Description-dependent Choices

Niels Boissonnet* Dino Borie† Dorian Jullien‡

Abstract

We propose a model of framing effects in which a choice reversal is due to differences in the descriptions of the same alternatives. Our contribution is to show under which conditions these framing effects can be seen as rational or not. Our results provide explanations that are aligned with existing empirical evidence not accounted for by other models and suggest experiments for further work.

JEL Classification: D89, D90, D91.

Keywords: Choice correspondence, framing effects, rational choice, description invariance, description-dependence.

*BIGSEM and IMW, Bielefeld Universität. E-mail:niels.boissonnet@uni-bielefeld.de

†Nantes Université, LEMNA, Nantes, France. E-mail:dino.borie@univ-nantes.fr.

‡Université Paris 1 Panthéon-Sorbonne, CNRS, Centre d'Économie de la Sorbonne, Paris, France. E-mail:dorian.jullien@univ-paris1.fr.

We thank Ariel Rubinstein, Efe Ok, Christian List, Peter Hammond, Massimo Marinacci, Éric Danan and especially Raphaël Giraud for their comments on early versions of the paper.

People's choices are influenced not only by their consequences but also by how these consequences are described. It has been observed in experiments that subjects are likelier to purchase a given yogurt when it is labeled as "80% fat-free" than when it is labeled as "20% fat" (Janiszewski et al., 2003); that they exhibit greater altruism when sharing "losses" than when sharing "gains", even though both conditions are equivalent in consequence (Antinyan et al., 2024); or that they are more productive when a bonus can be "lost" than when it can be "gained" (Hong et al., 2015). The most famous demonstration of such *framing effects* has been demonstrated for behavior under risk. In the unusual disease experiment from Tversky and Kahneman (1981), subjects are told that an unusual disease is about to kill 600 people if nothing is done. They are then asked to choose between two programs to combat the disease. In one frame, the programs present the number of lives saved:

> p: "200 people will be saved"
>
> q: "1/3 probability that 600 people will be saved and 2/3 probability that no people will be saved"

In the other frame, the (same) programs present the number of lives lost:

> p': "400 people will die"
>
> q': "1/3 probability that nobody will die and 2/3 probability that 600 people will die"

People tend to choose the sure program in the first frame ($p \succ q$) and the risky one in the second frame ($q' \succ p'$). In other words, the decision makers (DMs henceforth)'s preferences—e.g. their risk attitudes—depend on how the consequences of their choices are described. Tversky and Kahneman (1981) explain this behavior through prospect theory by postulating that subjects change their reference point: in the "saved" frame, they perceive the consequences as gains (implying risk-aversion in prospect theory) and in the "die" frame they perceive them as losses (implying risk-seeking). However, many psychologists have noted that prospect theory cannot account for regularities observed with descriptions such as "400 people will not be saved" or "200 people will not die" that push subjects to be risk seekers when only

the word "saved" is used or risk averse when only the word "die" is used (Kühberger, 1997; Takemura, 2014; Seta et al., 2017; Broniatowski and Reyna, 2018).

We argue that, to provide a unifying theoretical account of the regularities observed in framing experiments, we should revisit the axiom of rational choice theory called "Description Invariance" (DI). Indeed, in economics, framing effects have been interpreted as irrational behavior that violate this implicit axiom (Arrow, 1982, pp.6-8; Quiggin, 2014, p.713; Camerer, 1995, p.652; Bardsley et al., 2010, pp.130-1). Economic models of framing effects that have made DI explicit have the same explanatory limits as prospect theory regarding the full range of regularities observed using the unusual disease experiment (see, e.g., Dietrich and List, 2016; Giraud, 2005). More generally (i.e., beyond models that account for the original unusual disease experiment), the literature focuses on what we will call "External Description Invariance" (EDI), which emphasizes cross-menu consistency (e.g., making consistent choices between $\{p, q\}$ and $\{p', q'\}$. This paper introduces Internal Description Invariance (IDI) as an overlooked yet crucial aspect of DI, focusing on within-menu consistency (e.g., making consistent choices within $\{p, q, p', q'\}$). By integrating EDI and IDI, the paper (1) provides a comprehensive set-up to make DI fully explicit, (2) connects it to the Weak Axiom of Revealed Preference (WARP), and (3) proposes a weaker version of DI compatible with (3.a) prospect theory's explanation, (3.b) the full range of regularities observed in the unusual disease experiments and, more surprisingly, (3.c) the maximization of an expected utility. Therefore, our contribution can be seen as a formal input into the growing body of literature aiming to reconcile sensitivity to descriptions with rationality (see, e.g., Keren, 2011; Sher and McKenzie, 2011; Bermúdez, 2022).

For the sake of clarity, we shall illustrate our concepts and definitions with the unusual disease experiment. However, the reach of our contribution is more general because the structure of the unusual disease experiment is central in the experimental literature on framing effects. Indeed, it has been replicated many times and shown to be robust to changes in the scenario or the type of consequences (e.g., money versus lives); and one result of the most cited review and meta-analysis on framing effects in psychology is that "the likelihood of obtaining choice reversals [is] directly

related to the similarity between features of a given study and features of Tversky and Kahneman's (1981) original "[Unusual] disease problem."" (Levin et al., 1998, p.157).

The paper is organized as follows. Firstly, we make the axiom of DI and its behavioral implications formally explicit. Secondly, we show why violating this axiom can be considered irrational. Thirdly, we provide a simple weakening of the axiom under which behavior observed in the unusual disease experiment is compatible with the maximization of an expected utility. Fourthly, we discuss some empirical implications of our set-up. We then conclude by discussing how our contributions relate to other models of framing effects.

1 Description dependent choice

Whether in controlled laboratory environments or in the real-world, the consequences of a risky choice can always be presented through different descriptions. Thus, it is natural to conceptualize the alternatives under risk as *described lotteries*, denoted by $p \in L \subseteq \Delta(X \times D)$, where X represents a non-empty finite set of consequences with at least two elements, D a finite set of descriptions and $\Delta(X \times D)$ the set of lotteries over $X \times D$.[1] The class of all nonempty finite subsets of L, denoted by $\mathcal{L}$, is the set of menus of described lotteries. For each $(x, d) \in X \times D$, the notation $\delta_{(x,d)}$ stands for the lottery that assigns probability 1 to (x, d).

Each described lottery $p \in L$ generates a (standard) lottery over consequences, denoted by $\sigma^p \in \Delta(X)$, which gives the probability that p assigns to each consequence $x \in X$ regardless of its description, i.e., $\sigma^p(x) = \sum_{d \in D} p(x, d)$. Similarly, for all $Y \in \mathcal{L}$ the set $\sigma^Y = \{\sigma^p : p \in Y\}$ defines the set of lotteries generated from the described lotteries contained in Y. Two described lotteries $p, q \in L$ such that $\sigma^p = \sigma^q$ are said to be *consequentially equivalent*, and so are menus $T, T' \in \mathcal{L}$ such that $\sigma^T = \sigma^{T'}$. A choice correspondence $c : \mathcal{L} \to \mathcal{L}$ assigns to each menu $T \in \mathcal{L}$ the

[1] The definition of L restricts the variability of descriptions to the consequences of the lotteries, assuming that the distribution part of the lottery itself cannot be described in multiple ways. This assumption aligns with the experimental literature, where framing effects are not induced by different descriptions of the probabilities.

set of chosen described lotteries from T. We use the notation $\sigma^{c(T)}$ to denote the set of lotteries generated by the chosen described lotteries in T.

Example: Unusual Disease. To express the pattern of choices observed in the unusual disease experiment within our set-up, assume $X = \{B, M, W\}$, where B, M, and W are the best, the mixed, and the worst available alternatives, respectively. Descriptions of the best alternative could be "600 people will be saved", "nobody will die", "600 people will be saved and nobody will die", "nobody will die and 600 people will be saved", or "600 surviving individuals", and so on.[2] Similarly, M and W can correspond to "200 surviving individuals" and "0 surviving individuals", respectively. In the original unusual disease experiment, these consequences admit two possible descriptions, either in terms of the number of individuals which "will be saved" (s) or which "will die" (d), so that $D = \{d, s\}$. DMs have to choose from the menu $\{p, q\}$ of described lotteries $p := \delta_{(M,s)}$ and $q := \frac{1}{3}\delta_{(B,s)} + \frac{2}{3}\delta_{(W,s)}$ or from the menu $\{p', q'\}$ of described lotteries $p' := \delta_{(M,d)}$ and $q' := \frac{1}{3}\delta_{(B,d)} + \frac{2}{3}\delta_{(W,d)}$. Since p and p' are consequentially equivalent ($\sigma^p = \sigma^{p'}$), as well as q and q' ($\sigma^q = \sigma^{q'}$), then so are the two menus: $\sigma^{\{p,q\}} = \sigma^{\{p',q'\}}$. In Tversky and Kahneman's (1981) study, as well as in a number of subsequent replications, subjects behave such that $c(\{p, q\}) = \{p\}$ and $c(\{p', q'\}) = \{q'\}$. Hence, $\sigma^{\{p,q\}} = \sigma^{\{p',q'\}}$, but $\sigma^{c(p,q)} \neq \sigma^{c(p',q')}$

In this example, from two consequentially equivalent menus, the two alternatives chosen are not consequentially equivalent, thereby violating the following cross-menu consistency requirement.

External Description Invariance (EDI): For all $T, T' \in \mathcal{L}$, if $\sigma^T = \sigma^{T'}$, then $\sigma^{c(T)} = \sigma^{c(T')}$.

EDI is usually thought to fully characterize DI because most of the literature focuses on Tversky and Kahneman (1981) unusual disease experiments or other decision problems with the same structure. However, we argue that DI implicitly requires that the DM be indifferent between consequentially equivalent alternatives

[2]See Borie and Jullien (2020) for a theoretical discussion on the structure of the collection of all descriptions of an alternative.

within the same menu. The following example shows how such a property can be violated in experiments.

Example: Unusual Disease (joint evaluation). Kühberger and Gradle (2013, experiment 3) asked DMs to rate (between -3 and +3) their likeliness to choose the four alternatives of the unusual disease presented jointly, i.e., DMs are confronted with just *one* menu: $\{p, p', q, q'\}$. The result is that, on average, the sure program described in terms of lives saved is preferred to the two descriptions of the risky programs (indifferently), themselves preferred to the sure program described in terms of the number of people who will die: $p \succ q \sim q' \succ p'$. This violation of DI can be expressed with our notation as follows: we have $\sigma^p = \sigma^{p'}$, and we observe $p \in c(\{p, p', q, q'\})$ but not $p' \in c(\{p, p', q, q'\})$.[3]

This violates the following implicit property behind DI:

Internal Description Invariance (IDI): For all $T \in \mathcal{L}$ and $p, q \in T$ such that $\sigma^p = \sigma^q$, if $p \in c(T)$, then $q \in c(T)$.

We can easily show that EDI and IDI are independent through examples with the notations of the unusual disease. Assume the DM's choices are $c(\{p, p', q\}) = \{p, p'\}$ and $c(\{p, q, q'\}) = \{q, q'\}$. In that case IDI holds, but EDI is violated (hence IDI does not imply EDI). Assume now that her choices are $c(\{p, q\}) = \{p\}$ and $c(\{p, p', q\}) = \{p'\}$. In that case, EDI holds, but IDI is violated (hence EDI does not imply IDI).

We can now combine EDI and IDI to provide a full characterization of DI.

[3]Such experiments on framing effect with a joint evaluation mode are not very frequent. Another example is Janiszewski et al. (2003)'s study in which DMs have to evaluate and state their willingness to purchase at a given price 8 consumption items each presented under two descriptions. For example, a fruit yogurt is described as "25% fruit" and "75% yogurt" and a blueberry muffin is described as "15% blueberries" and "85% muffin batter". So while the choice set contains 8 alternatives, these are evaluated by the DMs from a menu that contains 16 described alternatives. Their results are consistent with a choice behavior such that, from the same menu, the yogurt described as "25% fruit" would be chosen and the one described as"75% yogurt" would not be chosen.

DI: For all $T, T' \in \mathcal{L}$, if $\sigma^T = \sigma^{T'}$, then $p \in c(T)$ implies $\{q \in T' : \sigma^q = \sigma^p\} \subseteq c(T')$.

DI thus states that if the description of an alternative is chosen from a menu then all its consequentially equivalent descriptions must be chosen from this menu and from consequentially equivalent menus. It is easy to see that IDI and EDI characterize DI.

Proposition 1. *IDI and EDI are satisfied if and only if DI is satisfied.*

Proposition 1 therefore shows that, although DI is often presented informally as *one* monolithic and implicit axiom, it can in fact be formally expressed explicitly as the conjunction of *two* axioms (IDI and EDI).

2 Description Invariance and Utility Maximization

When DI holds, framing effects cannot result from utility maximization. Yet it has been argued that framing effects are not necessarily instances of irrational behavior (e.g., Keren, 2011; Sher and McKenzie, 2011; Bermúdez, 2022). In what follows, we investigate the relationship between rationality and framing effects formally by exploring how DI relates (a) to the Weak Axiom of Revealed Preference (WARP) – i.e., utility maximization – or even (b) to the maximization of Expected Utility (EU). Our analysis uses the following (standard) formal definitions. Recall that c satisfies WARP when for any $T, T' \in \mathcal{L}$ and $p, q \in T \cap T'$ such that $p \in c(T)$, if $q \in c(T')$, then $p \in c(T')$. When WARP holds, c is said to be *rationalizable* by a complete and transitive preference relation. Furthermore, c maximizes Expected Utility (EU) when there exists $u : X \times D \to \mathbb{R}$ such that for all menus of described lotteries $T \in \mathcal{L}$ and all $p, q \in T$, $c(T) = \arg\max_{p \in T} \sum_{(x,d) \in X \times D} p(x,d)u(x,d)$. Finally, c is said to be *continuous* if (i) for all $\alpha \in (0,1), \alpha p + (1-\alpha)r \in L$ and (ii) for all $p, q, r \in L$, such that $\{p\} = c(p,q) = c(p,r)$, $\{q\} = c(p,r)$, there exists $\alpha, \beta \in (0,1)$ such that $\{\alpha p + (1-\alpha)r\} = c(\alpha p + (1-\alpha)r, q)$ and $\{q\} = c(\beta p + (1-\beta)r, q)$.

We start by investigating the relationship between WARP and the two subproperties of DI (EDI and IDI). Given that the definition of WARP is not dependent on whether the alternatives are the same under different descriptions or are different alternatives, a choice correspondence can perfectly be rationalizable without satisfying IDI, EDI, or both. Therefore, WARP does not imply DI.

On the other hand, a choice correspondence may satisfy IDI, EDI or both without being rationalizable. The typical scenario under which this happens consists in (at least) three classes of alternatives which are all consequentially equivalent within a class but not across classes. Then IDI requires that the DM be indifferent between all alternatives within the same class; and EDI imposes that if an alternative from one class —call it L_1—is preferred to an alternative from another class L_2, then this must be the case for all alternatives in L_1. However, this does not guarantee that transitivity holds between triples of alternatives that belong to three different classes. Indeed, under EDI, a preference for the alternatives in L_1 over those in L_2 and for the alternatives in L_2 over those in L_3 does not imply a preference between the alternatives in L_1 and L_3.[4] Therefore, DI does not imply WARP.

We now investigate the relation between EDI and IDI when WARP holds. The next proposition shows that, if c is rationalizable, then IDI and EDI are not independent anymore (as in the previous section) and IDI implies EDI.

Proposition 2. *Assume WARP holds. Then, IDI implies EDI. If c is continuous, then the converse also holds. Hence, if $L = \Delta(X \times D)$ and c is EU, then IDI and EDI are equivalent.*

[4]More formally, assume $X = \{x_1, x_2, x_3\}$ and D is a finite set. L is partitioned by $\{L_i\}_{i\in\{1,2,3\}} := \{\delta(x_i, d) : d \in D\}_{i\in\{1,2,3\}}$. Consider the choice correspondence c such that for each $T \in \mathcal{L}$

$$c(T) = \begin{cases} T \cap L_1 & \text{if } T \cap L_1 \neq \emptyset \text{ and } T \subseteq L_1 \cup L_2 \\ T \cap L_2 & \text{if } T \cap L_2 \neq \emptyset \text{ and } T \subseteq L_2 \cup L_3 \\ T \cap L_3 & \text{otherwise.} \end{cases}$$

It is easy to check that c satisfies both IDI (within each row) and EDI (across the three rows). However, for any $(p, q, r) \in L_1 \times L_2 \times L_3$, we have $p = c(\{p, q\})$, $q = c(\{q, r\})$, and $r = c(\{p, r\})$; so that no transitive preference relation can rationalize c.

This result demonstrates that a rational DM, who exhibits some sensitivity to descriptions, can only be modeled under very specific and restrictive conditions. Specifically, violations of IDI can only be captured if the function c is discontinuous. If c is continuous, however, it becomes impossible to model choices that violate either IDI or EDI while still exhibiting a form of sensitivity to description. As a result, the patterns of choice that are most relevant to our discussion—particularly the violation of EDI observed in the unusual disease example—cannot be modeled under the assumptions we have established so far. This impossibility aligns with the views of most commentators on the unusual disease example (and similar cases): these framing effects are instances of irrational behavior.

3 Relaxing Description Invariance

In this section, we show that a weaker version of DI allows to model rational choices that violate EDI and thus to rationalize the unusual disease example. The weaker version of DI consists in imposing that either EDI or IDI hold only for sure consequences. This has the advantage of limiting the richness of observations necessary to falsify the axioms. The intuition behind this weakening is that for the DM, realizing that sure alternatives are consequentially equivalent is easier in choice sets that contain only sure alternatives without risky ones.

Let δ_x denote the set of lotteries which yield the consequence $x \in X$ with probability 1, i.e., $L(x) := \{p \in L : \sigma^p(x) = 1\}$. Similarly, $\mathcal{L}^s \subset \mathcal{L}$ is the set of menus that contain only degenerate lotteries on consequences, i.e., if $A \in \mathcal{L}^s$, then $A \subseteq \bigcup_{x \in X} L(x)$. We now weaken IDI and EDI to restrict their applications to such sure lotteries.

Weak Internal Description Invariance (WIDI): For all $T \in \mathcal{L}^s$ and $p, q \in T \cap L(x)$, if $p \in c(T)$, then $q \in c(T)$.

Can this weakening leave sufficient room to represent interesting description-dependent choice behavior? The next proposition shows that, under full richness, this cannot be the case for choice correspondences that admit an expected utility

representation.

Proposition 3. *Assume EU holds and that $L = \Delta(X \times D)$, then WIDI implies IDI and EDI.*

In other words, as in the previous section, assuming WIDI under expected utility grants the full power of DI. Hence, WIDI is not a weakening capable of offering enough flexibility to represent the framing effects we are interested in.

Weak External Description Invariance (WEDI): For all $T, T' \in \mathcal{L}^s$, $\sigma^T = \sigma^{T'}$ implies $\sigma^{c(T)} = \sigma^{c(T')}$.

It seems at first sight that this axiom opens the possibility that, as observed in the unusual disease experiment, for some menus T, T' with $\{T, T'\} \notin \mathcal{L}^s$, $\sigma^T = \sigma^{T'}$ does not imply $\sigma^{c(T)} = \sigma^{c(T')}$. The following proposition displays a formulation of expected utility that is structured by this axiom.

We first introduce convenient notation. A pair of functions $(v, \{w_d\}_{d \in D})$ from $X \to \mathbb{R}$ is *scaled by consequences* if (i) for all $d \in D$ and for all $x, y \in X$, $|v(x) - v(y)| \geq |w_d(x) - w_d(y)|$ and (ii) for all $x \in X$, $\min_{d \in D} w_d(x) = 0$. Finally, a choice function c is *consistent over consequences* whenever there exists a pair $(v, \{w_d\}_{d \in D})$ scaled by consequences such that for all $T \in \mathcal{L}$

$$c(T) = \arg\max_{p \in T} \sum_{(x,d) \in X \times D} p(x, d) \Big(v(x) + w_d(x) \Big). \tag{1}$$

When a choice function is consistent over consequences, a description alters the value of a sure option in the DM's eyes, but not enough to reverse her preference (by (i)). The next proposition shows the conditions under which an expected utility representation can capture this kind of sensitivity to description.

Proposition 4. *c is represented by an expected utility and satisfies WEDI if and only if it is consistent over consequences. Moreover, if $(v, \{w_d\}_{d \in D})$ and $(v', \{w'_d\}_{d \in D})$ are both scaled by consequences satisfying (1), then there exist $\alpha, \beta \in \mathbb{R}$ with $\alpha > 0$ such that $v' = \alpha v + \beta$ and $w'_d = \alpha w_d$.*

Note that under EU, the utility of the DM can be decomposed into a part which is derived purely from the consequences regardless of their descriptions $v(x)$ and a part which is sensitive to the descriptions under which the consequences are presented $w_d(x)$. Proposition 4 establishes that when WEDI holds, the sensitivity to descriptions cannot reverse the preference of the DM over the consequences of the alternatives. In other words, it allows to rationally prefer some descriptions of an alternative over other descriptions of that alternative as long as all the alternatives are well-ordered.[5] Furthermore, w_d is a ratio scale (unique up to multiplication by a positive constant), which means that there is an absolute zero for this function (unlike with a cardinal scale). This implication is particularly well aligned with our interpretation of w_d as measuring the deviation (induced by framing) from the classical utility derived from the consequences.

However, this absence of reversal is restricted to choice sets that contain only sure alternatives (which is not the case in the unusual disease example). Applied to choice sets that contain sure alternatives and lotteries, the EU of Proposition 4 can account for the pattern of choices observed in the unusual disease example. Recall that $D = \{d, s\}$, where d and s are the descriptions of the consequences in terms of the number of people dead and saved, respectively; and $p = \delta_{(M,s)}$, $q = \frac{2}{3}\delta_{(W,s)} + \frac{1}{3}\delta_{(B,s)}$, $p' = \delta_{(M,d)}$, and $q' = \frac{2}{3}\delta_{(W,d)} + \frac{1}{3}\delta_{(B,d)}$. The DMs behavior $p = c(p, q)$ and $q' = c(p', q')$ can now be represented by the maximization of the expected utility criterion provided that

$$\frac{2}{3}w_d(W) + \frac{1}{3}w_d(B) + w_s(M) > \frac{2}{3}w_s(W) + \frac{1}{3}w_s(B) + w_d(M).$$

4 Implications for empirical work

One suggestion for further empirical work can be derived from our results. In the light of Proposition 4, testing WEDI seems particularly relevant. Doing so would require that subjects rank several descriptions of sure alternatives. To illustrate

[5] In a sense, this result can be seen as pushing the idea that the DM is the best judge of his or her own welfare within the constraints of a standard theory of choice that accounts for the existence of different descriptions of the objects of choice.

with the unusual disease example, the choice set could include "200 people will be saved", "400 people will die", "0 people will be saved", "600 people will die", "600 people will be saved", and "0 people will die". It would be particularly interesting to observe whether the commonly reported preference for "200 people will be saved" over "400 people will die" is still observed when no lotteries are involved in the choice set.

Such an experiment could serve as a robustness check for Kühberger and Gradle (2013, experiment 3), who observed this preference but within a choice set that also included descriptions of the risky programs. Note that the finiteness of the set of consequences is crucial for the model of proposition 4 to capture the unusual disease. If X were to be infinite, this EU representation would not allow for capturing the unusual disease example. This may be seen as a problem. However, since most of —if not all—the empirical literature on framing effects is confined to finite sets, this problem does not prevent suggestions for further empirical work to be derived from our set-up.

Furthermore, our model suggest that different behavioral mechanisms underpin two forms of consistency (i.e., EDI and IDI). Therefore, it could be interesting to test whether there are differences in the psychological or cognitive processes (e.g., emotional reaction and attention and) that govern how individuals evaluate cross-menu *versus* within-menu consistency. While our model is agnostic on such processes, we emphasize that EDI and IDI are not an invention specific to our model, but implicitly present in all models based on rational choice theory.

Besides suggesting experiments, our set-up also allows to construct tools for empirical analysis. For example, a relevant exercise of comparative statics for framing effects could be to compare the relative strength of different frames. As it transpired from our choice of terminology, and consistently with our analysis, the function w_d can be interpreted as a power index for the the frame d: The higher its image is, the more sensitive the individual is to the framing effect. We say that frame d' is less powerful than d if for all $x, y \in X$ and $l \in \Delta(X)$, $(l, d) \succsim (x, d)$ and $(x, d) \sim (y, d')$ implies $(l, d') \succsim (y, d')$. The comparison is made between consequences which display the same "total" utility and between frames which do not induce preference

reversal. We then immediately obtain:

Proposition 5. *Let d and d' be frames. Then d' is less powerful than d if and only if $w_{d'} = \phi(w_d)$ for some nondecreasing and convex function ϕ.*

Note that such a tool allows to rank the strength of frames in a spirit that is compatible with, and further develops, Kahneman and Tversky's original work on framing effects.

5 Related literature

Our theoretical account of framing effects differs from other existing contributions as follows. Firstly, many well-known choice theoretic works that rationalize preference reversals cannot (and were not developed to) account for framing effects triggered by different descriptions of the consequences (Masatlioglu and Ok, 2005; Salant and Rubinstein, 2008; Ok et al., 2015; Bernheim and Rangel, 2009). The choice theoretic framework of Dietrich and List (2016) can account for the unusual disease example, but only under the assumption that consequential equivalence between two alternatives implies the indifference of the DM between them. Accounting for framing effects without this assumption can be seen as a central contribution of this paper.

Secondly, there are several contributions that rationalize the framing effects we are interested in but they formalize only the preferences of the DM without characterizing his or her observable choice behavior. Furthermore, unlike what we did in this paper, most of these contributions use rather non-standard formalisms, such as predicate logic or programming languages (Bacharach, 2003; Gold and List, 2004; Ryan, 2005; Bourgeois-Gironde and Giraud, 2009; Lanzi, 2011; Blume et al., 2013; Lerner, 2014). Two exceptions are worth discussing. On the one hand, there is Giraud's work (2004b; 2004a; 2005). It account for framing effects by assuming that descriptions trigger positive or negative attitudes and he captures this in the DM's beliefs. As we explained in the introduction, this assumption has been rejected in psychological work on framing effects beyond the one from Tversky and Kahneman. Therefore, accounting for framing effects without this assumption can be seen as

a central contribution of this paper. On the other hand, there is the contribution of Borie and Jullien (2020). They rationalize framing effects in the same spirit as ours but use more theoretical (and non-standard) primitives as we do here. Notably their concatenation operation and equivalence relation on the descriptions make their account less straightforward than the one presented here.

A Appendix

Proof of Proposition 1. First assume that DI holds. IDI follows from applying DI with $T = T'$. To show EDI, assume we have $T, T' \in \mathcal{L}$ such that $\sigma^T = \sigma^{T'}$. Let $\gamma \in \sigma^{c(T)}$, i.e., there exists $p \in c(T)$ with $\sigma^p = \gamma$. By DI, $p \in c(T)$ entails $\{q \in T' : \sigma^q = \sigma^p\} \subseteq c(T')$. Given that $\sigma^T = \sigma^{T'}$, there exists $q \in c(T')$ such that $\sigma^p = \sigma^q$ and, thus, $\gamma = \sigma^q \in \sigma^{c(T')}$, proving $\sigma^{c(T)} = \sigma^{c(T')}$.

Assume now that IDI and EDI hold. Let $T, T' \in \mathcal{L}$ with $\sigma^T = \sigma^{T'}$ and $p \in c(T)$. Assume there exists $q \in \{q \in T' : \sigma^q = \sigma^p\} \backslash c(T')$. By IDI this implies that $\{q \in T' : \sigma^q = \sigma^p\} \cap c(T') = \emptyset$, which contradicts EDI. □

Proof of Proposition 2. Let $T, T' \in \mathcal{L}(D)$ such that $\sigma^T = \sigma^{T'}$ and $\gamma \in \sigma^{c(T)}$. Consider $p \in c(T)$ such that $\sigma^p = \gamma$. Given that $c(T')$ is nonempty, we can take $p' \in c(T')$. Since $\sigma^T = \sigma^{T'}$, there exist $q \in T$ and $q' \in T'$ such that $\sigma^{q'} = \sigma^p$ and $\sigma^q = \sigma^{p'}$. By IDI, $q' \notin c(T \cup \{q'\})$ would imply $p \notin c(T \cup \{q'\})$ which, together with WARP would imply $p \notin c(T)$. Hence, $p \in c(T)$ entails $q' \in c(T \cup \{q'\})$. A symmetric reasoning entails $q \in c(T' \cup \{q\})$. Given that $q \in T \cup \{q'\}$, WARP thus implies $q' \in c(T' \cup \{q\})$ and therefore $q' \in c(T')$. Hence, $\gamma = \sigma^p = \sigma^{q'} \in \sigma^{c(T')}$, which completes the proof.

To show that the converse holds under continuity, note that it is only needed to show that for all $p, q \in L$ such that $\sigma^p = \sigma^q$, $c(p, q) = \{p, q\}$—if so, then the result follows from WARP. To show this, assume by contradiction that $c(p, q) = \{p\}$. Consider r such that $\sigma^r \neq \sigma^q$. Such a r must exist because X is not a singleton. By EDI the only possible scenarios then are: (1) $r \notin c(p, r) \cup c(q, r)$, (2) $\{r\} = c(p, r) = c(q, r)$, and (3) $\{p, r\} = c(p, r)$ & $\{q, r\} = c(q, r)$. The last scenario is not compatible with WARP and so we can rule it out. Assume without loss of generality

that scenario (1) holds, then continuity implies that there exists $1 > \beta > 0$ such that $c(\beta p + (1-\beta)r, q) = \{\beta p + (1-\beta)\}$. Given that $\sigma^{\beta p+(1-\beta)} \neq \sigma^q$, EDU is violated. A contradiction.

□

Proof of proposition 3. In the light of Proposition 2, we only need to prove that WIDI implies IDI. Let $T \in \mathcal{L}$ and $p, q \in T$ such that $\sigma^p = \sigma^q$ and $p \in c(T)$. Hence, σ^p and σ^q have the same support. Note that we only to prove that $\{p, q\} = c(p, q)$, the result then follows from WARP.

We show this by induction on $|supp\ \sigma^p|$. If $|supp\ \sigma^p| = 1$, then $p \in \delta_x$ for some $x \in X$ and the result follows directly from WIDI. Assume now that the result holds for every $p', q' \in L$ with $\sigma^{p'} = \sigma^{q'}$ such that $|supp\ \sigma^{p'}| = n$ and let us show that it holds $p, q \in L$ with $\sigma^p = \sigma^q$ such that $|supp\ \sigma^p| = n + 1$. Let $p, q \in L$ such that $\sigma^p = \sigma^q$ and $x \in supp\ \sigma^p$. Given that $L = \Delta(X \times D)$, we can consider the lotteries

$$p' := \sum_{\substack{(x',d)\in X\times D \\ x'\in supp\ \sigma^p\setminus\{x\}}} \frac{p(x',d)}{\sigma^p(x)}\delta_{(x',d)} \text{ and } q' := \sum_{\substack{(x',d)\in X\times D \\ x'\in supp\ \sigma^q\setminus\{x\}}} \frac{q(x',d)}{\sigma^q(x)}\delta_{(x',d)}.$$

One can easily check that p', q' are described lotteries, that $\sigma^{p'} = \sigma^{q'}$, and that $supp\ p' = supp\ q' = n$. Hence, the induction hypothesis implies that $c(p', q') = \{p', q'\}$. Similarly, the lotteries

$$p'' = \frac{1}{\sigma^p(x)} \sum_{(x,d'):d'\in D} p(x,d')\delta_{(x,d)} = p \text{ and } q'' = \frac{1}{\sigma^p(x)} \sum_{(x,d'):d'\in D} q(x,d')\delta_{(x,d)}.$$

are described lotteries of support 1 and $\sigma^{p''} = \sigma^{q''}$. Hence, $c(p'', q'') = \{p'', q''\}$.

Noting now that $p = \sigma^p(x)p' + (1 - \sigma^p(x))p''$ and $q = \sigma^p(x)q' + (1 - \sigma^p(x))q''$, and given that c is representable by an expected utility (which is linear), $c(p', q') = \{p', q'\}$ and $c(p'', q'') = \{p'', q''\}$ imply that $c(p, q) = \{p, q\}$.

□

Proof of Proposition 4. Given that c is represented by an expected utility, there exists $u : X \times D \to \mathbb{R}$ such that $U(p) := \sum_{(x,d)\in X\times D} p(x,d)u(x,d)$ represents $\succsim$. Define $v : X \to \mathbb{R}$ by $v(x) := \min_{d\in D} u(x,d)$ and for all $d \in D$, $w^d : X \to \mathbb{R}$ by $w^d(x) = u(x,d) - v(x)$. Hence, for all $p \in T$, $U(p) = \sum_{(x,d)\in X\times D} p(x,d)[w^d(x) +$

$v(x)]$. This means that we only need to prove that $(v, \{w_d\}_{d\in D})$ is scaled by consequences. Let $d \in D$, $x, x' \in X$ such that $v(x) \geq v(x')$ and $d^x, d^{x'} \in D$ such that $d^x \in \arg\min_{d\in D} w^d(x) = 0$ and $d^{x'} \in \arg\min_{d\in D} w^d(x') = 0$. Hence, $\delta_{(x,d^x)} \in c(\delta_{(x,d^x)}, \delta_{(x',d^{x'})})$. Given that $\sigma^{\{\delta_{(x,d^x)},\delta_{(x',d^{x'})}\}} = \sigma^{\{\delta_{(x,d)},\delta_{(x',d)}\}}$, WEDI implies $\delta_{(x,d)} \in c(\delta_{(x,d)}, \delta_{(x',d)})$. Hence, $v(x) + w^d(x) \geq v(x') + w^d(x')$. A similar reasoning entails that $v(x') > v(x)$ implies $v(x) + w^d(x) > v(x') + w^d(x')$. This proves that $(v, \{w_d\}_{d\in D})$ is scaled by consequences. It remains to show the uniqueness part of the statement. Let $(v', \{w'_d\}_{d\in D})$ be an another pair scaled by consequences representing c. From routine arguments, there exists $u' : X \times D \to \mathbb{R}$ such that $u' = \alpha u + \beta$ for some $\alpha, \beta \in \mathbb{R}$ such that $\alpha > 0$ and $u'(x, d) = v'(x) + w'_d(x)$. The constraints on $(v', \{w'_d\}_{d\in D})$ yield $v'(x) = \min_{d\in D} u'(x, d)$. It follows that $v' = \alpha v + \beta$ and $w'_d = w_d$. This completes the proof. □

Proof of Proposition 5. The proof is an immediate application of Theorem 2 in Peters and Wakker (1986). □

Acknowledgments

Niels Boissonnet has been supported by the project JICE-InChange. The project JICE-InChange has received/is receiving funding from the program "Profilbildung 2020", an initiative of the Ministry of Culture and Science of the State of Northrheine Westphalia.

References

Antinyan, A., Corazzini, L., Fišar, M., and Reggiani, T. (2024). Mind the framing when studying social preferences in the domain of losses. *Journal of Economic Behavior & Organization*, 218:599–612.

Arrow, K. J. (1982). Risk Perception in Psychology and Economics. *Economic Inquiry*, XX:1–9.

Bacharach, M. O. (2003). Framing and Cognition in Economics: The Bad News and

the Good. In Dimitri, N., Basili, M., and Gilboa, I., editors, *Cognitive Processes and Economic Behaviour*, pages 63–74. Routledge, London.

Bardsley, N., Cubitt, R. P., Loomes, G., Moffatt, P., Starmer, C., and Sugden, R. (2010). *Experimental Economics: Rethinking the Rules*. Princeton University Press, Princeton.

Bermúdez, J. L. (2022). Rational framing effects: A multidisciplinary case. *Behavioral and Brain Sciences*, 45.

Bernheim, D. B. and Rangel, A. (2009). Beyond Revealed Preference: Choice-theoretic Foundations for Behavioral Welfare Economics. *Quarterly Journal of Economics*, 124(1):51–104.

Blume, L., Easley, D., and Halpern, J. Y. (2013). Constructive Decision Theory.

Borie, D. and Jullien, D. (2020). Description-dependent preferences. *Journal of Economic Psychology*, 81:102311.

Bourgeois-Gironde, S. and Giraud, R. (2009). Framing Effects as Violations of Extensionality. *Theory and Decision*, 67(4):385–404.

Broniatowski, D. A. and Reyna, V. F. (2018). A formal model of fuzzy-trace theory: Variations on framing effects and the allais paradox. *Decision*, 5(4):205–252.

Camerer, C. F. (1995). Individual Decision Making. In Kagel, J. H. and Roth, A. E., editors, *The Handbook of Experimental Economics*, pages 587–703. Princeton University Press, Princeton.

Dietrich, F. and List, C. (2016). Reason-based Choice and Context-dependence: An Explanatory Framework. *Economics and Philosophy*, (Forhcoming).

Giraud, R. (2004a). Framing under risk: Endogenizing the Reference Point and Separating Cognition and Decision. *Working Paper*.

Giraud, R. (2004b). *Une théorie de la décision pour les préférences imparfaites*. PhD thesis.

Giraud, R. (2005). Anomalies de la théorie des préférence. Une interprétation et une proposition de formalisation. *Revue Economique*, 56(4):829–854.

Gold, N. and List, C. (2004). Framing as Path Dependence. *Economics and Philosophy*, 20(2):253–277.

Hong, F., Hossain, T., and List, J. A. (2015). Framing manipulations in contests: A natural field experiment. *Journal of Economic Behavior & Organization*, 118:372–382.

Janiszewski, C., Silk, T., and Cooke, A. (2003). The role of subjective scales and experience in explaining attribute-framing effects. *Journal of Consumer Research*, 30(3):311–325.

Keren, G. (2011). *Perpectives on Framing*. Psychology Press, New York.

Kühberger, A. (1997). Theoretical conceptions of framing effects in risky decisions. In Ranyard, R., Crozier, R. W., and Svenson, O., editors, *Decision Making. Cognitive models and explanations*, chapter 8, pages 128–144. Routledge, London and New-york.

Kühberger, A. and Gradle, P. (2013). Choice, Rating, and Ranking: Framing Effects with Different Response Modes. *Journal of Behavioral Decision Making*, 26(2):109–117.

Lanzi, D. (2011). Frames as Choice Superstructures. *The Journal of Socio-Economics*, 40(2):115–123.

Lerner, S. (2014). A non-monotonic intensional framework for framing effects. *Journal of Economic Methodology*, 21(1):37–53.

Levin, I. P., Schneider, S. L., and Gaeth, G. J. (1998). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational behavior and human decision processes*, 76(2):149–188.

Masatlioglu, Y. and Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, 121(1):1–29.

Ok, E. A., Ortoleva, P., and Riella, G. (2015). Revealed (P)Reference Theory. *American Economic Review*, 105(1):299–321.

Peters, H. and Wakker, P. (1986). Convex functions on non-convex domains. *Economics Letters*, 22(2):251–255.

Quiggin, J. (2014). Non-Expected Utility Models Under Objective Uncertainty. In Machina, M. J. and Viscusi, K. W., editors, *Handbook of the Economics of Risk and Uncertainty*, pages 701–728. Elsevier B.V., Oxford.

Ryan, M. J. (2005). Framing, Switching and Preference Reversals. *Theory and Decision*, 57(3):181–211.

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

Seta, J. J., Seta, C. E., and McCormick, M. (2017). Commonalities and differences among frames: A unification model. *Journal of Behavioral Decision Making*, 30(5):1113–1130.

Sher, S. and McKenzie, C. R. M. (2011). Levels of Information: A Framing Hierarchy. In Keren, G., editor, *Perpectives on Framing*, chapter 2, pages 35–64. Psychology Press, New York.

Takemura, K. (2014). *Behavioral decision theory. Psychological and mathematical descriptions of human choice behavior*. Springer, Tokyo.

Tversky, A. and Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481):453–458.