cues: an individual  $B$  is more responsive to cues than an individual  $A$  if, when facing the same cue,  $B$  always

gives more weight to her hot self than does  $A$ . We prove that such an intuitive notion does not allow comparison of all individuals but only those with “similar” hot and cold preferences. We argue that the result helps to understand whether a heterogeneous behavioral response to a cue is determined by differential responsiveness to the cue or by different tastes.

Different from alternative approaches, we provide observable restrictions on cue-contingent choices that uniquely identify the preferences of the selves and the response to the cue. Our identification strategy relies on the observation that cues can often be ordered according to their intensity. For instance, a visual cue can be larger or brighter than another. We then exploit variations in the cue’s intensity to elicit the preferences of the selves and the effect of the cue on their relative decision weight.

Finally, we consider three extensions of the model. In the first, we relax the assumption that cues are perfectly observable, and we study a stochastic version of our cue-triggered model. We show that the derived stochastic choice rule is a particular case of the single-crossing random utility model (SCRUM) of Apesteguia et al. (2017). The relationship with the SCRUM allows us to make new predictions concerning the effect of managerial interventions and advertising campaigns. In the second, we relax the assumption that the selves are rational, and we provide a behavioral characterization of a generalized cue-triggered model. In the third, we allow the sensitivity to cues to be context dependent. This enables us to model the role of unchosen options in the presence of environmental cues.

## 2. The Model

### 2.1. Cues, Selves, and Choice

In our approach cues are features of the choice environment, hence distinct and independent from attributes of the objects of choice. For instance, a song is an acoustic cue when played as background music in a store, whereas it is not a cue when deciding what music to listen to. In the latter case, the song itself is an object of choice. Our model is based on two assumptions: (1) environmental cues can be ordered according to their intensity, and (2) choices are cue contingent. Concerning assumption (1), environmental cues are sensory stimuli (acoustic, visual, olfactory, haptic) and, as such, can have different intensities. For instance, acoustic cues can be ordered by their loudness or frequency and visual cues by their size or brightness (see Bernheim and Rangel (2009) for a similar assumption). We model different intensities of the cue with the set  $\Theta = [0, 1]$  endowed with its standard order  $\leq$ . For any given cue, we identify the cue with its intensity: if  $\theta \leq \theta'$ , the cue (associated with)  $\theta'$  is more

intense than the cue (associated with)  $\theta$ . We assume that the order  $\leq$  is objective and measurable on a physical scale, as is loudness or brightness. Clearly, the meaning of  $\leq$  depends on the application under consideration or, in the case of a laboratory experiment, the experimental setting.

We interpret  $\theta = 0$  as the absence of the cue and  $\theta = 1$  as the cue’s maximal intensity. The latter could be thought of as the upper limit of the human ability to discriminate between two physical stimuli. A stimulus that is physically more intense than  $\theta = 1$  is equivalent to  $\theta = 1$ , because physiological constraints do not allow a human to distinguish between the two. Cues can also be multidimensional; what matters for our model is the possibility of ordering cues according to their intensity. For this reason, the use of  $\Theta = [0, 1]$  is for expository purposes. Any (finite or infinite) set  $\Theta$  with  $|\Theta| > 1$  endowed with a weak order  $\leq$  and such that there exist two elements  $\theta_0, \theta_1 \in \Theta$  with  $\theta_0 \leq \theta \leq \theta_1$  for all  $\theta \in \Theta$  would be equivalent. Concerning assumption (2), indirect evidence supporting it comes from the high frequency of cue-triggered unplanned purchases observed during shopping occasions (see, e.g., Abratt and Goodey 1990).

Assumptions (1) and (2) translate to the model by a family of functions  $\{u_\theta\}_{\theta \in \Theta}$  representing cue-contingent utilities over a convex set  $X$  (i.e.,  $u_\theta : X \rightarrow \mathbb{R}$ ). The inequality  $u_\theta(p) \geq u_\theta(q)$  means that  $p$  has a (weakly) higher utility than  $q$  when the cue’s intensity is  $\theta$ . We can now introduce our model.

**Definition 1.** A cue-triggered (CT) model for  $\{u_\theta\}_{\theta \in \Theta}$  is a triple  $(u, v, \phi)$  such that, for each  $\theta \in \Theta$ ,

$$u_\theta = (1 - \phi(\theta))u + \phi(\theta)v, \quad (\text{CT})$$

where  $u : X \rightarrow \mathbb{R}$ ,  $v : X \rightarrow \mathbb{R}$  are affine<sup>2</sup> and  $\phi : \Theta \rightarrow [0, 1]$  is a weakly increasing function such that  $\phi(0) = 0$  and  $\phi(1) = 1$ .

Contingent on a cue’s intensity  $\theta$ , the utility  $u_\theta$  is expressed as a combination of the functions  $u$  and  $v$  with weights  $1 - \phi(\theta)$  and  $\phi(\theta)$ . In our interpretation, the function  $u = u_0$  represents the utility of the *cold* self:  $u(p) \geq u(q)$  if and only if  $p$  is (weakly) better than  $q$  in the absence of the cue. We interpret the function  $v = u_1$  as representing the utility of the *hot* self:  $v(p) \geq v(q)$  if and only if  $p$  is weakly better than  $q$  when the cue’s intensity is maximal. The function  $\phi : \Theta \rightarrow [0, 1]$  converts the cue’s intensity into decision weights. Exposure to the environmental cue arouses the hot self, giving it (more) decision weight at the moment of choice. Because  $\phi(\theta)$  is increasing, the more intense the cue, the larger the weight assigned to the hot self. As noted previously in this subsection, the use of  $\Theta = [0, 1]$  is for expository purposes. If cues are multidimensional and can be ordered according to their

intensity, the function  $\phi$  integrates all the dimensions into a single decision weight.

A possible interpretation of the mechanism linking cues to choice is that the cue activates a region of the brain that was inactive in the absence of the cue (see Berridge and Aldridge 2009). The more intense the cue, the more intense the activity of the region, and hence the easier it is to override the cold self. Empirical evidence that may be consistent with the above-mentioned interpretation is ample in the neuroeconomics literature. For intertemporal choices, McClure et al. (2004) observe the activation of different regions of the brain, the limbic system and the lateral prefrontal cortex, when evaluating immediate versus delayed rewards, respectively. A similar pattern was found by Kuhnen and Knutson (2005) for choices under risk. They observe the activation of the nucleus accumbens preceding a risky choice, whereas the activation of the anterior insula precedes a riskless choice. For addicted individuals, Bernheim and Rangel (2009) provide extensive support for the assumption that cues activate different regions of the brain.

Figure 1 illustrates the model in the particular case of  $\phi(\theta) = \theta$  and a choice between two options,  $p$  or  $q$ . Because  $u(q) > u(p)$ , the cold self prefers  $q$  to  $p$ . The hot self has the opposite preference,  $v(p) > v(q)$ . Suppose that  $p$  is a new energy drink and  $q$  is a traditional caffeinated drink. For a weak (intensity of the) environmental cue  $\theta$ , such as a small visual ad, the individual will buy the traditional caffeinated drink ( $u_\theta(q) > u_\theta(p)$ ). However, when facing a more intense cue  $\theta'$ —for instance, a large visual ad associating the drink with a sports car or a celebrity endorsing the product—the individual reverses her preference and decides to buy the new energy drink ( $u_{\theta'}(p) > u_{\theta'}(q)$ ).

It is clear from the previous example why the standard revealed preference analysis may fail in the presence of cues. Without information on the cue's intensity, revealed preferences say that  $p$  and  $q$  are both selected and hence are indifferent for the decision maker, a correct prediction in the absence of cues

but a potentially wrong conclusion if cues matter for choice behavior. We will devote Section 4 to showing how the use of enriched choice data can solve the previous identification issue.

The CT model has the following properties. First, it functions independently of dynamic aspects of consumption, such as habit formation or addiction, thus providing a unified treatment of cue-triggered behavior observed across choice domains (as we show in Section 3). Second, in the CT model, the mechanism converting cues into decision weights is independent of the cues' nature. Therefore, the CT model accounts for behavior elicited by cues that are apparently unrelated to the object of choice—for example, impatience, risk seeking, or unplanned purchases triggered by sexual stimuli.<sup>3</sup> If we interpret the hot self as a region of the brain that values immediate gratification, *any* cue that activates this region will trigger a desire to seek immediate gratification. Hence for the CT model, the decision to enter a bakery after smelling freshly baked cookies or to behave myopically after observing a sexually explicit image are driven by the same mechanism. Third, the dual-self structure of the CT model makes it suitable<sup>4</sup> to model customers' reaction to advertising cues, be they visual, olfactory (see Spangenberg et al. 1996), acoustic (see Milliman 1982), or haptic. Moreover, the model can be used to assess the welfare effect of non-informative advertising, as we discuss in more detail in Section 3.4. Finally, the CT model is static but can be extended to describe dynamic consumption patterns, such as those studied in Laibson (2001) and Bernheim and Rangel (2004). A simple dynamic extension consists in tying the current response to the cue (the function  $\phi$ ) to past behavior. Consider the following particular case of the CT model, where one of the selves acts as a dictator depending on the cue's intensity. This model is qualitatively similar to (a static version of) the cue-triggered addiction model of Bernheim and Rangel (2009) that we discuss in the next section. We call such a special case the 0-1 model.

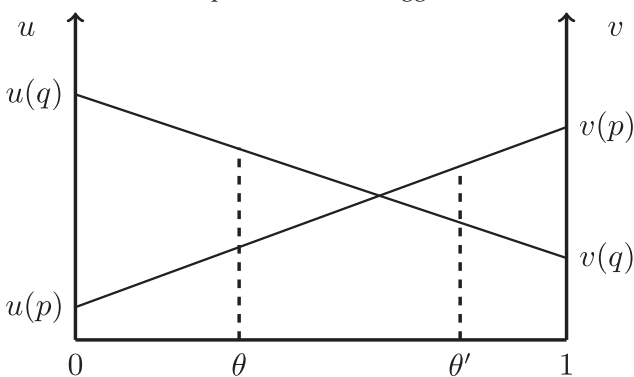
**Definition 2.** A 0-1 model for  $\{u_\theta\}_{\theta \in \Theta}$  is a CT model with  $\phi(\theta) = 0$  for all  $\theta \leq \hat{\theta}$  and  $\phi(\theta) = 1$  for all  $\hat{\theta} < \theta$ , for some  $\hat{\theta} \in \Theta$ , and exactly one strict inequality.

That is, for each  $\theta \in \Theta$ ,

$$u_\theta = \begin{cases} u & \theta \leq \hat{\theta}, \\ v & \hat{\theta} < \theta, \end{cases} \quad (0-1)$$

with exactly one strict inequality. In their theory of addiction, Bernheim and Rangel (2009) assume that the individual operates in either a hot or a cold "mode." The likelihood of switching modes is a function of past consumption, lifestyle, and a random

**Figure 1.** An Example of the Cue-Triggered Model



component. A dynamic extension of the 0-1 model in which the threshold  $\hat{\theta}$  is a function of past behavior would generate predictions similar to the model of Bernheim and Rangel (2009).<sup>5</sup> As for the CT, the 0-1 model is not limited to describing addiction. For instance, consider the results of Card and Dahl (2011), who found a positive relation between domestic violence and unexpected results of football matches. Suppose that  $\theta$  represents the emotional cue generated by winning or losing a football match. When the emotional cue is more intense than the threshold  $\hat{\theta}$ —for example, when the prematch probability of losing was low but the favorite team lost—the individual suddenly switches from a nonviolent to a violent self. The 0-1 model, for which we provide a behavioral characterization in Section 4, can easily describe this behavior.

## 2.2. Comparative Statics

In this section we exploit the structure of the CT model to perform a comparative static analysis aimed at ranking two individuals according to their *responsiveness to cues*. This exercise is fundamental in order to better understand the falsifiability of the model and to correctly target, for example, marketing strategies. Intuitively, we consider an individual  $B$  to be more responsive to cues than an individual  $A$  if, *ceteris paribus*,  $B$  always gives more decision weight to her hot self than does  $A$ —namely, when  $\phi^B(\theta) \geq \phi^A(\theta)$ , where  $\phi^B$  (respectively,  $\phi^A$ ) is the function transforming the cue to decision weight for individual  $B$  (respectively,  $A$ ). The following definition provides the behavioral characterization of the previous intuitive notion.

**Definition 3.** Given two individuals  $A, B$  with cue-contingent preferences represented by two CT models  $(u^A, v^A, \phi^A)$  and  $(u^B, v^B, \phi^B)$ , individual  $B$  is more responsive to cues (MRC) than individual  $A$  if

- (a)  $u^A \approx u^B$  and  $v^A \approx v^B$ , and
- (b) for all  $p, q \in X$ ,  $u^B(p) \geq u^B(q)$  and  $u_\theta^B(p) \geq u_\theta^B(q)$  for some  $\theta_0 \leq \theta$  implies  $u_\theta^A(p) \geq u_\theta^A(q)$ .

The first condition implies that the individuals have the same hot and the same cold preferences. The second condition says that whenever the cue-contingent preference of individual  $B$  is aligned with the preference of her cold self, the same is true for individual  $A$  (because by (a), the cold preferences of  $A$  and  $B$  are equivalent). Although the two conditions are expressed in terms of utilities, it is straightforward to formulate them using revealed cue-contingent preferences (see Section 4). The following proposition characterizes the parametric restrictions entailed by the definition of MRC.

**Proposition 1.** Given two individuals  $A, B$  with cue-contingent preferences represented by two CT models  $(u^A, v^A, \phi^A)$  and  $(u^B, v^B, \phi^B)$ , if  $B$  is MRC than  $A$ , then  $\phi^B(\theta) \geq \phi^A(\theta)$  for all  $\theta \in \Theta$ . The converse implication is true if  $u^A \approx u^B$  and  $v^A \approx v^B$ .

Consider again the effect of a visual cue on the decision to purchase a new energy drink  $p$  or a traditional caffeinated drink  $q$  and two individuals: one with  $\phi^A(\theta) = \theta$  and one with  $\phi^B(\theta) = \sqrt{\theta}$ . Suppose that  $u^A(p) = u^B(p) = 0$  and  $v^A(p) = v^B(p) = 1$  and  $u^A(q) = u^B(q) = v^A(q) = v^B(q) = 0.5$ . Then, for  $\theta = 0.3$ ,  $u_\theta^A(p) = 0.3 < 0.5 = u_\theta^A(q)$ , but  $u_\theta^B(p) = 0.54 > 0.5 = u_\theta^B(q)$ . Therefore, for the same intensity of the cue  $\theta = 0.3$ , the individuals  $A$  and  $B$  have the opposite behavior ( $B$  purchases the new energy drink and  $A$  sticks to the traditional one) even if their selves have identical preferences ( $u^A = u^B$  and  $v^A = v^B$ ). When both  $A$  and  $B$  have cue-contingent preferences represented by a 0-1 model, Proposition 1 implies that the individual  $B$  is MRC than  $A$  is if  $\hat{\theta}^B \leq \hat{\theta}^A$ . Hence the threshold determining the switch to the hot self is lower for the more responsive individual.

Proposition 1 proves that it is possible to order two individuals according to their responsiveness to the cue only if they have the same hot and cold preferences. Example 1 illustrates the possible consequences of allowing for a violation of assumption (a). The result of Proposition 1 helps to correctly interpret the effect of certain advertising campaigns. Suppose that an advertising campaign introduces a visual cue in a store. After the campaign, scanner data show that younger customers buy more of the advertised product than do older ones. This evidence is not sufficient to conclude that younger customers are more responsive to cues than older ones, as the following example shows.

**Example 1.** Suppose that  $p$  is the advertised product (the new energy drink) and  $q$  its competitor (the traditional caffeinated drink). Assume that there are two customers, one young and one old, for which  $u^Y = u^O \equiv u$  and  $\phi^Y = \phi^O \equiv \phi$ , so that both have the same cold preference and same response to the cue. However, their hot preferences diverge:  $v^Y(q) < v^Y(p)$ , but  $v^O(p) > v^O(q)$ . Suppose now that  $u(p) < u(q)$ . Hence before the introduction of the cue  $\theta = 0$ , both the young and the old buy the traditional drink  $q$ . After the introduction of the cue  $\theta > 0$ , the new energy drink  $p$  is preferred to  $q$  by the young if

$$\phi(\theta) > \frac{k^Y}{1 + k^Y},$$

where  $k^Y \equiv (u(q) - u(p)) / (v^Y(p) - v^Y(q)) > 0$ . Put differently,  $p$  will never be selected by the old customer (because  $u_\theta^O(q) > u_\theta^O(p)$  for all  $\theta$ ). Therefore, the young

may reverse her preferences because the difference  $v^Y(p) - v^Y(q)$  is large enough and not because she is more responsive to the cue. In fact, the young and the old customers respond identically to the cue;  $\phi^Y = \phi^O$ . However, the cue-contingent preferences of both satisfy point (b) of Definition 3, with  $B = Y$  and  $A = O$ . Hence, making clear the reason for the assumption in point (a) of Definition 3 is necessary.

It follows from Example 1 that a strategy targeting the older customer should raise her utility of consuming  $p$  (in particular, the value of  $p$  for the hot self of the older) rather than increasing her exposure to advertising cues. The intuition in Example 1 is also relevant for public policies. For example, Pollay et al. (1996) estimate that teenagers are three times “more responsive” to cigarette advertisements than are adults. According to Proposition 1, the conclusion is true *only if* teenagers and adults share the same hot and cold preferences, a condition that may be false in practice.

### 2.3. Related Literature

The two main models of cue-triggered behavior are Laibson (2001) and Bernheim and Rangel (2004). In the dynamic model of Laibson (2001), there is a (binary) random cue that occurs in each period. Each cue is associated with a compensatory process that is activated in the period only if the related cue occurs. Conditional upon the realization of the cue, the individual can take one of two actions. The instantaneous utility is concave and depends on each compensatory process (the habit) and the action. In turn, the current action influences future utilities by modifying the compensatory process. In particular, the repeated pairing of a cue and an action creates complementarity. First, our model allows for non-binary choices. Second, we do not require the presence of a compensatory process activated by the cue. The cue affects choices by arousing the hot self regardless of past behavior. In this respect, the Laibson (2001) model does not explain how the compensatory process *starts*—that is, what mechanism is at work the first time a cue is present. Our model can be thought of as describing such a mechanism. In any case, independence from habit makes our model applicable to contexts where habit is not well defined, such as risky and social choices. Finally, cues and consumption in the CT model are not necessarily complementary goods, as is the case in the Laibson (2001) model; hence the CT model allows for general substitution patterns between cues and consumption.

Bernheim and Rangel (2004) propose a model of cue-triggered addiction founded on three premises: consumption of addictive substances is often a

mistake, consumption of an addictive substance sensitizes the individual to environmental cues related to consumption, and addicts understand their condition and manage their exposure to cues. In their model, the individual operates in either a hot or a cold mode. Choices made in the hot mode are “mistakes.” In each period, with a probability that depends on past consumption, lifestyle, and a purely exogenous component, the individual is exposed to a cue. If the cue occurs, the individual will pass to the hot mode and will consume the substance. If the cue does not occur, the individual decides whether to consume the addictive substance or not. Addiction increases the future probability of entering the hot mode. In our main model, the cue is deterministic and independent of the individual’s behavior. Moreover, we allow for a possibly smooth transition from the cold to the hot mode (selves) as a function of the cue’s intensity. In addition, choices of the hot self are not necessarily mistakes (consider a cue-triggered donation). We believe that our assumptions are appropriate to model nonaddictive behavior. In any case, as we previously argued, our 0-1 model is in the spirit of the Bernheim–Rangel model. Both the Laibson (2001) and Bernheim and Rangel (2004) models are informed by psychological and neuroscientific evidence. However, they do not provide axiomatic foundations for their models, whereas we fully characterize our model by conditions on enriched revealed preferences (Section 4).

Our cue-triggered model is also related to the “animal spirit” approach of Loewenstein and O’Donoghue (2004). In their model, a deliberative self and an affective self determine choices. The ability to override the affective self depends on the decision maker’s willpower. In our model, it is the cue that determines the relative weight of the selves. Moreover, in the main model we assume that the preferences of the selves are both rational (we relax this assumption in Section 5.2), although they allow for more general preferences. For example, when modeling choice under risk, they allow the hot self to have a non-expected utility preference, whereas we assume expected utility preferences. Therefore, our CT model is more parsimonious than theirs, but it still accounts for a variety of behavioral anomalies. The present work is also related to the large literature on multiple selves models (e.g., Thaler and Shefrin 1981, Laibson 1997, Gul and Pesendorfer 2001, and Fudenberg and Levine 2006). Our approach is one of the first to exploit choice and nonchoice data to behaviorally characterize the model.

### 3. Applications

In the following sections, we apply the CT model to rationalize the empirical evidence of cue-triggered choice in three domains: intertemporal choice, social

choice, and choice under risk. Finally, we discuss how to use the CT model to assess the welfare effect of non-informative advertising.

### 3.1. Intertemporal Choice

Dual-self models are commonly used for describing the struggle between short-term and long-term goals. See, for example, Thaler and Shefrin (1981), Laibson (1997), Gul and Pesendorfer (2001), and Fudenberg and Levine (2006). In our dual-self model, an additional element mediates the tension between the hot and the cold self at the moment of choice: the cue. In this section, we assume that objects of choice are streams of consumption  $c = (c_0, c_1, \dots, c_T)$  for some  $T \geq 2$ . Each  $c_t$  belongs to a convex space  $C$ —for example,  $C = [0, M]$ —in which case we interpret  $c_t \in [0, M]$  as a monetary amount. The utilities  $u_\theta$  are defined over the space  $X = C^{T+1}$ . To focus our analysis on the effect of cues on intertemporal preferences, we assume that the hot and cold selves have different discount factors but the same evaluation of consumption:<sup>6</sup> namely,  $u(c) = \sum_{t=0}^T \delta^t U(c_t)$  and  $v(c) = \sum_{t=0}^T \beta^t U(c_t)$ . It follows that each  $u_\theta$  in the CT model becomes

$$u_\theta(c) = \sum_{t=0}^T U(c_t) ((1 - \phi(\theta))\delta^t + \phi(\theta)\beta^t) \quad (1)$$

for some  $\beta, \delta \in [0, 1]$ . This expression corresponds to the double exponential discounted utility (McClure et al. 2007). As is well known, a convex combination of different discount factors produces *diminishing impatience*, the hallmark of hyperbolic discounting. This happens because as the time horizon grows, the discount factor  $D_\theta(t) \equiv (1 - \phi(\theta))\delta^t + \phi(\theta)\beta^t$  converges to the highest discount factor ( $\delta$  or  $\beta$ ). In addition to diminishing impatience, the model in Equation (1) is consistent with cue-triggered spikes in impatience. Formally, we consider the following cue-contingent index of impatience (see Halevy 2008):  $I_\theta(t) \equiv D_\theta(t)/D_\theta(t+1)$ . We say that  $D_\theta(t)$  displays diminishing impatience if  $I_\theta(t) > I_\theta(t+1)$  for all  $0 \leq t \leq T-1$  and that  $D_\theta(t)$  displays the present bias if  $I_\theta(0) > I_\theta(t)$  for all  $0 \leq t \leq T-1$ . We obtain the following result.

#### Proposition 2.

(a) If  $\beta \neq \delta$ ,  $D_\theta(t)$  displays diminishing impatience for all  $\theta \in \Theta$  such that  $\phi(\theta) \neq \{0, 1\}$ .

(b) If  $\beta < \delta$  and  $\theta < \theta'$ ,  $I_\theta(t) \geq I_{\theta'}(t)$  for all  $0 \leq t \leq T-1$ .

Although both selves have stationary preferences, the realized discount function displays diminishing impatience. Moreover, impatience is cue triggered: if the hot self is more impatient than the cold self, the more intense the cue, the higher the impatience. Consider the extreme case  $\beta = 0$ : the hot self

is completely impatient as it cares only about immediate consumption. In this case,  $u_\theta$  is equal to the following:

$$u_\theta(c) = U(c_0) + (1 - \phi(\theta)) \sum_{t=1}^T \delta^t U(c_t), \quad (2)$$

a quasi-hyperbolic discounting model with a cue-dependent present bias factor  $\beta_\theta = 1 - \phi(\theta) \in [0, 1]$ . We will refer to  $u_\theta$  in the Equation (2) as a  $\beta_\theta$ - $\delta$  representation of  $\{u_\theta\}_{\theta \in \Theta}$ . Given the properties of  $\phi$ , the more intense the cue  $\theta \leq \theta'$ , the stronger the present bias factor  $\beta_{\theta'} \leq \beta_\theta$ , establishing in this way a direct link between the cue's intensity and the degree of present bias. The  $\beta_\theta$ - $\delta$  model (and also the CT model of Equation (1)) naturally describes cue-triggered myopic behavior, such as cue-triggered drug consumption after a period of abstinence or detoxification therapy (see, e.g., Laibson 2001). Moreover, under the interpretation of the selves as different regions of the brain, the CT model in Equation (1) accommodates the empirical evidence of impatience triggered by sexual cues (e.g., Van den Bergh et al. 2008 and Kim and Zauberman 2013), which is hard to explain with alternative models. Finally, a simple application of Proposition 1 yields the following result on comparative impatience.

**Lemma 1.** Given two individuals  $A, B$  with cue-contingent preferences having a  $\beta_\theta^A$ - $\delta$  and a  $\beta_\theta^B$ - $\delta$  representation, respectively, if  $B$  is MRC than  $A$ , then  $\beta_\theta^B \leq \beta_\theta^A$  for all  $\theta \in \Theta$ .

The more responsive individual  $B$  displays a higher degree of present bias than does  $A$  for all cue's intensities.

### 3.2. Social Preferences

Environmental cues affect social preferences. In organizations, managers employ cues to trigger cooperation. For instance, the use of identity-enforcing cues represents an efficient alternative to monetary incentives or strict supervision (Ashforth and Mael 1989, Akerlof and Kranton 2005, Goette et al. 2012). Environmental cues are also commonly used to foster charitable giving (Bennett 2009, Andreoni et al. 2017, Kessler and Milkman 2018, Charness and Holder 2019). Models of social preferences (e.g., Fehr and Schmidt 1999, Charness and Rabin 2002, and Cox and Sadiraj 2007) struggle to capture high-frequency choice variations at the individual level, because doing so would require high-frequency variations in social preferences parameters. The latter, however, are found to be rather stable at the individual level (Blanco et al. 2011). By explicitly modeling cues, the CT model can account for the observed within-subject variability while leaving social preference parameters unaltered.

We will consider a general model of cue-contingent social preferences that nests a version of the quasi-maximin model of Charness and Rabin (2002) and a version of the egocentric altruism model of Cox and Sadiraj (2007). In this section, elements of choice are allocations among  $N$  individuals (the group or society) indexed by  $i = 1, \dots, N$ . An allocation is an element  $p = (p_1, p_2, \dots, p_N) \in \mathbb{R}^N$ , where the  $i$ th coordinate  $p_i$  represents the resources allocated to individual  $i$ . We consider the following cue-contingent utility for the  $i$ th individual:

$$u_{\theta}^i(p) = (1 - \phi^i(\theta))u^i(p_i) + \phi^i(\theta)v^i(p). \quad (3)$$

We make the following assumptions:  $v^i$  is differentiable in  $p_j$  for all  $j = 1, \dots, N$  and  $\partial v^i / \partial p_j \geq 0$  for all  $j = 1, \dots, N$ , and  $u^i$  is differentiable in  $p_i$  and  $\partial u^i / \partial p_i > 0$ . The cold self is fully selfish; it cares only about the private payoff  $p_i$ . Put differently, the hot self considers also the payoff of (all) others. A variation in the cue's intensity  $\theta$  modifies the relative weight between the private and the other-regarding component of the utility. When  $p \in \mathbb{R}_+^N$ ,  $u^i(p_i) = p_i$  and  $v^i(p) = p_1 + p_2 + \dots + p_N$ , the cue-triggered utility  $u_{\theta}^i$  is a version of the quasi-maximin model of Charness and Rabin (2002). Put differently, if  $v^i(p) = \sum_{j \neq i} p_j$  and  $u^i(p_i) = p_i$ , the utility  $u_{\theta}^i$  is similar to a version of the egocentric altruism model of Cox and Sadiraj (2007) given by  $u_{\theta}^i(p) = (1 - \phi^i(\theta))p_i + \phi^i(\theta) \sum_{j \neq i} p_j$ .

When preferences are represented by the utility function in Equation (3), it is intuitive to define an individual more altruistic than another, if the former always gives more weight to the other-regarding component of her utility than the latter. The following comparative definition formalizes the intuition (a similar definition is introduced in Cox et al. (2008)).

**Definition 4.** A utility function  $u_{\theta}^i$  is more altruistic than  $u_{\theta'}^i$  if

$$\frac{\partial u_{\theta}^i / \partial p_j}{\partial u_{\theta}^i / \partial p_i} \geq \frac{\partial u_{\theta'}^i / \partial p_j}{\partial u_{\theta'}^i / \partial p_i}$$

for all  $j \neq i$ .

A preference represented by  $u_{\theta}^i$  is more altruistic than  $u_{\theta'}^i$  if the former is more willing to trade off a private payoff in favor of that of others.

**Proposition 3.** For all  $\theta \leq \theta'$ ,  $u_{\theta}^i$  is more altruistic than  $u_{\theta'}^i$ .

The previous proposition is consistent with the empirical evidence of Bennett (2009), Andreoni et al. (2017), and Kessler and Milkman (2018) that shows how charitable giving is successfully triggered by environmental cues, be they digital or physical.

Our cue-triggered theory of altruism also explains the “identifiable victim effect” (Small et al. 2007). A cue triggering empathy such as the image of a starving child can be more effective in promoting donations than providing potential donors with accurate statistical information about the recipients.

To further illustrate the scope of our approach, the following example presents an application of the cue-triggered egocentric altruism model to a public good contribution game.

**Example 2.** Consider  $N$  individuals with identical endowment  $e$  and deciding their contribution to a public good. Each individual can contribute  $c_i \in [0, e]$  to a public good with return  $1/N < r < 1$ . The monetary payoff of each individual is then  $m(c_i) = e - c_i + r \sum_{j=1}^N c_j$ . Suppose that each individual has preferences represented by the following model:  $u_{\theta}^i(p) = (1 - \phi^i(\theta))p_i + \phi^i(\theta) \sum_{j \neq i} p_j$ . Then, if  $\phi^i(\theta)/(1 - \phi^i(\theta)) > (1 - r)/(r(N - 1))$ , the optimal contribution for individual  $i$  is  $e$ . If the condition is satisfied for all individuals, then the full contribution is the unique Nash equilibrium of the game. The simple proof of the result is in the appendix.

Because the threshold  $(1 - r)/(r(N - 1))$  is independent of the cue's intensity, the presence of environmental cues increases the likelihood of (full) cooperation (by increasing the ratio  $\phi^i(\theta)/(1 - \phi^i(\theta))$ ). This rationalizes the use of identity-enforcing cues in place of monetary incentives as a means to trigger cooperation within organizations. Finally, by applying the comparative analysis developed in Section 2.2 to the public good contribution game, we know that replacing a player  $i$  with a more responsive player  $j$  makes full contribution more likely to occur. Indeed, given the same level of the cue, player  $j$  will give less weight to her selfish self compared with the replaced player  $i$ , hence making the necessary condition for optimality of full contribution  $\phi^i(\theta)/(1 - \phi^i(\theta)) > (1 - r)/(r(N - 1))$  easier to fulfill.

### 3.3. Choice Under Risk

The third application of the CT model is to choice under risk. Indeed, experimental evidence confirms that also attitude toward risk is affected by environmental cues. For instance, exposure to graphic images, such a scene from a horror movie, triggers risk aversion (Guiso et al. 2018). As with the case of social preferences, high-frequency variations in choice under risk are hard to rationalize with changing risk attitudes, whereas they can be captured by explicitly modeling the role of environmental cues. In this section,  $X$  is the set of lotteries over a finite prize space  $Z$ , denoted by  $\Delta(Z)$ . Affine utilities on  $X$  are expected utilities in the sense of von Neumann and Morgenstern (1953).



Therefore,  $u(p) = \sum_{z \in Z} p(z)U(z)$ , and  $v(p) = \sum_{z \in Z} p(z)V(z)$  for some  $U, V : Z \rightarrow \mathbb{R}$ , and the cue-contingent utility  $u_\theta$  becomes

$$u_\theta(p) = \sum_{z \in Z} p(z)[(1 - \phi(\theta))U(z) + \phi(\theta)V(z)].$$

We do not make assumptions on the form of  $U$  and  $V$ . For instance,  $U$  may be linear or mildly concave, whereas  $V$  may be loss averse (i.e., with a kink at a reference point). Because we are interested in analyzing how the cue affects the individual's risk attitude, we consider the following (comparative) definition of risk aversion: given two cue-contingent utilities  $u_\theta, u_{\theta'}$ , we say that  $u_\theta$  is more risk averse than  $u_{\theta'}$  if, for all  $p, \delta_z \in \Delta(Z)$   $u_\theta(p) \geq u_\theta(\delta_z)$  implies  $u_{\theta'}(p) \geq u_{\theta'}(\delta_z)$ , where  $\delta_z$  is a lottery paying  $z \in Z$  for sure. When  $U, V$  are twice differentiable, we can define the indexes of absolute risk aversion for the two selves  $r_U = -U''/U'$  and  $r_V = -V''/V'$  and for the cue-contingent utility

$$u_\theta : r_\theta = -\frac{(1 - \phi(\theta))U'' + \phi(\theta)V''}{(1 - \phi(\theta))U' + \phi(\theta)V'}$$

for a given  $\theta \in \Theta$ . Intuitively, a more intense cue moves the cue-contingent risk attitude toward the preferences of the hot self. The following result proves that the intuition is correct.

**Proposition 4.** *If  $u$  is more (less) risk averse than  $v$ , then  $u_{\theta'}$  is more (less) risk averse than  $u_\theta$  for all  $\theta \leq \theta'$ . If  $U, V$  are twice differentiable and  $r_U \geq r_V$  (respectively,  $r_U \leq r_V$ ), then  $r_\theta \geq r_{\theta'}$  (respectively,  $r_\theta \leq r_{\theta'}$ ) for all  $\theta \leq \theta'$ .*

An application of Proposition 4 explains the use of cues in reducing the so-called perceived risk, which inherently affects the adoption of new products (Bearden and Shimp 1982, Herzstein et al. 2007). Consider a decision to buy or not buy a newly released foldable smartphone: the buyer may be uncertain about its durability. In this case, the seller can introduce a visual cue associating the smartphone with a durable object as a way to trigger the decision to purchase the device (rather than offering an extended warranty, for example). For an alternative application, our model of cue-triggered risk sensitivity can explain risk seeking triggered by apparently unrelated cues, such as sexual images or weather conditions (Bassi et al. 2013). Neuroscientific evidence (e.g., Kuhn and Knutson 2005) demonstrates that risk seeking and risk aversion activate different regions of the brain, the nucleus accumbens and anterior insula, respectively. In our interpretation of the selves as regions of the brain, any cue arousing the nucleus accumbens—such as a sexually explicit image (see Knutson et al. 2008)—triggers risk seeking.

Finally, a straightforward application of the comparative analysis developed in Section 2.2 shows that

an individual  $B$  is more responsive to cues than an individual  $A$  is if  $A$  always displays less (more) risk aversion than does  $B$  when the common cold self is less (more) risk averse than the common hot self.

### 3.4. The Prescriptive Analysis of Advertising Cues

There are two main approaches to the welfare assessment of noninformative advertising. In the *persuasive* view of Braithwaite (1928) and Kaldor (1950), advertising cues are taste shifters. Therefore, advertising is unambiguously welfare decreasing. In the *complementary* view of Becker and Murphy (1993), advertising enters the utility function directly, and it can be a bad as well as a good. In any case, cues and consumption are complementary: the presence of an advertising cue raises the marginal utility of consumption.

Our model is hybrid, and the welfare assessment of advertising cues depends on which “self” represents the prescriptive ranking for the individual. There are two equally plausible approaches: one takes the cold self as a benchmark, and the other takes the overall utility  $u_\theta$  as a benchmark. If the prescriptive utility is that of the cold self  $u$ , advertising cues are (weakly) welfare decreasing. By moving the decision weight away from the cold self, the final choices could be different from the welfare maximizing choice. This approach is consistent with the view of Bernheim and Rangel (2009) that regards choices made by the hot self as mistakes. If the prescriptive utility is  $u_\theta$ , which corresponds to the *instant* utility of Kahneman et al. (1997), an advertising cue is a “good” when  $u(p) < v(p)$  and a “bad” when  $v(p) < u(p)$  (where  $p$  is the advertised good). This conclusion is shared with the complementary view of advertising (Becker and Murphy 1993). Different from the latter approach, however, in the CT model, cues and the advertised good  $p$  are *not necessarily* complementary. Indeed, they are so when (assuming that partial derivatives exist)  $\partial u_\theta(p)/\partial p \partial \theta \geq 0$ —namely, when  $\phi'(\theta)(v'(p) - u'(p)) \geq 0$ , corresponding to the case of the hot self having a higher marginal utility from  $p$  than the cold self has. Therefore, the CT model allows for a wide variety of substitution patterns between cues and consumption. In conclusion, the CT model offers a flexible tool to perform the prescriptive analysis of advertising that includes existing approaches as special cases.

## 4. Behavioral Characterization of the Cue-Triggered Model

In this section, we provide the testable restrictions characterizing the CT model. As argued in Section 2.1, the standard revealed preferences analysis may fail in the presence of environmental cues. To detect

cue-triggered choices, we adopt a rich set of primitives: a family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  of binary relations on  $X$  representing revealed preferences contingent on the cue's intensity. Choice data in the form of  $p \succsim_{\theta} q$  are observable in a controlled laboratory experiment with subjects deciding while being exposed to different cues (e.g., visual advertisements of different sizes). Contingent choice data can also be collected in the field; for instance, scanner data paired with information concerning the cues at the moment of choice are in the form  $p \succsim_{\theta} q$ .

Given our primitives, we can now state the behavioral conditions characterizing the CT model. The first condition subsumes the requirements that are necessary and sufficient for the existence of an affine representation of preferences on  $X$ . This condition is written directly in term of utilities because the revealed preference conditions are standard.

**Axiom 1** (Conditional Rationality). *The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies conditional rationality if, for each  $\theta \in \Theta$ , there exists a nonconstant and affine  $u_{\theta} : X \rightarrow \mathbb{R}$  such that*

$$p \succsim_{\theta} q \iff u_{\theta}(p) \geq u_{\theta}(q)$$

for all  $p, q \in X$ . Moreover, each  $u_{\theta}$  is cardinally unique; that is, if  $v_{\theta}$  represents  $\succsim_{\theta}$ ,  $u_{\theta} \approx v_{\theta}$ .

By Axiom 1 (conditional rationality), each preference  $\succsim_{\theta}$  has an affine and cardinally unique representation. Equivalently, each conditional preference  $\succsim_{\theta}$  satisfies axioms 1–3 of Herstein and Milnor (1953). However, conditional rationality does not restrict the representations  $u_{\theta}$  across the cue's intensities. This role is played by the next axiom, where  $p \succ_{\theta} q$  is a strict preference ( $p \succsim_{\theta} q$  and  $q \not\sucsim_{\theta} p$ ).

**Axiom 2** (No-cycles). *The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies no-cycles if, for all  $p, q \in X$ ,  $p \succsim_{\theta} q$  for some  $\theta \in \Theta$ , and  $q \succ_{\theta'} p$  for some  $\theta \leq \theta'$  imply  $q \succ_{\theta''} p$  for all  $\theta' \leq \theta''$ .*

The axiom is similar in spirit to the one-switch condition of Bell (1988) as it imposes a single-crossing condition on cue-contingent choices. Suppose we observe a preference reversal between  $p$  and  $q$ , following an increase of the cue's intensity from  $\theta$  to  $\theta'$ . Axiom 2 (no-cycles) rules out the possibility of reversing choice (from  $q$  to  $p$ ) for an even more intense cue  $\theta''$ . Therefore, the preference between two options  $p, q \in X$  can switch at most once as the cue's intensity increases.

The last condition is technical and imposes a minimal amount of heterogeneity in extreme preferences.

**Axiom 3** (Regularity). *The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  is regular if there exist  $p, q, r \in X$  such that  $p \succ_{>0} q \succ_{>0} r$  and  $p \succ_{>1} r \succ_{>1} q$ .*

Axiom 3 (regularity) allows us to identify the function  $\phi$  that transforms the cue into a decision weight.<sup>7</sup> We say that  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  has a CT representation  $(u, v, \phi)$  if each  $\succsim_{\theta}$  is represented by  $u_{\theta}$  and the

family  $\{u_{\theta}\}_{\theta \in \Theta}$  has a CT model  $(u, v, \phi)$  introduced in Definition 1. The next result proves that regularity, no-cycles, and conditional rationality completely characterize the CT representation.

**Theorem 1.** *The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  is regular and satisfies conditional rationality and no-cycles (Axioms 1 and 2) if and only if it has a CT representation  $(u, v, \phi)$ , such that  $u \neq v$  and  $u \neq -v$ . Given  $u$  and  $v$ , the function  $\phi$  is unique, and if  $(u', v', \phi')$  is another CT representation of  $\{\succsim_{\theta}\}_{\theta \in \Theta}$ , then  $u \approx u'$ ,  $v \approx v'$  and  $\phi(\theta) = \phi'(\theta)$  when  $\phi(\theta) \in \{0, 1\}$ , and there is  $k > 0$  such that  $\phi(\theta)/(1 - \phi(\theta)) = k\phi'(\theta)/(1 - \phi'(\theta))$ , when  $\phi(\theta) \in (0, 1)$ .*

The intuition behind the proof of Theorem 1 is the following: if we observe a preference switch as the intensity of the cue varies, we are eliciting the preference of the hot self. No-cycles bounds the number of switches to be 1 at most. Hence it establishes a monotone relationship between the intensity of the cue and the decision weight assigned to the hot self. Note that the no-cycles condition can be vacuously true: if there are no preference switches—namely, for all  $\theta, \theta' \in \Theta$  and all  $p, q \in X$ ,  $p \succsim_{\theta} q$ , and  $p \succsim_{\theta'} q$ —then all  $u_{\theta}$  represent the same preference. In this case, the CT model reduces to a cue-independent choice model,  $u_{\theta} \approx u'_{\theta}$  for all  $\theta, \theta' \in \Theta$ . In the presence of regularity, however, there exists at least one preference switch. It is also worth noting that no-cycles is falsifiable in an experimental setting involving only *three* levels of the cue's intensity  $\Theta = \{0, \theta, 1\}$  and two choice options  $p, q$ . If the experimenter observes two preferences switches—for example,  $p \succ_{>0} q$ ,  $q \succ_{>\theta} p$ , and  $p \succ_{>1} q$ —no-cycles is violated, and the model falsified.

We conclude this section by providing the behavioral characterization of the 0-1 model introduced in Section 2.1. We assume the following stronger version of the no-cycles axiom.

**Axiom 4** (Discontinuous Crossing). *There exists a  $\hat{\theta} \in \Theta$  such that  $\succsim_{\theta} = \succsim_{\theta'}$  for all  $\theta, \theta' \leq \hat{\theta}$  and  $\succsim_{\theta} = \succsim_{\theta'}$  for all  $\hat{\theta} \leq \theta, \theta'$ , with exactly one strict inequality.*

Axiom 4 (discontinuous crossing) implies Axiom 2 (no-cycles): if  $p \succsim_{\theta} q$  and  $q \succ_{\theta'} p$  for some  $\theta \leq \theta'$ , discontinuous crossing implies that  $\theta \leq \hat{\theta} \leq \theta'$  with at least one strict inequality. Therefore,  $\succsim_{\theta'} = \succsim_{\theta'}$  for all  $\theta' \leq \theta''$ , implying  $q \succ_{\theta''} p$  (i.e., no-cycles). The following result is immediate.

**Corollary 1.** *The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies Axiom 1 (conditional rationality) and Axiom 4 (discontinuous crossing) if and only if it has a 0-1 representation.*

Because all preferences indexed by  $\theta \leq \hat{\theta}$  are identical, their respective rationalizations are equivalent and can be normalized (if necessary) to be all equal. The same is true for all  $\hat{\theta} \leq \theta$ . The discontinuous crossing

condition is rather unsatisfactory. It requires the existence of an intensity  $\hat{\theta}$  that determines the transition from the cold to the hot self. Finding such a  $\theta$  through revealed preferences is difficult, especially with finitely many choice data. However, there is a case in which the threshold  $\hat{\theta}$  can be identified by simply observing the preferences of the two selves.

**Definition 5.** A completely conflicting CT representation of a family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  is a 0-1 representation with  $v = -u$ .

The completely conflicting (CC) representation of  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  can be rewritten as

$$u_{\theta} = \begin{cases} u & \theta \leq \hat{\theta}, \\ -u & \hat{\theta} \leq \theta, \end{cases} \quad (\text{CC})$$

with exactly one strict inequality. The next axiom immediately characterizes the conflicting preferences of the selves.

**Axiom 5 (Conflict).** For all  $p, q \in X$ ,  $p \succ_{\theta} q$  (respectively,  $p >_{\theta} q$ ) if and only if  $q \succeq_{\hat{\theta}} p$  (respectively,  $q >_{\hat{\theta}} p$ ).

**Proposition 5.** The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies Axiom 1 (conditional rationality), Axiom 2 (no-cycles), and Axiom 5 (conflict) if and only if it has a CC representation.

For example, a former heroin user may completely reverse her choice to abstain if she encounters a cue reminding her of past consumption of the substance. Different from the 0-1 representation, the axioms characterizing the CC model do not impose the existence of  $\hat{\theta}$ . The latter can be determined *endogenously* by observing revealed preferences of the selves.

As a final remark, the proof of Proposition 5 shows that the monotonicity of the function  $\phi$  is not granted without regularity, meaning that there are representations of  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  that satisfy conditional rationality and no-cycles but with a nonmonotone function transforming the cue into a decision weight. An example is a CC representation with an arbitrary  $\phi$  having the following properties:  $\phi(\theta) < \frac{1}{2}$  for all  $\theta \leq \hat{\theta}$  and  $\phi(\theta) > \frac{1}{2}$  for all  $\hat{\theta} < \theta$ .

## 5. Extensions

In this section we consider three extensions of the CT model. In the first, we study a random version arising from noisy perception of the cue. In the second, we propose a general version of the CT model that allows for more general preferences of the selves. In the third, we allow for context dependency in response to the cue.

### 5.1. Random Cues and Stochastic Choice

The rich primitives we have used so far—revealed choices plus the intensity of the cue at the moment of choice—can be collected in a controlled laboratory experiment. In different contexts, such as in field

studies, the available data are likely to be noisier than those collected through experiments. Indeed, even if the cue's intensity is known to the analyst, the individual may perceive it with some noise. Under such an assumption, the observed choices become stochastic from the analyst's point of view. We prove that the extension of the CT model accounting for a noisy cue's intensity is related to the SCRUM of Apesteguia et al. (2017). The link between the random CT model and the SCRUM turns out to be useful for estimating the effects of interventions involving cues.

To model the random intensity of the cue, we consider a random variable with values in  $\Theta$  and a cumulative distribution function  $F(k) = P(\theta \leq k)$  representing the distribution of the cue's intensity. The random cue-triggered (RCT) choice rule is a family of functions  $\mathbb{P}(\cdot|A) : A \times \mathcal{X} \rightarrow [0, 1]$  with  $\sum_{p \in A} \mathbb{P}(p|A) = 1$  for all  $A \in \mathcal{X}$ , where  $\mathcal{X}$  is the family of finite and nonempty subsets of  $X$  (the menus). The probability of selecting  $p$  from a menu  $A$ , denoted by  $\mathbb{P}(p|A)$ , is

$$\begin{aligned} \mathbb{P}(p|A) &= P\left(\theta \in \Theta : (1 - \phi(\theta))u(p) + \phi(\theta)v(p) \right. \\ &\quad \left. \geq \max_{q \in A} (1 - \phi(\theta))u(q) + \phi(\theta)v(q)\right). \end{aligned} \quad (\text{RCT})$$

More compactly,  $\mathbb{P}(p|A) = P(\theta \in \Theta : u_{\theta}(p) \geq u_{\theta}(q), \forall q \in A)$ . Because we allow for multiple maximizing elements in a set, we consider the following tie-breaking rule: if for some  $\theta \in \Theta$ ,  $|\arg \max_{p \in A} u_{\theta}(p)| = n \geq 1$  and  $P(\theta) > 0$ , then  $\mathbb{P}(p|A) = \frac{1}{n}P(\theta)$ . When menus are binary or ternary, we write  $\mathbb{P}(p|p, q) \equiv \mathbb{P}(p|\{p, q\})$  and  $\mathbb{P}(p|p, q, r) \equiv \mathbb{P}(p|\{p, q, r\})$ . For an illustrative example, we assume that  $\phi$  is strictly monotone, and we consider the probability of selecting  $p$  from a binary menu  $A = \{p, q\}$ , assuming that  $u(p) > u(q)$  and  $v(p) < v(q)$ :

$$\mathbb{P}(p|p, q) = P\left(\theta \in \Theta : \theta \geq \phi^{-1}\left(\frac{k_{p,q}}{1 + k_{p,q}}\right)\right),$$

where  $k_{p,q} \equiv (u(p) - u(q))/(v(q) - v(p)) > 0$ . The probability of choosing  $p$  over  $q$  is equal to the probability that the realization of the cue's intensity  $\theta$  is larger than  $\phi^{-1}(k_{p,q}/(1 + k_{p,q}))$ .

The RCT model includes as particular cases the following models that are known in the literature. The first is the random quasi-hyperbolic discounting (RQD) model (e.g., Duflo et al. 2011), which corresponds to the random version of the  $\beta\theta$ - $\delta$  model discussed in Section 3.1:

$$\begin{aligned} \mathbb{P}(c|A) &= P\left(\theta \in \Theta : U(c_0) + \beta_{\theta} \sum_{t=1}^T \delta^t U(c_t) \right. \\ &\quad \left. \geq \max_{c' \in A} U(c'_0) + \beta_{\theta} \sum_{t=1}^T \delta^t U(c'_t)\right). \end{aligned}$$

The second case is based on the random version of the cue-triggered mean-variance model introduced in Section 3.3 and corresponds to the random mean-variance model of Roberts and Urban (1988):

$$\mathbb{P}(p|A) = P\left(\theta \in \Theta : \bar{p} - ((1 - \phi(\theta))\gamma_u + \phi(\theta)\gamma_v)\sigma^2(p) \geq \max_{q \in A} \bar{q} - ((1 - \phi(\theta))\gamma_u + \phi(\theta)\gamma_v)\sigma^2(q)\right).$$

We consider the following two properties of the choice probabilities that are adapted from Apesteguia et al. (2017) to the present setting.

**Property 1** (Monotonicity). *If  $A \subseteq B$ , then  $\mathbb{P}(p|A) \geq \mathbb{P}(p|B)$*

**Property 2** (Centrality). *If  $p \succ_0 q \succ_0 r$  and  $\mathbb{P}(q|p, q, r) > 0$ , then  $\mathbb{P}(p|p, q, r) = \mathbb{P}(p|p, q)$  and  $\mathbb{P}(r|p, q, r) = \mathbb{P}(r|q, r)$ .*

We have the following result.

**Theorem 2.** *If  $\{\succ_\theta\}_{\theta \in \Theta}$  is a family of strict linear orders,<sup>8</sup> then the RCT model satisfies monotonicity and centrality.*

Example 3 in the appendix shows why the result requires  $\succ_\theta$  to be strict linear orders. In the remaining part of this section, we will use monotonicity and centrality to evaluate the effects of policy interventions in the presence of environmental cues. We call *menu interventions* the actions changing the choice set by removing or adding options. We call *nudging* the actions changing the cues' intensity but not the choice set (in the spirit of Thaler and Sunstein (2009)).

Consider a menu intervention. By centrality, the probability of selecting an option  $p$  from a menu  $A$  depends on two elements of  $A$ , at most: the one immediately dominating  $p$  (if any) and the one immediately dominated by  $p$  (if any) in the order generated by  $\succ_0$ . Consider, for example, a menu intervention designed by a store manager to increase the sales of an energy drink. Suppose that a shelf (a menu  $A = \{p, q, r\}$ ) contains the energy drink  $p$ , a traditional caffeinated drink  $q$ , and instant coffee  $r$ . The intervention consists in removing the traditional caffeinated drink from the shelf. Suppose that  $p \succ_0 q$ : according to centrality, the effect of removing  $q$  from  $A$  on the probability of selecting the new energy drink depends on the position of  $r$  (instant coffee) in the ranking generated by the cold self. If  $p \succ_0 q \succ_0 r$ , removing the “middle” option will *increase*, by monotonicity, the probability of buying the energy drink and will *increase* the probability of buying instant coffee, hence obtaining a partially successful intervention. On the contrary, if  $p \succ_0 r \succ_0 q$ , removing the traditional caffeinated drink has *no effect* on the probability of the new energy drink  $p$  being selected and will increase the consumption of instant coffee,

hence resulting in a clear failure of the intervention. This analysis follows from the following fact.

**Fact 1.** *If  $\mathbb{P}$  satisfies centrality and  $\mathbb{P}(q|A) > 0$ , then  $\mathbb{P}(p|A) = \mathbb{P}(p|A \setminus q)$  if and only if there exists  $r \in A$  such that  $p \succ_0 r \succ_0 q$  or  $q \succ_0 r \succ_0 p$ .*

In general, removing an option  $p$  affects the probability of selecting an option  $q$  only if there are no intermediate elements between  $p$  and  $q$  (in the order  $\succ_0$ ). This is because the transition from the cold to the hot self in the RCT is smooth. Removing an extreme option has no effect on the probability of choosing options that are distant in the ranking generated by the cold self. The opposite argument holds for adding options. Suppose that  $p \succ_0 q$  are smartphones with  $p$  having less storage than  $q$ . Suppose that the seller wants to foster buying  $q$  by adding a third type of smartphone,  $r$ , that has more storage than  $q$ , hence exploiting a compromise effect: the tendency of an option to be selected more often when it is a compromise between two alternatives (Simonson 1989). In the RCT model, adding an extreme option may leave the probability of choosing the middle option  $q$  unchanged (if  $p \succ_0 r \succ_0 q$ ) and may reduce the probability of choosing  $q$  (by monotonicity), which is the opposite of the desired effect.

We now turn to nudging. These interventions leave the menu unchanged but decrease or increase the exposure to cues, for instance, by implementing an advertising campaign. Additional examples of nudging interventions are banning advertising of junk food (Dubois et al. 2017) or introducing graphic images on cigarette packages (McCool et al. 2012). Note that in the RCT model, the probability of selecting an option  $q$  in a menu  $A$ , when strictly positive, can be written as (assuming for simplicity that  $\phi$  is strictly monotone)

$$\mathbb{P}(q|A) = P(\theta \in \Theta : \theta_q \leq \theta \leq \theta^q),$$

where  $\theta_q \equiv \phi^{-1}(k_{p,q}/(1+k_{p,q}))$  and  $p = \max\{r \in A : r \succ_0 q\}$  and  $\theta^q \equiv \phi^{-1}(k_{p',q}/(1+k_{p',q}))$  and  $p' = \min\{r \in A : q \succ_0 r\}$ . The probability of selecting  $q$  is equal to the probability that the realization of  $\theta$  is larger than  $\theta_q$  and smaller than  $\theta^q$ , leading to the following inequality:

$$\mathbb{P}(q|A) \leq \frac{\mathbb{E}_p[\theta]}{\theta_q} \quad (4)$$

that follows from  $\mathbb{E}_p[\theta] = \int_0^1 \theta dP(\theta) \geq \int_{\theta_q}^{\theta^q} \theta dP(\theta) \geq \theta_q \int_{\theta_q}^{\theta^q} dP(\theta) = \theta_q \mathbb{P}(q|A)$ . The upper bound in Equation (4) is binding if  $\mathbb{E}_p[\theta] < \theta_q$ . Because  $\theta_q$  is independent of the distribution of  $\theta$ , the inequality (4) offers an estimate for the effect of modifying the distribution of the cue's intensity. Consider a managerial intervention aimed at fostering cooperation among workers in an organization—for example, by employing an identity-enforcing cue. After the intervention, the cue

is uniformly more intense than before, meaning that the distribution of the cue's intensity first-order stochastic dominates the preintervention distribution (i.e.,  $P(\theta \leq k) \leq P'(\theta \leq k)$ ). Suppose that  $u(p) > u(q)$  and  $v(q) > v(p)$ ; then  $\mathbb{P}(q|p, q) \leq \mathbb{P}'(q|p, q) \leq \mathbb{E}_{P'}[\theta]/\theta_q$ . Uniformly increasing the level of the cue increases the probability of selecting the options that are optimal for the hot self  $v$  (e.g., cooperation in the social dilemma). However, the effect can be bounded if the threshold  $\theta_q$  is large enough. Hence nudging can have a limited effect. A large  $\theta_q$  means that the difference in the utility of  $p$  and  $q$  for the hot self is small compared with the same difference for the cold self. In this case, the preference switch from  $p$  to  $q$  happens for very intense cues, and the effect of nudging is limited.

We conclude the section noting that both menu and nudging interventions can be used to affect choice probabilities. The RCT model provides clear predictions in both cases; hence it is able to guide marketing strategists and managers in the choice of the most effective intervention.

## 5.2. General Cue-Triggered Model

In this section, we characterize a generalization of the CT model in which the selves are “less rational.” In the CT model, the assumption that the preferences of the selves are represented by affine utilities may be unduly strong. For example, when choice is over lotteries (i.e.  $X = \Delta(Z)$ ), such an assumption implies that the cue-contingent preferences are represented by expected utilities, à la von Neumann and Morgenstern (1953). To allow for a more general preference structure, we consider the following weakening of the conditional rationality axiom.

**Axiom 6** (Conditional Rationality\*). *The family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies conditional rationality\* if, for all  $\theta \in \Theta$ ,  $\succsim_{\theta}$  is represented by a nonconstant function  $w_{\theta} : X \rightarrow \mathbb{R}$ .*

Different from conditional rationality, no assumptions are made concerning the structure of the utilities  $w_{\theta}$ . The axiom is equivalent to assuming that each  $\succsim_{\theta}$  is a complete, nontrivial, and transitive preference and that there exists, for each  $\theta$ , a countable order-dense subset of  $X$ . Namely, there is a set  $Y \subseteq X$  that may depend on  $\theta$  such that for all  $p, q \in X$  with  $p \succ_{\theta} q$ , there is a  $r \in Y$  with  $p \succ_{\theta} r \succ_{\theta} q$ . We have the following result, in which we denote by  $f(X) \subseteq \mathbb{R}$  the image of a function  $f : X \rightarrow \mathbb{R}$ .

**Theorem 3.** *If the family  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies Axiom 6 (conditional rationality\*) and Axiom 2 (no-cycles), then there are  $\tilde{u}, \tilde{v} : X \rightarrow \mathbb{R}$  such that for each  $\theta \in \Theta$ ,*

$$w_{\theta} = F(\tilde{u}, \tilde{v}, \theta),$$

where  $F : \tilde{u}(X) \times \tilde{v}(X) \times \Theta \rightarrow \mathbb{R}$  is weakly monotone in the first two arguments. Moreover,  $F(\tilde{u}, \tilde{v}, 0) = \tilde{u}$  and  $F(\tilde{u}, \tilde{v}, 1) = \tilde{v}$ .

As for uniqueness, if  $F', \tilde{u}', \tilde{v}'$  also represents  $\{\succsim_{\theta}\}_{\theta \in \Theta}$ , there is a monotone increasing function  $G$  such that  $F'(\tilde{u}', \tilde{v}', \theta) = G(F(\tilde{u}, \tilde{v}, \theta))$ . Hence we can only identify preferences up to ordinal transformations.

Some remarks following Theorem 3 are in order. First, the CT model corresponds to  $F(r, s, \theta) = (1 - \phi(\theta))r + \phi(\theta)s$ . Second, the no-cycles axiom (weakly) characterizes the dual-self model also in the case of general preferences. Third, as in the CT model, the preferences of the selves are identified at the boundary of the cue's intensity spectrum. Finally, cue-independent choices correspond to the case of the selves having ordinally equivalent preferences (i.e., when there is a strictly monotonic function  $g$  such that  $\tilde{u} = g(\tilde{v})$ ). Despite its generality, the model in Proposition 3 can be still used to analyze the welfare analysis of advertising (as we did in Section 3.4) and to study the complementary properties of cues and consumption (by studying the cross-derivatives of  $F$  when they exist).

## 5.3. Context-Dependent Cue-Triggered Choice

In the last section, we discuss an extension of the CT model that allows for context dependence of the effect of cues on choices. This is needed to model the case of responses to environmental cues that are affected by the availability of various options, or the lack thereof. For instance, a cue triggering the desire to smoke may be weaker or stronger when the smoking option is unavailable (e.g., a smoking cue while on board a plane). To model context dependence, we consider the menus introduced in Section 5.1—namely, the finite and nonempty subsets of  $X$ . The following menu-dependent cue-triggered model generalizes the CT model by allowing for a menu-dependent function  $\phi$ :

$$u_{A, \theta}(p) = (1 - \phi_A(\theta))u(p) + \phi_A(\theta)v(p) \quad (5)$$

with  $\phi_A(0) = 0$  and  $\phi_A(1) = 1$  for all  $A \in \mathcal{X}$ . To gain intuition on the scope of model (5), consider  $p, q, r$  with  $u(r) = u(q)$  and  $v(q) > v(r)$ : the hot self prefers  $q$  to  $p$ , whereas the cold self is indifferent. Suppose that  $\phi_{\{p, q\}}(\theta) \leq \phi_{\{p, q, r\}}(\theta)$  for all  $\theta \in \Theta$ : when the  $r$  is available, the effect of environmental cues is uniformly stronger. The menu-dependent CT model allows for the following behavior:  $u_{\{p, q\}, \theta}(p) > u_{\{p, q\}, \theta}(q)$  and  $u_{\{p, q, r\}, \theta}(q) > u_{\{p, q, r\}, \theta}(p)$ . When  $r$  is not available, the individual selects  $p$  over  $q$  for a given cue's intensity  $\theta$ . However, when the  $r$  is available and for the same cue's intensity  $\theta$ , she may reverse her preferences, choosing  $q$  over  $p$ . The simple presence of the  $r$  induces a differential response to the same environmental cue and generates a behavior akin to the asymmetric dominance effect (being  $r$  dominated by  $q$  for all cue-contingent utilities). The context-dependent model (5) can be used to study the unexplored interplay of context dependency and environmental cues.

It is worth noting that also in the context-dependent CT model (5), for a fixed menu, the no-cycles axiom is satisfied, suggesting once again that no-cycles is the condition characterizing dual-self models of cue-triggered choices.

## 6. Conclusion

We proposed and characterized a dual-self model of cue-triggered behavior that accommodates the large empirical evidence on the effect of cues across choice domains. The mechanism linking cues to choice in the CT model is simple and informed by neuroscientific evidence. Beyond its descriptive power, our model makes new predictions on the effect of common policies. It also clarifies under what conditions empirical evidence allows us to ascertain when two individuals have differential responsiveness to cues. When applied to advertising, our model goes in the direction outlined by Bagwell (2007), who suggests the use of a dual-self approach to model individuals' response to advertising cues.

In general, the present work represents a first step toward the development of a comprehensive approach to cue-triggered behavior. By embedding our model in a dynamic setting, future research may capture additional features of cue-triggered behavior such as cue management or learning. For instance, cue management refers to the active choice of how much exposition to cue an individual is willing to accept. Cue management is treated in the consumption-based models of Laibson (2001) and Bernheim and Rangel (2009), but its revealed preference characterization is still missing. By providing a first foundation for a model of cue-triggered behavior, we believe that our approach represents a building block for future developments in the field.

## Acknowledgments

The author thanks the department editor, one associate editor, and two referees for exceptional comments. The author thanks Leonardo Boncinelli and Francesco Cerigioni for providing invaluable suggestions. The author also thanks seminar participants at Bologna and at the XIII GRASS in Milan.

## Appendix A. Proofs

The next proposition is from Fishburn (1984) and is used repeatedly in the proofs.

**Proposition A1.** *Suppose  $X$  is a convex space and  $f_1, f_2, g : X \rightarrow \mathbb{R}$  are affine; then we have the following.*

(1) *For all  $p, q \in X$ ,  $f_1(p) \geq f_1(q)$  and  $f_2(p) \geq f_2(q)$  imply  $g(p) \geq g(q)$  if and only if there exist  $a, b \geq 0$  and  $c \in \mathbb{R}$  such that  $g = af_1 + bf_2 + c$ .*

(2) *For all  $p, q \in X$ ,  $f_1(p) = f_1(q)$  and  $f_2(p) = f_2(q)$  imply  $g(p) = g(q)$  if and only if there exist  $a, b, c \in \mathbb{R}$  such that  $g = af_1 + bf_2 + c$ .*

**Proof of Proposition 1.** By part (1),  $u^A$  and  $u^B$  represent the same preferences, and the same is true for  $v^A$  and  $v^B$ . By the

cardinal uniqueness property of the CT model, we can normalize, if necessary, the utilities to have  $u^A = u^B$  and  $v^A = v^B$ , and we write  $u \equiv u^A = u^B$  and  $v \equiv v^A = v^B$ . By part (2), for all  $p, q \in X$ , if  $u(p) \geq u(q)$  and  $u_\theta^B(p) \geq u_\theta^B(q)$  for some  $\theta \in (0, 1]$ , then  $u_\theta^A(p) \geq u_\theta^A(q)$ . Then, by Proposition A1 there are  $\hat{a}(\theta)$ ,  $\hat{b}(\theta) \geq 0$  and  $\hat{c}(\theta) \in \mathbb{R}$  such that

$$u_\theta^A(p) = \hat{a}(\theta)u(p) + \hat{b}(\theta)u_\theta^B + \hat{c}(\theta).$$

We can divide both sides by  $z_\theta \equiv \hat{a}(\theta) + \hat{b}(\theta) > 0$  (because  $u_\theta^A$  is nontrivial), so  $(1/z_\theta)u_\theta^A(p) = a(\theta)u(p) + (1 - a(\theta))u_\theta^B + c(\theta)$ , where  $a(\theta) = \hat{a}(\theta)/z_\theta$  and  $c(\theta) = \hat{c}(\theta)/z_\theta$ . By definition,

$$\begin{aligned} & \frac{1}{z_\theta}(\phi^A(\theta)(v(p) - u(p)) + u(p)) \\ &= a(\theta)u(p) + (1 - a(\theta))(\phi^B(\theta)(v(p) - u(p)) + u(p)) \\ & \quad + c(\theta), \end{aligned}$$

or equivalently,

$$\begin{aligned} \phi^A(\theta)(v(p) - u(p)) &= z_\theta a(\theta)u(p) + z_\theta(1 - a(\theta)) \\ & \quad \times (\phi^B(\theta)(v(p) - u(p)) + u(p)) \quad (\text{A.1}) \\ & \quad - u(p) + z_\theta c(\theta) \end{aligned}$$

$$\begin{aligned} &= z_\theta(1 - a(\theta))\phi^B(\theta)(v(p) - u(p)) \quad (\text{A.2}) \\ & \quad + z_\theta u(p) - u(p) + z_\theta c(\theta) \end{aligned}$$

for all  $p, q \in X$ . Then, by nontriviality, there is a  $p \in X$  with  $v(p) \neq u(p)$ ; hence,

$$\begin{aligned} \phi^A(\theta) &= z_\theta(1 - a(\theta))\phi^B(\theta) + z_\theta\phi^B(\theta) \\ & \quad + \frac{z_\theta u(p) - u(p) + z_\theta c(\theta)}{v(p) - u(p)}. \end{aligned}$$

Because this equality is true for all  $p, q \in X$  with  $u(p) \neq v(p)$ , and the equality in Equation (A.1) is true for all  $p, q \in X$ , it must be that  $z_\theta = 1$  and  $c(\theta) = 0$ . So  $\phi^A(\theta) = (1 - a(\theta))\phi^B(\theta)$ , implying  $\phi^B \geq \phi^A$ .

For the opposite implication, we can normalize, if necessary, the utilities to have  $u^A = u^B$  and  $v^A = v^B$ , and we write  $u \equiv u^A = u^B$  and  $v \equiv v^A = v^B$ . Suppose that  $u(p) \geq u(q)$  and  $(1 - \phi^B(\theta))(u(p) - u(q)) + \phi^B(\theta)(v(p) - v(q)) \geq 0$ . There are two cases:  $v(p) - v(q) \geq 0$  and  $v(p) - v(q) < 0$ . In the first case, the conclusion  $u_\theta^A(p) \geq u_\theta^A(q)$  is trivial. In the second case,  $\phi^B \geq \phi^A$  and  $u(p) \geq u(q)$  imply  $(1 - \phi^A(\theta))(u(p) - u(q)) + \phi^B(\theta)(v(p) - v(q)) \geq (1 - \phi^B(\theta))(u(p) - u(q)) + \phi^B(\theta)(v(p) - v(q)) \geq 0$ . Moreover,  $v(p) - v(q) < 0$  and  $\phi^B \geq \phi^A$  imply  $(1 - \phi^A(\theta))(u(p) - u(q)) + \phi^A(\theta)(v(p) - v(q)) \geq (1 - \phi^A(\theta))(u(p) - u(q)) + \phi^B(\theta)(v(p) - v(q)) \geq 0$ —hence the conclusion  $u_\theta^A(p) \geq u_\theta^A(q)$ .

**Proof of Proposition 2.** The first inequality is standard. To see the second inequality,  $I_\theta(t)$  can be written as  $I_\theta(t) = \alpha_\theta \delta^t + (1 - \alpha_\theta)\beta^t$ , where  $\alpha_\theta = (1 - \phi(\theta))/(1 - \phi(\theta))\delta^{t+1} + \phi(\theta)\beta^{t+1}$ . As  $\theta$  increases to  $\theta'$ ,  $\alpha_\theta$  decreases. In the case of  $\phi(\theta) = 1$ , the result is trivial. Suppose that  $\phi(\theta) < 1$ :

$$\begin{aligned} \alpha_\theta &= \frac{1 - \phi(\theta)}{(1 - \phi(\theta))\delta^{t+1} + \phi(\theta)\beta^{t+1}} = \frac{1}{\delta^{t+1} + \frac{\phi(\theta)}{1 - \phi(\theta)}\beta^{t+1}} \\ &\leq \frac{1}{\delta^{t+1} + \frac{\phi(\theta')}{1 - \phi(\theta')}\beta^{t+1}} = \alpha_{\theta'}. \end{aligned}$$

**Proof of Proposition 3.** It follows from the fact that the expression  $(\partial u_{\theta}^i / \partial p_j) / (\partial u_{\theta}^i / \partial p_i) = (\phi^i(\theta) v_j^i) / ((1 - \phi^i(\theta)) u_i^i + \phi^i(\theta) v_i^i)$ , where  $v_k^i = \partial v_i / \partial p_k$  for  $k = i, j$  and  $u_i^i = \partial u_i / \partial p_i$ , is increasing in  $\theta$ . The proof of this result is similar to that of Proposition 4.

**Proof of the Result in Example 2.** Consider the CT model when  $(p_1, \dots, p_N) \in \mathbb{R}_+^N$  and  $u_{\theta}^i(p) = (1 - \phi^i(\theta)) p_i + \phi^i(\theta) \sum_{j \neq i} p_j$ . The utility of contributing  $c_i$  for individual  $i$  is  $u_{\theta}^i = (1 - \phi^i(\theta))(e - c_i + r \sum_j c_j) + \phi^i(\theta) \sum_{j \neq i} m(c_j)$ . By deriving  $u_{\theta}^i$  with respect to  $c_i$ , we obtain  $\partial u_{\theta}^i / \partial c_i = (1 - \phi^i(\theta))(r - 1) + \phi^i(\theta) \times (r(N - 1))$ . The previous expression is strictly greater than 0 if  $\phi^i(\theta) / (1 - \phi^i(\theta)) > (1 - r) / r(N - 1)$ .

**Proof of Proposition 4.** Suppose, by contradiction, that  $u$  is more risk averse than  $v$  (hence  $u(p) \geq u(\delta_z)$  implies  $v(p) \geq v(\delta_z)$ ) and there exist  $p, \delta_z \in \Delta(Z)$  such that  $u_{\theta}(p) \geq u_{\theta}(\delta_z)$  and  $u_{\theta'}(\delta_z) > u_{\theta'}(p)$  for some  $\theta < \theta'$ . By no-cycles,  $v(\delta_z) > v(p)$  and also  $u(p) \geq u(\delta_z)$  (otherwise,  $u(\delta_z) > u(p)$ ,  $u_{\theta}(p) \geq u_{\theta}(\delta_z)$ , and  $u_{\theta'}(\delta_z) > u_{\theta'}(p)$  for some  $\theta < \theta'$  would contradict the weak monotonicity of  $\phi(\theta)$ ). Therefore,  $u$  is not more risk averse than  $v$  as assumed, which is a contradiction to the initial assumption. The case of  $v$  being more risk averse than  $u$  is similar. If  $U, V$  are twice differentiable, the risk aversion index  $r_{\theta}$  can be rewritten as (omitting the argument)  $r_{\theta} = \alpha_{\theta} r_U + (1 - \alpha_{\theta}) r_V$ , a convex combination of the absolute risk aversion indexes of the first self,  $r_U$ , and the second self,  $r_V$ , with weight  $\alpha_{\theta} = ((1 - \phi(\theta))U') / ((1 - \phi(\theta))U' + \phi(\theta)V')$ . The fact that  $\alpha_{\theta}$  is decreasing can be proved as before. If  $\phi(\theta) = 1$ , the result is trivial. Suppose that  $\phi(\theta) < 1$ ; then

$$\alpha_{\theta} = \frac{U'}{U' + \frac{\phi(\theta)}{1-\phi(\theta)} V'} \leq \frac{U'}{U' + \frac{\phi(\theta')}{1-\phi(\theta')} V'} = \alpha_{\theta'}$$

where the inequality follows from  $U', V' > 0$ .

**Proof of Theorem 1.** We prove necessity first. Conditional rationality holds because, for each  $\theta \in \Theta$ ,  $u_{\theta}(p) = (1 - \phi(\theta))u(p) + \phi(\theta)v(p)$  is a nonconstant affine utility rationalizing  $\succsim_{\theta}$ . Regularity follows from the fact that both  $u \neq v$  and  $u \neq -v$ . Indeed, because  $u \neq v$ , there are  $p', q' \in X$  satisfying  $u(p') = u(q')$  and  $v(q') > v(p')$ . Moreover, because  $u \neq -v$ , there are  $p'', q'' \in X$  such that  $u(p'') > u(q'')$  and  $v(p'') > v(q'')$ . Let  $p = (1 - \epsilon)q' + \epsilon p''$ ,  $r = (1 - \epsilon)q' + \epsilon q''$ , and  $q = (1 - \epsilon)p' + \epsilon[\frac{1}{2}p'' + \frac{1}{2}q'']$ . Then,  $u(p) = (1 - \epsilon)u(q') + \epsilon u(p'') > (1 - \epsilon)u(p') + \epsilon[\frac{1}{2}u(p'') + \frac{1}{2}u(q'')] = u(q) > (1 - \epsilon)u(q') + \epsilon u(q'') = u(r)$  and  $v(p) = (1 - \epsilon)v(q') + \epsilon v(p'') > (1 - \epsilon)v(q') + \epsilon v(q'') = v(r) > (1 - \epsilon)v(p') + \epsilon[\frac{1}{2}v(p'') + \frac{1}{2}v(q'')] = v(q)$ , where the last inequality holds for a sufficiently small  $\epsilon$ . Therefore, there are  $p, q, r \in X$  such that  $u(p) > u(q) > u(r)$  and  $v(p) > v(r) > v(q)$ .

To see no-cycles, if  $p \succsim_{\theta} q$ , then  $(1 - \phi(\theta))u(p) + \phi(\theta)v(p) \geq (1 - \phi(\theta))u(q) + \phi(\theta)v(q)$ , whereas  $q \succ_{\theta'} p$  for some  $\theta \leq \theta'$  implies  $(1 - \phi(\theta'))u(q) + \phi(\theta')v(q) > (1 - \phi(\theta'))u(p) + \phi(\theta')v(p)$ . There are various cases to check:

- Suppose that  $0 = \phi(\theta) = \phi(\theta')$ , but this implies at the same time that  $u(p) \geq u(q)$  and  $u(q) > u(p)$ , a contradiction. A similar reasoning holds for  $1 = \phi(\theta) = \phi(\theta')$ .

- Suppose that  $0 = \phi(\theta) < \phi(\theta') < 1$ , then  $0 \leq u(p) - u(q) < \phi(\theta') / (1 - \phi(\theta'))(v(q) - v(p))$ , and the fact that  $\phi(\theta') \in (0, 1)$  implies that  $v(q) - v(p) > 0$ ; then for all  $\theta' < \theta''$ ,  $\phi(\theta') / (1 - \phi(\theta')) \leq \phi(\theta'') / (1 - \phi(\theta''))$ . Hence  $0 \leq u(p) - u(q) < \phi(\theta'') / (1 - \phi(\theta''))(v(q) - v(p))$  and  $u_{\theta'}(q) > u_{\theta'}(p)$  for all  $\theta' \leq \theta''$ .

- Suppose that  $0 = \phi(\theta) < \phi(\theta') = 1$ , then  $u(p) \geq u(q)$  and  $v(q) > v(p)$ . Because  $\phi$  is weakly increasing,  $\phi(\theta'') = 1$  for all  $\theta' \leq \theta''$ ; then  $v(q) > v(p)$  implies  $u_{\theta'}(q) > u_{\theta'}(p)$  for all  $\theta' \leq \theta''$ .

- Suppose that  $0 < \phi(\theta) < \phi(\theta') = 1$ , then  $\phi(\theta) / (1 - \phi(\theta)) (u(q) - u(p)) \leq v(p) - v(q)$  and  $v(q) > v(p)$  imply that  $\phi(\theta) / (1 - \phi(\theta)) (u(q) - u(p)) \leq v(p) - v(q) < 0$ ; then  $u(p) > u(q)$ . Because  $\phi$  is weakly increasing,  $\phi(\theta'') = 1$  for all  $\theta' \leq \theta''$ ; then  $v(q) > v(p)$  implies  $u_{\theta'}(q) > u_{\theta'}(p)$  for all  $\theta' \leq \theta''$ .

- Suppose that  $0 < \phi(\theta) < \phi(\theta') < 1$ , then  $(1 - \phi(\theta))u(p) + \phi(\theta)v(p) \geq (1 - \phi(\theta))u(q) + \phi(\theta)v(q)$  and  $(1 - \phi(\theta'))u(q) + \phi(\theta')v(q) > (1 - \phi(\theta'))u(p) + \phi(\theta')v(p)$  imply that  $\phi(\theta) / (1 - \phi(\theta)) (v(q) - v(p)) \leq u(p) - u(q) < \phi(\theta') / (1 - \phi(\theta')) (v(q) - v(p))$ . Because  $0 < \phi(\theta) < \phi(\theta') < 1$ ,  $v(q) \neq v(p)$ . Because  $\phi$  is weakly increasing,  $\phi(\theta) / (1 - \phi(\theta)) \leq \phi(\theta') / (1 - \phi(\theta'))$  implies  $v(q) > v(p)$  and  $u(p) - u(q) < \phi(\theta') / (1 - \phi(\theta')) (v(q) - v(p))$  for all  $\theta' \leq \theta''$ ; then  $u_{\theta'}(q) > u_{\theta'}(p)$  for all  $\theta' \leq \theta''$ .

We now turn to sufficiency. By conditional rationality, for each  $\theta \in \Theta$ ,  $\succsim_{\theta}$  is represented by a nonconstant and affine utility  $u_{\theta} : X \rightarrow \mathbb{R}$ . We now show that for all  $p, q \in X$ , if  $u_{\theta}(p) \geq u_{\theta}(q)$  and  $u_1(p) \geq u_1(q)$ , then  $u_{\theta}(p) \geq u_{\theta}(q)$  for all  $0 \leq \theta \leq 1$ . Suppose not, then, there is  $\theta \in \Theta$  such that  $u_{\theta}(p) \geq u_{\theta}(q)$ ,  $u_1(p) \geq u_1(q)$  and  $u_{\theta}(q) > u_{\theta}(p)$ . This implies  $p \succ_{\theta} q$ ,  $q \succ_1 p$ , and  $p \succ_1 q$ —a contradiction to the no-cycles condition. By Proposition A1 there exist scalars  $a_{\theta}, b_{\theta} \geq 0$  and  $c_{\theta} \in \mathbb{R}$  such that  $u_{\theta} = a_{\theta}u_0 + b_{\theta}u_1 + c_{\theta}$ . Fix an arbitrary  $\bar{q} \in X$  and normalize  $u_{\theta}(\bar{q}) = 0$  for all  $\theta \in \Theta$ . Then  $0 = u_{\theta}(\bar{q}) = a_{\theta}u_0(\bar{q}) + b_{\theta}u_1(\bar{q}) + c_{\theta} = c(\theta)$ ; hence  $c(\theta) = 0$  for all  $\theta \in \Theta$ . We can also divide both sides by  $a_{\theta} + b_{\theta}$  (by non-triviality,  $a_{\theta} + b_{\theta} > 0$  for all  $\theta \in \Theta$ ) to have  $u_{\theta} \approx (1 - \phi(\theta))u_0 + \phi(\theta)u_1$  with  $\phi(\theta) \equiv b_{\theta} / (a_{\theta} + b_{\theta})$ . Clearly,  $\phi(0) = 0$  and  $\phi(1) = 1$ . Let us define  $u, v : X \rightarrow \mathbb{R}$  as  $u \equiv u_0$  and  $v \equiv u_1$ .

We need to prove that  $\phi : \Theta \rightarrow [0, 1]$  is weakly increasing in the order  $\leq$ . We will use the following result.

**Lemma 2.** *If  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies Axiom 1 (conditional rationality) and Axiom 3 (regularity), there exist  $p, q, p', q' \in X$  such that  $p \succ_0 q$  and  $p \sim_1 q$  and  $p' \succ_1 q'$  and  $p' \sim_0 q'$ .*

**Proof of Lemma 2.** By regularity, there are  $p, q, r \in X$  with  $p \succ_0 q \succ_0 r$  and  $p \succ_1 r \succ_1 q$ . By conditional rationality, there is  $\alpha_0 \in (0, 1)$  such that  $u_0(q) = \alpha_0 u_0(p) + (1 - \alpha_0)u_0(r) = u_0(\alpha_0 p + (1 - \alpha_0)r)$  (or  $q \sim_0 \alpha_0 p + (1 - \alpha_0)r$ ). By conditional rationality,  $u_1(\alpha_0 p + (1 - \alpha_0)r) > u_1(q)$ ; hence  $\alpha_0 p + (1 - \alpha_0)r \succ_1 q$ . For the remaining condition, by conditional rationality, there is  $\alpha_1 \in (0, 1)$  such that  $u_1(r) = \alpha_1 u_1(p) + (1 - \alpha_1)u_1(q) = u_1(\alpha_1 p + (1 - \alpha_1)r)$  (or  $r \sim_1 \alpha_1 p + (1 - \alpha_1)r$ ). By conditional rationality,  $u_0(\alpha_1 p + (1 - \alpha_1)r) > u_0(r)$ ; hence  $\alpha_1 p + (1 - \alpha_1)r \succ_0 r$ .

The condition in Lemma 2 is known in the literature on preferences aggregation as the independent prospects condition (see, e.g., Weymark 1993).

**Lemma 3.** *Suppose that  $\{\succsim_{\theta}\}_{\theta \in \Theta}$  satisfies Axiom 1 (conditional rationality) and Axiom 2 (no-cycles), if it satisfies Axiom 3*

(regularity), then for all  $\theta \in \Theta$ , there exist  $p_\theta, q_\theta \in X$  such that  $p_\theta \sim_\theta q_\theta$  and  $p_\theta \succ_{\theta} q_\theta$ ,  $q_\theta \succ_{\theta} p_\theta$  with at least one strict preference.

**Proof of Lemma 3.** We generalize an argument of Borgers and Choo (2017). By conditional rationality and regularity, Lemma 2 implies that there are  $p, q, p', q'$  such that  $p \succ_{\theta} q$ ,  $p \sim_1 q$ , and  $p' \succ_1 q'$ , and  $p' \sim_{\theta} q'$ . For  $\alpha \in [0, \frac{1}{4}]$ , let

$$p(\alpha) \equiv \left(\frac{1}{2} - \alpha\right)p + \alpha q + \left(\frac{1}{4} - \alpha\right)p' + \left(\frac{1}{4} + \alpha\right)q',$$

$$q(\alpha) \equiv \alpha p + \left(\frac{1}{2} - \alpha\right)q + \left(\frac{1}{4} + \alpha\right)p' + \left(\frac{1}{4} - \alpha\right)q'.$$

Then, for  $\alpha = 0$ ,  $u(p(0)) = \frac{1}{2}u(p) + \frac{1}{4}u(p') + \frac{1}{4}u(q')$  and  $u(q(0)) = \frac{1}{2}u(q) + \frac{1}{4}u(p') + \frac{1}{4}u(q')$ . Because  $u(p) > u(q)$ ,  $u(p(0)) > u(q(0))$ . Moreover,  $v(p(0)) = \frac{1}{2}v(p) + \frac{1}{4}v(p') + \frac{1}{4}v(q')$  and  $v(q(0)) = \frac{1}{2}v(q) + \frac{1}{4}v(p') + \frac{1}{4}v(q')$ . Because  $v(p) = v(q)$ ,  $v(p(0)) = v(q(0))$ . By no-cycles, for all  $\theta \in \Theta \setminus \{0, 1\}$ ,  $u_\theta(p(0)) \geq u_\theta(q(0))$  (otherwise, we would have  $p(0) \succ_{\theta} q(0)$ ,  $q(0) \succ_{\theta} p(0)$ , and  $p(0) \succ_1 q(0)$ , a contradiction to no-cycles). For  $\alpha = \frac{1}{4}$ ,  $v(p(\frac{1}{4})) = \frac{1}{4}v(p) + \frac{1}{4}v(q) + \frac{1}{2}v(q')$  and  $v(q(\frac{1}{4})) = \frac{1}{4}v(p) + \frac{1}{4}v(q) + \frac{1}{2}v(p')$ . Because  $v(q') > v(p')$ ,  $v(q(\frac{1}{4})) > v(p(\frac{1}{4}))$ . Moreover,  $u(p(\frac{1}{4})) = u(q(\frac{1}{4}))$ . By no-cycles, for all  $\theta \in \Theta \setminus \{0, 1\}$ ,  $u_\theta(p(\frac{1}{4})) \geq u_\theta(q(\frac{1}{4}))$  (otherwise, we would have  $p(\frac{1}{4}) \succ_{\theta} q(\frac{1}{4})$ ,  $p(\frac{1}{4}) \succ_{\theta} q(\frac{1}{4})$ , and  $q(\frac{1}{4}) \succ_1 p(\frac{1}{4})$ , a contradiction to no-cycles). By continuity, for all  $\theta \in \Theta$ , there exists  $\alpha_\theta \in [0, \frac{1}{4}]$  such that  $p(\alpha_\theta) \sim_\theta q(\alpha_\theta)$ . Let  $p_\theta = p(\alpha_\theta)$  and  $q_\theta = q(\alpha_\theta)$ ; then  $p_\theta \sim_\theta q_\theta$  (by definition). Moreover, for all  $\alpha_\theta \in (0, \frac{1}{4})$ ,  $p_\theta \succ_{\theta} q_\theta$  and  $q_\theta \succ_{\theta} p_\theta$ . If  $\alpha_\theta = 0$ , then  $u(p_\theta) > u(q_\theta)$  and  $v(q_\theta) = v(p_\theta)$ . If  $\alpha_\theta = \frac{1}{4}$ ,  $v(q_\theta) > v(p_\theta)$  and  $u(p_\theta) = u(q_\theta)$ .

To conclude the proof of Theorem 1, it remains to be shown that  $\phi$  is weakly monotone. Take  $p, q \in X$  with  $p \succ_{\theta} q$  and  $q \succ_1 p$  (they exist by regularity) with at least one strict preference, and define  $D_{p,q}(\theta) = u_\theta(p) - u_\theta(q) = (1 - \phi(\theta))[u(p) - u(q)] + \phi(\theta)[v(p) - v(q)]$ . Suppose, by contrapositive, that there exists  $\theta' < \theta$  such that  $1 - \phi(\theta) > 1 - \phi(\theta') \geq 0$  (implying  $\phi(\theta) < 1$ ). Because  $p, q \in X$  with  $p \succ_{\theta} q$  and  $q \succ_1 p$  with at least one strict preference,

$$D_{p,q}(\theta) > D_{p,q}(\theta'). \quad (\text{A.3})$$

To see inequality (A.3), consider three cases: (1)  $u(p) = u(q)$ ,  $v(q) > v(p)$ ; (2)  $u(p) > u(q)$ ,  $v(q) = v(p)$ ; and (3)  $u(p) > u(q)$ ,  $v(q) > v(p)$ .

If  $u(p) = u(q)$  and  $v(q) > v(p)$ ,  $D_{p,q}(\theta) = \phi(\theta)(v(p) - v(q)) > \phi(\theta')(v(p) - v(q)) = D_{p,q}(\theta')$ .

If  $u(p) > u(q)$  and  $v(q) = v(p)$ ,  $D_{p,q}(\theta) = (1 - \phi(\theta))(u(p) - u(q)) > (1 - \phi(\theta'))(u(p) - u(q)) = D_{p,q}(\theta')$ .

If  $u(p) > u(q)$  and  $v(q) > v(p)$ ,  $D_{p,q}(\theta) = (1 - \phi(\theta))(u(p) - u(q)) + \phi(\theta)(v(p) - v(q)) > (1 - \phi(\theta'))(u(p) - u(q)) + \phi(\theta')(v(p) - v(q)) = D_{p,q}(\theta')$ .

Now consider  $p_\theta, q_\theta$  defined in Lemma 3. We show that inequality (A.3) leads to a contradiction in all three cases  $\alpha_\theta = 0$ ,  $\alpha_\theta = \frac{1}{4}$  and  $\alpha_\theta \in (0, \frac{1}{4})$ :

- $\alpha_\theta = 0$ . Then  $p_\theta = p(\alpha_\theta) = p(0)$  and  $q_\theta = q(\alpha_\theta) = q(0)$ . By definition,  $p(0) \succ_{\theta} q(0)$ ,  $p(0) \sim_{\theta} q(0)$ , and  $p(0) \sim_1 q(0)$ . Therefore  $D_{p_\theta, q_\theta}(\theta) = (1 - \phi(\theta))[u(p(0)) - u(q(0))] + \phi(\theta)[v(p(0)) - v(q(0))] = 0$ . Because  $u(p(0)) - u(q(0)) > 0$  and  $v(p(0)) - v(q(0)) = 0$ ,  $0 = D_{p_\theta, q_\theta}(\theta) = (1 - \phi(\theta))[u(p(0)) - u(q(0))]$  implies  $\phi(\theta) = 1$ —a contradiction to  $1 - \phi(\theta) > 1 - \phi(\theta') \geq 0$ .

- $\alpha_\theta = \frac{1}{4}$ . Then  $p_\theta = p(\alpha_\theta) = p(\frac{1}{4})$  and  $q_\theta = q(\alpha_\theta) = q(\frac{1}{4})$ . By definition,  $p(\frac{1}{4}) \sim_{\theta} q(\frac{1}{4})$ ,  $p(\frac{1}{4}) \sim_{\theta} q(\frac{1}{4})$ , and  $q(\frac{1}{4}) \succ_1 p(\frac{1}{4})$ . Therefore

$D_{p_\theta, q_\theta}(\theta) = (1 - \phi(\theta))[u(p(\frac{1}{4})) - u(q(\frac{1}{4}))] + \phi(\theta)[v(p(\frac{1}{4})) - v(q(\frac{1}{4}))] = 0$ . Because  $u(p(\frac{1}{4})) - u(q(\frac{1}{4})) = 0$  and  $v(p(\frac{1}{4})) - v(q(\frac{1}{4})) < 0$ ,  $0 = D_{p_\theta, q_\theta}(\theta) = \phi(\theta)[v(p(\frac{1}{4})) - v(q(\frac{1}{4}))] = 0$  implies  $\phi(\theta) = 0$ . Because  $\phi(\theta') > \phi(\theta)$ ,  $D_{p_\theta, q_\theta}(\theta') = \phi(\theta')[v(p(\frac{1}{4})) - v(q(\frac{1}{4}))] < 0$ ; hence  $u_{\theta'}(q(\frac{1}{4})) > u_{\theta'}(p(\frac{1}{4}))$ —a contradiction to no-cycles, because we have  $p(\frac{1}{4}) \sim_{\theta} q(\frac{1}{4})$ ,  $q(\frac{1}{4}) \succ_{\theta'} p(\frac{1}{4})$ , and  $p(\frac{1}{4}) \sim_{\theta} q(\frac{1}{4})$  with  $\theta' < \theta$ .

- $\alpha_\theta \in (0, \frac{1}{4})$ . Then we have  $u_\theta(p_\theta) - u_\theta(q_\theta) = 0$ , or

$$D_{p_\theta, q_\theta}(\theta) = u_\theta(p_\theta) - u_\theta(q_\theta) = (1 - \phi(\theta))[u(p_\theta) - u(q_\theta)] + \phi(\theta)[v(p_\theta) - v(q_\theta)] = 0;$$

by inequality (A.3),

$$0 = D_{p_\theta, q_\theta}(\theta) > D_{p_\theta, q_\theta}(\theta') = u_{\theta'}(p_\theta) - u_{\theta'}(q_\theta).$$

Then we have  $p_\theta \succ_{\theta} q_\theta$ ,  $q_\theta \succ_{\theta'} p_\theta$ , and  $p_\theta \sim_{\theta} q_\theta$  for some  $\theta' < \theta$ —a contradiction to no-cycles.

Therefore,  $\phi(\theta) \geq \phi(\theta')$ . This concludes the proof of the main part of Theorem 1.

For the uniqueness part, given the utilities  $u, v : X \rightarrow \mathbb{R}$ , if  $\phi(\theta) \in \{0, 1\}$ , then it must be unique. Consider the cases  $\theta \in \Theta$  with  $\phi(\theta) \in (0, 1)$ . Then, the following is true: for all  $p, q \in X$ ,  $u(p) \geq u(q)$  and  $v(p) \geq v(q)$  with at least one strict inequality imply  $u_\theta(p) > u_\theta(q)$ . This condition plus regularity implies that the  $\phi(\theta)$  are unique, by Fishburn (1984, corollary 1). For the second part, suppose that  $(u_2, v_2, \phi_2)$  is another representation; then  $u_2 = \alpha_u u_2 + \beta_u$  and  $v_2 = \alpha_v v_2 + \beta_v$  for some  $\alpha_u, \alpha_v > 0$  and  $\beta_u, \beta_v \in \mathbb{R}$ . By Lemma 3, there are  $p_\theta, q_\theta \in X$  such that  $u_1(p_\theta) \geq u_1(q_\theta)$ ,  $v_1(q_\theta) \geq v_1(p_\theta)$  with at least one strict inequality, and  $(1 - \phi_1(\theta))[u_1(p_\theta) - u_1(q_\theta)] = \phi_1(\theta)[v_1(q_\theta) - v_1(p_\theta)]$  for  $\theta \in \Theta \setminus \{0, 1\}$ . If either  $u_1(p_\theta) = u_1(q_\theta)$  or  $v_1(q_\theta) = v_1(p_\theta)$ , then  $\phi_1(\theta) = 1$  or  $\phi_1(\theta) = 0$ . Similarly,  $u_2(p_\theta) = u_2(q_\theta)$  or  $v_2(q_\theta) = v_2(p_\theta)$  implies that  $\phi_2(\theta) = 1$  or  $\phi_2(\theta) = 0$ . If both  $u_1(p_\theta) > u_1(q_\theta)$  and  $v_1(q_\theta) > v_1(p_\theta)$ , then  $\phi_1(\theta) \in (0, 1)$ . In this case,  $(1 - \phi_2(\theta))[u_2(p_\theta) - u_2(q_\theta)] = \phi_2(\theta)[v_2(q_\theta) - v_2(p_\theta)]$  implies  $(1 - \phi_2(\theta))\alpha_u[u_1(p_\theta) - u_1(q_\theta)] = \phi_2(\theta)\alpha_v[v_1(q_\theta) - v_1(p_\theta)]$ . Then  $(\phi_2(\theta)/(1 - \phi_2(\theta)))(\alpha_v/\alpha_u) = \phi_1(\theta)/(1 - \phi_1(\theta))$ . Because it must be true for all  $\theta \in \Theta \setminus \{0, 1\}$ , the result holds with  $k = \alpha_v/\alpha_u$ .

**Proof of Proposition 5.** We first prove sufficiency: by conditional rationality, no-cycles, and an application of Proposition A1 for each  $\theta \in \Theta$ , there is  $\phi(\theta) \in [0, 1]$  such that  $u_\theta \approx (1 - \phi(\theta))u + \phi(\theta)v$ . Because  $u = -v$ , we have  $u_\theta \approx (1 - 2\phi(\theta))u$  for all  $\theta \in \Theta$ . Because  $u_\theta$  is nontrivial,  $\phi(\theta) \neq \frac{1}{2}$ . Let  $\hat{\theta} = \inf\{\theta \in \Theta : 1 - 2\phi(\theta) < 0\}$ . If  $\phi(\theta) > \frac{1}{2}$ , no-cycles implies that  $\phi(\theta') > \frac{1}{2}$  for all  $\theta \leq \theta'$ . Indeed, suppose that  $\phi(\theta) > \frac{1}{2}$ , and for some  $\theta \leq \theta'$ ,  $\phi(\theta') < \frac{1}{2}$ . Then, there are, by nontriviality,  $p, q \in X$  for which  $u(p) > u(q)$ ; hence  $u_\theta(q) \approx -u(q) > -u(p) \approx u_\theta(p)$  and  $u_{\theta'}(p) \approx u(p) > u(q) \approx u_{\theta'}(q)$ —a contradiction to no-cycles. The symmetric property follows immediately:  $\phi(\theta) < \frac{1}{2}$  implies that  $\phi(\theta') < \frac{1}{2}$  for all  $\theta' \leq \theta$ . Therefore, there exists a unique  $\hat{\theta}$  such that for all  $\hat{\theta} \leq \theta$ ,  $u_\theta \approx -u$ , and for all  $\theta < \hat{\theta}$ ,  $u_\theta \approx u$ .

For necessity, if  $\{\succ_{\theta}\}_{\theta \in \Theta}$  has a CC representation, let  $\phi(\theta) = 0$  if  $\theta \leq \hat{\theta}$  and let  $\phi(\theta) = 1$  if  $\hat{\theta} \leq \theta$  with exactly one strict inequality that (exists and) depends on the representation. Conditional rationality follows immediately. No-cycles follows from the monotonicity of the function  $\phi$  defined previously.



**Proof of Theorem 2.** Monotonicity follows immediately. For centrality, consider  $p \succ_0 q \succ_0 r$  with  $\mathbb{P}(q|p, q, r) > 0$ . Let  $\theta_q = \min_{\theta \in \Theta} \{q \succ_\theta p, r\}$ , which is well defined because  $\mathbb{P}(q|p, q, r) > 0$ , and  $\succ_\theta$  is a strict order. Because  $p \succ_0 q$  and  $q \succ_{\theta_q} p$ , no-cycles implies  $q \succ_\theta p$  for all  $\theta_q < \theta$ . Moreover, the definition of  $\theta_q$  and  $q \succ_0 r$  imply  $q \succ_\theta r$  for all  $\theta < \theta_q$ . Therefore,  $\mathbb{P}(p|p, q) = P(\theta \in \Theta : \theta < \theta_q) = \mathbb{P}(p|p, q, r)$ . A similar argument holds for  $\mathbb{P}(r|p, q, r)$ .

**Example 3.** Take  $\Theta = \{0, \frac{1}{2}, 1\}$  and  $p \succ_0 q \succ_0 r$  with  $u(p) = 1$ ,  $u(q) = 0.5$ , and  $u(r) = 0$ . Assume  $v(p) = 0$ ,  $v(q) = 0.5$ , and  $v(r) = 1$ ; then if  $P(\theta) = \frac{1}{3}$  for all  $\theta \in \Theta$  and  $\phi(\theta) = 1 - \theta$ ,  $\mathbb{P}(p|p, q) = \frac{1}{2}$ , but  $\mathbb{P}(p|p, q, r) = \frac{4}{9}$ , a violation of centrality. To see this, by the tiebreaking rule,  $\mathbb{P}(p|p, q) = P(\theta \in \Theta : u_\theta(p) \geq u_\theta(q)) = P(0) + 0.5P(0.5) = \frac{1}{2}$ . But  $\mathbb{P}(p|p, q, r) = P(0) + \frac{1}{3}P(0.5)$ .

**Proof of Fact 1.** Consider  $A = \{p, q, r\}$  with  $\mathbb{P}(p|p, q, r) = \mathbb{P}(p|p, r)$ , and suppose without loss of generality that  $p \succ_0 q$ . Then if  $r \succ_0 p \succ_0 q$ ,  $\mathbb{P}(p|p, q, r) = 1 - \mathbb{P}(q|p, q, r) - \mathbb{P}(r|p, q, r) > 1 - \mathbb{P}(r|p, q, r)$ . By centrality,  $1 - \mathbb{P}(r|p, q, r) = 1 - \mathbb{P}(r|p, r) = \mathbb{P}(p|p, r)$ ; hence  $\mathbb{P}(p|p, q, r) > \mathbb{P}(p|p, r)$ —a contradiction to the initial hypothesis. If  $p \succ_0 q \succ_0 r$ ,  $\mathbb{P}(p|p, q, r) = 1 - \mathbb{P}(q|p, q, r) - \mathbb{P}(r|p, q, r) > 1 - \mathbb{P}(r|p, q, r) = \mathbb{P}(q|p, q, r) + \mathbb{P}(p|p, q, r)$ . By centrality,  $\mathbb{P}(q|p, q, r) + \mathbb{P}(p|p, q, r) = \mathbb{P}(p|p, q) + \mathbb{P}(q|p, q, r) \geq \mathbb{P}(p|p, q)$ , which is a contradiction to the initial hypothesis.

**Proof of Theorem 3.** The proof is a simple adaptation of the technique developed by Abbas and Bell (2015, theorem 7). By Axiom 6 (conditional rationality\*), for each  $\theta$ , there exists  $w_\theta : X \rightarrow \mathbb{R}$  such that  $w_\theta(p) \geq w_\theta(q)$  if and only if  $p \succ_\theta q$ . If for some  $p, q \in X$ ,  $w_0(p) = w_0(q)$ , and  $w_1(p) = w_1(q)$ , then  $w_\theta(p) = w_\theta(q)$  for all  $\theta \in \Theta$ . Suppose not, then, there exists  $\theta \in \Theta$  such that  $w_\theta(p) > w_\theta(q)$  or  $w_\theta(q) > w_\theta(p)$ . In both cases, no-cycles is violated. Let  $w_\theta(p) \equiv F(w_0(p), w_1(p), \theta)$ . The representation is well defined because  $w_0(p) = w_0(q)$  and  $w_1(p) = w_1(q)$  imply  $w_\theta(p) = w_\theta(q)$  for all  $\theta \in \Theta$ . If no-cycles holds,  $\theta < \theta''$  and  $w_\theta(p) \geq w_\theta(q)$  and  $w_{\theta''}(p) \geq w_{\theta''}(q)$  imply  $w_{\theta'}(p) \geq w_{\theta'}(q)$  for all  $\theta < \theta' < \theta''$ . Suppose not, then, there exists  $\theta < \theta' < \theta''$  such that  $w_{\theta'}(q) > w_{\theta'}(p)$ ; then  $p \succ_{\theta'} q$ ,  $q \succ_{\theta'} p$ , and  $p \succ_{\theta'} q$ , a contradiction to no-cycles. Hence the representation is weakly monotone in the first two arguments. By defining  $\tilde{u} \equiv w_0$  and  $\tilde{v} \equiv w_1$ , we have the result.

## Endnotes

<sup>1</sup> Such a neuroscientific interpretation is indirectly supported by some empirical evidence (e.g., Berridge and Aldridge 2009). For specific choice domains, the results of McClure et al. (2004) and Knutson and Knutson (2005) are consistent with the neuroscientific interpretation for intertemporal and risky choice, respectively. The works of Bernheim and Rangel (2004) and Bechara (2005) contain an exhaustive neuroscientific foundation for cue-triggered addiction.

<sup>2</sup> Given two elements  $p, q \in X$ , we denote by  $\alpha p + (1 - \alpha)q$  their convex combination with weight  $\alpha \in [0, 1]$ . We say that a function  $f : X \rightarrow \mathbb{R}$  is affine if  $f(\alpha p + (1 - \alpha)q) = \alpha f(p) + (1 - \alpha)f(q)$  for all  $p, q \in X$  and all  $\alpha \in [0, 1]$ .

<sup>3</sup> See Knutson et al. (2008) and Kim and Zauberman (2013) for evidence on risk and impatience, respectively, and see Reichert (2002) for a review of sex in advertising.

<sup>4</sup> The use of dual-self models in advertising is advocated in Bagwell (2007, section 10.2). He suggests the use of dual-self models as a potential way to describe the mechanism regarding the relationship

between advertising cues and behavior. This aspect is often left unmodelled in the advertising literature.

<sup>5</sup> In Bernheim and Rangel (2009), the exposure to a cue is random and depends on past behavior, a lifestyle, and an exogenous component. In our model cues are contingent and cannot be controlled by the individual. However, in Section 5.1 we discuss random cues.

<sup>6</sup> Because each  $C$  is convex,  $X$  is also a convex set. For any two elements  $c, c' \in X$ , their convex combination with weight  $\gamma \in (0, 1)$ ,  $\gamma c + (1 - \gamma)c' \in X$  is defined componentwise,  $\gamma c + (1 - \gamma)c' = (\gamma c_0 + (1 - \gamma)c'_0, \gamma c_1 + (1 - \gamma)c'_1, \dots, \gamma c_T + (1 - \gamma)c'_T)$ . In the case of  $C = [0, M]$ , an affine utility over  $C$  must be linear (i.e.,  $U(c_i) = ac_i + b$  for some  $a, b \in \mathbb{R}$ ). In this case, we can interpret the streams  $c = (c_0, c_1, \dots, c_T)$  directly as streams of utils.

<sup>7</sup> Regularity guarantees the monotonicity and uniqueness of the function  $\phi$ . For monotonicity, suppose that  $\succ_0$  and  $\succ_1$  are identical. In this case, choices are cue independent, and any function  $\phi$  would rationalize the cue-contingent preferences. A similar consideration is valid if  $\succ_0$  and  $\succ_1$  are completely opposed (see Equation (CC)), as we discuss after Proposition 5. For uniqueness, we show that regularity and conditional rationality imply (see Lemma 3) a condition known in the literature on preferences aggregation as independent prospects (see, e.g., Weymark 1993), which guarantees the uniqueness of the weights assigned to the selves.

<sup>8</sup> A strict linear order is a binary relation  $>$  on  $X$  that is irreflexive ( $x \not> x$ ), transitive ( $x > y$  and  $y > z$  implies  $x > z$ ), and complete (either  $x > y$  or  $y > x$  or  $x = y$ ).

## References

- Abbas AE, Bell DE (2015) Ordinal one-switch utility functions. *Oper. Res.* 63(6):1411–1419.
- Abratt R, Goodey SD (1990) Unplanned buying and in-store stimuli in supermarkets. *Management Decision Econom.* 11(2):111–121.
- Akerlof GA, Kranton RE (2005) Identity and the economics of organizations. *J. Econom. Perspect.* 19(1):9–32.
- Andreoni J, Rao JM, Trachtman H (2017) Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *J. Political Econom.* 125(3):625–653.
- Apestegua J, Ballester MA, Lu J (2017) Single-crossing random utility models. *Econometrica* 85(2):661–674.
- Ashforth BE, Mael F (1989) Social identity theory and the organization. *Acad. Management Rev.* 14(1):20–39.
- Bagwell K (2007) The economic analysis of advertising. Armstrong M, Porter R, eds. *Handbook of Industrial Organization*, Vol. 3 (North-Holland, Amsterdam), 1701–1844.
- Bassi A, Colacito R, Fulghieri P (2013) 'O sole mio: An experimental analysis of weather and risk attitudes in financial decisions. *Rev. Financial Stud.* 26(7):1824–1852.
- Bearden WO, Shimp TA (1982) The use of extrinsic cues to facilitate product adoption. *J. Marketing Res.* 19(2):229–239.
- Bechara A (2005) Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat. Neurosci.* 8(11):1458–1463.
- Becker GS, Murphy KM (1993) A simple theory of advertising as a good or bad. *Quart. J. Econom.* 108(4):941–964.
- Bell DE (1988) One-switch utility functions and a measure of risk. *Management Sci.* 34(12):1416–1424.
- Bennett R (2009) Impulsive donation decisions during online browsing of charity websites. *J. Consumer Behav.* 8(2–3):116–134.
- Bernheim BD, Rangel A (2004) Addiction and cue-triggered decision processes. *Amer. Econom. Rev.* 94(5):1558–1590.
- Bernheim BD, Rangel A (2009) Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quart. J. Econom.* 124(1):51–104.

- Berridge KC, Aldridge JW (2009) Decision utility, incentive salience, and cue-triggered “wanting”. *Oxford Ser. Soc. Cognition Soc. Neurosci.* 2009:509–533.
- Blanco M, Engelmann D, Normann HT (2011) A within-subject analysis of other-regarding preferences. *Games Econom. Behav.* 72(2):321–338.
- Börgers T, Choo Y-M (2017) *Revealed relative utilitarianism*. Working paper, University of Michigan, Ann Arbor.
- Braithwaite D (1928) The economic effects of advertisement. *Econom. J.* 38(149):16–37.
- Card D, Dahl GB (2011) Family violence and football: The effect of unexpected emotional cues on violent behavior. *Quart. J. Econom.* 126(1):103–143.
- Charness G, Holder P (2019) Charity in the laboratory: Matching, competition, and group identity. *Management Sci.* 65(3):1398–1407.
- Charness G, Rabin M (2002) Understanding social preferences with simple tests. *Quart. J. Econom.* 117(3):817–869.
- Cox JC, Sadiraj V (2007) On modeling voluntary contributions to public goods. *Public Finance Rev.* 35(2):311–332.
- Cox JC, Friedman D, Sadiraj V (2008) Revealed altruism. *Econometrica* 76(1):31–69.
- Dubois P, Griffith R, O’Connell M (2017) 04. The effects of banning advertising in junk food markets. *Rev. Econom. Stud.* 85(1):396–436.
- Duflo E, Kremer M, Robinson J (2011) Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *Amer. Econom. Rev.* 101(6):2350–2390.
- Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quart. J. Econom.* 114(3):817–868.
- Fishburn PC (1984) On Harsanyi’s utilitarian cardinal welfare theorem. *Theory Decision* 17(1):21–28.
- Fudenberg D, Levine DK (2006) A dual-self model of impulse control. *Amer. Econom. Rev.* 96(5):1449–1476.
- Goette L, Huffman D, Meier S, Sutter M (2012) Competition between organizational groups: Its impact on altruistic and antisocial motivations. *Management Sci.* 58(5):948–960.
- Guiso L, Sapienza P, Zingales L (2018) Time varying risk aversion. *J. Financial Econom.* 128(3):403–421.
- Gul F, Pesendorfer W (2001) Temptation and self-control. *Econometrica* 69(6):1403–1435.
- Halevy Y (2008) Strotz meets Allais: Diminishing impatience and the certainty effect. *Amer. Econom. Rev.* 98(3):1145–1162.
- Herstein IN, Milnor J (1953) An axiomatic approach to measurable utility. *Econometrica* 21(2):291–297.
- Herzenstein M, Posavac SS, Brakus JJ (2007) Adoption of new and really new products: The effects of self-regulation systems and risk salience. *J. Marketing Res.* 44(2):251–260.
- Kahneman D, Wakker PP, Sarin R (1997) Back to Bentham? Explorations of experienced utility. *Quart. J. Econom.* 112(2):375–405.
- Kaldor N (1950) The economic aspects of advertising. *Rev. Econom. Stud.* 18(1):1–27.
- Kessler JB, Milkman KL (2018) Identity in charitable giving. *Management Sci.* 64(2):845–859.
- Kim BK, Zauberman G (2013) Can Victoria’s Secret change the future? A subjective time perception account of sexual-cue effects on impatience. *J. Experiment. Psych. General* 142(2):328–335.
- Knutson B, Wimmer GE, Kuhnen CM, Winkielman P (2008) Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *Neuroreport* 19(5):509–513.
- Krishna A (2012) An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *J. Consumer Psych.* 22(3):332–351.
- Kuhnen CM, Knutson B (2005) The neural basis of financial risk taking. *Neuron* 47(5):763–770.
- Laibson D (1997) Golden eggs and hyperbolic discounting. *Quart. J. Econom.* 112(2):443–477.
- Laibson D (2001) A cue-theory of consumption. *Quart. J. Econom.* 116(1):81–119.
- Loewenstein G, O’Donoghue T (2004) Animal spirits: Affective and deliberative processes in economic behavior. Working paper, Carnegie Mellon University, Pittsburgh.
- McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306(5695):503–507.
- McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD (2007) Time discounting for primary rewards. *J. Neurosci.* 27(21):5796–5804.
- McCool J, Webb L, Cameron LD, Hoek J (2012) Graphic warning labels on plain cigarette packs: Will they make a difference to adolescents? *Soc. Sci. Medicine* 74(8):1269–1273.
- Milliman RE (1982) Using background music to affect the behavior of supermarket shoppers. *J. Marketing* 46(3):86–91.
- Pollay RW, Siddarth S, Siegel M, Haddix A, Merritt RK, Giovino GA, Eriksen MP (1996) The last straw? Cigarette advertising and realized market shares among youths and adults, 1979–1993. *J. Marketing* 60(2):1–16.
- Reichert T (2002) Sex in advertising research: A review of content, effects, and functions of sexual information in consumer advertising. *Annual Rev. Sex Res.* 13(1):241–273.
- Roberts JH, Urban GL (1988) Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Sci.* 34(2):167–185.
- Simonson I (1989) Choice based on reasons: The case of attraction and compromise effects. *J. Consumer Res.* 16(2):158–174.
- Small DA, Loewenstein G, Slovic P (2007) Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organ. Behav. Human Decision Processes* 102(2):143–153.
- Spangenberg ER, Crowley AE, Henderson PW (1996) Improving the store environment: Do olfactory cues affect evaluations and behaviors? *J. Marketing* 60(2):67–80.
- Thaler RH, Shefrin HM (1981) An economic theory of self-control. *J. Political Econom.* 89(2):392–406.
- Thaler RH, Sunstein CR (2009) *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Penguin Books, London).
- Van den Bergh B, Dewitte S, Warlop L (2008) Bikinis instigate generalized impatience in intertemporal choice. *J. Consumer Res.* 35(1):85–97.
- von Neumann J, Morgenstern O (1953) *Theory of games and economic behavior*. (Princeton University Press, Princeton, NJ).
- Weymark JA (1993) Harsanyi’s social aggregation theorem and the weak Pareto principle. *Soc. Choice Welfare* 10(3):209–221.