Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 193 (2021) 105221
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Pairwise normalization: A neuroeconomic theory of multi-attribute choice ☆

Peter Landry [a,b,*], Ryan Webb [a,c,*]

[a] Rotman School of Management, University of Toronto, Canada
[b] Department of Management, University of Toronto, Mississauga, Canada
[c] Department of Economics, University of Toronto, Canada

Received 18 July 2019; final version received 12 January 2021; accepted 8 February 2021
Available online 19 February 2021

Abstract

We present a theory of multi-attribute choice founded in the neuroscience of perception. Valuation is formed through a series of pairwise, attribute-level comparisons implemented by *(divisive) normalization* — a form of relative value coding observed across sensory modalities and in species ranging from honeybees to humans. Such "pairwise normalization" captures a broad range of behavioral regularities in multi-attribute choice, including the compromise and asymmetric dominance effects, the diversification bias in allocation decisions, and majority-rule preference cycles.

JEL classification: D01; D11; D87; D91

Keywords: Neuroeconomics; Consumer theory; Multi-attribute choice; Decoy effects

☆ We thank Pedro Bordalo, David Freeman, Kenway Louie, Antonio Rangel, Colin Stewart, and especially Anthony Marley for many helpful comments on this paper. We also received valuable feedback from seminar participants at Stanford Graduate School of Business, California Institute of Technology, McGill, Carnegie-Mellon, Simon Fraser University, Maastricht University, University of Zurich, Virginia Tech, the Barcelona Summer Forum, BRIC 2018, the Sloan-Nomis Workshop, Peking University Business School, Harbin Institute of Technology, and Chinese University of Hong Kong. Fiorella Wever provided excellent research assistance. We also thank Alireza Soltani, Benedetto de Martino, and Colin Camerer for providing experimental data.
* Corresponding authors.
E-mail addresses: peter.landry@rotman.utoronto.ca (P. Landry), ryan.webb@rotman.utoronto.ca (R. Webb).

https://doi.org/10.1016/j.jet.2021.105221

1. Introduction

Standard choice theories presume that an individual's valuation of an alternative does not depend on the set of alternatives under consideration. However, a large empirical literature has revealed several violations of such "context-independence." For example, simply adding an alternative to a choice set can alter preferences among existing alternatives (see Rieskamp et al., 2006, for a review). Empirical demonstrations of context effects can be found in both laboratory experiments (beginning with Huber et al., 1982, and Simonson, 1989) and in field data (e.g. Doyle et al., 1999; Geyskens et al., 2010), and extend to many types of decisions — including consumer choice, choices among lotteries, doctors' prescription decisions, perceptual decisions, and mate selection, to name just a few.[1]

Though less familiar to behavioral researchers, context-independence is also challenged by an established neuroscience literature (beginning with Hartline and Wagner, 1952) demonstrating that the brain encodes information in relative, not absolute terms. For example, the neural activity encoding the value of an alternative decreases (indicating a reduced valuation) as the value of another alternative rises (Louie et al., 2011; Holper et al., 2017). This pattern of neural activity is consistent with *divisive normalization*, a well-documented neural computation that, in its simplest conceivable form, merely re-expresses some input value a — which may represent the utility of an alternative, or the intensity of sensory stimuli (such as the brightness of a pixel) — relative to another input b as $\frac{a}{a+b}$ (see Rangel and Clithero, 2012; Carandini and Heeger, 2012, and Louie et al., 2015, for reviews). Indeed, the prevailing neuroscience literature conceptualizes such "division by neurons" as an arithmetic operation that is actually performed in the brain.[2]

Why would our brains not just encode a independently of b? The reason is thought to stem from biological constraints. The brain has a limited number of neurons, each with a bounded response range. Thus, information must be compressed within these bounds. A relative value encoding is then needed to ensure this compression is well-calibrated to the choice environment (a point previously noted in the economics literature by Robson, 2001, and Rayo and Becker, 2007; also see Netzer, 2009; Woodford, 2012; Robson and Whitehead, 2018, and Frydman and Jin, 2020). A relative encoding using the divisive normalization computation has been shown to optimally mitigate choice mistakes subject to these biological constraints (Steverson et al., 2019; Webb et al., 2020a). Put simply, divisive normalization efficiently facilitates the perception of both large and small differences on a common scale — e.g. helping to distinguish "one dollar from two dollars and one million dollars from two million dollars" (Carandini and Heeger, 2012).

In this paper, we explore whether divisive normalization — an inherently context-dependent computation — might relate to context-dependent behavior. To do so, we adapt the "$\frac{a}{a+b}$" divisive normalization model to the multi-attribute choice setting where behavioral research on context-dependence is mainly focused. Let $\mathbf{x}$ be an alternative with $x_1, \ldots, x_N$ denoting its N attribute values. These may be observable values or utility-transformed values, as discussed in Section 3.1. The decision-maker's valuation of $\mathbf{x}$ according to our basic *pairwise normalization (PN)* model is normalized relative to other alternatives in the choice set X as:

[1] For examples, see Huber et al. (2014), Soltani et al. (2012), Schwartz and Chapman (1999), Trueblood et al. (2013), and Lea and Ryan (2015).

[2] See Carandini and Heeger (1994), who coined the phrase in quotes, as well as related work by Wilson et al. (2012). At the neurocomputational level, the divisive functional form can be derived as the equilibrium solution to the dynamics that govern neural activity in a stylized neural circuit (Louie et al., 2014).

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} \sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{x_n}{x_n + y_n}.$$

This formulation is "pairwise" in the sense that each term reflects an attribute-level comparison (normalization) of **x** to some other alternative **y**. Pairwise comparisons have long been a feature of multi-attribute choice theories (e.g. Tversky and Simonson, 1993) and have substantial empirical support from eye-tracking research in multi-attribute choice environments; in one such study, Noguchi and Stewart (2014) find that "alternatives are repeatedly compared in pairs on single dimensions."[3]

Our modeling approach demonstrates how neuroscience may prove useful to economists as a source of candidate functional form representations to consider in model selection (as suggested by Bernheim, 2009). Arguably the simplest, standard multi-attribute choice model is an additive model, $V(\mathbf{x}) = \sum_n x_n$. This additive model provides a common foundation for many leading multi-attribute choice theories that address context-dependence (e.g. Tversky and Simonson, 1993; Kivetz et al., 2004a; Koszegi and Szeidl, 2013; Bordalo et al., 2013; and Bushong et al., 2019). These theories typically replace each term in the summation with a function of x_n that also depends on the set of alternatives. Similarly, our theory modifies the additive model by applying pairwise normalization to each attribute value (effectively replacing x_n with $\sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{x_n}{x_n + y_n}$). This formalizes pairwise normalization in its most elemental form, isolated from other factors that may influence choice, and with minimal parametric freedom. Despite its simplicity, the model's predictions capture a broad range of context-dependent behavioral regularities that are only partially captured by prevailing multi-attribute choice theories. These predictions are summarized in Table 1, where '**Y**' indicates that the model robustly captures the associated behavior (i.e. never predicts the opposite or no effect under conditions for which it would be expected), 'S' means the model sometimes predicts the behavior and sometimes predicts the opposite effect, and 'N' means the model does not predict the behavior.[4]

The rest of this paper proceeds as follows. Section 2 provides background on the neurobiological basis of our model and the behavioral regularities that it captures. Section 3 presents the theoretical model. Section 4 examines how a preference between two alternatives can be affected by a third alternative, and relates these effects to the notion advanced by Tversky and Russo (1969) and Natenzon (2019) that similar alternatives are "easy to compare." Section 5 considers choices among alternatives defined on three dimensions. Section 6 considers various allocation problems. Section 7 explores a generalization of the model. Section 8 clarifies the behavioral role of pairwise comparisons in our model and elaborates on the varying representations of attribute-level comparisons in the relevant theoretical literature.

[3] Also see Russo and Dosher (1983), Arieli et al. (2011), Noguchi and Stewart (2018), Turner et al. (2017), as well as Russo and Rosen (1975), who emphasize that the use of pairwise comparisons may stem from cognitive constraints, as even ternary comparisons (which they observed roughly 2 percent as often as pairwise comparisons) can stretch working memory to its limits. That said, whether individuals use pairwise comparisons can be sensitive to various aspects of the choice environment, including the manner in which alternatives are presented and the cost of learning attribute information (see, for example, Payne et al., 1993, and Reeck et al., 2017). Our model is only intended to explain behavior in situations where individuals do use pairwise comparisons. For a lengthier discussion of pairwise comparisons in relation to other theoretical representations of attribute-level comparisons, see Section 8.

[4] Appendix C provides detailed explanations of how each model's predictions were classified in Table 1. Additional empirical tests of pairwise normalization are reported by Sullivan et al. (2019) and Dumbalska et al. (2020).

Table 1
Behavioral patterns generated by pairwise normalization*.

			Pairwise Normalization	Bordalo et al. (2013)	Bushong et al. (2019)	Kivetz et al. (2004a)	Soltani et al. (2012)	Tversky and Simonson (1993)
(I)	Compromise Effect		**Y**	S	**Y**	**Y**	S	**Y**
(II)	Dominance Effect	Weak	**Y**	S	N	N	N	**Y**
		Strict	**Y**	S	**Y**	**Y**	**Y**	**Y**
	...Decoy-Range Effect		**Y**	S	**Y**	**Y**	**Y**	N
(III)	Relative Difference Effect		**Y**	S	N	N	N	N
(IV)	Majority-Rule Pref. Cycles		**Y**	S	**Y**	**Y**	-	N
(V)	Splitting Bias		**Y**	S	**Y**	**Y**	-	N
(VI)	Alignability Effect		**Y**	S	**Y**	**Y**	-	N
(VII)	Diversification Bias		**Y**	S	N	N	N	N
(VIII)	Feature Bias		**Y**	**Y**	N	N	S	N

Key behavioral regularities predicted by the basic PN model as compared to several prevailing multi-attribute choice theories. Here, '-' means the model's predictions cannot be classified. The '**Y**,' 'S,' and 'N' classifications are defined in the text. See Appendix C for detailed explanations and Fig. 2 for illustrations of items (I) and (II).
* This table only includes theories that are directly comparable to the basic PN model in that the domains of their analyses have sufficient overlap with ours. Notable models addressing the compromise and/or dominance effects in somewhat different domains include Kamenica's (2008) contextual inference theory (which, unlike the theories listed above, models a market with both consumers and a firm), De Clippel and Eliaz's (2012) dual-self intrapersonal bargaining theory, Ok et al.'s (2015) endogenous reference point theory, Guo's (2016) contextual deliberation theory, and Natenzon's (2019) Bayesian probit model. In Appendix C, we elaborate on how these models differ. That said, a static version of Koszegi and Szeidl's (2013) dynamic 'focusing' theory is directly comparable to our model. While this theory does not seek to address the context-dependent phenomena addressed by other theories, for completeness we derive its predictions (all 'N') for each item in Appendix C.

2. Additional background

2.1. Review of divisive normalization

The divisive normalization computation was initially conceived as an explanation for the non-linear responses of cortical neurons to properties of visual stimuli such as contrast and brightness (Heeger, 1992; Carandini et al., 1997). The computation serves as a form of gain control, re-scaling its inputs into a bounded output range, and has been shown to yield an efficient coding of visual information in a constrained neural system.[5] Divisive normalization has since been observed in the neural activity of other sensory systems including olfaction and auditory perception (e.g. Olsen et al., 2010; Rabinowitz et al., 2011), as well as multi-sensory integration (Ohshiro et al., 2011). Even more recently, divisive normalization has been observed in regions of the primate brain thought to guide value-based decision-making (e.g. Louie et al., 2014; Yamada et al., 2018; Zimmermann et al., 2018) and the comparison of attribute values in multi-attribute choice (Hunt et al., 2014).

Complementing the neural evidence, divisive normalization has been shown to predict perceptual and value-based judgments (e.g. Xing and Heeger, 2001; Furl, 2016), as well as empirically-observed aspects of choice behavior, particularly as it relates to variability in choice and violations of the axiom of Independence of Irrelevant Alternatives (e.g. Louie et al., 2013; Itthipuripat et al., 2015; Khaw et al., 2017; Webb et al., 2020a, 2020b). From a normative standpoint, Steverson et al. (2019) demonstrate that divisive normalization optimizes the tradeoff between the cost of stochastic choice errors and information processing costs (modeled through Shannon Entropy) within a rational inattention framework. Unlike the present study, however, these previous behavioral investigations of divisive normalization have focused on choice environments where alternatives are represented by a single, integrated value, as opposed to multiple attribute values.[6] It has therefore remained an open question whether divisive normalization — an inherently context-dependent computation — might explain the patterns of context-dependent behavior in multi-attribute choice settings where behavioral research on context-dependence has primarily focused.

A distinct but related notion of "range normalization" has been applied to two-attribute choice (Soltani et al., 2012), and its behavioral predictions are shown in Table 1. In its simplest form, the range normalization computation expresses the value of a relative to a set of reference values as $\frac{a}{a_{\max}-a_{\min}}$, so that a scales linearly between the maximum ($a_{\max}$) and minimum ($a_{\min}$) values in the set. Range normalization was originally proposed to account for evidence that the neural response to a choice alternative varies with the range of values among all available alternatives (Padoa-Schioppa, 2009), though more recent studies suggest this response is not strictly linear (Rustichini et al., 2017). While range-dependence is undoubtedly a feature of neural processing, it can also arise from other computational approaches, including those founded on the principle that neural processing resources should be allocated to prioritize the ability to distinguish stimuli that are encountered more frequently (Rayo and Becker, 2007; Woodford, 2012; Robson and Whitehead, 2018; Frydman and Jin, 2020). As shown in Appendix B, (pairwise) divisive normal-

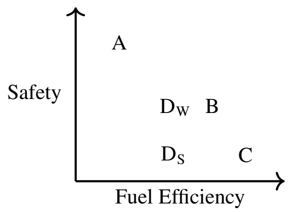
[5] See, for example, Schwartz and Simoncelli (2001), Wainwright et al. (2002), Sinz and Bethge (2013), and Qamar et al. (2013). Efficiency in sensory coding, dating back to Barlow (1961), is defined as the reduction of mutual information in the activity of two neurons. In visual processing, neural activity tends to be correlated for a simple reason: stimuli (photons) from nearby regions of visual space tend to be correlated. By scaling the activity of a neuron in relation to the activity of a neighboring neuron, divisive normalization reduces these correlations. In image compression, for example, an algorithm based on divisive normalization outperforms the standard JPEG and JPEG2000 algorithms (Ballé et al., 2017).

[6] Another difference between our model and previous work is that we implement a *pairwise* form of divisive normalization, as opposed to normalizing with respect to the sum of all relevant values (e.g. $\frac{a}{a+b}+\frac{a}{a+c}$ instead of $\frac{a}{a+b+c}$). That said, the simple "$\frac{a}{a+b+c}$" computation is considered in Section 8.1, though it does not fare well in explaining observed multi-attribute choice behavior (e.g. predicting the opposite of the compromise and dominance effects). More elaborate forms of this computation with richer empirical properties are examined by Daviet (2018) and Dumbalska et al. (2020). See Webb et al. (2020a) for a discussion of why these two different forms of normalization might arise at different stages of the choice process in the human brain.

ization can similarly adapt to a uni-modal distribution of stimuli and accommodate behavioral evidence of range-dependence.

2.2. *Review of behavioral patterns in Table 1*

(I) *Compromise Effect.* The 'compromise effect' refers to the tendency for decision-makers to show a stronger preference for an alternative if it is presented as the middle option on each dimension (Simonson, 1989; Geyskens et al., 2010). For example, if car A is safer but less efficient than car B, someone who chooses A to B in binary choice may instead choose B when a third car C is included that is even less safe and more efficient than B (see Fig. 1).



While car A may be chosen over car B in binary choice, car B may be chosen with car C in the choice set, reflecting a compromise effect, or with some car D, whether weakly (D_W) or strictly (D_S) dominated by car B, reflecting a dominance effect.

Fig. 1. Illustration of compromise and dominance effects.

(II) *Dominance Effect.* The '(asymmetric) dominance effect,' also known as the 'attraction' or 'decoy' effect, refers to the tendency to show a stronger preference for an alternative when presented with a 'decoy' that is worse on each dimension (e.g. Huber et al., 1982; Doyle et al., 1999). Though sometimes demonstrated with weakly-dominated decoys that match the dominant alternative on its weaker attribute (e.g. Kivetz et al.'s, 2004b, economist subscription study), the dominance effect appears stronger when the range on this dimension is expanded (Huber et al., 1982; Soltani et al., 2012; Dumbalska et al., 2020). This 'decoy-range effect' suggests a preference reversal from car A to car B is more likely with the strictly dominated decoy D_S than with the weakly dominated D_W.

(III) *Relative Difference Effect.* The 'relative difference effect' refers to the tendency to treat a difference between small values as if it were greater than an equal-sized difference between large values. For example, people are often willing to drive twenty minutes to save \$5 on a \$15 calculator, but not to save \$5 on a \$125 jacket (Kahneman and Tversky, 1984) — a finding that has since been confirmed and generalized (see, for example, Mowen and Mowen, 1986, and Frisch, 1993, as well as Azar, 2008, for a more extensive review).

(IV) *Majority-Rule Preference Cycles.* Suppose each of three potential alternatives is best on one dimension, second best on another, and worst on a third as follows:

	Attribute 1	Attribute 2	Attribute 3
Alternative A	Best	Middle	Worst
Alternative B	Middle	Worst	Best
Alternative C	Worst	Best	Middle

As first shown by May (1954), binary choices among three such alternatives often exhibit a 'majority-rule preference cycle' (e.g. A preferred to B, B to C, yet C to A) whereby each alternative is preferred to that for which it is better on two of three attributes.[7]

(V) *Splitting Bias.* The 'splitting bias' refers to the tendency to place more (cumulative) weight on an attribute when it is split into two subattributes. As one example from Weber et al.'s (1988) study, job applicants weighted "income" of a potential job more heavily if the attribute was decomposed into "starting salary" and "future salary increases."[8]

(VI) *Alignability Effect.* The 'alignability effect' refers to the tendency to place more weight on an attribute that is 'alignable' in the sense that it is present (though not necessarily equal) for all alternatives (Markman and Medin, 1995; Zhang and Markman, 1998; Gourville and Soman, 2005). For example, when considering a 1000-watt microwave or a 1100-watt microwave, one of which has a moisture sensor and the other an adjustable-speed turntable, the alignability effect implies that the wattage difference may be overweighted relative to the other, nonalignable features.[9]

(VII) *Diversification Bias.* The 'diversification bias' refers to the tendency to disproportionately favor equal allocations of an asset or resource across its components. For instance, trick-or-treaters often select a mixed bundle with two different candy bars over a bundle with two of the same kind, despite selecting the same kind of candy bar in two sequential choices (Read and Loewenstein, 1995). Analogously, diversified gambles are often preferred to undiversified gambles that stochastically dominate the former (Rubinstein, 2002), while investors often favor savings plans that allocate contributions equally across possible funds (Benartzi and Thaler, 2001; Bardolet et al., 2007).[10]

(VIII) *Feature Bias.* The '(extra) feature bias' refers to the tendency to overvalue products with the most features. For example, demand for a video game rises substantially after the development of a new "button" or "scrollbar" control, despite buyers' negligible usage of the new feature (Meyer et al., 2008), while buyers commonly report dissatisfaction, stress, and anxiety with many-feature products after purchase (Thompson et al., 2005; Mick and Fournier, 1998). Supply-side responses to the feature bias appear to be common in light of the widely-noted proliferation of products with an excessive number of features — a trend known as "feature bloat" or "feature creep" (Thompson and Norton, 2011).

[7] In May's experiment, 17 of 62 subjects exhibited this preference ordering among hypothetical spouses, while no subjects exhibited the opposite 'minority-rule' cycle. In a recent study with alternatives designed to put subjects on the cusp of indifference, Tsetsos et al. (2016) show that majority-rule preference cycles can even be more common than transitive preferences. Also see Russo and Dosher (1983) and Zhang et al. (2006) for similar evidence.

[8] See Weber and Borcherding (1993) for a brief review of this literature, as well as Jacobi and Hobbs (2007) and Hamalainen and Alaja (2008) for more recent evidence. Two direct analogs (or special cases) of the splitting bias are the 'event-splitting' (or 'coalescing') effect, which refers to the tendency to value a probabilistic reward more if the event for which the reward is attained is described as two sub-events (Starmer and Sugden, 1993; Humphrey, 1995; Birnbaum and Bahra, 2007), and the 'part-whole bias,' which refers to the tendency to value a good more when its components are evaluated separately than when evaluated holistically (Kahneman and Knetsch, 1992; Bateman et al., 1997).

[9] Similarly, individuals tend to weight alignable attributes more heavily when alternatives are evaluated jointly rather than separately. For example, Hsee et al. (1999) find that a complete 24-piece dinnerware set is often rated more favorably than an incomplete 31-piece set when the sets are separately rated, but not when they are jointly rated.

[10] As noted in these studies, investors favoring equal allocations across funds will end up investing more (or less) in stocks than in bonds simply because the available plans happen to include a greater (smaller) number of stock funds than bond funds.

3. Basic theoretical model

The basic PN model is cast in a standard multi-attribute choice setting, featuring a single decision-maker (DM) who faces a choice set X. Each alternative $\mathbf{x} \in \mathbb{R}^N_+$ is defined on $N > 0$ attribute dimensions, where $x_n \geq 0$ denotes $\mathbf{x}$'s unnormalized attribute value on dimension n, with $x_n = 0$ for the case in which $\mathbf{x}$ provides nothing on dimension n.[11] While we primarily interpret each x_n as the utility-transformed value of $\mathbf{x}$ on dimension n, these attribute values can alternatively be thought of as observable values (with little bearing on our results). The DM's valuation of $\mathbf{x}$ is given by:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} \sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{x_n}{x_n + y_n}, \tag{1}$$

where the DM chooses $\mathbf{x}$ from X if $V(\mathbf{x}; X) > V(\mathbf{y}; X)$ for all $\mathbf{y} \in X \setminus \mathbf{x}$, and is indifferent between $\mathbf{x}$ and $\mathbf{y}$ if $V(\mathbf{x}; X) = V(\mathbf{y}; X)$. For ease of exposition, whenever we specify that either $\mathbf{x}$ or $\mathbf{y}$ is chosen from a choice set that includes both alternatives, it will be implicit that the DM is *not* indifferent (unless otherwise stated). In addition, since $\frac{x_n}{x_n+y_n}$ is undefined if $x_n = y_n = 0$ (a case that will not be relevant to our analysis), we assume that, for all $n \leq N$, there is at most one $\mathbf{x} \in X$ with $x_n = 0$.

The valuation in (1) can be thought of as arising from a series of pairwise comparisons, where each of $\mathbf{x}$'s attribute values are normalized in relation to the corresponding attribute value of each other alternative $\mathbf{y} \in X \setminus \mathbf{x}$. That is, when 'compared' to $\mathbf{y}$, the normalized value of $\mathbf{x}$ on dimension n is simply $\frac{x_n}{x_n+y_n}$, while the overall valuation of $\mathbf{x}$ is the sum of all such terms.[12] Here it is implicit that the DM attends to all attributes of all alternatives when computing $V(\mathbf{x}; X)$.[13] In a setting where there are too many attributes and/or alternatives to realistically attend to all attributes of all alternatives, X and $n = 1, \dots, N$ could instead represent the subsets of alternatives and attributes that are attended to (e.g. Noguchi and Stewart, 2018), though our model is agnostic as to how attention would be allocated in such environments.

We begin with the model in (1) — which does not entail any additional parameters in relation to the standard additive model, $V^+(\mathbf{x}) = \sum_n x_n$ — to provide a clear illustration of the behavioral consequences of pairwise normalization in its most elemental form. Notably, the basic PN model lacks additional degrees of freedom found in prevailing multi-attribute choice models that either

---

11 While this restriction implies a cardinal scale, the model *is* invariant to the rescaling of attribute values by a positive constant (since $\frac{k \cdot a}{k \cdot a + k \cdot b} = \frac{a}{a+b}$).

12 Note, unnormalized attribute values are separable across dimensions in that the total valuation $V(\mathbf{x}, X)$ is the sum (over n) of the attribute-level valuations (each $\sum_y \frac{x_n}{x_n+y_n}$). This assumption aligns with additive separability assumptions throughout the theoretical literature (e.g. Tversky and Simonson, 1993; Bordalo et al., 2013; Koszegi and Szeidl, 2013; Bushong et al., 2019), and allows us to consider the standard additive model, $V^+(\mathbf{x}) = \sum_n x_n$, as a benchmark in the absence of pairwise normalization (V^+ is also typically regarded as a candidate for representing welfare, but this interpretation is not necessary for our analysis). Also note, the model would imply equivalent choice behavior if the right side of (1) was multiplied by $\frac{1}{||X||-1}$, in which case $V(\mathbf{x}; X)$ would reflect the average (rather than the total) valuation of $\mathbf{x}$ arising from pairwise comparisons to each of the $||X|| - 1$ other alternatives in the choice set. Furthermore, the model can readily accommodate an attribute, such as price, for which larger values are less desirable by subtracting (instead of adding) the normalized attribute value.

13 This does not prevent the model from addressing the context-dependent behaviors described in Section 2.2. For example, in Noguchi and Stewart's (2014) analysis of the dominance and compromise effects, all attributes of all alternatives are typically attended to in two-attribute, three-alternative choice sets.

use general attribute weighting functions or parameterized functional forms (see Section 8 and Appendix C). Of course, the broader use of this (or any other) functional-form model that does not add parameters to the standard additive model will be limited by its rigidity. We therefore consider an extension of the model with one additional parameter in Section 7, while Daviet and Webb (2020) propose a parametric specification of subjective attributes for empirical analysis.

3.1. Attribute values: units and observation

While each x_n may be an observable value in a specific dimension, we primarily interpret x_n as the utility-transformed value of alternative $\mathbf{x}$ on attribute n. Other theories often make this more explicit by expressing attribute values through utility functions. Importantly, the behavioral propositions of the model do *not* require attribute values to be fully observable. The propositions do, however, take as given the DM's *ranking* of alternatives on each dimension.[14] For an attribute that is observed as a numerical consumption level, these rankings would naturally be inferred from the rankings of the observed consumption levels. For non-numerical attributes, the rankings of attribute values are typically considered self-evident in motivating behavioral studies. As one example, in Weber et al.'s (1988) study of the splitting bias, subjects presumably favored "very high job security" over a "very high risk of losing the job" in their evaluations of hypothetical jobs. In cases where they are not so obvious, these rankings may be elicited by offering binary choices among alternatives that differ solely on this dimension.

3.2. Preliminary analysis: binary choice with two attributes

Our analysis of the basic PN model first considers two-attribute binary choice. In this setting, the DM's choice can be represented by a multiplicative model, $V^{\mathrm{CD}}(\mathbf{x}) = x_1 x_2$, reminiscent of symmetric Cobb-Douglas preferences (with attribute values as inputs).

Observation 1. *Given $N = 2$, (1) implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) = \frac{2(x_1 x_2 - y_1 y_2)}{(x_1 + y_1)(x_2 + y_2)}$. Therefore the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}\}$ if and only if $x_1 x_2 > y_1 y_2$.*

Much of our subsequent analysis builds on the two-attribute binary-choice problem addressed in Observation 1. Except where otherwise noted, we will assume that $\mathbf{x}$ is stronger on the first attribute and $\mathbf{y}$ is stronger on the second, $x_1 > y_1$ and $x_2 < y_2$, ensuring the choice between $\mathbf{x}$ and $\mathbf{y}$ is nontrivial. It will also be useful to work from a benchmark of binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$, as in the next result.

Proposition 1. *Suppose the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$. Then, given $\mathbf{x}' = (x_1, x_2 + k)$ and $\mathbf{y}' = (y_1, y_2 + k)$ with $k > 0$, the DM chooses $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{y}'\}$.*

[14] Some propositions will feature conditions on attribute values besides those that specify the DM's rankings of alternatives on a given dimension. These additional conditions all serve the same general purpose: to ensure that the unnormalized value of a given "entity" is invariant to how it is *framed* or otherwise represented (in terms of the alternative and/or attributes on which it is expressed). Appendix B.7 presents alternate versions of our main behavioral propositions in which the original conditions on attribute values are instead placed on observable consumption levels, and provides a detailed summary of the additional conditions mentioned above.

Proof. From Observation 1, $\mathbf{x}'$ is chosen from $X = \{\mathbf{x}', \mathbf{y}'\}$ if and only if $x_1(x_2+k) > y_1(y_2+k)$. Noting $x_1x_2 = y_1y_2$ given binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$, this inequality reduces to $x_1k > y_1k$, which must hold since $x_1 > y_1$ and $k > 0$. $\square$

Identical improvements to $\mathbf{x}$ and $\mathbf{y}$ on the same dimension (formalized as equal increases in their unnormalized attribute values) therefore break the DM's initial indifference in favor of the alternative that is weaker on that dimension. This captures previously-cited evidence of the relative difference effect.[15] In contrast, most prevailing theories predict that the DM would remain indifferent. The exception is Bordalo et al.'s (2013) theory, which allows the DM's indifference to break in either direction, depending on a variety of factors such as the extent to which the alternatives are improved (i.e. the magnitude of $k > 0$, in our notation; see Appendix C.3 for details).

To aid our interpretation of Proposition 1 (and several later results), we define

$$\Delta(a, b) \equiv \left|\frac{a-b}{a+b}\right|, \tag{2}$$

which provides a metric of the perceptual "distance" or *contrast* between two values, after each has been normalized in relation to the other.[16]

Observation 2. *Given $N = 2$, (1) implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) = \Delta(x_1, y_1) - \Delta(y_2, x_2)$. Therefore the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}\}$ if and only if $\Delta(x_1, y_1) > \Delta(y_2, x_2)$.*

Thus, for our two-attribute binary-choice problem, $\mathbf{x}$ will be chosen over $\mathbf{y}$ if and only if there is greater contrast on the first dimension (where $\mathbf{x}$ has an advantage) than on the second dimension (where $\mathbf{y}$ has an advantage).

Given this link between choice and contrast, Proposition 1 can be understood as arising from a key property of Δ, *diminishing sensitivity*, whereby increasing two input values by the same amount decreases the perceived distance between them: in this case, $\Delta(y_2 + k, x_2 + k) < \Delta(y_2, x_2)$. The notion that diminishing sensitivity may be important in understanding perceptions of value in multi-attribute choice was previously highlighted by Bordalo et al. (2013).[17] Along similar lines, we may also regard Proposition 1 as a choice analog of Weber's (1834) Law of Perception, which describes how increasing the intensities of two stimuli diminishes the perception of their difference — for example, a one-gram difference in the weights of two rocks is more easily detected if the rocks weigh 1 gram and 2 grams than if they weigh 101 grams and 102 grams.

---

[15] Here, the magnitude of $k > 0$ does not need to be observable, as the analyst only needs to know that x_2 and y_2 increase by the same amount, or equivalently, that the differences on this dimension are the same among the modified alternatives and among the original alternatives ($y_2' - x_2' = y_2 - x_2 > 0$). This condition adheres to standard interpretations of relative difference effects, which are considered anomalous precisely because they suggest differential weighting of equal-sized differences in value. See Appendix B.7 for another (somewhat trivial) version of Proposition 1, in which the identical improvements are represented as equal increases in numerical consumption levels (as opposed to utility-transformed values).

[16] This definition of Δ — which parallels the standard conceptualization of contrast in the visual perception literature (Carandini and Heeger, 2012) — qualifies as a metric (distance) function because it satisfies: (a) $\Delta(a, b) \geq 0$ for all a, b; (b) $\Delta(a, b) = 0$ if and only if $a = b$; (c) $\Delta(a, b) = \Delta(b, a)$; and (d) $\Delta(a, c) \leq \Delta(a, b) + \Delta(b, c)$ (this last property, the Triangle Inequality, is addressed in Section 5).

[17] Since Δ exhibits "ordering" — i.e. $\Delta(a', b') > \Delta(a, b)$ given $a' > a > b > b' \geq 0$ — in addition to diminishing sensitivity, it constitutes a salience function as defined by Bordalo et al. (2013).

4. Adding a third alternative to the choice set

We now examine how a choice between $\mathbf{x}$ and $\mathbf{y}$ may be impacted by a third alternative $\mathbf{z}$. To aid our understanding, let $\mathbf{m}^{\mathrm{xy}} \equiv \left(\frac{x_1+y_1}{2}, \frac{x_2+y_2}{2}\right)$ denote the midpoint between $\mathbf{x}$ and $\mathbf{y}$. Noting $\frac{x_1}{x_2} > \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}} > \frac{y_1}{y_2}$ since $\mathbf{x}$ is better than $\mathbf{y}$ on the first dimension and worse on the second, we will say that $\mathbf{z}$ is more *similar* to $\mathbf{x}$ than to $\mathbf{y}$ if and only if $\frac{z_1}{z_2} > \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$, in which case $\mathbf{z}$'s attribute values are tilted towards $\mathbf{x}$ and away from $\mathbf{y}$ in relation to $\mathbf{m}^{\mathrm{xy}}$.

Lemma 1. *Suppose the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$, and $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$. Then:*

$$|V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\})| \geq |V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\})|, \tag{3}$$

which only binds when both sides are zero.

Proof of Lemma 1. See Appendix.

Given $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$, and with indifference between $\mathbf{x}$ and $\mathbf{y}$ in binary choice, Lemma 1 indicates that the magnitude of the perceived value difference will be larger between $\mathbf{z}$ and $\mathbf{x}$ than between $\mathbf{z}$ and $\mathbf{y}$. In this way, pairwise normalization makes it "easier to compare" more similar alternatives, as proposed by Tversky and Russo (1969) and operationalized by Natenzon (2019).

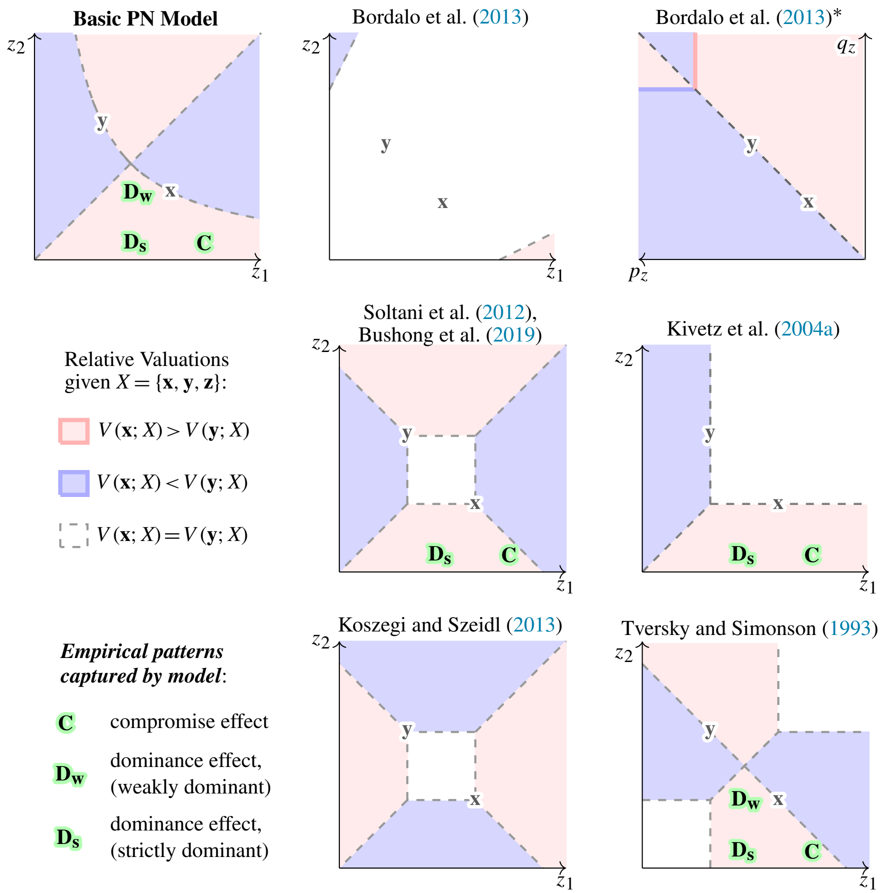
Since the relative ranking of $\mathbf{x}$ and $\mathbf{y}$ is generally not observed in cases where $\mathbf{z}$ is chosen in trinary choice, the effect of an *inferior* $\mathbf{z}$ that is not chosen over $\mathbf{x}$ and $\mathbf{y}$ is of particular interest.[18] In such cases, both value differences in (3) would be positive even without taking absolute values. Thus, with an inferior $\mathbf{z}$ that is more similar to $\mathbf{x}$ than to $\mathbf{y}$, and given binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$, Lemma 1 implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) > 0$, which itself is equivalent to $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$, implying the DM would choose $\mathbf{x}$ over $\mathbf{y}$ with $\mathbf{z}$ in the choice set.[19] In effect, through the relative ease of comparing more similar alternatives, an inferior $\mathbf{z}$ enhances the DM's perception of $\mathbf{x}$ more than it enhances the perception of $\mathbf{y}$, causing $\mathbf{x}$ to be chosen. This yields the well-known compromise and dominance effects.

Proposition 2. *Suppose the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$. Then the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ in each of the following scenarios:*

(i) $\mathbf{x}$ *is a compromise between* $\mathbf{y}$ *and* $\mathbf{z}$ *(i.e. $z_1 > x_1 > y_1$ and $y_2 > x_2 > z_2$), and* $\mathbf{z}$ *is not chosen.*
(ii) $\mathbf{x}$ *asymmetrically dominates $\mathbf{z} \neq \mathbf{x}$ (i.e. $x_1 \geq z_1 > y_1$ and $y_2 > x_2 \geq z_2$).*

[18] Formally, $\mathbf{z}$ is inferior if $V(\mathbf{z}; X) < V(\mathbf{x}'; X)$ for $\mathbf{x}' \in \{\mathbf{x}, \mathbf{y}\}$ and $X = \{\mathbf{x}', \mathbf{z}\}$, meaning $\mathbf{x}$ and $\mathbf{y}$ are chosen over $\mathbf{z}$ in binary-choice. In Appendix B.4, we show that, given indifference between $\mathbf{x}$ and $\mathbf{y}$ in binary choice, $\mathbf{z}$'s status as inferior would be the same if it were instead defined in terms of trinary choice with $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$. A stochastic choice variant of the "ease of comparison" result is formalized in Appendix B.1.

[19] This equivalence follows because $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{z}\}) = \sum_n \frac{x_n'}{x_n'+z_n} = \frac{1}{2} \cdot \sum_n \frac{2x_n'+z_n-z_n}{x_n'+z_n} = \frac{1}{2} \cdot \sum_n \left(\frac{x_n'}{x_n'+z_n} + 1 - \frac{z_n}{x_n'+z_n}\right) = \frac{1}{2}(V(\mathbf{x}'; \{\mathbf{x}', \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{x}', \mathbf{z}\})) + 1$ under (1), along with $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = (V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) + V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})) - (V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) + V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})) = V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$ given indifference between $\mathbf{x}$ and $\mathbf{y}$ in binary choice.



Each graph shows the effect of **z** on the relative valuation of **x** and **y**, as predicted by the indicated model (note, this comparison determines the DM's choice, as long as **z** is not chosen from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$). With one exception (see *), the graphs were created using $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$ as a simple illustration that ensures binary-choice indifference in all models considered. Additional parametric and functional form restrictions needed to create the graphs are described in Appendix D.

* While Bordalo et al.'s (2013) model can be analyzed for alternatives defined by two quality attributes (Bordalo et al., 2013, Appendix B), it is primarily analyzed for alternatives defined by their price and a single quality attribute. For this reason, both model variations are considered here, where we use $\mathbf{x} = (p_x, q_x) = (1, 1)$ and $\mathbf{y} = (p_y, q_y) = (2, 2)$ to create the 'price-quality' graph.

Fig. 2. The effect of adding **z** to the choice set. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Proof of Proposition 2. See Appendix.

To understand how pairwise normalization generates the compromise effect in part (i) of Proposition 2, consider the top-left graph in Fig. 2. If **z** is on the same side of the dashed line projecting from the origin as **x**, it is more similar to **x** than to **y**. This must be true in cases where **z** makes **x** a compromise between **y** and **z**, because such a **z** would be below and to the right of **x**; we can also algebraically verify that **z** must be more similar to **x** than to **y** from the fact that $z_1 > x_1$ and $z_2 < x_2$ imply $\frac{z_1}{z_2} > \frac{x_1}{x_2} > \frac{m_1^{xy}}{m_2^{xy}}$. Suppose, in addition, **z** is inferior — i.e. below the dashed, binary-choice indifference curve on which **x** and **y** reside, as in the point labeled 'C' — and hence *not* chosen in trinary choice. Then **z** will enhance the DM's perception of the compromise alternative **x** more than it enhances the perception of **y**, due to the relative ease of comparing more similar alternatives (Lemma 1). This leads the DM to choose **x** with **z** in the choice set.

Likewise, if **z** is asymmetrically dominated by **x** (e.g. the point labeled 'D_s'), it too must be more similar to **x** than to **y** — since $z_1 > y_1$ and $z_2 < x_2$ imply $\frac{z_1}{z_2} > \frac{y_1}{x_2} = \frac{y_1(x_2+y_2)}{x_2(x_2+y_2)} = \frac{y_1x_2+y_1y_2}{x_2(x_2+y_2)} = \frac{y_1x_2+x_1x_2}{x_2(x_2+y_2)} = \frac{x_2(x_1+y_1)}{x_2(x_2+y_2)} = \frac{x_1+y_1}{x_2+y_2} = \frac{m_1^{xy}}{m_2^{xy}}$, given $x_1x_2 = y_1y_2$ with binary-choice indifference (Observation 1) — and inferior to both. Again, **z** enhances the perception of the more similar (and in this case, dominant) alternative **x** more than it enhances the perception of **y**, leading the DM to choose **x**. The dominance effect, as captured in part (ii) of Proposition 2, thus arises through the same ease-of-comparison property as the compromise effect. As illustrated in Fig. 2, Tversky and Simonson's (1993) model likewise captures the compromise effect and the dominance effect with both weakly and strictly dominated decoys, while the remaining models do not.

4.1. The "strength" of the compromise and dominance effects

Our next result describes how changing **z**'s position in attribute space can strengthen or weaken its effect on the DM's choice between **x** and **y**.

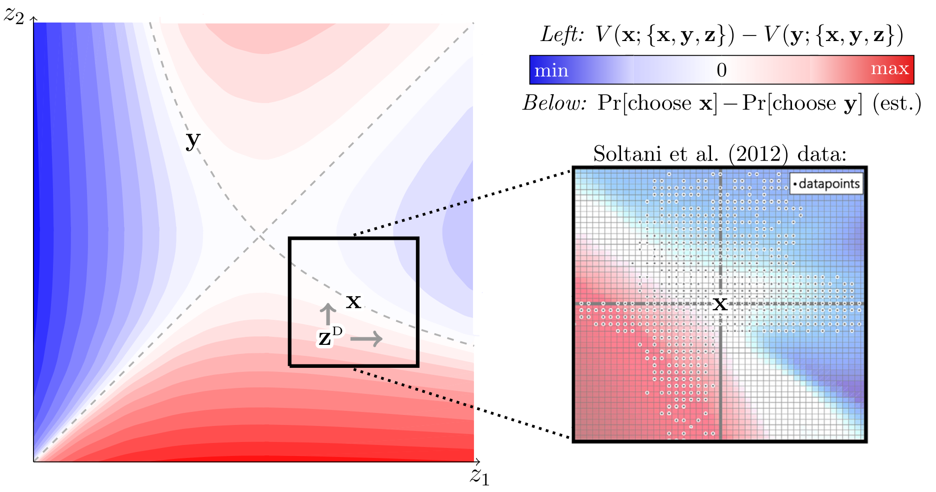
Proposition 3. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *with* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$*, but chooses* **y** *from* $X = \{\mathbf{x}, \mathbf{y}\}$*. Then:*

(i) *if* $x_1 > z_1 = z_1' > y_1$ *and* $y_2 > x_2 \geq z_2 > z_2'$*, the DM chooses* **x** *from* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}$*;*
(ii) *if* $z_1 > x_1 > z_1' > y_1$ *and* $y_2 > x_2 > z_2 = z_2'$*, the DM chooses* **x** *from* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}$*;*
(iii) *if* $x_1 > z_1 > y_1 = z_1'$ *and* $y_2 > x_2 > z_2 = z_2'$*, the DM chooses* **y** *from* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}$*.*

Proof of Proposition 3. Using $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ and (1), $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}$ implies $V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{(z_2-z_2')(y_2-x_2)(x_2y_2-z_2z_2')}{(x_2+z_2)(x_2+z_2')(y_2+z_2)(y_2+z_2')} > 0$ in (i), $V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{(z_1-z_1')(x_1-y_1)(z_1z_1'-x_1y_1)}{(x_1+z_1)(x_1+z_1')(y_1+z_1)(y_1+z_1')} > 0$ in (ii), and $V(\mathbf{y}; X) - V(\mathbf{x}; X) = \frac{(z_1-y_1)y_1(x_1-y_1)(x_1-z_1)}{2y_1(y_1+z_1)(x_1+y_1)(x_1+z_1)} > 0$ in (iii). Since these expressions are positive (given the indicated restrictions), **x** is chosen from X in (i) and (ii), while **y** is chosen from X in (iii). □

Part (i) of Proposition 3 first considers a variation of the dominance effect whereby the DM chooses **y** over **x** in binary choice, but becomes indifferent when an asymmetrically dominated alternative **z** is included in the choice set. A modified decoy **z**′, which is worse than **z** on its second dimension (and the same on its first dimension), will then create a stronger dominance effect as **x** becomes strictly favored (and chosen) over **y** with **z**′ in the choice set. This prediction fits with evidence of a decoy-range effect (see Section 2.2), in which the dominance effect becomes more prominent when the decoy becomes worse on the dimension for which it is the weakest.

Part (ii) analogously considers a variation of the compromise effect whereby the DM again chooses **y** over **x** in binary choice, but is indifferent when the choice set includes a third alternative **z** that, in this case, makes **x** a compromise (i.e. intermediate on both dimensions). In turn,



The graph on left re-illustrates the predicted effect of **z** on the relative valuations of **x** and **y**, as originally shown in Fig. 2, except here the regions are shaded based on the magnitude of the normalized value difference between **x** and **y** with **z** in the choice set. The inset on right provides a visualization of choice data from Soltani et al.'s (2012) experiment, showing **z**'s effect on the relative likelihood of choosing the more similar alternative **x** relative to the less similar alternative **y** (which were calibrated to reside on the same binary-choice indifference curve). The choice data at each location (pooled over all subjects) is smoothed using a locally-weighted linear regression with smoothing parameter $1/3$.

Fig. 3. The "strength" of **z**'s effect.

an asymmetrically dominated decoy $\mathbf{z}'$, which is the same as **z** on the second dimension but now worse than both **x** and **z** on the first dimension, causes the DM to choose **x** over **y**. In this sense, the model predicts that the dominance effect is "stronger" than the compromise effect. While additional tests would be useful, a recent experiment by Noguchi and Stewart (2018) provides evidence for this prediction.

Part (iii) of Proposition 3 then considers the variation of the dominance effect from part (i), whereby the addition of an asymmetrically dominated decoy **z** makes the DM indifferent between **x** and **y** in trinary choice, despite choosing **y** over **x** in binary choice. A modified decoy $\mathbf{z}'$ is then considered, which is equal to **y** on its first dimension. With $\mathbf{z}'$ now weakly dominated by **y** (and still strictly dominated by **x**), the DM reverts to choosing **y** in trinary choice. The effect of this "symmetrically dominated" decoy is therefore weaker than the dominance effect, which is broadly consistent with experimental evidence from Wedell (1991). That said, Wedell does not observe a significant effect of adding a symmetrically dominated alternative of this form, as would be predicted by Tversky and Simonson's (1993) model, while the basic PN model still predicts a shift in the DM's valuation in favor of **x** relative to **y**.

To help illustrate how **z**'s location in attribute space determines the direction and extent to which it affects the DM's perception of **x** and **y**, Fig. 3 reproduces the graph in Fig. 2 depicting the predictions of the basic PN model, except the regions are now shaded based on the magnitude of the difference between the normalized valuations of **x** and **y** in trinary choice. The gray arrows indicate that an asymmetrically-dominated decoy $\mathbf{z}^D$ enhances the perception of **x** more than it enhances the perception of **y**, but this effect weakens as z_2^D increases — effectively shrinking the range $(y_2 - z_2^D)$ of values on this dimension, as in the decoy-range effect — and also as z_1^D increases to a point where **x** no longer dominates $\mathbf{z}^D$, becoming a compromise instead.

For comparison, the inset in Fig. 3 depicts the estimated difference in the choice probabilities of $\mathbf{x}$ and $\mathbf{y}$ as a function of $\mathbf{z}$'s location in attribute space using choice data from Soltani et al. (2012). In this experiment, the locations of $\mathbf{x}$ and $\mathbf{y}$ were chosen so that each subject was just indifferent between them in binary choice. The introduction of $\mathbf{z}$ necessarily revealed subjects' rankings among $\mathbf{x}$ and $\mathbf{y}$ because $\mathbf{z}$ was only a "phantom" alternative that was presented with $\mathbf{x}$ and $\mathbf{y}$, but could not actually be chosen. While only suggestive, the observed patterns align with the model's predictions that increasing z_1^{D} and z_2^{D} would weaken the effect of $\mathbf{z}^{\mathrm{D}}$ in shifting perceptions in favor of $\mathbf{x}$ relative to $\mathbf{y}$.[20] The phantom design also allows us to consider the effect of a superior $\mathbf{z}$, which would presumably be chosen over $\mathbf{x}$ and $\mathbf{y}$ if it were feasible. In this case, subjects' perceptions appear to shift in favor of $\mathbf{y}$ instead of $\mathbf{x}$ (see the blue region above and to the right of $\mathbf{x}$), which is qualitatively consistent with the prediction of the basic PN model. This prediction, as well as the effects of changing z_1^{D} and z_2^{D}, is also generated by the multi-attribute choice models of Soltani et al. (2012) and Bushong et al. (2019).

As a caveat to this discussion, it is important to note a discrepancy in the representation of attributes, as the choice alternatives in Soltani et al.'s experiment were lotteries defined by the probability of winning a monetary reward and the amount of the reward. Considering these attributes are *not* separable, it is reasonable to question the value of using this type of lottery data to test the predictions of multi-attribute choice models — including our model, the aforementioned models of Soltani et al. (2012) and Bushong et al. (2019), and the other models in Table 1 — that assume separability (see footnote 12).

5. Binary choice with three attributes

So far, our analysis has only considered choices with alternatives defined on two attribute dimensions. We now consider binary choices among alternatives that vary on three attribute dimensions. Our first such example shows that, with three-attribute choice alternatives, choices can now be intransitive.

Example 1. Suppose $\mathbf{x} = (a, b, c)$, $\mathbf{x}' = (b, c, a)$, and $\mathbf{x}'' = (c, a, b)$ with $a > b > c$. Then:

$$V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) - V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\})$$
$$= V(\mathbf{x}''; \{\mathbf{x}, \mathbf{x}''\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}''\}) = \Delta(a, b) + \Delta(b, c) - \Delta(a, c) = \tfrac{(a-b)(b-c)(a-c)}{(a+b)(b+c)(a+c)} > 0.$$

Thus, the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$, $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{x}''\}$, and $\mathbf{x}''$ from $X = \{\mathbf{x}, \mathbf{x}''\}$.

In Example 1, $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ satisfy a 'cyclical majority-dominance' property whereby $\mathbf{x}$ is better than $\mathbf{x}'$ on two of three attributes, $\mathbf{x}'$ is better than $\mathbf{x}''$ on two of three attributes, and $\mathbf{x}''$ is better than $\mathbf{x}$ on two of three attributes. In turn, the DM exhibits a majority-rule preference cycle, as each alternative is preferred to that for which it is better on two out of three attributes. This particular cycle arises directly from the fact that, as a metric of perceptual distance, the contrast function satisfies the triangle inequality (see footnote 16). That is, if $a > b > c > 0$, then

[20] The representations of these effects in Fig. 3 do not map exactly to the formal statements in Proposition 3. This is because Proposition 3 presumes the DM favors $\mathbf{y}$ over $\mathbf{x}$ in binary choice, while Fig. 3 illustrates strength effects under binary-choice indifference (which facilitates comparisons to Soltani et al.'s experimental data). Nonetheless, the behavioral effects formalized in Proposition 3 are driven by the effects of changing the positions of z_1 and z_2 on the relative value difference between $\mathbf{x}$ and $\mathbf{y}$, as depicted in Fig. 3. A stronger context effect in the sense of Proposition 3 therefore corresponds to the illustration of effect strength based on value differences in Fig. 3.

$$\Delta(a,b) + \Delta(b,c) > \Delta(a,c).$$

This relation implies that, for any two alternatives among $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ in Example 1, the total contrast on the two dimensions for which the majority-dominant alternative has an advantage will be greater than the contrast on the dimension for which the minority-dominant alternative has an advantage.

With three potential alternatives that satisfy the cyclical majority-dominance property, the basic PN model does **not** imply that binary choices *must* be intransitive. Instead, the model predicts that *if* choices are intransitive, they will only be intransitive in one direction. That is, while intransitivity could in principle arise in one of two forms — a majority-rule cycle (as in Example 1), or an opposite 'minority-rule' cycle — the next result clarifies that only majority-rule cycles can arise under pairwise normalization (in line with previously discussed evidence).

Proposition 4. *Given $N = 3$, suppose $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ satisfy $x_1 > x_1' > x_1''$, $x_2'' > x_2 > x_2'$, and $x_3' > x_3'' > x_3$. Then, if binary-choice preferences among $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ are intransitive, it must be the case that the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$, $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{x}''\}$, and $\mathbf{x}''$ from $X = \{\mathbf{x}, \mathbf{x}''\}$.*

Proof of Proposition 4. See Appendix.

Our next result considers the effect of splitting an attribute into two subattributes, effectively re-framing a choice between two-attribute alternatives, $\mathbf{x}$ and $\mathbf{y}$, as a choice between three-attribute alternatives, $\mathbf{x}'$ and $\mathbf{y}'$.

Proposition 5. *Suppose the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ given $X = \{\mathbf{x}, \mathbf{y}\}$. Also suppose $\mathbf{x}' = (x_{1a}, x_{1b}, x_2)$ and $\mathbf{y}' = (y_{1a}, y_{1b}, y_2)$ with $x_{1a} + x_{1b} = x_1$, $y_{1a} + y_{1b} = y_1$, $x_{1a} \geq y_{1a}$, and $x_{1b} \geq y_{1b}$. Then the DM chooses $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{y}'\}$.*

Proof of Proposition 5. The DM chooses $\mathbf{x}'$ if and only if $\Delta(x_{1a}, y_{1a}) + \Delta(x_{1b}, y_{1b}) > \Delta(y_2, x_2)$. Since $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$, $\Delta(y_2, x_2) = \Delta(x_1, y_1) = \Delta(x_{1a} + x_{1b}, y_{1a} + y_{1b})$, which implies the previous condition is equivalent to $\Delta(x_{1a}, y_{1a}) + \Delta(x_{1b}, y_{1b}) > \Delta(x_{1a} + x_{1b}, y_{1a} + y_{1b})$, which itself is equivalent to $\frac{(x_{1a}-y_{1a})(x_{1b}+y_{1b})^2+(x_{1b}-y_{1b})(x_{1a}+y_{1a})^2}{(x_{1a}+y_{1a})(x_{1b}+y_{1b})(x_{1a}+y_{1a}+x_{1b}+y_{1b})} > 0$ and must hold since both terms in the numerator are non-negative and at least one is strictly positive given $x_{1a} \geq y_{1a}$ and $x_{1b} \geq y_{1b}$ (with at most one inequality binding). □

Consistent with evidence of the splitting bias (see Section 2.2), attribute-splitting tilts choice in favor of the alternative that is stronger on the split attribute, in this case $\mathbf{x}$, provided its advantage is maintained on each subattribute.[21] In effect, $\mathbf{x}$'s advantage over $\mathbf{y}$ on attribute 1 is perceived to be larger when spread over two subattributes as result of the fact that the contrast function Δ satisfies the triangle inequality.[22]

[21] Here, the (sub)attribute values do not need to be fully observable. Rather, the analyst only needs to know that the unnormalized value of each alternative does not change if an attribute is re-framed as two attributes (which is consistent with standard interpretations of the splitting bias) and that the ranking of alternatives on the unsplit attribute dimension is not reversed on either subattribute dimension.

[22] This effect is also amplified by diminishing sensitivity in Δ. To illustrate, suppose $x_1 = 6$, $y_1 = 4$, $x_{1a} = x_{1b} = 3$, and $y_{1a} = y_{1b} = 2$. The triangle inequality implies $\Delta(6,4) < \Delta(6,5) + \Delta(5,4)$, while diminishing sensitivity implies $\Delta(6,5) < \Delta(5,4) < \Delta(3,2)$. Thus, $\Delta(6,4) < \Delta(3,2) + \Delta(3,2)$, which means the total contrast between x_{1a} and y_{1a} and between x_{1b} and y_{1b} exceeds the contrast between x_1 and y_1.

Next, we examine the effect of attribute alignability in binary choice. Here, an attribute is considered 'alignable' if the corresponding attribute values are strictly positive for both alternatives. To isolate the effect of alignability, we will work from our benchmark of binary-choice indifference between **x** and **y**, while presuming that both attributes are alignable. We then consider a choice among two modified alternatives with only one alignable attribute.

Proposition 6. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *given* $X = \{\mathbf{x}, \mathbf{y}\}$*. Also suppose* $\min\{x_1, x_2, y_1, y_2\} > 0$ *and let* $\mathbf{x}' = (x_1, x_2, 0)$ *and* $\mathbf{y}' = (y_1, 0, y_2)$*. Then the DM chooses* $\mathbf{x}'$ *from* $X = \{\mathbf{x}', \mathbf{y}'\}$*.*

Proof of Proposition 6. $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{x_1 - y_1}{x_1 + y_1} + \frac{x_2}{x_2} - \frac{y_2}{y_2} = \frac{x_1 - y_1}{x_1 + y_1} > 0.$ □

The choice of $\mathbf{x}'$ over $\mathbf{y}'$ described by Proposition 6 (along with indifference between **x** and **y**) indicates that the advantage $y_2 > x_2$ is weighted more heavily if y_2 and x_2 exist on the same attribute dimension than if they exist on separate (i.e. non-alignable) dimensions. This matches evidence of the alignability effect described in Section 2.2.

6. (Binary) allocation problems

We now consider a choice between two different allocations of an asset (or resource), with total value $A > 0$, across N dimensions, so that a given allocation **x** satisfies $\sum_{n \leq N} x_n = A$. While stylized, this formulation provides a simple baseline that can be related to a variety of allocation problems. For example, A could represent an investor's recurring contribution to a savings plan that includes N different funds or a budget that is spent on consumption bundles defined over N goods. For simplicity, the setup implicitly presumes that allocations generate the same rate of return on all dimensions, though the model's implications with unequal returns will also be discussed.

Proposition 7. *Given* $N > 1$ *and* $A > 0$*, suppose* $x_n = \frac{A}{N}$ *for all* $n \leq N$*. Then, for any* $\mathbf{x}' \neq \mathbf{x}$ *with* $\sum_{n \leq N} x'_n = A$*, the DM chooses* **x** *from* $X = \{\mathbf{x}, \mathbf{x}'\}$*.*

Proof of Proposition 7. See Appendix.

From Proposition 7, a balanced allocation with an equal $\frac{1}{N}$ share of the asset on each dimension will always be chosen over an unbalanced allocation of the asset. This result aligns with evidence of a diversification bias, such as Benartzi and Thaler's (2001) finding that investors often follow a "$\frac{1}{N}$ heuristic" by selecting a balanced savings plan that allocates contributions equally across the N available funds. Note, since we abstract from the possibility of uncertain returns, the elevated valuation of a balanced allocation cannot be rationalized as a variance-reduction strategy and thus represents a "bias" in relation to a standard additive preference model ($V^+(\mathbf{x}) = \sum_n x_n$), which would predict indifference between any two allocations of the same asset.

In Appendix B.2, we show that Proposition 7 still applies in the case where the attribute values of both alternatives on a given dimension n are scaled by a constant "return" $R_n > 0$, with $R_n \neq R_{n'}$ for $n \neq n'$. Thus, if allocations yield higher returns on some dimension than others, the DM will still always choose a balanced allocation over an unbalanced allocation of the asset. This applies even if the latter alternative involves a complete allocation to the dimension that yields the highest return — as would be optimal according to the standard additive model — in which case the interpretation of the predicted diversification bias as a "bias" may be especially evident.

Considering the DM's aversion to unbalanced allocations, it is natural to suspect that the DM would generally favor allocations for which all dimensions receive a positive share of the asset. To explore this idea, suppose two firms previously offered identical products defined on $N - 1 > 0$ dimensions, each of which may be thought of as representing a distinct product feature. However, both firms have since invested $q > 0$ in research and development to improve their products. One firm improved the quality (i.e. unnormalized attribute value) of an existing product feature from $x_{n'} > 0$ to $x_{n'} + q$ on dimension $n' \leq N - 1$. The other firm innovated a Nth product feature, attaining a quality level of $x_N = q$ on this new dimension. As our next result shows, the product with the new feature will now be chosen over the product with the improvement to an existing feature.

Proposition 8. *Given $N > 1$ and $q > 0$, suppose $x_N = q$, $x'_N = 0$, $x'_{n'} = x_{n'} + q$ for some $n' < N$, and $x'_n = x_n > 0$ for all other $n < N$. Then the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$.*

Proof of Proposition 8. $V(\mathbf{x}; X) - V(\mathbf{x}'; X) = \Delta(q, 0) - \Delta(x_{n'} + q, x_{n'}) = \frac{2x_{n'}}{q+2x_{n'}} > 0.$ □

This prediction that a new product feature will be valued more than an otherwise-equivalent improvement to an existing product feature fits with evidence of the feature bias (see Section 2). Here, the feature bias can be understood as a consequence of diminishing sensitivity in Δ. Since the mean attribute value between the two products is higher for the existing feature than for the new feature (i.e. $x_{n'} + \frac{q}{2} > \frac{q}{2}$), the value difference on the new dimension N will, as a result of diminishing sensitivity, be perceived as greater than the equal-sized value difference on the existing dimension n'.

Like Proposition 7, Proposition 8 applies even if allocations yield higher returns on some dimensions than others (see Appendix B.2). This is especially noteworthy if the return to investing q (in terms of the increase in the corresponding attribute value) is lower on the new dimension. In this case, the product with the new feature would be chosen despite its lower overall quality. Thus, in product-level investment decisions, firms would naturally have an incentive to allocate research and development resources towards innovating new features, even if they add little actual value to the product. In this way, pairwise normalization offers a potential explanation for the proliferation of products with an excessive number of features ("feature creep") as well as the related observation that developing "irrelevant" new product features can foster a sustained competitive advantage (Thompson and Norton, 2011; Carpenter et al., 1994).

7. A parametric generalization

Next, we consider a generalization of our model based on a common formulation of the normalization computation in neuroscience whereby an input value a is normalized relative to b as $\frac{a}{\sigma+a+b}$ (e.g. Shevell, 1977; Heeger, 1992; Louie et al., 2011; Carandini and Heeger, 2012). The parameter $\sigma \geq 0$ is a "semi-saturation" constant (in the sensory perception literature) that represents the minimum input intensity at which neural responses are half-saturated (i.e. $\frac{a}{\sigma+a+b} = \frac{1}{2}$ for $a = \sigma$ and $b = 0$). It also appears in recent adaptations of divisive normalization to (single-attribute) economic choice (Steverson et al., 2019; Webb et al., 2020a).

As with the $\frac{a}{a+b}$ model, we adapt the $\frac{a}{\sigma+a+b}$ model to multi-attribute choice through our notion of pairwise, attribute-level comparisons as

$$V^*(\mathbf{x}; X) = \sum_{n \le N} \sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{x_n}{\sigma + x_n + y_n}. \tag{4}$$

In the analysis that follows, we assess the extent to which the predictions of the basic PN model are maintained under (4) with $\sigma > 0$, while also revealing some new behaviors that can arise. In Appendix B.5, we provide a more detailed discussion of how σ affects the perception of attributes and highlight an interpretation of σ as a dynamic reference point based on recent evidence that σ may vary with the intensity of previously-encountered stimuli (LoFaro et al., 2014; Louie et al., 2014; Khaw et al., 2017; Tymula and Glimcher, 2019; Guo and Tymula, 2020).

7.1. Binary choice with two attributes

With two attributes and two alternatives, the DM's choice can now be represented using a composite of the (symmetric) Cobb-Douglas and additive preference models.

Observation 3. *Given $N = 2$, $V^+(\mathbf{x}'') = x''_1 + x''_2$, and $V^{\text{CD}}(\mathbf{x}'') = x''_1 x''_2$ for $\mathbf{x}'' \in \{\mathbf{x}, \mathbf{x}'\}$, $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V^*(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \frac{\sigma(V^+(\mathbf{x}) - V^+(\mathbf{x}')) + 2(V^{\text{CD}}(\mathbf{x}) - V^{\text{CD}}(\mathbf{x}'))}{(\sigma + x_1 + x'_1)(\sigma + x_2 + x'_2)}$. Therefore:*

(i) *Given $V^+(\mathbf{x}) \ge V^+(\mathbf{x}')$, $V^{\text{CD}}(\mathbf{x}) \ge V^{\text{CD}}(\mathbf{x}')$, and $\sigma \ge 0$ with at least two of these inequalities non-binding, the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$.*
(ii) *Given $V^+(\mathbf{x}) > V^+(\mathbf{x}')$ and $V^{\text{CD}}(\mathbf{x}) < V^{\text{CD}}(\mathbf{x}')$, the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$ if and only if $\sigma > \frac{2(V^{\text{CD}}(\mathbf{x}') - V^{\text{CD}}(\mathbf{x}))}{V^+(\mathbf{x}) - V^+(\mathbf{x}')}$.*

Thus, if the Cobb-Douglas (V^{CD}) and additive (V^+) models agree in their rankings among two alternatives, the DM's choice will align with this ranking. If there is disagreement, the DM's choice will align with the additive model's ranking if and only if σ is sufficiently large. Thus, a larger σ effectively implies a larger weight on additive relative to Cobb-Douglas preferences. Compared to the basic PN model (equivalently, Cobb-Douglas), the model with $\sigma > 0$ predicts flatter binary-choice indifference curves (Fig. 4, top right). In the large-σ limit, the binary-choice indifference curves become arbitrarily flat, converging to those generated by the additive model (Fig. 4, bottom right), which also represents binary choices in most prevailing multi-attribute choice theories.[23]

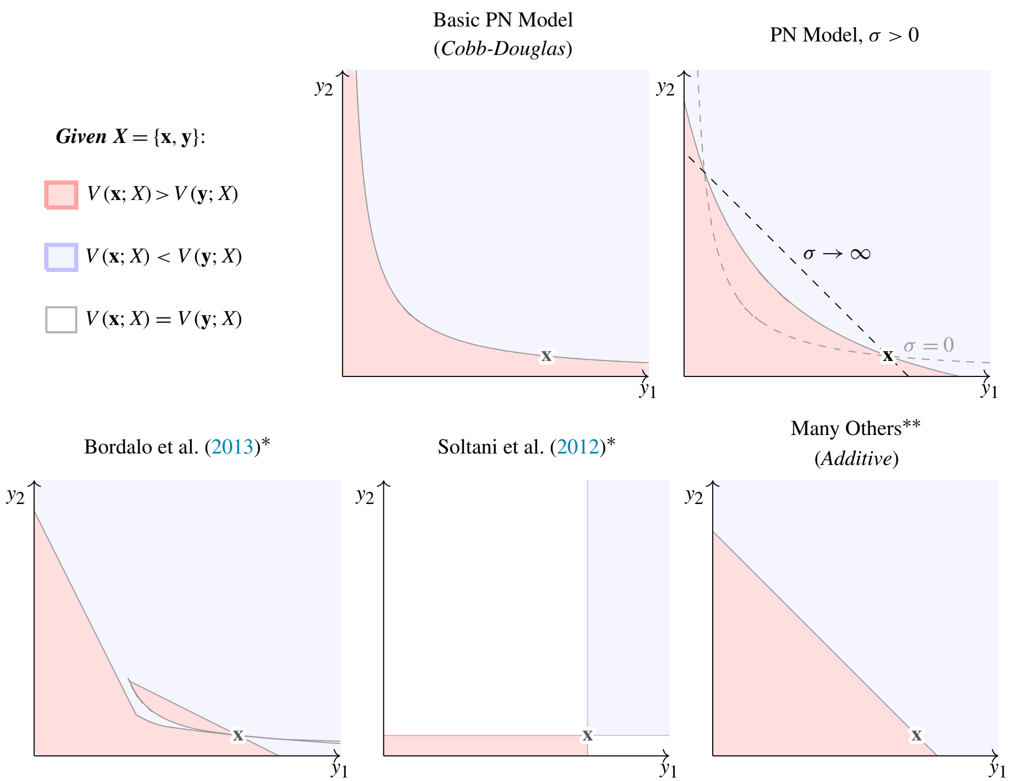
7.2. Robustness of key behavioral predictions

Next, we see that many key predictions of the basic PN model are maintained with $\sigma > 0$.

Proposition 9. *For all $\sigma \ge 0$, the following results still hold under (4):*

(i) *relative difference effect (Prop. 1);*
(ii) *majority-rule cycles (Proposition 4);*
(iii) *the splitting bias (Proposition 5);*
(iv) *the alignability effect (Proposition 6);*
(v) *the diversification bias (Proposition 7);*
(vi) *the feature bias (Proposition 8).*

[23] In addition to Cobb-Douglas and additive preferences, pairwise normalization can — through a second parametric variation of the model — also be related to constant elasticity of substitution (CES) preferences and rank-based lexicographic preferences (Appendix B.5).



The shaded region(s) of each graph indicate the region(s) where $\mathbf{x}$ is chosen from $X = \{\mathbf{x}, \mathbf{y}\}$ with $N = 2$, as predicted by the indicated model(s) with $\mathbf{x} = (1, .1)$.

* Bordalo et al. (2013) and Soltani et al.'s (2012) models can generate various geometric configurations of the binary-choice preference regions. For example, in Bordalo et al.'s model, the shapes of the regions can vary with a 'salience distortion' parameter $\delta \in (0, 1)$ (this illustration uses $\delta = .5$) as well as the choice of $\mathbf{x}$ (e.g. using a different $\mathbf{x}$ on the boundary between the shaded and unshaded regions can yield different preference regions than those shown here). In Soltani et al.'s model, the regions can vary with four (alternative- and attribute-specific) weighting parameters as well as two 'representation factors.' While this model does not (under other parameterizations) necessarily predict indifference in the region above and to the left of $\mathbf{x}$ (where $\mathbf{x}$ is better on the first dimension and $\mathbf{y}$ is better on the second), it does predict that the value difference $V(\mathbf{x}; X) - V(\mathbf{y}; X)$, and hence the relative ranking of $\mathbf{x}$ and $\mathbf{y}$, is fixed throughout this region. For an exact description of the models used to generate these graphs, see Appendix C and Appendix D.

** Includes Tversky and Simonson (1993), Kivetz et al. (2004a), Koszegi and Szeidl (2013), Bushong et al. (2019), as well as the model in (4), in the limiting case as $\sigma \to \infty$. Again, see Appendix C and Appendix D for additional details.

Fig. 4. Binary choice with two attributes.

Proof of Proposition 9. See Appendix.

The compromise and dominance effects are not included in Proposition 9 because they depend on the magnitude of σ. To illustrate these relationships, we again work from a benchmark of indifference between $\mathbf{x}$ and $\mathbf{y}$ in binary choice (with $x_1 > y_1$ and $y_2 > x_2$). An added technical complication, however, is that allowing σ to vary may undo binary-choice indifference (in light of Observation 3). Therefore, the following results use a stronger condition that ensures binary-choice indifference is preserved even as σ varies.

Lemma 2. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *given* $X = \{\mathbf{x}, \mathbf{y}\}$ *for all* $\sigma \geq 0$ *under* (4) *(equivalently,* $x_1 = y_2$ *and* $y_1 = x_2$*). Then, if* $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ *with* $\sigma = 0$*,* $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ *will still hold with* $\sigma > 0$ *unless* $V^{\text{CD}}(\mathbf{z}) < V^{\text{CD}}(\mathbf{x}) < V^{\text{CD}}(\mathbf{z}')$*, where* $\mathbf{z}' \equiv (z_1 + \sigma, z_2 + \sigma)$*.*

Proof of Lemma 2. See Appendix.

Lemma 2 implies that the effect of an inferior $\mathbf{z}$ on the relative valuation of $\mathbf{x}$ and $\mathbf{y}$ predicted by the basic PN model is maintained with $\sigma > 0$, unless $\mathbf{z}$ is only "modestly inferior" in the sense that the Cobb-Douglas model ranks $\mathbf{x}$ below some $\mathbf{z}'$ featuring a magnitude-σ improvement to $\mathbf{z}$ on each dimension (see Fig. 5). This suggests that the compromise and dominance effects will not always appear with $\sigma > 0$.

Proposition 10. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *given* $X = \{\mathbf{x}, \mathbf{y}\}$ *for all* $\sigma \geq 0$ *under* (4)*, and that* $\mathbf{z}$ *satisfies the conditions for either the compromise or dominance effect in Proposition 2, implying* $\mathbf{x}$ *is chosen from* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *with* $\sigma = 0$*. Then, with* $\sigma > 0$ *and* $\mathbf{z}' = (z_1 + \sigma, z_2 + \sigma)$*,* $\mathbf{x}$ *is still chosen from* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *if* $V^{\text{CD}}(\mathbf{x}) > V^{\text{CD}}(\mathbf{z}')$*, but not if* $V^{\text{CD}}(\mathbf{x}) < V^{\text{CD}}(\mathbf{z}')$*.*

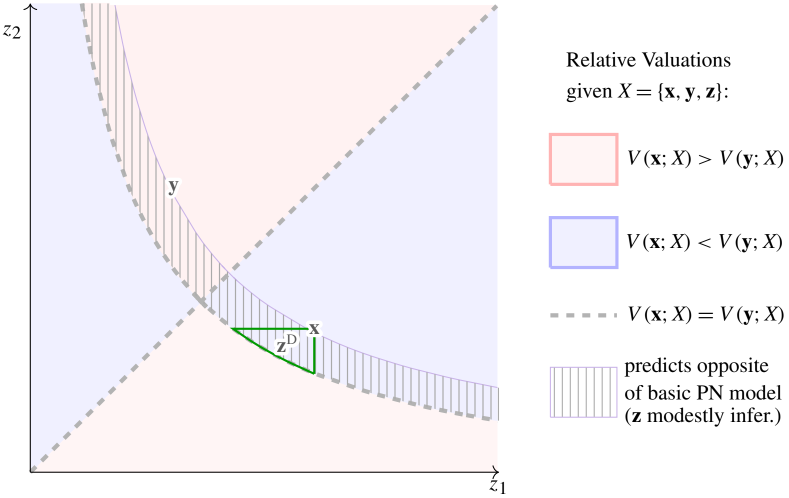
Proof of Proposition 10. See Appendix.

As illustrated in Fig. 5, Proposition 10 implies that the dominance effect can be reversed — leading to a choice of $\mathbf{y}$ instead of $\mathbf{x}$ — with a decoy $\mathbf{z}^{\text{D}}$ that is only modestly inferior. Note that a modestly inferior decoy is not just more similar to $\mathbf{x}$ than to $\mathbf{y}$ (as described in Section 4), it is also similar to $\mathbf{x}$ in an absolute sense, referring to its proximity in attribute space. In Fig. 5, for example, a dominance effect will be reversed for any decoy in the region bounded by green lines in close proximity of $\mathbf{x}$.[24]

The implication that the DM would choose the *dis*similar alternative $\mathbf{y}$ instead of $\mathbf{x}$ with $\mathbf{z}^{\text{D}}$ in the choice set is consistent with Tversky's (1972) "similarity hypothesis." Though the dominance effect is robustly observed *on average* within many experimental samples (e.g. Huber et al., 2014), recent experimental evidence from Król and Król (2019) suggests that decoys in close proximity (in attribute space) to a dominant alternative tend to generate a "repulsion effect," shifting choice *away* from the dominant alternative, while a dominance effect is reliably observed with more distant decoys. While the boundaries of the dominance effect are only beginning to be explored, Król and Król's finding is consistent with the prediction of the model in (4) with $\sigma > 0$.[25]

[24] Unlike the dominance effect, the absence of a compromise effect with $\sigma > 0$ and a modestly inferior $\mathbf{z}$ is not necessarily observable. This is because "modest inferiority" is defined based on the Cobb-Douglas model, which does not always agree with (4). Consequently, a modestly inferior $\mathbf{z}$ may be chosen over $\mathbf{x}$ and $\mathbf{y}$ in trinary choice with $\sigma > 0$, in which case the relative ranking of $\mathbf{x}$ and $\mathbf{y}$ is unobserved.

[25] Other experimental work highlights substantial within-sample heterogeneity in the incidence of the dominance effect, even if the effect is robust on average (e.g. Trueblood et al., 2015; Liew et al., 2016; Castillo, 2020). In a follow-up empirical analysis of pairwise normalization, Daviet and Webb (2020) show how allowing individual variation in σ may be useful for capturing such observed heterogeneity.



This graph shows the effect of adding **z** to the choice set on the DM's relative valuations of **x** and **y**, as predicted by the model in (4) with $\sigma > 0$. The green lines denote the boundaries of the 'similarity' subregion, as described in the text.

Fig. 5. The effect of a third alternative **z** when $\sigma > 0$.

The next result illustrates how the presence of a dominance (or repulsion) effect with $\sigma > 0$ can also depend on the overall value of the alternatives.

Corollary 1. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *given* $X = \{\mathbf{x}, \mathbf{y}\}$ *for all* $\sigma \geq 0$ *under* (4)*. Also suppose the DM is indifferent between* **x** *and* **y** *given* $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *with* $\sigma > 0$ *and* **z** *asymmetrically dominated by* **x***. Then, letting* $\mathbf{w}' = (\gamma \cdot w_1, \gamma \cdot w_2)$ *for each* $\mathbf{w} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$*,* $\mathbf{x}'$ *is chosen from* $X = \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\}$ *if* $\gamma > 1$*, while* $\mathbf{y}'$ *is chosen from* $X = \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\}$ *if* $\gamma < 1$*.*

Proof of Corollary 1. See Appendix.

Corollary 1 first considers a benchmark in which the DM is indifferent between **x** and **y** in both binary and trinary choice — here, **z** neither helps nor hurts the perception of **x** relative to **y**. In turn, $\mathbf{x}'$, $\mathbf{y}'$, and $\mathbf{z}'$ are defined as analogs to **x**, **y**, and **z**, except with their attribute values scaled by a constant $\gamma > 0$. As seen, $\mathbf{x}'$ will then be chosen in trinary choice with $\mathbf{z}'$ — consistent with the dominance effect — if (and only if) $\gamma > 1$, in which case $\mathbf{x}'$, $\mathbf{y}'$, and $\mathbf{z}'$ represent higher-value alternatives than **x**, **y**, and **z**.

Corollary 1's implication that the dominance effect will be more prominent for higher-value alternatives has support in the empirical literature. In fact, Malkoc et al. (2013) directly manipulated alternatives' desirability within each product class and found a robust dominance effect with more desirable alternatives, but not with less desirable alternatives. As the authors conclude, the results "establish (un)desirability as an important boundary condition" for the dominance effect, as Corollary 1 would suggest.[26]

[26] As additional support for this idea, in Frederick et al.'s (2014) study — which revealed several reversals of the dominance effect — the lone product class in which the decoy created a non-negligible shift in subjects' choices towards

8. Additional discussion

This paper presented a theory of multi-attribute choice based on pairwise attribute-level comparisons through divisive normalization — a well-documented form of relative value encoding in the brain. In this section, we take a closer look at the behavioral consequences of modeling divisive normalization through *pairwise* comparisons, as opposed to other potential approaches (Section 8.1). We also elaborate on the representations of attribute-level comparisons in other multi-attribute choice theories (Section 8.2).

8.1. Pairwise versus other forms of divisive normalization

To clarify the role of pairwise comparisons in our model, we now operationalize various alternatives to pairwise normalization (within an otherwise equivalent modeling framework) and compare their predictions to those of the basic PN model. The alternate normalization models we consider can all be represented as special cases of the following form:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} \frac{x_n}{x_n + r(x_n)}. \tag{5}$$

Under (5), a given attribute value x_n is no longer normalized through a series of pairwise attribute-level comparisons with the corresponding attribute values of other alternatives in the choice set. Instead, x_n is normalized through a single comparison, where $r(x_n)$ represents the value to which x_n is compared. We then consider the five alternate normalization models defined in Table 2, which vary based on their specification of $r(x_n)$.

Table 2
Alternate (divisive) normalization models.

Model	x_n "compared" to... (on same dimension)
Joint Normalization	sum of other attribute values, $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x'_n$
Average Normalization	average attribute value, $r(x_n) = \lVert X \rVert^{-1} \sum_{\mathbf{x}' \in X} x'_n$
Maximum Normalization	maximum attribute value, $r(x_n) = \max_{\mathbf{x}' \in X} \{x'_n\}$
Minimum Normalization	minimum attribute value, $r(x_n) = \min_{\mathbf{x}' \in X} \{x'_n\}$
Max-Min Normalization	range, $r(x_n) = \max_{\mathbf{x}' \in X} \{x'_n\} - \min_{\mathbf{x}' \in X} \{x'_n\}$

The first model listed in Table 2, referred to as the "joint normalization" model, effectively compares each x_n to the sum of the corresponding attribute values of all other alternatives in the choice set. For example, if $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, the normalized value of $\mathbf{x}$ on dimension n becomes $\frac{x_n}{x_n+y_n+z_n}$ under joint normalization, as opposed to $\frac{x_n}{x_n+y_n} + \frac{x_n}{x_n+z_n}$ under pairwise normalization. The remaining models in Table 2 embed attribute-level comparisons to the average, the

the dominant option was also the highest-value product class considered (apartments, as compared to bottled water, fruit, hotel rooms, jelly beans, kool-aid, mints, movies, and popcorn). In Huber et al.'s (1982) original study documenting the dominance effect, the decoy shifted subjects' choices in favor of the dominant alternative in all six product classes considered, but the largest effects were similarly observed in the two highest-value product classes (cars and televisions, as compared to beer, photographic film, restaurant meals, and lotteries with expected payoffs on the order of $20).

Table 3
Behavioral patterns generated by other forms of normalization.

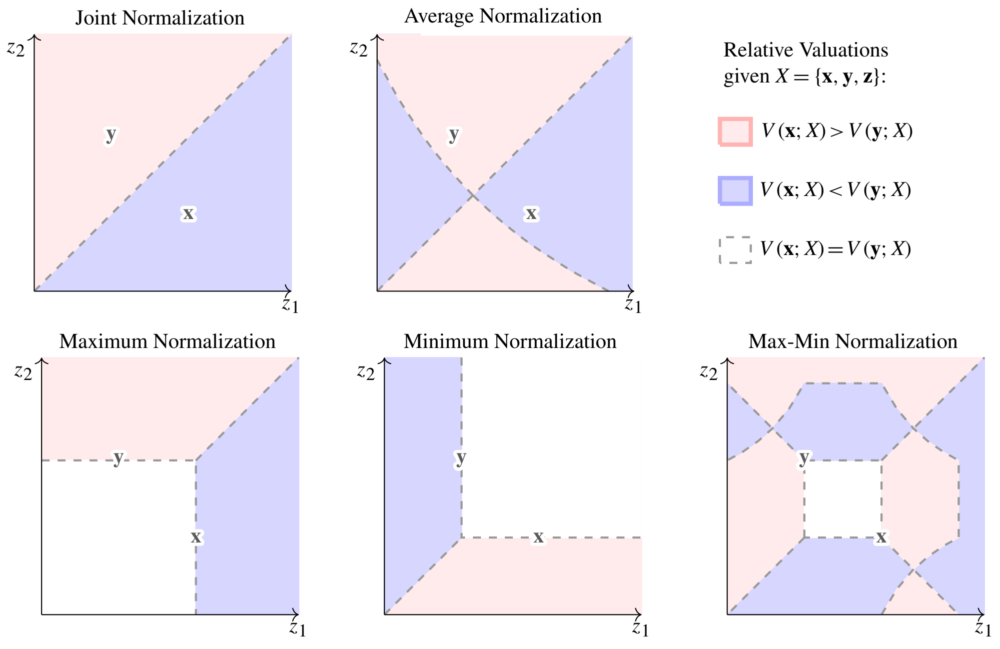
		Joint Normalization	Average Normalization	Maximum Normalization	Minimum Normalization	Max-Min Normalization
(I)	Compromise Effect	N	S	N	**Y**	S
(II)	Dominance Effect: Weak	N	S	N	N	N
	Dominance Effect: Strict	N	S	N	**Y**	S
	...Decoy-Range Effect	N	**Y**	N	**Y**	**Y**
(III)	Relative Difference Effect	**Y**	**Y**	**Y**	**Y**	**Y**
(IV)	Majority-Rule Preference Cycles	**Y**	**Y**	**Y**	**Y**	S
(V)	Splitting Bias	**Y**	**Y**	**Y**	**Y**	**Y**
(VI)	Alignability Effect	**Y**	**Y**	**Y**	**Y**	**Y**
(VII)	Diversification Bias	**Y**	**Y**	**Y**	**Y**	S
(VIII)	Feature Bias	**Y**	**Y**	**Y**	**Y**	**Y**

Behavioral predictions of the alternate (non-pairwise) normalization models defined in Table 2. As in Table 1, here '**Y**' means the model robustly predicts the behavior (i.e. never predicts the opposite or no effect under conditions for which it would be expected), 'N' means the model does not predict the behavior, and 'S' means the model sometimes predicts the behavior and sometimes predicts the opposite effect. See Appendix C for detailed explanations.

maximum, the minimum, and the range of attribute values in the choice set. Attribute-level comparisons to summary statistics such as these are common in prevailing multi-attribute choice theories (as will be discussed in Section 8.2), though their present implementation through divisive normalization is unique.

Table 3 classifies the predictions of the five alternate normalization models using the same criteria used to classify the predictions of the basic PN model in Table 1. In all but two cases, the alternate normalization models robustly capture the behavioral patterns listed in items (III) through (VIII), which — unlike items (I) and (II) — are formalized exclusively in terms of binary choices. The ability of these models to capture these same behaviors is not surprising considering their commonalities when applied to binary choice. In fact, for binary choice with two attributes, all of the alternate normalization models are equivalent to the basic PN model (see Appendix B.6).

Despite the behavioral similarities in binary choice, the alternate normalization models are less effective in capturing the compromise and dominance effects. In contrast to the basic PN model, none of the alternate normalization models are able to capture all of the behaviors in Table 3. That said, the "minimum normalization" model comes close, though there are several behavioral differences between minimum normalization and pairwise normalization that are not revealed in Table 3. Of note, the predictions of the minimum normalization model shown in the



Each graph shows the effect of $\mathbf{z}$ on the relative valuations of $\mathbf{x}$ and $\mathbf{y}$, as predicted by the indicated model. As in Fig. 2, the graphs were created using $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$.

Fig. 6. The effect of $\mathbf{z}$ under alternate normalization models.

bottom center panel of Fig. 6 often differ from the corresponding predictions of the basic PN model in Fig. 2. Under minimum normalization, for instance, the third alternative $\mathbf{z}$ can only affect the relative ranking of $\mathbf{x}$ and $\mathbf{y}$ in cases where $\mathbf{z}$ is strictly worse than both $\mathbf{x}$ and $\mathbf{y}$ on at least one dimension, while this is not true under pairwise normalization. Similarly, the valuations of $\mathbf{x}$ and $\mathbf{y}$ cannot vary with z_n in cases where z_n is at least as large as x_n or y_n. As a result, the minimum normalization model does not share the basic PN model's predictions from parts (ii) and (iii) of Proposition 3 that the dominance effect is stronger than the compromise effect, and that the dominance effect weakens when the decoy becomes symmetrically dominated.

8.2. "Comparisons" in multi-attribute choice theories

Like our model, other multi-attribute choice theories typically suggest that an alternative's attributes are "compared" (or otherwise valued in relation) to the corresponding attributes of other alternatives. While formal representations of such attribute-level comparisons vary from model to model, the use of the divisive normalization computation for this general purpose is not unique to our theory.

For instance, Bordalo et al. (2013)'s proposed form of their "salience function" (eq. 4, p. 809) is identical to our contrast function $\Delta(a, b) = \left|\frac{a-b}{a+b}\right|$, though their implementation differs. While we use Δ to express the perceived, decision-relevant value difference between two attribute values, Bordalo et al.'s salience function is used in an intermediate step that determines which attribute receives a smaller weight by comparing an attribute value to the average of all alternatives in the choice set. More precisely, for the case with alternatives defined on two quality dimensions, Bordalo et al.'s (2013) model can be represented as:

$$V(\mathbf{x};X)=\begin{cases}\beta x_1+x_2, & \Delta(x_1,\bar{x}_1)<\Delta(x_2,\bar{x}_2),\\ \frac{1+\beta}{2}\big(x_1+x_2\big), & \Delta(x_1,\bar{x}_1)=\Delta(x_2,\bar{x}_2),\\ x_1+\beta x_2, & \Delta(x_1,\bar{x}_1)>\Delta(x_2,\bar{x}_2),\end{cases} \tag{6}$$

given $0\le\beta<1$ and where $\bar{x}_n\equiv||X||^{-1}\sum_{\mathbf{x}'\in X}x'_n$ is the average attribute value on dimension n. Note, under (6), x_1 receives more (less) weight than x_2 if there is more (less) contrast between x_1 and $\bar{x}_1$ than between x_2 and $\bar{x}_2$.

Tversky and Simonson's (1993) model also uses divisive normalization to express the total "advantage" of $\mathbf{x}$ over $\mathbf{y}$, $A(\mathbf{x},\mathbf{y})=\sum_n\max\{x_n-y_n,0\}$, relative to its "disadvantage," $D(\mathbf{x},\mathbf{y})=A(\mathbf{y},\mathbf{x})$, as $\frac{A(\mathbf{x},\mathbf{y})}{A(\mathbf{x},\mathbf{y})+D(\mathbf{x},\mathbf{y})}$ in its simplest form (eq. 8, p. 1185). In addition, Tversky and Simonson conceptualize the advantage and disadvantage functions as arising from pairwise comparisons of attribute values.[27] Unlike our use of divisive normalization, however, attribute information is first aggregated across dimensions since attribute values from *all* dimensions enter each $\frac{A(\mathbf{x},\mathbf{y})}{A(\mathbf{x},\mathbf{y})+D(\mathbf{x},\mathbf{y})}$ computation.

Next, the simplest version of Soltani et al.'s (2012) "range normalization" model can be represented by:

$$V(\mathbf{x};X)=\frac{x_1}{x_1^{\max}-x_1^{\min}}+\frac{x_2}{x_2^{\max}-x_2^{\min}}, \tag{7}$$

where $x_n^{\max}=\max\{x'_n:\mathbf{x}'\in X\}$ and $x_n^{\min}=\min\{x'_n:\mathbf{x}'\in X\}$ denote the maximum and minimum attribute values on dimension n.[28] Thus, each x_n is divided by the range $(x_n^{\max}-x_n^{\min})$ of attribute values on that dimension. Note, however, the attribute value itself (x_n) does not enter the denominator unless it represents the maximum or minimum on that dimension. This distinguishes range normalization from standard formulations of divisive normalization in neuroscience.

Table 4 summarizes the varying representations of attribute-level comparisons in these and other prevailing multi-attribute choice theories. Attribute-level comparisons are either implemented through some form of normalization, through attribute weights, or through subtraction (with possible additional transformations). In these comparisons, attribute values are either compared to each other (in pairs), or compared to a summary statistic, such as the average, minimum, or range of attribute values on that dimension.[29]

For binary-choice problems, pairwise normalization is certainly simple, as it only requires a single computation to express the perceived value of an attribute, i.e. $\frac{x_n}{x_n+y_n}$, while comparisons to summary statistics would require at least two distinct computations — the computation of the summary statistic itself and the computation used to implement the comparison between the attribute value and that summary statistic. With many alternatives, however, the use of a summary statistic could certainly simplify the problem. This observation reinforces the sentiment (expressed in Section 3) that our model is not tailored to choice environments where there are

[27] Also see Marley (1991) and Tserenjigmid (2015) for axiomatizations of pairwise comparisons.

[28] Soltani et al.'s (2012) more general model features $2||X||+3$ additional parameters. For clarity, we show a version without these extra parameters and also omit a term, $-\big(\frac{x_1^{\min}}{x_1^{\max}-x_1^{\min}}+\frac{x_2^{\min}}{x_2^{\max}-x_2^{\min}}\big)$, that is common to all alternatives and therefore not choice-relevant. See Appendix C for details.

[29] In addition to Kivetz et al.'s (2004a) model, Tserenjigmid's (2019) reference-dependent model also features comparisons between an attribute value and the minimum on its dimension.

Table 4
"Comparisons" in multi-attribute choice models.

	Attribute-level "comparisons"...		Inter-attribute "comparisons" of attribute-level outputs?
	Computation used in each comparison	What is each attribute value compared to?	
Pairwise Normalization	***normalization***	***other attribute values, in pairs***	no
Bordalo et al. (2013)	***normalization******	average of attribute values	yes, outputs are ranked
Bushong et al. (2019)	weight by decr. function of...	range of attribute values (max - min)	no
Kivetz et al. (2004a)	subtraction**	minimum of attribute values	no
Koszegi and Szeidl (2013)	weight by incr. function of...	range of attribute values (max - min)	no
Soltani et al. (2012)	*(range) normalization**	range of attribute values (max - min)	no
Tversky and Simonson (1993)	subtraction	***other attribute values, in pairs***	yes, through ***normalization******

* Here, "normalization" is used to encompass any computation involving division. That said, Soltani et al.'s (2012) use of "range normalization" may be interpreted as distinct from the divisive normalization computation, in which an input value is typically divided by a term that includes itself. See text for relevant caveats and Appendix C for technical details.

** Additional transformations of the difference between two attribute values may be applied.

too many alternatives to realistically carry out every possible pairwise comparison. Therefore the reported results should be interpreted as applying to settings when all attributes and alternatives are considered.[30]

Models that entail attribute-level comparisons to a summary statistic are naturally equipped to address evidence that choices can be sensitive to that particular statistic. For instance, behavior dependent on the average attribute value is evident from empirical evidence of the relative difference effect (see Section 2.2). Experimental research has also revealed range-dependence, whereby a fixed difference between two attribute values is weighted less when the range of attribute values on that dimension is wider (Mellers and Cooke, 1994; Yeung and Soman, 2005). Even though pairwise normalization does not compare attributes to the average or range of attribute values in the choice set, it nonetheless captures both average- and range-dependence in choice. Proposition 1 established average-dependence (in the form the relative difference effect) while range-dependence is demonstrated in Appendix B.3.

[30] Of course, with arbitrarily many alternatives and/or attributes, it may also be impractical to compute summary statistics across all alternatives on every dimension, not to mention carrying out the additional intra-attribute comparisons embedded in some models (see Table 4). With that said, Koszegi and Szeidl's (2013) focusing theory may be regarded as providing a reduced-form description of how decision-makers allocate attention across attributes when there are many possible attributes to consider.

As described in Section 4, pairwise normalization also implies that more similar alternatives will be "easier to compare" than less similar alternatives. This idea is also prominent in Natenzon's (2019) model, in which an imperfectly-informed (yet Bayesian rational) decision-maker can exhibit the compromise and dominance effects due to the relative ease of comparing an inferior third alternative to the existing alternative to which it is more similar. In Natenzon's model, ease-of-comparison is operationalized as an assumption about value correlations among alternatives that may be encountered in one's environment. Pairwise normalization provides a potential foundation for Natenzon's assumption, while suggesting that the relative ease of comparing more similar alternatives does not need to reflect an inherent feature of the alternatives or environment. Instead, it may arise due to the manner in which our brains encode sensory information.

Appendix A. Additional proofs

A.1. Proof of Lemma 1

Suppose $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) < V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ for $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$ (i.e. that $z_1z_2 < x_1x_2 = y_1y_2$, from Observation 1). Noting $\frac{a-b}{a+b} = \frac{2a}{a+b} - 1$, we can see that $|V(\mathbf{x};\{\mathbf{x},\mathbf{z}\}) - V(\mathbf{z};\{\mathbf{x},\mathbf{z}\})| > |V(\mathbf{y};\{\mathbf{y},\mathbf{z}\}) - V(\mathbf{z};\{\mathbf{y},\mathbf{z}\})|$ is equivalent to $\frac{x_1}{x_1+z_1} + \frac{x_2}{x_2+z_2} > \frac{y_1}{y_1+z_1} + \frac{y_2}{y_2+z_2}$, which we can equivalently re-express as $\frac{\tilde{x}_1}{\tilde{x}_1+\tilde{z}_1} + \frac{\tilde{x}_2}{\tilde{x}_2+\tilde{z}_2} > \frac{\tilde{y}_1}{\tilde{y}_1+\tilde{z}_1} + \frac{\tilde{y}_2}{\tilde{y}_2+\tilde{z}_2}$ with $\tilde{w}_n \equiv \frac{w_n}{\sqrt{x_1x_2}}$ for $\mathbf{w} \in \{\mathbf{x},\mathbf{y},\mathbf{z}\}$ and $n \in \{1,2\}$. Using $\tilde{x}_1\tilde{x}_2 = \tilde{y}_1\tilde{y}_2 = 1$ to substitute out $\tilde{x}_2$ and $\tilde{y}_2$, cross-multiplying and collecting terms, then factoring out $(1-\tilde{z}_1\tilde{z}_2)(\tilde{x}_1 - \tilde{y}_1) > 0$, we see this condition holds if and only if $\tilde{z}_1 > \tilde{x}_1\tilde{y}_1\tilde{z}_2$. Substituting out each $\tilde{w}_n = \frac{w_n}{\sqrt{x_1x_2}}$, then multiplying both sides by $\frac{\sqrt{x_1x_2}}{z_2} > 0$, and then substituting out $\frac{y_1}{x_2} = \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$ (which holds since $y_1m_2^{\mathrm{xy}} = \frac{y_1x_2+y_1y_2}{2} = \frac{y_1x_2+x_1x_2}{2} = x_2m_1^{\mathrm{xy}}$), we see this is equivalent to $\frac{z_1}{z_2} > \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$, i.e. that $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$.

Now suppose $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) > V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ for $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$ (implying $z_1z_2 > x_1x_2 = y_1y_2$, from Observation 1). Noting $\frac{a-b}{a+b} = \frac{2a}{a+b} - 1$, we can see that $|V(\mathbf{x};\{\mathbf{x},\mathbf{z}\}) - V(\mathbf{z};\{\mathbf{x},\mathbf{z}\})| > |V(\mathbf{y};\{\mathbf{y},\mathbf{z}\}) - V(\mathbf{z};\{\mathbf{y},\mathbf{z}\})|$ is equivalent to $\frac{x_1}{x_1+z_1} + \frac{x_2}{x_2+z_2} < \frac{y_1}{y_1+z_1} + \frac{y_2}{y_2+z_2}$, which we can equivalently re-express as $\frac{\tilde{x}_1}{\tilde{x}_1+\tilde{z}_1} + \frac{\tilde{x}_2}{\tilde{x}_2+\tilde{z}_2} < \frac{\tilde{y}_1}{\tilde{y}_1+\tilde{z}_1} + \frac{\tilde{y}_2}{\tilde{y}_2+\tilde{z}_2}$ with $\tilde{w}_n \equiv \frac{w_n}{\sqrt{x_1x_2}}$ for $\mathbf{w} \in \{\mathbf{x},\mathbf{y},\mathbf{z}\}$ and $n \in \{1,2\}$. Using $\tilde{x}_1\tilde{x}_2 = \tilde{y}_1\tilde{y}_2 = 1$ to substitute out $\tilde{x}_2$ and $\tilde{y}_2$, cross-multiplying and collecting terms, then factoring out $(1-\tilde{z}_1\tilde{z}_2)(\tilde{x}_1 - \tilde{y}_1) < 0$, we see this condition holds if and only if $\tilde{z}_1 > \tilde{x}_1\tilde{y}_1\tilde{z}_2$. The rest of the proof then follows our work for the case with $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) > V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ with $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$.

Next, if $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) = V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ for $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$, then $|V(\mathbf{x};\{\mathbf{x},\mathbf{z}\}) - V(\mathbf{z};\{\mathbf{x},\mathbf{z}\})| = |V(\mathbf{y};\{\mathbf{y},\mathbf{z}\}) - V(\mathbf{z};\{\mathbf{y},\mathbf{z}\})| = 0$. Proposition 12 (see Appendix B.4) then implies that either $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) = V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ for $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$, $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) < V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ for $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$, $V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\}) > V(\mathbf{z};\{\mathbf{x}',\mathbf{z}\})$ for $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$ must hold. This means the desired result has been established for all applicable cases. □

A.2. Proof of Proposition 2

First, observe $\frac{z_1}{z_2} > \frac{x_1}{x_2} > \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$ given $z_1 > x_1$ and $z_2 < x_2$ in part (i), and $\frac{z_1}{z_2} > \frac{y_1}{x_2} = \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$ given $z_1 > y_1$ and $z_2 < x_2$ in part (ii), where $\frac{y_1}{x_2} = \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$ is verified by cross multiplication with $x_1x_2 =$

$y_1 y_2$ (which holds from Observation 1 with binary-choice indifference). Thus, $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$ in both cases. Next, given $\mathbf{z}$ is not chosen in (i) and is asymmetrically dominated by $\mathbf{x}$ in (ii), Proposition 12 (see Appendix B.4) implies $V(\mathbf{z}; X') < V(\mathbf{x}'; X')$ for all $\mathbf{x}' \in \{\mathbf{x}, \mathbf{y}\}$ and $X' \in \{\{\mathbf{x}', \mathbf{z}\}, \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}\}$. With Lemma 1 and $\mathbf{z}$'s relative similarity to $\mathbf{x}$, this implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\})$, which is equivalent (after adding 2 to both sides, then dividing through by 2, while noting $\frac{a-b}{a+b} = \frac{2a}{a+b} - 1$) to $\frac{x_1}{x_1+z_1} + \frac{x_2}{x_2+z_2} > \frac{y_1}{y_1+z_1} + \frac{y_2}{y_2+z_2}$, i.e. $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$. Noting $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}'', \mathbf{z}\}) = V(\mathbf{x}'; \{\mathbf{x}', \mathbf{z}\}) + V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\})$ for $\mathbf{x}', \mathbf{x}'' \in \{\mathbf{x}, \mathbf{y}\}$ with $\mathbf{x}' \neq \mathbf{x}''$, and given $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$, we then see that $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ must hold in both (i) and (ii), as desired. □

A.3. *Proof of Proposition 4*

We proceed by contradiction. If there is a minority-rule preference cycle, then $\mathbf{x}'$ is preferred to $\mathbf{x}$, $\mathbf{x}''$ to $\mathbf{x}'$, and $\mathbf{x}$ to $\mathbf{x}''$. Let $\lambda_1 \equiv \frac{1}{x'_1}$, $\lambda_2 \equiv \frac{1}{x_2}$, $\lambda_3 \equiv \frac{1}{x''_3}$, and $\tilde{w}_n \equiv \lambda_n w_n$ for all $\mathbf{w} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\}$ and $n = 1, 2, 3$. Also define $k_n \equiv \max\{\tilde{w}_n\} - 1 > 0$ and $q_n \equiv 1 - \min\{\tilde{w}_n\} > 0$ so that the ordered, rescaled attribute values are $(1 + k_n, 1, 1 - q_n)$ for each n. Noting normalized valuations are invariant to scaling all attribute-n values by $\lambda_n > 0$, our preference cycle implies:

$$
\begin{aligned}
V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) \quad &\Rightarrow \quad \tfrac{k_1}{2+k_1} + \tfrac{q_2}{2-q_2} < \tfrac{k_3+q_3}{2+k_3-q_3}, \\
V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}''; \{\mathbf{x}, \mathbf{x}''\}) \quad &\Rightarrow \quad \tfrac{k_2}{2+k_2} + \tfrac{q_3}{2-q_3} < \tfrac{k_1+q_1}{2+k_1-q_1}, \\
V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}'\}) > V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) \quad &\Rightarrow \quad \tfrac{k_3}{2+k_3} + \tfrac{q_1}{2-q_1} < \tfrac{k_2+q_2}{2+k_2-q_2}.
\end{aligned}
$$

Summing these conditions yields $\sum_{n=1}^{3}\left(\frac{k_n}{2+k_n} + \frac{q_n}{2-q_n}\right) < \sum_{n=1}^{3}\left(\frac{k_n+q_n}{2+k_n-q_n}\right)$. Thus, $\frac{k_n}{2+k_n} + \frac{q_n}{2-q_n} < \frac{k_n+q_n}{2+k_n-q_n}$ for at least one $n \in \{1, 2, 3\}$. Combining the fractions on the left-side, we get $\frac{2(k_n+q_n)}{(2+k_n)(2-q_n)} < \frac{k_n+q_n}{2+k_n-q_n}$, which holds if and only if $2(2 + k_n - q_n) < (2 + k_n)(2 - q_n)$, i.e. if and only if $-q_n k_n > 0$, a contradiction. □

A.4. *Proof of Proposition 7*

Follows from Proposition 7* with $R_n = 1$ for all $n \leq N$ (see Appendix B.2). □

A.5. *Proof of Proposition 9*

Part (i). Using the notation in Proposition 1, $\frac{\partial[V^*(\mathbf{x}';\{\mathbf{x}',\mathbf{y}'\}) - V^*(\mathbf{y}';\{\mathbf{x}',\mathbf{y}'\})]}{\partial k} = \frac{-2(y_2-x_2)}{(\sigma+y_2+x_2+2k)^2} < 0$ since $y_2 > x_2$, which ensures the relative difference effect holds under (4) for all $\sigma \geq 0$.

Part (ii). Let $\widehat{w}_n = w_n + \frac{\sigma}{2}$ for $w \in \{x, x', x''\}$ and $n = 1, 2, 3$. Thus, a majority dominance relationship among $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ exists if and only if it also exists among $\widehat{\mathbf{x}}$, $\widehat{\mathbf{x}}'$, and $\widehat{\mathbf{x}}''$. From Proposition 4, if binary-choice preferences among $\widehat{\mathbf{x}}$, $\widehat{\mathbf{x}}'$, and $\widehat{\mathbf{x}}''$ are intransitive, they must follow a majority-rule cycle in the basic PN model. Noting $V(\widehat{\mathbf{w}}; \{\widehat{\mathbf{w}}, \widehat{\mathbf{w}}'\}) = V^*(\mathbf{w}; \{\mathbf{w}, \mathbf{w}'\})$, if preferences among $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ are intransitive under (4) with $\sigma \geq 0$, they must also follow the same majority-rule cycle.

Part (iii). Using the notation in Proposition 5, $\frac{x_{1a}-y_{1a}}{\sigma+x_1+y_1} < \min\left\{\frac{x_{1a}-y_{1a}}{\sigma+x_{1a}+y_{1a}}, \frac{x_{1b}-y_{1b}}{\sigma+x_{1b}+y_{1b}}\right\}$ holds since $x_1 + y_1 > \max\{x_{1a} + y_{1a}, x_{1b} + y_{1b}\}$. Thus, $\frac{x_{1a}-y_{1a}}{\sigma+x_{1a}+y_{1a}} + \frac{x_{1b}-y_{1b}}{\sigma+x_{1b}+y_{1b}} > \frac{x_{1a}-y_{1a}+x_{1b}-y_{1b}}{\sigma+x_1+y_1} = \frac{x_1-y_1}{\sigma+x_1+y_1}$, which ensures the splitting bias holds under (4).

Part (iv). Using the notation in Proposition 6 and given $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$, we get $V^*(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V^*(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = V^*(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) - V^*(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) + V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})) = \frac{x_2}{\sigma+x_2} - \frac{y_2}{\sigma+y_2} - \frac{x_2-y_2}{\sigma+x_2+y_2} = \frac{(y_2-x_2)x_2y_2}{(\sigma+x_2)(\sigma+y_2)(\sigma+x_2+y_2)} > 0$ for any $\sigma \geq 0$ since $y_2 > x_2$. Thus, the alignability effect holds under (4).

Part (v). Using the notation in Proposition 7, we see $V^*(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) - V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = \sum_{n\leq N} \frac{x'_n - x_n}{\sigma + x'_n + x_n} = \sum_{n<N} \frac{x'_n - AN^{-1}}{\sigma + x'_n + AN^{-1}} + \frac{A - \sum_{n<N} x'_n - AN^{-1}}{\sigma + A - \sum_{n<N} x'_n + AN^{-1}}$. Differentiating by x'_n, $n \leq N-1$ and substituting x'_N back in using $\sum_{n\leq N} x'_n = A$ gives $\frac{\sigma+2AN^{-1}}{(\sigma+AN^{-1}+x'_n)^2} = \frac{\sigma+2AN^{-1}}{(\sigma+AN^{-1}+x'_N)^2}$. Thus, the system of $N-1$ first-order conditions is solved by $x'_n = x'_N$, implying $x'_n = \frac{A}{N} = x_n$ for all $n \leq N$, thus ensuring $V^*(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) < V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\})$ for $\mathbf{x}' \neq \mathbf{x}$.

Part (vi). Using the notation in Proposition 8, $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V^*(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \frac{q}{\sigma+q} - \frac{q}{\sigma+2x_{n'}+q} > 0$ since $x_{n'} > 0$. Thus, the feature bias holds for any $\sigma \geq 0$ under (4). □

A.6. *Proof of Lemma 2*

Using (1) and (4), $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ if and only if $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\})$. Since $\frac{z'_1}{z'_2} = \frac{z_1+\sigma}{z_2+\sigma}$ and $\frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}} = 1$ given $x_1 = y_2$ and $y_1 = x_2$, $\mathbf{z}'$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$ if and only if $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$. From Observation 1, Lemma 1, and Proposition 12 (see Appendix B.4), it then follows that $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) < V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ if and only if $z_1z_2 < x_1x_2 = y_1y_2 < z'_1z'_2$. □

A.7. *Proof of Proposition 10*

Since $x_1x_2 > z_1z_2$ must hold in the case of the compromise and dominance effects with $\sigma = 0$, the desired result then follows from Lemma 2. □

A.8. *Proof of Corollary 1*

Given $V^*(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V^*(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ with $x_1 = y_2$ and $y_1 = x_2$, Proposition 10 implies $V^{\mathrm{CD}}(\mathbf{x}) = V^{\mathrm{CD}}(\mathbf{z}^a)$, with $\mathbf{z}^a = (z_1 + \sigma, z_2 + \sigma)$. Since $\gamma x_1 = \gamma y_2$ and $\gamma y_1 = \gamma = x_2$, Proposition 10 also implies that $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\}) > V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\})$ holds if and only if $V^{\mathrm{CD}}(\mathbf{x}') > V^{\mathrm{CD}}(\mathbf{z}^b)$, with $\mathbf{z}^b = (\gamma z_1 + \sigma, \gamma z_2 + \sigma)$. We can then compute $V^{\mathrm{CD}}(\mathbf{x}') - V^{\mathrm{CD}}(\mathbf{z}^b) = \gamma^2 x_1x_2 - \gamma^2 z_1z_2 + \gamma\sigma(z_1 + z_2) + \sigma^2 = \gamma^2 V^{\mathrm{CD}}(\mathbf{x}) - \gamma^2 V^{\mathrm{CD}}(\mathbf{z}^a) - \gamma(1-\gamma)\sigma(z_1 + z_2) - (1-\gamma^2)\sigma = -\gamma(1-\gamma)\sigma(z_1 + z_2) - (1-\gamma^2)\sigma$ since $V^{\mathrm{CD}}(\mathbf{x}) = V^{\mathrm{CD}}(\mathbf{z}^a)$. It is then readily apparent that $-\gamma(1-\gamma)\sigma(z_1+z_2) - (1-\gamma^2)\sigma \gtrless 0$ for $\gamma \gtrless 1$. This, along with the fact that $V(\mathbf{z}'; \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\}) < \max\{V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\}), V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}', \mathbf{z}'\})\}$ must hold given $\mathbf{x}$ ($\mathbf{x}'$) dominates $\mathbf{z}$ ($\mathbf{z}'$), yields the desired result. □

Appendix B. Additional results

B.1. *Ease of comparisons*

As demonstrated by Lemma 1, pairwise normalization implies that it is “easier to compare” more similar alternatives than less similar alternatives. Formally, given $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$, with the DM indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice, the magnitude of the perceived value difference will be larger between $\mathbf{z}$ and $\mathbf{x}$ than between $\mathbf{z}$ and $\mathbf{y}$: $|V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})| > |V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})|$.

However, the implication that $\mathbf{z}$ is easier to compare to $\mathbf{x}$ than to $\mathbf{y}$ in this sense is not directly observable. This is because $\mathbf{x}$ and $\mathbf{y}$ would either both be chosen over $\mathbf{z}$ with certainty in binary choice, or $\mathbf{z}$ would be chosen over both $\mathbf{x}$ and $\mathbf{y}$. With this in mind, the following corollary shows how an adaptation of the basic PN model to a stochastic choice environment captures the ease of comparison concept in an observable form:

Corollary 2. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *with* $X = \{\mathbf{x}, \mathbf{y}\}$*, and that* $\mathbf{z}$ *is more similar to* $\mathbf{x}$ *than to* $\mathbf{y}$ *in that* $\frac{z_1}{z_2} > \frac{m_1^{\mathrm{xy}}}{m_2^{\mathrm{xy}}}$*. Next, consider a stochastic extension of the deterministic basic PN model given in (1), with binary-choice probabilities given by*

$$\Pr[\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}] = f(V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}), V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\})), \tag{8}$$

where f *is strictly increasing in its first argument and strictly decreasing in its second argument. Then* $\left|\Pr[\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}] - \frac{1}{2}\right| \leq \left|\Pr[\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}] - \frac{1}{2}\right|$ *(which only binds when both sides of the inequality are zero).*

Proof. $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\})$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$ and $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \lesseqgtr V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\})$ (Proposition 12). Noting $\Pr[\mathbf{z}; \{\mathbf{w}, \mathbf{z}\}] = f(V(\mathbf{z}; \{\mathbf{w}, \mathbf{z}\}), V(\mathbf{w}; \{\mathbf{w}, \mathbf{z}\}))$ for $\mathbf{w} \in \{\mathbf{x}, \mathbf{y}\}$, $\Pr[\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}] \lesseqgtr \Pr[\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}]$ is assured for $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\})$ since f is increasing in its first argument and decreasing in its second argument. Thus, either $\Pr[\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}] < \Pr[\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}] < \frac{1}{2}$ or $\Pr[\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}] > \Pr[\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}] > \frac{1}{2}$ must hold, ensuring $\left|\Pr[\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}] - \frac{1}{2}\right| < \left|\Pr[\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}] - \frac{1}{2}\right|$, unless $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) = (\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\})$ for all $\mathbf{x}', \mathbf{x}'' \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ with $\mathbf{x}' \neq \mathbf{x}''$, in which case $\left|\Pr[\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}] - \frac{1}{2}\right| = \left|\Pr[\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}] - \frac{1}{2}\right| = 0$. □

Corollary 2 can be understood as follows. Suppose $\mathbf{x}$ and $\mathbf{y}$ are equally likely to be chosen in binary choice (as indirectly implied by (8)). Also suppose that the probabilities of choosing $\mathbf{z}$ in binary choices with $\mathbf{x}$ and with $\mathbf{y}$ are both less than one half, suggesting $\mathbf{z}$ is inferior to $\mathbf{x}$ and $\mathbf{y}$. Then, if $\mathbf{z}$ is more similar to $\mathbf{x}$ than to $\mathbf{y}$, the likelihood of choosing $\mathbf{z}$ in a binary choice is lower with $X = \{\mathbf{x}, \mathbf{z}\}$ than with $X = \{\mathbf{y}, \mathbf{z}\}$. That is, $\mathbf{z}$ is easier to compare to the similar alternative $\mathbf{x}$ than to the less similar alternative $\mathbf{y}$ in the sense that there is a lower probability that the DM will "mistakenly" choose the inferior alternative $\mathbf{z}$ with $\mathbf{x}$ than with $\mathbf{y}$.

B.2. Allocation and investment results with unequal returns

As mentioned in Section 6, the results capturing the diversification and feature biases still hold in the basic PN model even if the returns to allocations along each dimension are not equal. We will now formalize and prove these results.

To begin, we now distinguish between the amount of an allocation to a given dimension and the attribute value generated by that allocation. In particular, we now let a_n denote the allocation of $A > 0$ to dimension $n \leq N$, where the (unnormalized) attribute value generated by this allocation is now $R_n a_n$ given $R_n > 0$ is the (gross) rate of return on dimension n. In the following generalization of Proposition 7, we will assume that $\mathbf{x}$ and $\mathbf{x}'$ are the alternatives associated with the allocations $a_1, \ldots, a_N$ and $a'_1, \ldots, a'_N$, respectively (with $\sum_{n\le N} a_n = \sum_{n\le N} a'_n = A$), implying $x_n = R_n a_n$ and $x'_n = R_n a'_n$ for all $n \le N$.

Proposition 7*. *Given $N > 1$, $A > 0$, and $R_n > 0$ for all $n \le N$, suppose $x_n = R_n a_n$ with $a_n = \frac{A}{N}$ for all $n \le N$. Then, for any $\mathbf{x}' \ne \mathbf{x}$ satisfying $x'_n = R_n a'_n$ with $\sum_{n\le N} a'_n = A$, $\mathbf{x}$ is chosen from $X = \{\mathbf{x}, \mathbf{x}'\}$.*

Proof. Using $\sum_{n\le N} a'_n = A$ and $x_n = R_n \frac{A}{N}$ to substitute out $x'_N = R_N a'_N$ and each x_n from (1) while canceling all R_n terms gives $V(\mathbf{x}'; X) = \sum_{n=1}^{N-1} \frac{a'_n}{A\, N^{-1}+a'_n} + \frac{A-\sum_{n=1}^{N-1} a'_n}{A\, N^{-1}+A-\sum_{n=1}^{N-1} a'_n}$. Differentiating by a'_n, $n \le N-1$, setting each derivative to zero, and substituting a'_N back in using $\sum_{n\le N} a'_n = A$ gives $\frac{A\,N}{(A+a'_n N)^2} = \frac{A\,N}{(A+a'_N N)^2}$. Thus, the system of $N-1$ first-order conditions is solved by $a'_n = a'_N$, implying $a'_n = \frac{A}{N}$ (and $x'_n = x_n$) for all $n \le N$, ensuring $V(\mathbf{x}'; X) < V(\mathbf{x}; X)$ for $\mathbf{x}' \ne \mathbf{x}$. □

Thus, the diversification bias captured in Proposition 7 still holds with unequal returns.

To formalize the feature bias with unequal returns, we now assume that an investment of $q > 0$ on dimension $n \le N$ yields a $R_n q$ increase in the unnormalized attribute value on dimension n. We can then generalize Proposition 8 as:

Proposition 8*. *Given $N > 1$, $q > 0$, and $R_n > 0$ for $n = 1, \ldots, N$, suppose $x_N = R_N \cdot q$, $x'_N = 0$, $x'_{n'} = x_{n'} + R_{n'} \cdot q$ for some $n' < N$, and $x'_n = x_n > 0$ for all $n < N$. Then $\mathbf{x}$ is chosen from $X = \{\mathbf{x}, \mathbf{x}'\}$.*

Proof. $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \Delta(R_N \cdot q, 0) - \Delta(x_{n'} + R_{n'} \cdot q, x_{n'}) = \frac{2x_{n'}}{R_{n'} \cdot q + 2x_{n'}} > 0$ given $q > 0$, $R_{n'} > 0$, and $x_{n'} > 0$. □

Thus, the feature bias captured in Proposition 8 also still holds with unequal returns.

B.3. Range-dependent preferences

The following result shows how the perceived value difference between two attribute values decreases with the range of values on that dimension (holding the average fixed):

Proposition 11. *Suppose the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ when $X = \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\}$, and $x'_2 < x_2 < y_2 < y'_2$. Also suppose $x''_1 = x'_1$, $y''_1 = y'_1$, $x''_2 = x'_2 - k$, and $y''_2 = y'_2 + k$ for some $k > 0$. Then $V(\mathbf{x}; X) > V(\mathbf{y}; X)$ with $X = \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\}$.*

Proof. Using $x''_1 = x'_1$, $y''_1 = y'_1$, $x''_2 = x'_2 - k$, and $y''_2 = y'_2 + k$, we can express $V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\}) = \frac{z_2}{z_2+x'_2-k} + \frac{z_2}{z_2+y'_2+k} - \frac{z_2}{z_2+x'_2} - \frac{z_2}{z_2+y'_2}$ for each $z_2 \in \{x_2, y_2\}$. Hence, $\frac{\partial^2}{\partial z_2 \partial k}\big[V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\})\big]_{k=0} = \frac{x'_2 - z_2}{(z_2+x'_2)^3} - \frac{y'_2 - z_2}{(z_2+y'_2)^3} < 0$ given $x'_2 < z_2 < y'_2$. Thus, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\})$ since $y_2 > x_2$, implying $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}'', \mathbf{y}''\})$ given $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'\})$. □

While Proposition 1 demonstrated how an increase in the average attribute value shifted perceptions in favor of the alternative that was weaker on that dimension, Proposition 11 demonstrates how an increase in the range of attribute values has the same effect, in line with evidence from Mellers and Cooke (1994) and Yeung and Soman (2005).

B.4. Superiority/inferiority result

The following result shows that, when the DM is indifferent between **x** and **y** in binary choice, the superiority or inferiority of **z** relative to **x** and **y** does not depend on whether alternatives are evaluated in binary or trinary choice.

Proposition 12. *Suppose the DM is indifferent between* $\mathbf{x} = (x_1, x_2)$ *and* $\mathbf{y} = (y_1, y_2)$ *with* $X = \{\mathbf{x}, \mathbf{y}\}$. *Then the following conditions are equivalent:*

(i-a) $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$*;*
(i-b) $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$*,*
(ii-a) $V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$*,*
(ii-b) $V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$*.*

Proof. From Observation 1, $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$ (i-a) is equivalent to $z_1 z_2 \gtreqless x_1 x_2 = y_1 y_2$ given $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$. Thus, (i-a) must be equivalent to (i-b). We then see $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) + V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\}) = \frac{x_1'+x_1''}{x_1'+x_1''} + \frac{x_2'+x_2''}{x_2'+x_2''} = 2$. Therefore, $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) \gtreqless V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\})$, $1 \gtreqless V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\})$, and $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) \gtreqless 1$, are equivalent. Since $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$ and $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$ are equivalent, $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$ implies $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) + V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) + 1$. Since $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = 1$. Thus, $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$ implies $V(\mathbf{z}; \{\mathbf{x}, \mathbf{y})\} = V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) + V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) + V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ (ii-a). In turn, we see $V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) + V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) + V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ implies $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$ or $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = 1$ or both. Since $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless 1$ is equivalent to $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$, which itself is equivalent to $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$, since at least one among $V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$ and $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) \gtreqless V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = 1$ are true, they must both hold. Thus, (ii-a) and (i-a) are equivalent. By switching **x** and **y** in the arguments outlined above, we can likewise establish the equivalence of (ii-b) and (i-b). □

B.5. A generalization with two additional parameters

Next, we provide additional results arising from a variation of the pairwise normalization model (based on a common formulation of the normalization computation studied in neuroscience) that includes the parameter $\sigma \geq 0$ from the model in (4), as well as an additional parameter $\alpha > 0$. This version of the model is given by:

$$V^{**}(\mathbf{x}; X) = \sum_{n=1}^{N} \sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{x_n^\alpha}{\sigma^\alpha + x_n^\alpha + y_n^\alpha}. \tag{9}$$

When considered for an attribute that is observed as a numerical consumption level c_n, the model in the model (4) nests (9) by taking $x_n = (u_n(c_n(\mathbf{x})))^\alpha$. The following result shows how the model in (9) nests some classic microeconomic preference representations when applied to two-attribute binary choice.

Proposition 13. *Given $X = \{\mathbf{x}, \mathbf{x}'\}$ and $N = 2$ under (9). For each of the following specifications of $\widetilde{V}(a, b)$ with the indicated parametric restrictions, $\mathbf{x}$ is chosen if $\widetilde{V}(\mathbf{x}) > \widetilde{V}(\mathbf{x}')$:*

(i) $\widetilde{V}(a, b) = V^{\text{CD}}(a, b) \equiv ab$, *with $\sigma = 0$ and any $\alpha > 0$.*
(ii) $\widetilde{V}(a, b) = V^{\text{CES}}(a, b) \equiv (a^\alpha + b^\alpha)^{1/\alpha}$, *with $\sigma > 0$ sufficiently large and any $\alpha > 0$.*
(iii) $\widetilde{V}(a, b) = \max\{a, b\}$, *with $\sigma > 0$ and $\alpha > 0$ both sufficiently large; if $\widetilde{V}(\mathbf{x}) = \widetilde{V}(\mathbf{x}')$, $\mathbf{x}$ is then preferred to $\mathbf{x}'$ if and only if $\widetilde{V}_0(\mathbf{x}) > \widetilde{V}_0(\mathbf{x}')$, where $\widetilde{V}_0(a, b) = \min\{a, b\}$.*

Proof. Given $X = \{\mathbf{x}, \mathbf{x}'\}$, $\mathbf{x}$ is chosen if and only if $V^{**}(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V^{**}(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \sum_{n=1}^{2} \frac{x_n^\alpha - x_n'^\alpha}{\sigma^\alpha + x_n^\alpha + x_n'^\alpha} > 0$. Combining terms and factoring out the denominator yields:

$$\sigma^\alpha (x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha > \sigma^\alpha (x_1'^\alpha + x_2'^\alpha) + 2x_1'^\alpha x_2'^\alpha, \tag{10}$$

so that $\mathbf{x}$ is chosen given $\sigma = 0$ if and only if $x_1^\alpha x_2^\alpha > x_1'^\alpha x_2'^\alpha$, which is equivalent to $x_1 x_2 > x_1' x_2'$. This establishes part (i).

For part (ii), $\widetilde{V}(\mathbf{x}) > \widetilde{V}(\mathbf{x}')$ if and only if $\widetilde{V}(\mathbf{x})^\alpha > \widetilde{V}(\mathbf{x}')^\alpha$, which is equivalent to $x_1^\alpha + x_2^\alpha > x_1'^\alpha + x_2'^\alpha$ given $\widetilde{V}(a, b) = (a^\alpha + b^\alpha)^{1/\alpha}$. Let $\sigma_0 = \left(\frac{2(y_1^\alpha y_2^\alpha - x_1^\alpha x_2^\alpha)}{x_1^\alpha + x_2^\alpha - x_1'^\alpha - x_2'^\alpha} \right)^{1/\alpha} < \infty$. Observe $\sigma_0^\alpha (x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha = \sigma_0^\alpha (x_1'^\alpha + x_2'^\alpha) + 2x_1'^\alpha x_2'^\alpha$. Thus, $\sigma^\alpha (x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha > \sigma^\alpha (x_1'^\alpha + x_2'^\alpha) + 2x_1'^\alpha x_2'^\alpha$ for all $\sigma > \sigma_0$, implying from (10) that $\mathbf{x}$ is chosen. The converse is established by contradiction. Namely, suppose $\mathbf{x}$ is chosen but $\widetilde{V}(\mathbf{x}) < \widetilde{V}(\mathbf{x}')$, or equivalently, $x_1^\alpha + x_2^\alpha < x_1'^\alpha + x_2'^\alpha$. From (10), we see, together, these conditions require $x_1^\alpha x_2^\alpha > x_1'^\alpha x_2'^\alpha$, so that $x_1^\alpha x_2^\alpha - x_1'^\alpha x_2'^\alpha > 0$. By inspection, we can now see $\sigma > \sigma_0$ with $\sigma_0 > 0$ as defined above implies $\sigma^\alpha (x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha < \sigma^\alpha (x_1'^\alpha + x_2'^\alpha) + 2x_1'^\alpha x_2'^\alpha$, which from (10) implies $\mathbf{x}'$ is chosen. Hence, we have a contradiction, so that choosing $\mathbf{x}$ over $\mathbf{x}'$ necessarily requires $\widetilde{V}(\mathbf{x}) > \widetilde{V}(\mathbf{x}')$ for sufficiently large $\sigma > 0$.

For part (iii), given $\widetilde{V}(a, b) = \max\{a, b\}$, letting $x = \max\{x_1, x_2\}$ and $x' = \max\{x_1', x_2'\}$, without loss of generality, we see $\widetilde{V}(\mathbf{x}) > \widetilde{V}(\mathbf{x}')$ holds if and only if $x > x'$. Observe, $\sigma^\alpha (x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha \geq \sigma^\alpha x^\alpha$. Given any $\sigma > x'$, we also see $\sigma^\alpha (x_1'^\alpha + x_2'^\alpha) + 2x_1'^\alpha x_2'^\alpha \leq 2\sigma^\alpha x_1'^\alpha + 2x_1'^{2\alpha} < 4\sigma^\alpha x'^\alpha$. From (10), we can then see that a sufficient condition for $\mathbf{x}$ to be chosen given any $\sigma > x'$ is $\sigma^\alpha x^\alpha > 4\sigma^\alpha x'^\alpha$. Factoring out $\sigma^\alpha > 0$ then taking the natural log, we see this condition is equivalent to $\alpha \ln(x) > \alpha \ln(x') + \ln(4)$. Taking $\alpha_0 \equiv \frac{\ln(4)}{\ln(x) - \ln(x')} > 0$, we see $\alpha \ln(x) > \alpha \ln(x') + \ln(4)$ holds for any $\alpha > \alpha_0$ and $\sigma > x'$, so that $\mathbf{x}$ must be chosen for sufficiently large α and σ. The converse is established by contradiction. Suppose $\mathbf{x}$ is chosen but $\widetilde{V}(\mathbf{x}) < \widetilde{V}(\mathbf{x}')$, or equivalently, $x' > x$. Using (10) while applying the logic outlined above (except switching the roles of $\mathbf{x}$ and $\mathbf{x}'$), it must be the case that, for any $\sigma > x$, $\alpha \ln(x) + \ln(4) > \alpha \ln(x')$ by virtue of the choice of $\mathbf{x}$ over $\mathbf{x}'$. Defining $\alpha_0' \equiv \frac{\ln(4)}{\ln(x') - \ln(x)} > 0$ (positive because $x' > x$), we can see $\alpha > \alpha_0'$ implies $\alpha \ln(x) + \ln(4) < \alpha \ln(x')$. Hence, we have a contradiction, so that choosing $\mathbf{x}$ (with $\widetilde{V}(\mathbf{x}) \neq \widetilde{V}(\mathbf{x}')$) must require $\widetilde{V}(\mathbf{x}) > \widetilde{V}(\mathbf{x}')$ for sufficiently large $\sigma > 0$ and $\alpha > 0$. In the case of $\widetilde{V}(\mathbf{x}) = \widetilde{V}(\mathbf{x}')$, i.e., $x = x'$, we see from (10) that, in this case, $\mathbf{x}$ will be chosen if and only if $\sigma^\alpha (x^\alpha + \ell^\alpha) + 2x^\alpha \ell^\alpha > \sigma^\alpha (x^\alpha + \ell'^\alpha) + 2x^\alpha \ell'^\alpha$ with $\ell \equiv \min\{x_1, x_2\}$ and $\ell' \equiv \min\{x_1', x_2'\}$. Subtracting $\sigma^\alpha x^\alpha$ from both sides, then factoring out $\sigma^\alpha + 2x^\alpha > 0$, we see this is equivalent to $\ell > \ell'$. Given $\widetilde{V}_0(a, b) \equiv \min\{a, b\}$ with $x \geq \ell$ and $x' \geq \ell'$, we see that $\ell > \ell'$ is equivalent to $\widetilde{V}_0(\mathbf{x}) > \widetilde{V}_0(\mathbf{x}')$. $\square$

Part (i) of Proposition 13 shows that the previously-established equivalence between the basic PN model and the (symmetric) Cobb-Douglas model in two-attribute binary choice (Observation 1) extends with any $\alpha > 0$, provided $\sigma = 0$ is maintained. Part (ii) shows that preferences converge to those represented by a constant elasticity of substitution (CES) preference model in the large-σ limit of the model in (9). In this case, $(1-\alpha)^{-1}$ represents the effective elasticity of substitution across attributes, implying preferences are nonconvex if $\alpha > 1$ (i.e. if $(1-\alpha)^{-1} < 0$). Lastly, part (iii) shows that when σ and α are both arbitrarily large, the choice is equivalently represented by a rank-based lexicographic model, in which the choice between $\mathbf{x}$ and $\mathbf{x}'$ is determined by each alternative's larger attribute value ($\max\{x_1, x_2\}$, $\max\{x_1', x_2'\}$). In the event of a tie, the choice is then determined by their smaller attribute values ($\min\{x_1, x_2\}$, $\min\{x_1', x_2'\}$).

More generally, in two-attribute binary choice, the model in (9) is effectively a composite of the Cobb-Douglas and CES preference models, with σ determining the relative weight of each representation:

Proposition 14. *Given $N = 2$ and $X = \{\mathbf{x}, \mathbf{x}'\}$ under (9):*

(i) *If $V^{\text{CD}}(\mathbf{x}) \geq V^{\text{CD}}(\mathbf{x}')$, $V^{\text{CES}}(\mathbf{x}) \geq V^{\text{CES}}(\mathbf{x}')$, and $\sigma \geq 0$ with at least two of these inequalities non-binding, then $\mathbf{x}$ is chosen from $X = \{\mathbf{x}, \mathbf{x}'\}$.*
(ii) *If $V^{\text{CD}}(\mathbf{x}) > V^{\text{CD}}(\mathbf{x}')$ and $V^{\text{CES}}(\mathbf{x}') > V^{\text{CES}}(\mathbf{x})$, there exists a $\sigma_0 > 0$ (determined by x_1, x_2, x_1', x_2', and α) for which $\mathbf{x}$ is chosen from $X = \{\mathbf{x}, \mathbf{x}'\}$ if and only if $\sigma < \sigma_0$.*

Proof. Using (10), $\mathbf{x}$ is chosen if and only if $(\sigma V^{\text{CES}}(\mathbf{x}))^\alpha + 2(V^{\text{CD}}(\mathbf{x}))^\alpha > (\sigma V^{\text{CES}}(\mathbf{x}'))^\alpha + 2(V^{\text{CD}}(\mathbf{x}'))^\alpha$, from which the result in part (i) is readily verified. Part (ii) is also readily verifiable from this condition, where $\sigma_0 \equiv \left(\frac{2((V^{\text{CD}}(\mathbf{x}))^\alpha - (V^{\text{CD}}(\mathbf{x}'))^\alpha)}{(V^{\text{CES}}(\mathbf{x}'))^\alpha - (V^{\text{CES}}(\mathbf{x}))^\alpha}\right)^{1/\alpha}$ is derived from the implied indifference condition. □

Thus, if the Cobb-Douglas and CES models agree in their rankings among the two alternatives, the DM's choice will align with this ranking. Otherwise, the choice will coincide with Cobb-Douglas if $\sigma < \sigma_0$ and with CES if $\sigma > \sigma_0$, for some $\sigma_0 > 0$.

The next result shows that, unlike the contrast function Δ from the basic PN model, the analogous contrast function under (9), denoted as Δ^{**}, does not exhibit diminishing sensitivity over its full domain:

Proposition 15. *Given $x_n \geq y_n$ (without loss of generality), $\hat{\sigma}(y_n) \equiv \left(\frac{2}{\alpha-1}\right)^{1/\alpha} y_n$, and $\Delta^{**}(x_n, y_n) \equiv \left|\frac{x_n^\alpha - y_n^\alpha}{\sigma^\alpha + x_n^\alpha + y_n^\alpha}\right|$:*

(i) *$\Delta^{**}(x_n, y_n)$ satisfies diminishing sensitivity if and only if $\sigma = 0$ or $\alpha \leq 1$ (or both).*
(ii) *If $\alpha > 1$ and $\sigma \leq \hat{\sigma}(y_n)$, $\Delta^{**}(x_n, y_n)$ exhibits diminishing sensitivity (locally) and is concave in x_n for all $x_n \geq y_n$.*
(iii) *If $\alpha > 1$ and $\sigma > \hat{\sigma}(y_n)$, there exist increasing functions $\check{x}(\sigma) > y_n$ and $\hat{x}(\sigma) > y_n$ such that $\Delta^{**}(x_n, y_n)$ exhibits diminishing sensitivity if and only if $x_n > \check{x}(\sigma)$, and is concave in x_n if and only if $x_n > \hat{x}(\sigma)$.*

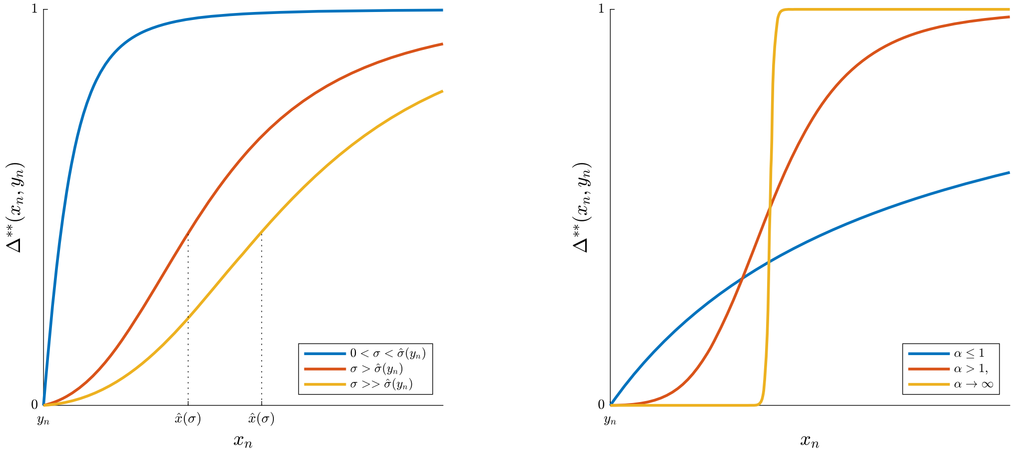
Proof. For part (i), note $\frac{d[\Delta^{**}(x_n+\epsilon, y_n+\epsilon)]}{d\epsilon} = -\frac{\alpha x_n^\alpha y_n^\alpha (2(x_n^\alpha - y_n^\alpha) + \sigma^\alpha (x_n^{1-\alpha} - y_n^{1-\alpha}))}{x_n y_n (x_n^\alpha + y_n^\alpha + \sigma^\alpha)^2}$ given $x_n \geq y_n$ (without loss of generality). Thus, $\frac{d[\Delta^{**}(x_n+\epsilon, y_n+\epsilon)]}{d\epsilon} < 0$ if and only if $2(x_n - y_n) + \sigma^\alpha (x_n^{1-\alpha} -$

$y_n^{1-\alpha}) > 0$. With $x_n \geq y_n$, this clearly holds for $\sigma = 0$ and also for $\alpha \leq 1$ because, together, $\alpha \leq 1$ and $x_n \geq y_n$ guarantee $x^{1-\alpha} - y^{1-\alpha} \geq 0$. Thus, $\Delta^{**}(x_n + \epsilon, y_n + \epsilon) < \Delta^{**}(x_n, y_n)$ for all $\epsilon > 0$ given $\sigma = 0$ or $\alpha \leq 1$ (or both). Thus, to complete the proof, we only need to show that for any $\sigma > 0$ and $\alpha > 1$, there exist a $x_n \geq 0$ and $y_n \geq 0$ with $x_n \geq y_n$ such that $2(x_n - y_n) + \sigma^\alpha(x^{1-\alpha} - y^{1-\alpha}) < 0$. Take $y_n = \frac{\sigma(\alpha-1)^{1/\alpha}}{2}$ and let $x_n = y_n + \delta$. Substituting these into $2(x_n - y_n) + \sigma^\alpha(x^{1-\alpha} - y^{1-\alpha})$ then differentiating with respect to δ, we get $2 - 2^\alpha < 0$ for $\alpha > 1$. Also note $2(x_n - y_n) + \sigma^\alpha(x^{1-\alpha} - y^{1-\alpha}) = 0$ given $x_n = y_n$, i.e., given $\delta = 0$. Together, these imply $2(x_n - y_n) + \sigma^\alpha(x^{1-\alpha} - y^{1-\alpha}) < 0$ for $y_n = \frac{\sigma(\alpha-1)^{1/\alpha}}{2}$ and $x_n = y_n + \delta$, provided $\delta > 0$ is sufficiently small, as desired.

For part (ii), let $h(x_n|y_n, \sigma, \alpha) \equiv 2(x_n - y_n) + \sigma^\alpha(x_n^{1-\alpha} - y_n^{1-\alpha})$. From our above work, we can see that $\Delta^{**}(x_n, y_n)$ exhibits diminishing sensitivity for all $x_n \geq y_n$ if and only if $h(x_n|y_n, \sigma, \alpha) > 0$ for all $x_n \geq y_n$. Note $h'(x_n|y_n, \sigma, \alpha) = 2 - \frac{(\alpha-1)\sigma^\alpha}{x_n^\alpha}$, so that $h'(x_n|y_n, \sigma, \alpha) = 0$ if and only if $x_n = x_n^* \equiv \sigma\left(\frac{\alpha-1}{2}\right)^{1/\alpha}$. Also note, $h''(x_n|y_n, \sigma, \alpha) = \frac{\alpha(\alpha-1)\sigma^\alpha}{x^{\alpha+1}} > 0$. Thus, $x_n = x_n^*$ uniquely maximizes $h(x_n|y_n, \sigma, \alpha)$. Given $h'(x_n|y_n, \sigma, \alpha) > 0$ for all $x_n > x_n^*$ and $h(y_n|y_n, \sigma, \alpha) = 0$, $\Delta^{**}(x_n, y_n)$ satisfies diminishing sensitivity for all $x_n \geq y_n$ if and only if $x_n^* \leq y_n$ or $h(x_n^*|y_n, \sigma, \alpha) \geq 0$ (or both). Given $h'(x_n|y_n, \sigma, \alpha) < 0$ for all $x_n < x_n^*$ and $h(y_n|y_n, \sigma, \alpha) = 0$, $x_n^* > y_n$ implies $h(x_n^*|y_n, \sigma, \alpha) < 0$. Taken together, these last two observations imply $\Delta^{**}(x_n, y_n)$ satisfies diminishing sensitivity for all $x_n \geq y_n$ if and only if $x_n^* \leq y_n$, which, using the definitions of x_n^* and of $\hat{\sigma}(y_n)$, we can see this is equivalent to $\hat{\sigma}(y_n) = \left(\frac{2}{\alpha-1}\right)^{1/\alpha} y_n$. Computing $\frac{\partial^2\Delta^{**}(x_n, y_n)}{\partial x_n^2}$, multiplying through by $x_n^2(x_n^\alpha + y_n^\alpha + \sigma^\alpha) > 0$, diving by $\alpha x_n^\alpha(2y_n^\alpha + \sigma^\alpha) > 0$, and rearranging, we see $\Delta^{**}(x_n, y_n)$ is concave in x_n for all $x_n \geq y_n$ if and only if $x_n^\alpha(1 + \alpha) \geq (\alpha - 1)(y_n^\alpha + \sigma^\alpha)$ for all $x_n \geq y_n$. Since the left-side of this inequality is clearly increasing in x_n, $\Delta^{**}(x_n, y_n)$ is concave in x_n for all $x_n \geq y_n$ if and only if the inequality holds at $x_n = y_n$, i.e., if and only if $y_n^\alpha(1 + \alpha) \geq (\alpha - 1)(y_n^\alpha + \sigma^\alpha)$. Solving for σ, we see this condition is equivalent to $\sigma \leq \hat{\sigma}(y_n) = \left(\frac{2}{\alpha-1}\right)^{1/\alpha} y_n$, as desired.

For part (iii), let $\check{x}(\sigma) \equiv \left\{x_n : \sigma^\alpha = \frac{2(x_n - y_n)}{x_n^{1-\alpha} - y_n^{1-\alpha}}\right\} > y_n$, and $\hat{x}(\sigma) \equiv \left(\frac{(\alpha-1)(y_n^\alpha + \sigma^\alpha)}{\alpha+1}\right)^{1/\alpha} > y_n$. Using our definitions of $h(x_n|y_n, \sigma, \alpha)$ and $\check{x}(\sigma)$, $h(\check{x}(\sigma)|y_n, \sigma, \alpha) = 0$ is readily verifiable. Given $x_n^* > y_n$ for $\sigma > \hat{\sigma}(y_n)$ from part (i), $h(y_n|y_n, \sigma, \alpha) = 0$, $h'(x_n|y_n, \sigma, \alpha) < 0$ for all $x_n < x_n^*$, and $h'(x_n|y_n, \sigma, \alpha) > 0$ for all $x_n > x_n^*$, it follows that $\check{x}(\sigma) > x_n^*$, implying $h(x_n|y_n, \sigma, \alpha) < 0$ for $y_n < x_n < \check{x}_n$ and $h(x_n|y_n, \sigma, \alpha) > 0$ for $x_n > \check{x}_n$. Recalling from part (i) that $\Delta^{**}(x_n, y_n)$ is concave in x_n if and only if $x_n^\alpha(1 + \alpha) \geq (\alpha - 1)(y_n^\alpha + \sigma^\alpha)$, we can rearrange this inequality to see that it binds at $\hat{x}(\sigma)$. By inspection, we can then see that $x_n < \hat{x}(\sigma)$ implies $x_n^\alpha(1 + \alpha) < (\alpha - 1)(y_n^\alpha + \sigma^\alpha)$ and $x_n > \hat{x}(\sigma)$ implies $x_n^\alpha(1 + \alpha) > (\alpha - 1)(y_n^\alpha + \sigma^\alpha)$, implying the desired result. Expressing $\tilde{h}(\check{x}, \sigma, y) \equiv h(\check{x}(\sigma)|y_n, \sigma, \alpha) = 2(\check{x} - y_n) + \sigma^\alpha(\check{x}^{1-\alpha} - y_n^{1-\alpha}) = 0$, we see $\frac{\partial\tilde{h}(\check{x}, \sigma, y)}{\partial\check{x}} = 2 - (\alpha - 1)\check{x}^{-\alpha}\sigma^\alpha$, $\frac{\partial\tilde{h}(\check{x}, \sigma, y_n)}{\partial y_n} = -2 + (\alpha - 1)y_n^{-\alpha}\sigma^\alpha$, and $\frac{\partial\tilde{h}(\check{x}, \sigma, y_n)}{\partial\sigma} = \alpha\sigma^{\alpha-1}\left(\frac{1}{\check{x}^{\alpha-1}} - \frac{1}{y_n^{\alpha-1}}\right) < 0$. Next, observe $\frac{\partial\tilde{h}(\check{x}, \sigma, y)}{\partial\check{x}} = 2 + (1 - \alpha)\sigma^\alpha\check{x}^{-\alpha} > 0$. Together, from the implicit function theorem, these inequalities imply $\check{x}(\sigma)$ is increasing in σ. By inspection, we can also readily verify that $\hat{x}(\sigma)$ is increasing in σ since, holding $\alpha > 1$ fixed, $\hat{x}(\sigma)$ is clearly increasing in $(y^\alpha + \sigma^\alpha)$ and $(y^\alpha + \sigma^\alpha)$ is clearly increasing in σ. □

To help convey key features of $\Delta^{**}(x_n, y_n)$, Proposition 15 effectively fixes the smaller attribute value, taken here to be y_n, while allowing the larger attribute value x_n to vary. Of particular relevance, if σ is sufficiently small in relation to y_n, $\Delta^{**}(x_n, y_n)$ will exhibit diminishing sen-



Left: increasing σ leads to the emergence and then expansion of a convex region of the contrast function, with σ determining the point at which contrast is maximally responsive to changes in x_n (for fixed $\alpha > 1$). Right: when $\alpha \leq 1$ the contrast function is concave, while its responsivity becomes more concentrated over a smaller range for larger α (for fixed $\sigma > 0$).

Fig. 7. The effects of σ and α on contrast.

sitivity and strict concavity (in x_n) for all $x_n \geq y_n$. If σ is large in relation to y_n, however, $\Delta^{**}(x_n, y_n)$ will instead exhibit increasing sensitivity and convexity for values of x_n that are sufficiently close to y_n.

Therefore σ determines where the direct contrast function is maximally responsive to a change in x_n relative to y_n. Since $\Delta^{**}(x_n, y_n)$ is most responsive to changes in x_n at the threshold $\hat{x}(\sigma)$, the effect of increasing σ can also be understood here as shifting this point of maximum responsiveness further to the right (Fig. 7, left).[31] As noted by Rayo and Becker (2007), a bounded value function with such properties is optimal when agents are limited in their ability to discriminate small differences. While σ has been typically treated as a constant in the neuroscience literature (e.g. Shevell, 1977; Heeger, 1992; Louie et al., 2011), recent work suggests σ may arise dynamically in neural systems from the history of stimuli (LoFaro et al., 2014; Louie et al., 2014; Khaw et al., 2017; Tymula and Glimcher, 2019), thus acting as a dynamic reference point.

The constant $\alpha > 1$ determines the extent to which the responsiveness of $\Delta^{**}(x_n, y_n)$ is concentrated over a small range of x_n, as opposed to being dispersed over a large range. That is, as $\alpha > 1$ increases, $\Delta^{**}(x_n, y_n)$ becomes more responsive to changes in x_n near $\hat{x}(\sigma)$, but becomes less responsive for x_n further from $\hat{x}(\sigma)$. For example, in the limit as $\alpha \to \infty$, $\Delta^{**}(x_n, y_n)$ assumes the shape of a step-function that is infinitely responsive at $\hat{x}(\sigma)$ but unresponsive to changes in x_n everywhere else (Fig. 7, right).

B.6. Binary-choice equivalence with alternate normalization models

Recall, in two-attribute binary choice, the basic PN model can be equivalently represented by a symmetric Cobb-Douglas preference model, $V^{CD}(\mathbf{x}) = x_1 x_2$ (Observation 1). The next result

[31] This interpretation of $\hat{x}(\sigma)$ follows because $\frac{\partial \Delta^{**}(x_n, y_n)}{\partial x_n} > 0$ is increasing where $\Delta^{**}(x_n, y_n)$ is convex and decreasing where $\Delta^{**}(x_n, y_n)$ is concave.

shows that this equivalence generalizes to all of the alternate normalization models considered in Section 8.1.

Proposition 16. *Given* $\mathbf{x} = (x_1, x_2)$*,* $\mathbf{y} = (y_1, y_2)$*, and* $X = \{\mathbf{x}, \mathbf{y}\}$*, the DM chooses* $\mathbf{x}$ *if and only if* $x_1 x_2 > y_1 y_2$ *in all of the following models (as defined in Section 8.1):*

(i) *the joint normalization model;*
(ii) *the average normalization model;*
(iii) *the maximum normalization model;*
(iv) *the minimum normalization model;*
(v) *the max-min normalization model.*

Proof. It is readily verifiable that, given $||X|| = 2$, the joint normalization model and the basic PN model are equivalent. Thus, part (i) follows from Observation 1. For part (ii), it is readily verifiable that, under average normalization, $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ is strictly increasing in y_2 and that $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ given $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = \left(y_1, \frac{x_1 x_2}{y_1}\right)$. Together, these properties imply $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ if and only if $x_1 x_2 > y_1 y_2$. For parts (iii) and (iv), we can consider any $N \geq 2$. We can then compute $\Delta(a, b | a > b) = \frac{a}{a+a} + \frac{b}{a+b} = \frac{a-b}{2(a+b)}$ with $\Delta(a, b | a < b) = -\Delta(a, b | a > b)$ under maximum normalization and $\Delta(a, b | a > b) = \frac{a}{a+b} + \frac{b}{b+b} = \frac{a-b}{2(a+b)}$ with $\Delta(a, b | a < b) = -\Delta(a, b | a > b)$ under minimum normalization. Now $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ if and only if $\sum_{n=1}^{N} \Delta(x_n, y_n) > 0$, which itself is equivalent to $\sum_{n=1}^{N} 2\Delta(x_n, y_n) > 0$. Since $2\Delta(a, b) = \frac{|a-b|}{a+b}$ under both maximum and minimum normalization is the same as $\Delta(a, b)$ under (1), $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ in the former models must hold if and only if $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ holds under (1). The desired result then follows from Observation 1. For part (v), it is readily verifiable that, under max-min normalization and with $N = 2$, $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ is strictly increasing in y_2 and that $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ given $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = \left(y_1, \frac{x_1 x_2}{y_1}\right)$. Together, these properties imply $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ if and only if $x_1 x_2 > y_1 y_2$. □

B.7. Analysis with attributes in observable units

As discussed in Section 3.1, for attributes that are observed as numerical consumption levels, alternatives' attribute values can implicitly be equated to a representation that expresses the value of $\mathbf{x}$ on dimension n through a function u_n that maps an observed consumption level $c_n(\mathbf{x}) \geq 0$ to utilities — that is, with $x_n = u_n(c_n(\mathbf{x}))$. This appendix presents versions of our main behavioral propositions in which the original conditions on attributes' utility values are instead placed on their associated consumption levels, which are presumably expressed in an observable unit, such as dollars. To start, we consider the case in which the underlying utility functions are the same for all attributes, i.e. $u_n = u$ for all n, so that the basic PN model can be re-expressed as:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} \sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{u(c_n(\mathbf{x}))}{u(c_n(\mathbf{x})) + u(c_n(\mathbf{y}))}. \tag{11}$$

Proposition 17. *Suppose* u *is strictly increasing and weakly concave with* $u(0) = 0$*. Then, if the restrictions on each* x_n *in the following results are instead imposed on each corresponding* $c_n(\mathbf{x})$*, the result still holds under (11):*

(i) *the relative difference effect (Proposition 1);*
(ii) *the compromise effect (Proposition 2, part i);*
(iii) *the dominance effect (Proposition 2, part ii, which covers weak and strict versions);*
(iv) *the decoy-range effect (Proposition 3, part i);*
(v) *majority-rule preference cycles (Proposition 4);*
(vi) *the splitting bias (Proposition 5);*
(vii) *the alignability effect (Proposition 6);*
(viii) *the diversification bias (Proposition 7);*
(ix) *the feature bias (Proposition 8).*

Proof. For part (i), Observation 1 implies $\mathbf{x}'$ is chosen with $X = \{\mathbf{x}', \mathbf{y}'\}$ if and only if $u(c_1(\mathbf{x}))u(c_2(\mathbf{x}) + k) > u(c_1(\mathbf{y}))u(c_2(\mathbf{y}) + k)$. Since $u(c_1(\mathbf{x}))u(c_2(\mathbf{x})) = u(c_1(\mathbf{y}))u(c_2(\mathbf{y}))$ with binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$, we can re-express the desired condition as $u(c_1(\mathbf{x}))(u(c_2(\mathbf{x}) + k) - u(c_2(\mathbf{x}))) > u(c_1(\mathbf{y}))(u(c_2(\mathbf{y}) + k) - u(c_2(\mathbf{y})))$. Since u is increasing and $c_1(\mathbf{x}) > c_1(\mathbf{y})$, $u(c_1(\mathbf{x})) > u(c_1(\mathbf{x}))$. Since u is concave and $c_2(\mathbf{x}) < c_2(\mathbf{y})$, $u(c_2(\mathbf{x}) + k) - u(c_2(\mathbf{x})) \geq u(c_2(\mathbf{y}) + k) - u(c_2(\mathbf{y}))$. Together, these properties guarantee $u(c_1(\mathbf{x}))(u(c_2(\mathbf{x}) + k) - u(c_2(\mathbf{x}))) > u(c_1(\mathbf{y}))(u(c_2(\mathbf{y}) + k) - u(c_2(\mathbf{y})))$, as desired.

For parts (ii) and (iii), it suffices to show that the conditions on x_1, x_2, y_1, y_2, z_1, and z_2 in parts (i) and (ii) of Proposition 2 are necessarily satisfied given the conditions are satisfied for $c_1(\mathbf{x})$, $c_2(\mathbf{x})$, $c_1(\mathbf{y})$, $c_2(\mathbf{y})$, $c_1(\mathbf{z})$, and $c_2(\mathbf{z})$. This must be true because, for any $\mathbf{x}', \mathbf{x}''$ and n, $u(c_n(\mathbf{x}')) \gtreqless u(c_n(\mathbf{x}''))$ (i.e. $x_n' \gtreqless x_n''$) if and only if $c_n(\mathbf{x}') \gtreqless c_n(\mathbf{x}'')$ since u is strictly increasing. For part (iv), $c_1(\mathbf{x}) > c_1(\mathbf{z}) = c_1(\mathbf{z}') > c_1(\mathbf{y})$ and $c_2(\mathbf{y}) > c_2(\mathbf{x}) \geq c_2(\mathbf{z}) > c_2(\mathbf{z}')$ similarly implies $x_1 > z_1 = z_1' > y_1$ and $y_2 > x_2 \geq z_2 > z_2'$ since u is strictly increasing. For part (v), the result likewise follows because $x_1 > x_1' > x_1''$, $x_2'' > x_2 > x_2'$, and $x_3' > x_3'' > x_3$ all hold if and only if $c_1(\mathbf{x}) > c_1(\mathbf{x}') > c_1(\mathbf{x}'')$, $c_2(\mathbf{x}'') > c_2(\mathbf{x}) > c_2(\mathbf{x}')$, and $c_3(\mathbf{x}') > c_3(\mathbf{x}'') > c_3(\mathbf{x})$ with u strictly increasing.

For part (vi), we want $\Delta(u(c_{1a}(\mathbf{x})), u(c_{1a}(\mathbf{y}))) + \Delta(u(c_{1b}(\mathbf{x})), u(c_{1b}(\mathbf{y}))) > \Delta(u(c_1(\mathbf{x})), u(c_1(\mathbf{y})))$. Let $\delta_i \equiv c_{1i}(\mathbf{x}) - c_{1i}(\mathbf{y})$ for $i \in \{a, b\}$ and $\delta \equiv c_1(\mathbf{x}) - c_1(\mathbf{y})$, implying $\delta_a + \delta_b = \delta$. It suffices to show $\Delta(u(c_{1i}(\mathbf{x})), u(c_{1i}(\mathbf{y}))) > \frac{\delta_i}{\delta} \cdot \Delta(u(c_1(\mathbf{x})), u(c_1(\mathbf{y})))$ for $i \in \{a, b\}$, i.e. $Z(c_{1i}(\mathbf{y}), c_1(\mathbf{y}), \delta_i, \delta) \equiv \Delta(u(c_{1i}(\mathbf{y}) + \delta_i), u(c_{1i}(\mathbf{y}))) - \frac{\delta_i}{\delta} \cdot \Delta(u(c_1(\mathbf{y}) + \delta), u(c_1(\mathbf{y}))) > 0$. Next, $\frac{\partial Z(c_{1i}(\mathbf{y}), c_1(\mathbf{y}), \delta_i, \delta)}{\partial c_{1i}(\mathbf{y})} = \frac{2(u(c_{1i}(\mathbf{y}))u'(c_{1i}(\mathbf{y})+\delta_i) - u(c_{1i}(\mathbf{y})+\delta_i)u'(c_{1i}(\mathbf{y}))))}{(u(c_{1i}(\mathbf{y})+\delta_i)+u(c_{1i}(\mathbf{y}))))^2} < 0$ since $u(c_{1i}(\mathbf{y}) + \delta_i) > u(c_{1i}(\mathbf{y})) > 0$ and $u'(c_{1i}(\mathbf{y})) \geq u'(c_{1i}(\mathbf{y}) + \delta_i) > 0$ given u is strictly increasing and weakly concave. This implies $Z(c_{1i}(\mathbf{y}), c_1(\mathbf{y}), \delta_i, \delta)$ is minimized at $c_{1i}(\mathbf{y}) = c_1(\mathbf{y})$, so that $Z(c_1(\mathbf{y}), c_1(\mathbf{y}), \delta_i, \delta) > 0$ is sufficient for $Z(c_{1i}(\mathbf{y}), c_1(\mathbf{y}), \delta_i, \delta) > 0$. Using the definition of Δ and multiplying through by $\delta_i^{-1}(u(c_1(\mathbf{y}) + \delta) + u(c_1(\mathbf{y})))(u(c_1(\mathbf{y}) + \delta_i) + u(c_1(\mathbf{y}))) > 0$, we see $Z(c_1(\mathbf{y}), c_1(\mathbf{y}), \delta_i, \delta) > 0$ is equivalent to $\frac{u(c_1(\mathbf{y})+\delta_i)-u(c_1(\mathbf{y}))}{\delta_i} \cdot (u(c_1(\mathbf{y}) + \delta) + u(c_1(\mathbf{y}))) > \frac{u(c_1(\mathbf{y})+\delta)-u(c_1(\mathbf{y}))}{\delta} \cdot (u(c_1(\mathbf{y}) + \delta_i) + u(c_1(\mathbf{y})))$. Since u is strictly increasing and $\delta_i < \delta$, $u(c_1(\mathbf{y}) + \delta) + u(c_1(\mathbf{y})) > u(c_1(\mathbf{y}) + \delta_i) + u(c_1(\mathbf{y}))$. Therefore, it suffices to show $\frac{u(c_1(\mathbf{y})+\delta_i)-u(c_1(\mathbf{y}))}{\delta_i} \geq \frac{u(c_1(\mathbf{y})+\delta)-u(c_1(\mathbf{y}))}{\delta}$, which must hold since u is concave.

For part (vii), with $\mathbf{x}'$ and $\mathbf{y}'$ defined as in Proposition 6, $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \Delta(u(c_1(\mathbf{x})), u(c_1(\mathbf{y}))) + \Delta(u(c_2(\mathbf{x})), u(0)) - \Delta(0, u(c_3(\mathbf{y}')))$, noting $c_3(\mathbf{y}') = c_2(\mathbf{y})$. Since $\Delta(u(c_2(\mathbf{x})), u(0)) = \Delta(0, u(c_3(\mathbf{y}'))) = 1$ and $\Delta(u(c_1(\mathbf{x})), u(c_1(\mathbf{y}))) = \frac{u(c_1(\mathbf{x}))-u(c_1(\mathbf{y}))}{u(c_1(\mathbf{x}))+u(c_1(\mathbf{y}))} > 0$ given $c_1(\mathbf{x}) > c_1(\mathbf{y})$ and u strictly increasing, $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) > 0$ is assured, implying the alignability effect still holds.

For part (viii), with $\mathbf{x}$ and $\mathbf{x}'$ as defined in Proposition 7, define $\mathbf{z} \neq \mathbf{x}$ such that $c_n(\mathbf{z}) \equiv u^{-1}\big(\frac{A}{N} - \big(c_n(\mathbf{x}') - \frac{A}{N}\big)u'\big(\frac{A}{N}\big)\big)$. Then $\sum_{n=1}^{N} u(c_n(\mathbf{z})) = \sum_{n=1}^{N} u(c_n(\mathbf{x})) = \sum_{n=1}^{N} u\big(\frac{A}{N}\big)$, implying $V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\}) > V(\mathbf{z}; \{\mathbf{x}, \mathbf{z}\})$ from Proposition 7. Next, since u is concave, $c_n(\mathbf{x}') \leq \frac{A}{N} - \big(c_n(\mathbf{x}') - \frac{A}{N}\big)u'\big(\frac{A}{N}\big)$ for all n, implying $u(c_n(\mathbf{x}')) \leq u(c_n(\mathbf{z}))$ for all n, which ensures $V(\mathbf{x}', \{\mathbf{x}, \mathbf{x}'\}) < V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\})$ since $V(\mathbf{z}, \{\mathbf{x}, \mathbf{z}\}) < V(\mathbf{x}; \{\mathbf{x}, \mathbf{z}\})$.

For part (ix), with $\mathbf{x}$ and $\mathbf{x}'$ as defined in Proposition 8, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \Delta(u(q), 0) - \Delta(u(c_n(\mathbf{x}) + q), u(c_n(\mathbf{x}))) = \frac{2u(c_n(\mathbf{x}))}{u(c_n(\mathbf{x})+q)+u(c_n(\mathbf{x}))} > 0$. □

Thus, under (11), all of our main behavioral propositions (i.e. those corresponding to the behaviors listed in Table 1) still apply when the original restrictions on attributes' utility values (i.e. each x_n) are placed on the observable consumption levels (each $c_n(\mathbf{x})$) instead, provided u is strictly increasing and weakly concave, with $u(0) = 0$. It is indeed common to assume that utility functions over monetary payoffs (among other observable measures of consumption) are increasing and concave. However, such utility functions are generally only defined up to a positive affine transformation, and thus do not require $u(0) = 0$. Here, the $u(0) = 0$ restriction is necessary — and in fact already implicit in our original formulation, which presumed $x_n = 0$ for the case in which $\mathbf{x}$ provides nothing on dimension n.

Of note, parts (ii), (iii), (iv), and (v) of Proposition 17 all follow directly from the original results to which they correspond (and in fact do not require u to be concave). This is because the only conditions on attribute values in the original results were those that specified the DM's rankings of alternatives on a given attribute dimension (e.g. $x_n > x_n'$), and these conditions are in fact equivalent to the new conditions on the associated consumption levels (e.g. $c_n(\mathbf{x}) > c_n(\mathbf{x}')$) given u is strictly increasing. Parts (i), (vi), (vii), (viii), and (ix), however, do not follow as directly because their original results featured additional conditions on attribute values. As noted in footnote 14, these additional conditions were all used to ensure that a given entity had the same unnormalized value regardless of how it was framed or otherwise represented in terms of the particular alternative and/or attribute dimensions on which it was expressed. Table 5 summarizes these additional conditions — which are consistent with standard interpretations of the behavioral phenomena they are intended to address — as they appear in each original result.

By re-expressing the relevant conditions in Table 5 in terms of numerical consumption levels, the standard interpretations of these conditions (as formally translated in the right-most column) may no longer apply under (11). For example, in the version of the splitting bias from part (vi) of Proposition 17, the new condition no longer ensures that the unnormalized value of an alternative stays the same when one of its attributes is reframed as two subattributes, i.e. $c_{1a}(\mathbf{w}) + c_{1b}(\mathbf{w}) = c_1(\mathbf{w})$ does not imply $u(c_{1a}(\mathbf{w})) + u(c_{1b}(\mathbf{w})) = u(c_1(\mathbf{w}))$ for $\mathbf{w} \in \{\mathbf{x}, \mathbf{y}\}$. Despite this issue — and similar issues for the other behaviors highlighted in Table 5 — Proposition 17 demonstrates that these behaviors are still captured under (11).

Aside from the unnatural interpretations of the conditions in Table 5 when re-expressed in terms of numerical consumption levels, the setup underlying Proposition 17 is also limited by the fact that the same utility function u is used on all dimensions. After all, different attributes could conceivably be evaluated using different utility functions, i.e. with a unique u_n for each n, as is common in the theoretical literature (e.g. Koszegi and Szeidl, 2013; Bushong et al., 2019). In light of this, we now modify (11) as follows:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} \sum_{\mathbf{y} \in X \setminus \mathbf{x}} \frac{u_n(c_n(\mathbf{x}))}{u_n(c_n(\mathbf{x})) + u_n(c_n(\mathbf{y}))}. \tag{12}$$

Table 5
Additional conditions on attribute values.

Prop.	Behavior	Relevant Condition(s)	Interpretation
1	Rel. Diff. Effect	$x_2' - x_2 = y_2' - y_2$	A given improvement on attribute 2 provides the same increase in unnormalized value regardless of whether it is applied to **x** or **y**.
5	Splitting Bias	$w_1 = w_{1a} + w_{1b}$, $\mathbf{w} \in \{\mathbf{x}, \mathbf{y}\}$	The unnormalized value of an alternative does not change if attribute 1 is separated into subattributes 1a and 1b.
6	Align. Effect	$y_2 = y_3'$	The unnormalized value of **y** on attribute 2 does not change if it is re-expressed on (non-alignable) attribute 3.
7	Divers. Bias	$x_n = \frac{A}{N}$ for all n, $\sum_n x_n' = A$	An asset's (total) unnormalized value does not depend on how it is allocated across attribute dimensions (that are presumed to generate equal rates of return).
8	Feature Bias	$x_N = x_{n'}' - x_{n'} > 0$	A given improvement provides the same increase in unnormalized value regardless of whether it is applied to (new) attribute N or to (existing) attribute n'.

One complication of allowing attribute-specific utility functions is that the conditions in Table 5 — i.e. those used to ensure that the unnormalized value of a given entity does not depend on how it is framed or otherwise represented — are even less naturally expressed than with a single utility function, as unnormalized attribute values can vary in even more arbitrary ways. When considering the splitting bias, for instance, there are no restrictions as to how the subattribute utility functions u_{1a} and u_{1b} relate to the composite utility function u_1 (or to each other) under (12), despite the fact that the subattributes themselves are simply two components of the composite attribute. As a result, even when u_1 and $c_1(\mathbf{x}) = c_{1a}(\mathbf{x}) + c_{1b}(\mathbf{x})$ are given, each of the subattribute utilities $u_{1a}(c_{1a}(\mathbf{x}))$ and $u_{1b}(c_{1b}(\mathbf{x}))$ can be *any* positive real number, which makes the relationship between $u_1(c_1(\mathbf{x}))$ and $u_{1a}(c_{1a}(\mathbf{x})) + u_{1b}(c_{1b}(\mathbf{x}))$ impossible to pin down. In our analysis of (12), we therefore impose additional restrictions (in two cases) to ensure u_n does not differ across dimensions on which the same underlying component of an alternative may exist.

Proposition 18. *Suppose u_n is strictly increasing and weakly concave with $u_n(0) = 0$ for all $n = 1, \ldots, N$. Then, if the restrictions on each x_n in the following results are instead imposed on each corresponding $c_n(\mathbf{x})$, the result still holds under (12):*

(i) *the relative difference effect (Proposition 1);*
(ii) *the compromise effect (Proposition 2, part i);*
(iii) *the dominance effect (Proposition 2, part ii, which covers weak and strict versions);*
(iv) *the decoy-range effect (Proposition 3, part i);*
(v) *majority-rule preference cycles (Proposition 4);*
(vi) *the splitting bias (Proposition 5), provided $u_1 = u_{1a} = u_{1b}$ also holds;*
(vii) *the alignability effect (Proposition 6);*
(viii) *the diversification bias (Proposition 7), provided $u_n = u$ for all n;*
(ix) *the feature bias (Proposition 8).*

Proof. By substituting $u_n(c_n(\mathbf{x}))$ for each $u(c_n(\mathbf{x}))$ in the corresponding proofs in Proposition 17, it is readily verifiable that parts (i), (ii), (iii), (iv), (v), (vii), and (ix) still hold. For part (vi), the proof is identical to the proof of part (vi) of Proposition 17, provided we substitute u_1 for each u (note, u_2 does not appear in the original proof, so it does not need to be further adapted to allow $u_1 \neq u_2$). Part (viii) is equivalent to part (viii) of Proposition 17, and therefore follows from this earlier result. □

Thus, the implications of Proposition 17 extend to the case in which observed consumption levels are translated through attribute-specific utility functions. As seen, however, this conclusion requires qualification when considered for the splitting bias and for the diversification bias. For the splitting bias (part vi), the additional $u_1 = u_{1a} = u_{1b}$ restriction simply asserts that the utility function associated with the composite attribute is maintained after the attribute has been split. That said, this restriction does not guarantee that composite attribute values are equal to the sum of their corresponding subattribute values, i.e. it is still possible that $u_1(c_1(\mathbf{w})) \neq u_{1a}(c_{1a}(\mathbf{w})) + u_{1b}(c_{1b}(\mathbf{w}))$ for $\mathbf{w} \in \{\mathbf{x}, \mathbf{y}\}$. Indeed, it may have been more reasonable to impose $u_1(c_1(\mathbf{w})) = u_{1a}(c_{1a}(\mathbf{w})) + u_{1b}(c_{1b}(\mathbf{w}))$ for $\mathbf{w} \in \{\mathbf{x}, \mathbf{y}\}$, but this would simply bring us back to our original result in Proposition 5. As for the diversification bias (part viii), the additional restriction simply ensures that, in keeping with our original specification in which allocations of the asset were presumed to generate equal returns on all dimensions, the utility functions on each dimension must be the same.[32] Of course, however, this additional restriction just gets us back to the version of the diversification bias in part (viii) of Proposition 17. With all that said, even though allowing attribute-specific utility functions in (12) exacerbates the issues of interpretation for *all* of the conditions listed in Table 5 under (11), Proposition 18 shows that the other three affected behavioral predictions — i.e. the relative difference effect (Proposition 1), the alignability effect (Proposition 6), and the feature bias (Proposition 8) — still hold under (12) without any additional restrictions on the attribute-specific utility functions.

Appendix C. Classifying other models' predictions

This appendix explains how other models' predictions were classified in Tables 1 and 2. For each of the comparable models listed in Table 1, we will describe the value function $V(\mathbf{x}; X)$ used to classify the model's predictions and demonstrate that it generates the corresponding predictions listed in the table (these value functions were also used to generate the corresponding graphs shown in Figs. 2 and 4). The models listed in the footnote of Table 1 but not classified in the main table will be addressed at the end of this appendix. For the alternate normalization models considered in Section 8.1, we will characterize each model's predictions using (5) along with the corresponding definition of $r(x_n)$ in Table 2.

For clarity and to facilitate consistent comparisons across models, certain restrictions were applied to some models. For instance, we only considered deterministic versions of each model and, with one exception (as explained in Section C.6) presumed that attributes are ex-ante symmetric, so that any attribute-specific parameters or functions were taken to be the same across dimensions. These and other model-specific restrictions (discussed below) may lead to a classification of '**Y**' (robustly captures the behavior) or 'N' (predicts no effect or the opposite effect)

[32] Without this restriction, we could (as an example) take $u_1(c_1(\mathbf{x})) = 10{,}000 \cdot c_1(\mathbf{x})$ and $u_2(c_2(\mathbf{x})) = .1 \cdot \sqrt{c_2(\mathbf{x})}$, in which case an allocation with $c_1(\mathbf{x}) = c_2(\mathbf{x}) = 100$ would generate grossly unequal returns on dimensions 1 and 2, with $u_1(c_1(\mathbf{x})) = 1{,}000{,}000$ and $u_2(c_2(\mathbf{x})) = 1$.

when a more general version of the model would imply 'S' (captures the behavior in some cases, but predicts the opposite in other cases). However, these restrictions can never prevent a '**Y**' or 'N' classification. Thus, if we re-created Table 1 using more general versions of each model, each classification would either remain the same or change to 'S'.

The rules used to classify each prediction are then based on whether or not $V(\mathbf{x}; X)$ as given for that model predicts the corresponding result as formalized in this paper. In particular, the operative definitions used in this analysis are as follows:

- As in part (i) of Proposition 2, the *compromise effect* is captured if the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$; $\mathbf{x}$ is a compromise between $\mathbf{y}$ and $\mathbf{z}$ in that $z_1 > x_1 > y_1$ and $y_2 > x_2 > z_2$; and $\mathbf{z}$ is not chosen from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$.
- As in part (ii) of Proposition 2 with $z_2 = x_2$, the *"weak" dominance effect* is captured if the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$; and $\mathbf{x}$ asymmetrically dominates $\mathbf{z} \neq \mathbf{x}$ with $x_1 > z_1 > y_1$ and $y_2 > x_2 = z_2$.
- As in part (ii) of Proposition 2 with $z_2 < x_2$, the *"strict" dominance effect* is captured if the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$; and $\mathbf{x}$ asymmetrically dominates $\mathbf{z} \neq \mathbf{x}$ with $x_1 > z_1 > y_1$ and $y_2 > x_2 > z_2$.
- As in part (i) of Proposition 3, the *decoy-range effect* is captured if the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$; the DM chooses $\mathbf{y}$ from $X = \{\mathbf{x}, \mathbf{y}\}$; $x_1 > z_1 = z_1' > y_1$; and $y_2 > x_2 \geq z_2 > z_2'$.
- As in Proposition 1, the *relative difference effect* is captured if the DM chooses $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{y}'\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$; $\mathbf{x}' = (x_1, x_2 + k)$; and $\mathbf{y}' = (y_1, y_2 + k)$; and $k > 0$.
- As in Proposition 4, *majority-rule preference cycles* are captured if the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$, $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{x}''\}$, and $\mathbf{x}''$ from $X = \{\mathbf{x}, \mathbf{x}''\}$ given: $N = 3$; $x_1 > x_1' > x_1''$; $x_2'' > x_2 > x_2'$; $x_3' > x_3'' > x_3$; and binary-choice preferences among $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ are intransitive.
- As in Proposition 5, the *splitting bias* is captured if the DM chooses $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{y}'\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$; $\mathbf{x}' = (x_{1a}, x_{1b}, x_2)$; $\mathbf{y}' = (y_{1a}, y_{1b}, y_2)$; $x_{1a} + x_{1b} = x_1$; $y_{1a} + y_{1b} = y_1$; $x_{1a} \geq y_{1a}$; and $x_{1b} \geq y_{1b}$.
- As in Proposition 6, the *alignability effect* is captured if the DM chooses $\mathbf{x}'$ from $X = \{\mathbf{x}', \mathbf{y}'\}$ given: the DM is indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ with $X = \{\mathbf{x}, \mathbf{y}\}$; $\min\{x_1, x_2, y_1, y_2\} > 0$; $\mathbf{x}' = (x_1, x_2, 0)$; and $\mathbf{y}' = (y_1, 0, y_2)$.
- As in Proposition 7, the *diversification bias* is captured if, for any $\mathbf{x}' \neq \mathbf{x}$ with $\sum_{n \leq N} x_n' = A$, the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$ given: $N > 1$; $A > 0$; and $x_n = \frac{A}{N}$ for all $n \leq N$.
- As in Proposition 8, the *feature bias* is captured if the DM chooses $\mathbf{x}$ from $X = \{\mathbf{x}, \mathbf{x}'\}$ given: $N > 1$; $q > 0$; $x_N = q$; $x_N' = 0$; $x_{n'}' = x_{n'} + q$ for some $n' < N$; and $x_n' = x_n > 0$ for all other $n < N$.

C.1. Tversky and Simonson (1993)

For Tversky and Simonson's (1993) model, we use the following value function (for consistency, we will express other models using the notation of the basic PN model, except where new notation is needed)[33]:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} x_n + \theta \cdot \sum_{\mathbf{y}\in X\setminus\mathbf{x}} \frac{\sum_n \max\{x_n - y_n, 0\}}{\sum_n \max\{x_n - y_n, y_n - x_n\}}, \quad \theta > 0. \tag{13}$$

Compromise Effect (**Y**). In (13), $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ if and only if $x_1 + x_2 = y_1 + y_2$. Taking $x_1 + x_2 = y_1 + y_2 = 1$ (without loss of generality), we know $z_1 + z_2 = 1 - \omega$ for some $\omega \in (0, 1)$ since $\mathbf{z}$ is inferior. In turn, if $\mathbf{z}$ makes $\mathbf{x}$ a compromise, it is readily verifiable that $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ if and only if $\frac{z_1 - x_1 + \omega}{2(z_1 - x_1) + \omega} > \frac{z_1 - y_1 + \omega}{2(z_1 - y_1) + \omega}$, which must hold since $z_1 > x_1 > y_1$.

Dominance Effect, Weak (**Y**) *and Strict* (**Y**). Given $x_1 > z_1 > y_1$ and $y_2 > x_2 \geq z_2$, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ must hold since $\frac{x_1 - z_1 + x_2 - z_2}{x_1 - z_1 + x_2 - z_2} = 1 > \frac{y_2 - z_2}{y_2 - z_2 + z_1 - y_1}$.

Decoy-Range Effect (N). It is also verifiable that the decoy-range effect is captured if $\frac{x_1 - z_1 + x_2 - z_2}{x_1 - z_1 + x_2 - z_2} - \frac{y_2 - z_2}{y_2 - z_2 + z_1 - y_1} = 1 - \frac{y_2 - z_2}{y_2 - z_2 + z_1 - y_1} > 0$ is increasing in z_2. However, this expression is decreasing in z_2 since $y_2 > z_2$.

Relative Difference Effect (N). Using (13) and with $\mathbf{x}'$ and $\mathbf{y}'$ as defined in Proposition 1, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ implies $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) = V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\})$. Thus, the relative difference effect is not captured.

Majority-Rule Preference Cycles (N). Using (13), $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ if and only if $\sum_{n=1}^{N} x_n > \sum_{n=1}^{N} y_n$. Since $\sum_{n=1}^{N} x_n > \sum_{n=1}^{N} y_n > \sum_{n=1}^{N} z_n > \sum_{n=1}^{N} x_n$ is a contradiction, intransitive choice (majority-rule or otherwise) is not possible.

Splitting Bias (N). Using the notation in Proposition 5, $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) > V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\})$ under (13) if and only if $\frac{(x_{1a} - y_{1a}) + (x_{1b} - y_{1b}) - (y_2 - x_2)}{x_{1a} - y_{1a} + x_{1b} - y_{1b} + y_2 - x_2} > \frac{(x_1 - y_1) - (y_2 - x_1)}{x_1 - y_1 + y_2 - x_2}$, but these expressions are equal given $x_{1a} + x_{1b} = x_1$ and $y_{1a} + y_{1b} = y_1$. Thus, the splitting bias is not captured.

Alignability Effect (N). Using the notation in Proposition 6, the alignability effect is captured if $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) > V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\})$, which holds under (13) if and only if $\frac{(x_1 - y_1) + (x_2 - 0) - (y_2 - 0)}{(x_1 - y_1) + (x_2 - 0) + (y_2 - 0)} > 0$. However, $x_1 + x_2 = y_1 + y_2$ given $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$, implying $\frac{(x_1 - y_1) + (x_2 - 0) - (y_2 - 0)}{(x_1 - y_1) + (x_2 - 0) + (y_2 - 0)} = 0$.

Diversification Bias (N). $V(\mathbf{x}; X) - V(\mathbf{x}'; X) \propto \sum_{n \leq N} \max\left\{\frac{A}{N} - x_n', 0\right\} - \sum_{n \leq N} \max\left\{x_n' - \frac{A}{N}, 0\right\} = 0$ given $X = \{\mathbf{x}, \mathbf{x}'\}$, $\sum_{n \leq N} x_n' = A$, and $x_n = \frac{A}{N}$ for $n \leq N$. Thus, the DM is indifferent between $\mathbf{x}$ and $\mathbf{x}'$, implying the diversification bias is not captured.

Feature Bias (N). Using the notation in Proposition 8, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \frac{q}{2q} + \sum_{n \notin \{n', N\}} x_n$ under (13). Thus, the DM is indifferent between $\mathbf{x}$ and $\mathbf{x}'$.

[33] For simplicity, here we take $\delta_i(t) = t$ (in their notation, see page 1885) and omit any influence of "background context" (besides the choice set).

C.2. Kivetz et al. (2004a)

For Kivetz et al.'s (2004a) model, we use:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} (x_n - \min_{\mathbf{x}' \in X}\{x'_n\})^c, \quad 0 < c < 1. \tag{14}$$

Compromise Effect (**Y**). Noting $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ if and only if $x_1 + x_2 = y_1 + y_2$ under (14), the compromise effect is captured since $(x_1 - y_1)^c + (x_2 - z_2)^c - (y_2 - z_2)^c > 0$ given $(x_1 - y_1) + (x_2 - z_2) - (y_2 - z_2) = x_1 + x_2 - y_1 - y_2 = 0$ and $0 < c < 1$.

Dominance Effect, Weak (N) *and Strict* (**Y**). The dominance effect likewise holds as a result of $(x_1 - y_1)^c + (x_2 - z_2)^c - (y_2 - z_2)^c > 0$, provided $z_2 < x_2$. If $z_2 = x_2$, however, $(x_1 - y_1)^c + (x_2 - z_2)^c - (y_2 - z_2)^c = (x_1 - y_1)^c - (y_2 - x_2)^c = 0$.

Decoy-range Effect (**Y**). The decoy-range effect is captured if $\frac{\partial}{\partial z_2}\left[(x_2 - z_2)^c - (y_2 - z_2)^c\right] = c((x_2 - z_2)^{c-1} - (y_2 - z_2)^{c-1}) > 0$, which must hold since $y_2 > x_2 > z_2$ and $c < 1$.

Relative Difference Effect (N). Under (14), the relative difference effect holds if $((x_1 + k) - (y_1 + k))^c > (x_1 - y_1)^c$ for $k > 0$. However, both expressions are clearly equal.

Majority-Rule Preference Cycles (**Y**). We first provide an example that shows majority-rule preference cycles are possible under (14). Namely, if $\mathbf{x} = (3, 2, 1)$, $\mathbf{x}' = (2, 1, 3)$, $\mathbf{x}'' = (1, 3, 2)$, and $c = \frac{1}{2}$, we can see from $\sqrt{3-2} + \sqrt{2-1} - \sqrt{3-1} = 2 - \sqrt{2} > 0$ that, in any binary choice, the DM chooses the alternative that is superior on two out of three dimensions. To show that minority-rule preference cycles are not possible, suppose $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{x}''$ satisfying the cyclical majority-dominance property where, without loss of generality, $x_1 + x_2 + x_3 = \min_{\mathbf{y} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\}}\{y_1 + y_2 + y_3\}$, $\mathbf{x}'$ is superior to $\mathbf{x}$ on two out of three attribute dimensions, and $x_1 > x'_1$. Hence, if a minority-rule preference cycle exists among $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$, we must have $\delta_1^c - \delta_2^c - \delta_3^c > 0$ for $\delta_n = |x_n - x'_n|$. Since $x_1 + x_2 + x_3 = \min_{\mathbf{y} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\}}\{y_1 + y_2 + y_3\}$, $\delta_1^c - \delta_2^c - \delta_3^c \leq (\delta_2 + \delta_3)^c - \delta_2^c - \delta_3^c$ while $(\delta_2 + \delta_3)^c - \delta_2^c - \delta_3^c < 0$ with $0 < c < 1$, a minority-rule preference cycle is impossible.

Splitting Bias (**Y**). The splitting bias is likewise captured under (14) since $(x_{1a} - y_{1a})^c + (x_{1b} - y_{1b})^c > (x_{1a} + x_{1b} - y_{1a} - y_{1b})^c = (x_1 - y_1)^c$ with $0 < c < 1$.

Alignability Effect (N). The alignability effect is likewise captured under (14) since $(y_2 - x_2)^c + x_2^c > y_2^c$ with $0 < c < 1$, where (as is readily verifiable) this condition ensures $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) > V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\})$ with $\mathbf{x}'$ and $\mathbf{y}'$ as defined in Proposition 6.

Diversification Bias (N). To show that (14) does not predict the diversification bias, it suffices to show an example for which the balanced allocation $\mathbf{x}$ is not chosen over $\mathbf{x}' \neq \mathbf{x}$ given $X = \{\mathbf{x}, \mathbf{x}'\}$. Take $x'_2 = \frac{2A}{N}$, $x'_1 = 0$, and $x'_n = \frac{A}{N}$ for all $n > 2$. Under (14), we then have $V(\mathbf{x}; X) = V(\mathbf{x}'; X) = \frac{A^c}{N^c}$, implying indifference between $\mathbf{x}$ and $\mathbf{x}'$.

Feature Bias (N). Under (14) and using the notation in Proposition 8, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = q^c$, implying there is no bias in favor of $\mathbf{x}$ over $\mathbf{x}'$.

C.3. Bordalo et al. (2013)

For Bordalo et al.'s (2013) model, we use:

$$V(\mathbf{x};X) = \frac{\sum_{n=1}^{N} \delta^{\sum_{m\neq n} \mathbb{I}[\rho_m(\mathbf{x};X)\geq \rho_n(\mathbf{x};X)]} \cdot x_n}{\sum_{n=1}^{N} \delta^{\sum_{m\neq n} \mathbb{I}[\rho_m(\mathbf{x};X)\geq \rho_n(\mathbf{x};X)]}}, \quad \rho_n(\mathbf{x};X) = \left|\frac{x_n - \bar{x}_n}{x_n + \bar{x}_n}\right|, \quad 0<\delta<1, \tag{15}$$

where $\bar{x}_n \equiv ||X||^{-1} \sum_{\mathbf{x}\in X} x_n$ is the mean attribute value in X on dimension n. This formulation uses the degree-zero homogeneous salience function given in equation (4) of Bordalo et al. (2013), which is denoted here as $\rho_n(\mathbf{x};X)$. For the special case of binary choice (taking $X = \{\mathbf{x}, \mathbf{y}\}$) with two attributes, the model in (15) reduces to:

$$V(\mathbf{x};X) = \begin{cases} \frac{\delta x_1 + x_2}{1+\delta}, & \rho_1(\mathbf{x};X) < \rho_2(\mathbf{x};X), \\ x_1 + x_2, & \rho_1(\mathbf{x};X) = \rho_2(\mathbf{x};X), \\ \frac{x_1 + \delta x_2}{1+\delta}, & \rho_1(\mathbf{x};X) > \rho_2(\mathbf{x};X), \end{cases} \qquad \rho_n(\mathbf{x};X) = \left|\frac{x_n - y_n}{3x_n + y_n}\right|. \tag{16}$$

As mentioned in the footnote of Table 4, Bordalo et al.'s model could be evaluated using a version in which one attribute is the price of the alternative or using a version in which all attributes represent different quality dimensions. To facilitate consistent comparisons across models, here we consider the latter version.[34]

Compromise Effect (S). To show that (15) sometimes predicts the compromise effect and sometimes predicts the opposite, it suffices to use examples. For instance, with $\delta = .5$, the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice but chooses $\mathbf{x}$ in trinary choice if $\mathbf{x} = (3, .5)$, $\mathbf{y} = (2, 1)$, and $\mathbf{z} = (3.2, 0)$, in which case a compromise effect is predicted, while the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice but chooses $\mathbf{y}$ in trinary choice if $\mathbf{x} = (1, 2)$, $\mathbf{y} = (.5, 3)$, and $\mathbf{z} = (1.2, 0)$, in which case the opposite effect is predicted.

Dominance Effect, Weak (S) *and Strict* (S). Maintaining $\delta = .5$, it is similarly verifiable that the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice but chooses $\mathbf{x}$ in trinary choice if $\mathbf{x} = (3, .5)$, $\mathbf{y} = (2, 1)$, and $\mathbf{z} = (2.8, 0)$, in which case a dominance effect with a strictly dominated decoy is predicted, while the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice but chooses $\mathbf{y}$ in trinary choice if $\mathbf{x} = (1, 2)$, $\mathbf{y} = (.5, 3)$, and $\mathbf{z} = (.8, 0)$, in which case the opposite effect is predicted. In turn, the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice but chooses $\mathbf{x}$ in trinary choice if $\mathbf{x} = (2, 1)$, $\mathbf{y} = (.5, 3)$, and $\mathbf{z} = (.75, 1)$, in which case a dominance effect with a weakly dominated decoy is predicted, while the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice but chooses $\mathbf{y}$ in trinary choice if $\mathbf{x} = (3, .5)$, $\mathbf{y} = (1, 2)$, and $\mathbf{z} = (1.5, .5)$, in which case the opposite effect is predicted.

[34] Following very similar arguments and examples as those used here, it is readily verifiable that all of the Table 1 classifications for Bordalo et al.'s model would be the same for the version of their model with price as an attribute, with the possible exception of the diversification bias, which would (depending on how a model with price as an attribute was translated to the formal setting considered in Proposition 7) either: (a) no longer be testable, since allocating an equal share of an asset A to a price dimension — formally, allocating more to this dimension would mean a higher price paid — would be unnatural and in violation of the "equal returns" assumption (i.e. there would be a negative return on this dimension and a positive return on others), or (b) would be unchanged if we presume that both allocations have the same price (which may naturally be the case if the asset represents a consumption budget or a monthly contribution to a savings plan, as examples) and where the asset itself can only be allocated to the remaining quality dimensions. In this case, the salience of each alternative's price would be zero, according to ρ as defined in (15), so that the salience rankings of the quality dimensions for each alternative would be the same as the rankings with price omitted from the model.

Decoy-Range Effect (S). Take $\mathbf{x} = (5, 1)$, $\mathbf{y} = (3, 2)$, $\mathbf{z} = (4.6, 1)$, and $\mathbf{z}' = (4.6, .5)$ with $\delta = .5$. We can then compute $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}) = \frac{7}{3}$ and $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}) = \frac{11}{3}$, implying (15) predicts the decoy-range effect in this scenario. Now take $\mathbf{x} = (4, 3)$, $\mathbf{y} = (1, 9)$, $\mathbf{z} = (1.25, 3)$, and $\mathbf{z}' = (1, 25, 0)$, with $\delta = .5$. We can then compute $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}) = \frac{11}{3}$ and $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}) = \frac{19}{3}$, implying (15) predicts the opposite of the decoy-range effect in this scenario.

Relative Difference Effect (S). Maintaining $\delta = .5$ while taking $\mathbf{x} = (2.5, .75)$ and $\mathbf{y} = (2, 1)$, so that $\mathbf{x}' = (2.5, .75 + k)$, and $\mathbf{y}' = (2, 1 + k)$, we can see that the DM is indifferent in a binary choice between $\mathbf{x}$ and $\mathbf{y}$. In particular, we can use (16) to compute $\rho_1(\mathbf{x}; X) = \frac{1}{19} < \frac{1}{13} = \rho_2(\mathbf{x}; X)$ and $\rho_1(\mathbf{y}; X) = \frac{1}{17} < \frac{1}{15} = \rho_2(\mathbf{y}; X)$, implying $V(\mathbf{x}; X) = \frac{.5 \cdot 2.5 + .75}{1.5} = \frac{4}{3} = \frac{.5 \cdot 2 + 1}{1.5} = V(\mathbf{y}; X)$. Next, we can verify that in a binary choice between $\mathbf{x}'$ and $\mathbf{y}'$, $\mathbf{y}'$ is chosen if $k = .25$ since (16) implies $\rho_1(\mathbf{x}'; X') = \frac{1}{19} < \frac{1}{17} = \rho_2(\mathbf{x}'; X')$ and $\rho_1(\mathbf{y}'; X') = \frac{1}{17} > \frac{1}{19} = \rho_2(\mathbf{y}'; X')$, implying $V(\mathbf{x}'; X') = \frac{.5 \cdot 2.5 + .75}{1.5} = \frac{4}{3} < \frac{5}{3} = \frac{2 + .5 \cdot 1}{1.5} = V(\mathbf{y}'; X')$ with $k = .25$. If $k = .5$, however, $\mathbf{x}'$ is chosen since (16) implies $\rho_1(\mathbf{x}'; X') = \frac{1}{19} > \frac{1}{21} = \rho_2(\mathbf{x}'; X')$ and $\rho_1(\mathbf{y}'; X') = \frac{1}{17} > \frac{1}{23} = \rho_2(\mathbf{y}'; X')$, implying $V(\mathbf{x}'; X') = \frac{2.5 + .5 \cdot .75}{1.5} = \frac{23}{12} > \frac{5}{3} = \frac{2 + .5 \cdot 1}{1.5} = V(\mathbf{y}'; X')$ with $k = .5$. Thus, the model captures the relative difference effect with $k = .5$ as well as its opposite with $k = .25$.

Majority-Rule Preference Cycles (S). Take $\mathbf{x} = (2, 1, 0)$, $\mathbf{x}' = (1, 0, 2)$, $\mathbf{x}'' = (0, 2, 1)$ and $\delta = .5$. Then a minority-rule preference cycle will exist where $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) = V(\mathbf{x}''; \{\mathbf{x}, \mathbf{x}''\}) = \frac{\delta + 2\delta^2}{1 + \delta + \delta^2} = \frac{4}{7}$ and $V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\}) = V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}''\}) = \frac{2\delta + \delta^2}{1 + \delta + \delta^2} = \frac{5}{7} > \frac{4}{7}$. If we instead take $\mathbf{x} = (20, 4, 1)$, $\mathbf{x}' = (4, 1, 20)$, and $\mathbf{x}'' = (1, 20, 4)$ while maintaining $\delta = .5$, then $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) = V(\mathbf{x}''; \{\mathbf{x}, \mathbf{x}''\}) = \frac{1 + 20\delta + 4\delta^2}{1 + \delta + \delta^2} = \frac{48}{7}$ and $V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\}) = V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}''\}) = \frac{4 + \delta + 20\delta^2}{1 + \delta + \delta^2} = \frac{38}{7} < \frac{48}{7}$, in which case a majority-rule preference cycle will exist.

Splitting Bias (S). Maintaining $\delta = .5$, take $\mathbf{x} = (3, .5)$ and $\mathbf{y} = (2, 1)$, implying $\mathbf{x}' = (x_{1a}, 3 - x_{1a}, .5)$ and $\mathbf{y}' = (y_{1a}, 2 - y_{1a}, 1)$. Then it is readily verifiable that the DM is indifferent in a binary choice between $\mathbf{x}$ and $\mathbf{y}$ under (15) and that, in a binary choice between $\mathbf{x}'$ and $\mathbf{y}'$, $\mathbf{x}'$ is chosen if $x_{1a} = .5$, which captures the splitting bias, while $\mathbf{y}'$ is chosen if $k = .25$, which captures the opposite.

Alignability Effect (S). Maintaining $\delta = .5$, we can verify that the alignability effect is captured under (15) if $\mathbf{x} = (6, 1)$ and $\mathbf{y} = (3, 2)$ but the opposite effect is predicted if $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$.

Diversification Bias (S). Here, we show that the diversification bias is always captured for $N = 2$ and never captured for $N > 2$. Given $X = \{\mathbf{x}, \mathbf{x}'\}$, with $N = 2$ it is readily verifiable that $V(\mathbf{x}; X) = \frac{A}{2}$. Without loss of generality, suppose $x'_1 < x'_2$, implying $x'_2 = A - x'_1$. Then $\rho_1(\mathbf{x}'; X) = \frac{A - 2x'_1}{A + 6x'_1}$ and $\rho_2(\mathbf{x}'; X) = \frac{A - 2x'_1}{7A - 6x'_1}$, implying $\rho_1(\mathbf{x}'; X) > \rho_2(\mathbf{x}'; X)$ since $7A - 6x'_1 > A - 2x'_1$ given $x'_1 < x'_2 = A - x'_1$. Hence, $V(\mathbf{x}; X) = \frac{x'_1 + \delta(A - x'_1)}{1 + \delta} < \frac{A}{2} = V(\mathbf{x}; X)$, as we can verify through cross-multiplication given $2x'_1 < A$, implying the diversification bias is captured. For $N > 2$, suppose $x'_1 = x'_2 = \frac{3A}{4N}$, $x'_3 = \frac{3A}{2N}$, and $x'_n = x_n$ for all $n > 3$. We then have $\rho_1(\mathbf{x}'; X) = \rho_2(\mathbf{x}'; X) = \frac{1}{13} < \frac{1}{11} = \rho_3(\mathbf{x}'; X)$, implying $V(\mathbf{x}'; X) = \frac{A}{N} \cdot \frac{3(1 + \delta^2) + 2(N - 3)\delta^{N-1}}{2 + 4\delta^2 + 2(N - 3)\delta^{N-1}} > \frac{A}{N} = V(\mathbf{x}; X)$, with the inequality holding for all $\delta < 1$ since $3(1+\delta^2) > 2+4\delta^2$, contradicting the diversification bias.

Feature Bias (Y). To show that the feature bias is robustly captured under (15), note that $\rho_{n'}(\mathbf{x}; X) = \frac{q}{4x_{n'}+q}$ given $X = \{\mathbf{x}, \mathbf{x}'\}$, $\rho_N(\mathbf{x}; X) = \frac{1}{3}$, $\rho_{n'}(\mathbf{x}'; X) = \frac{q}{4x_{n'}+3q}$, and $\rho_N(\mathbf{x}'; X) = 1$, where $\mathbf{x}$ and $\mathbf{x}'$ are defined as in Proposition 8. We can also see $\rho_n(\mathbf{x}; X) = \rho_n(\mathbf{x}'; X) = 0$ for all $n \notin \{n', N\}$. From this, we get $V(\mathbf{x}'; X) = \frac{\delta(x_{n'}+q)+\delta^{N-1}\sum_{n\notin\{n',N\}} x_n}{1+\delta+(N-2)\delta^{N-1}}$. If $\frac{q}{4x_{n'}+q} = \frac{1}{3}$, then $V(\mathbf{x}; X) = \frac{\delta(x_{n'}+q)+\delta^{N-1}\sum_{n\notin\{n',N\}} x_n}{2\delta+(N-2)\delta^{N-1}}$, ensuring $V(\mathbf{x}; X) > V(\mathbf{x}'; X)$ given $\delta < 1$. If $\frac{q}{4x_{n'}+q} \neq \frac{1}{3}$, then $V(\mathbf{x}; X) \geq \frac{\min\{x_{n'}+\delta q, \delta x_{n'}+q\}+\delta^{N-1}\sum_{n\notin\{n',N\}} x_n}{1+\delta+(N-2)\delta^{N-1}}$, which also ensures $V(\mathbf{x}; X) > V(\mathbf{x}'; X)$ since $\min\{x_{n'}+\delta q, \delta x_{n'}+q\} > \delta(x_{n'}+q)$. Thus, $\mathbf{x}$ must be chosen over $\mathbf{x}'$, capturing the feature bias.

C.4. Koszegi and Szeidl (2013)

For Koszegi and Szeidl's (2013) model, we use:

$$V(\mathbf{x}; X) = \sum_{n=1}^{N} h\Big(\max_{x'\in X}\{x'_n\} - \min_{x'\in X}\{x'_n\}\Big) \cdot x_n, \tag{17}$$

where h is strictly increasing.

Compromise Effect (N). We can readily verify that the compromise effect holds under (17) if $h(z_1-y_1)\cdot(x_1-y_1) > h(y_2-z_2)\cdot(y_2-x_2)$. Given $V(\mathbf{z}; \{\mathbf{y},\mathbf{z}\}) < V(\mathbf{y}; \{\mathbf{y},\mathbf{z}\})$, $y_2-z_2 > z_1-y_1$ must hold, while binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$ implies $x_1-y_1 = y_2-x_2$. Thus, the above condition is violated since $h(y_2-z_2) > h(z_1-y_1)$ with h increasing.

Dominance Effect — Weak (N) *and Strict* (N). The dominance effect holds under (17) if $h(x_1-y_1)\cdot(x_1-y_1) > h(y_2-z_2)\cdot(y_2-x_2)$. Since $z_2 \leq x_2$, $x_1-y_1 = y_2-x_2$, and h is increasing, this condition cannot hold.

Decoy-Range Effect (N). Given $y_2 > x_2 \geq z_2$, $\frac{\partial}{\partial z_2}\big[V(\mathbf{x}, \{\mathbf{x},\mathbf{y},\mathbf{z}\}) - V(\mathbf{y}, \{\mathbf{x},\mathbf{y},\mathbf{z}\})\big] = (y_2-x_2)h'(y_2-z_2) > 0$ under (17). Thus, if the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ given $X = \{\mathbf{x},\mathbf{y},\mathbf{z}\}$, $\mathbf{y}$ must be chosen from $X = \{\mathbf{x},\mathbf{y},\mathbf{z}'\}$ with $z'_1 = z_1$ and $z'_2 < z_2$, which is the opposite of the decoy-range effect.

Relative Difference Effect (N). The relative difference effect holds if $h((y_2+k)-(x_2+k))\cdot((y_2+k)-(x_2+k))$ is decreasing in $k \geq 0$. Since $h((y_2+k)-(x_2+k))\cdot((y_2+k)-(x_2+k)) = h(y_2-x_2)\cdot(y_2-x_2)$, the expression is clearly independent of k, implying the relative difference effect is not captured.

Majority-Rule Preference Cycles (N). To allow majority-rule preference cycles while precluding minority-rule preference cycles, it suffices to show that $V(\mathbf{x}, \{\mathbf{x},\mathbf{z}\}) > V(\mathbf{z}, \{\mathbf{x},\mathbf{z}\})$ for any $\mathbf{x}$ and $\mathbf{z}$ satisfying $x_1 > z_1$, $x_2 > z_2$, and $x_1+x_2+x_3 = z_1+z_2+z_3$ with $N = 3$. Noting these conditions imply $z_3 > x_3$ and $z_3 - x_3 = x_1+x_2-z_1-z_2$ while letting $\delta_n \equiv x_n - z_n > 0$ for $n = 1, 2$, under (17) we have $V(\mathbf{x}, \{\mathbf{x},\mathbf{z}\}) - V(\mathbf{z}, \{\mathbf{x},\mathbf{z}\}) = (h(\delta_1) - h(\delta_1+\delta_2))\cdot\delta_1 + (h(\delta_2) - h(\delta_1+\delta_2))\cdot\delta_2$. Thus, $V(\mathbf{x}, \{\mathbf{x},\mathbf{z}\}) - V(\mathbf{z}, \{\mathbf{x},\mathbf{z}\}) > 0$ cannot hold since $h(\delta_1+\delta_2) > h(\delta_n)$ for

$n = 1, 2$ with $h(\cdot)$ increasing, implying majority-rule preference cycles are not captured under (17).

Splitting Bias (N). Letting $\delta^a = x_{1a} - y_{1a}$ and $\delta^b = x_{1b} - y_{1b}$, we can see that the splitting bias is captured under (17) if $(h(\delta^a) - h(\delta^a + \delta^b)) \cdot \delta^a + (h(\delta^b) - h(\delta^a + \delta^b)) \cdot \delta^b > 0$, but this condition cannot hold since $h(\delta^a + \delta^b) > \max\{h(\delta^a), h(\delta^b)\}$ with h increasing.

Alignability Effect (N). Since binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$ implies $x_1 + x_2 = y_1 + y_2$ under (17), $x_1 - y_1 = y_2 - x_2$. We can then see that the alignability effect is captured if $(h(y_2 - x_2) - h(y_2)) \cdot (y_2 - x_2) + (h(x_2) - h(y_2)) \cdot x_2 > 0$, which cannot hold since $h(y_2) > \max\{h(x_2), h(y_2 - x_2)\}$ with $h(\cdot)$ increasing.

Diversification Bias (N). To show that (17) does not predict the diversification bias, suppose $x'_1 = 0$, $x'_2 = \frac{2A}{N}$, and $x'_n = \frac{A}{N}$ for all $n > 2$. Then, with $x_n = \frac{A}{N}$ for all $n \leq N$ and $X = \{\mathbf{x}, \mathbf{x}'\}$, we get $V(\mathbf{x}; X) - V(\mathbf{x}'; X) = h\big(\frac{A}{N} - 0\big) \cdot \big(\frac{A}{N} - 0\big) - h\big(\frac{2A}{N} - \frac{A}{N}\big) \cdot \big(\frac{2A}{N} - \frac{A}{N}\big) = h\big(\frac{A}{N}\big) \cdot \frac{A}{N} - h\big(\frac{A}{N}\big) \cdot \frac{A}{N} = 0$, implying indifference between $\mathbf{x}$ and $\mathbf{x}'$.

Feature Bias (N). With the new feature, $\mathbf{x}$ has an effective advantage of $h(q) \cdot q$ on dimension N. With the improved existing feature, $\mathbf{x}'$ has an effective advantage of $h(x_{n'} + q - x_{n'}) \cdot (x_{n'} + q - x_{n'}) = h(q) \cdot q$ on dimension n'. Since these advantages are equal under (17), $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ must hold, implying the feature bias is not captured.

C.5. Bushong et al. (2019)

For Bushong et al.'s (2019) model, we use $V(\mathbf{x}; X)$ as given in (17), except now h is strictly decreasing and $h(z) \cdot z$ is strictly increasing in z.

Compromise Effect (**Y**). As before, the compromise effect holds if $h(z_1 - y_1) \cdot (x_1 - y_1) > h(y_2 - z_2) \cdot (y_2 - x_2)$. Given $V(\mathbf{z}; \{\mathbf{y}, \mathbf{z}\}) < V(\mathbf{y}; \{\mathbf{y}, \mathbf{z}\})$, $y_2 - z_2 > z_1 - y_1$ must hold, while binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$ implies $x_1 - y_1 = y_2 - x_2$. Thus, the above condition holds (ensuring a compromise effect) since $h(y_2 - z_2) < h(z_1 - y_1)$ given h is decreasing.

Dominance Effect, Weak (N) *and Strict* (**Y**). The dominance effect holds if $h(x_1 - y_1) \cdot (x_1 - y_1) > h(y_2 - z_2) \cdot (y_2 - x_2)$. Since $x_1 - y_1 = y_2 - x_2$ and h is decreasing, the condition holds for $z_2 < x_2$ but not for $z_2 = x_2$. Thus, the dominance effect is captured for a strictly dominated decoy $\mathbf{z}$ but not if $\mathbf{z}$ is only weakly dominated.

Decoy-Range Effect (**Y**). As seen, $\frac{\partial}{\partial z_2}\big[V(\mathbf{x}, \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}, \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})\big] = (y_2 - x_2)h'(y_2 - z_2)$ given $y_2 > x_2 \geq z_2$ under (17), except now $(y_2 - x_2)h'(y_2 - z_2) < 0$ since $h'(y_2 - z_2) < 0$. Thus, with indifference between $\mathbf{x}$ and $\mathbf{y}$ given $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, $\mathbf{x}$ will be chosen from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}$ with $z'_1 = z_1$ and $z'_2 < z_2$, capturing the decoy-range effect.

Relative Difference Effect (N). Same as Koszegi and Szeidl (2013) — see above.

Majority-Rule Preference Cycles (**Y**). To allow majority-rule preference cycles while precluding minority-rule preference cycles, it suffices to show $V(\mathbf{x}, \{\mathbf{x}, \mathbf{z}\}) > V(\mathbf{z}, \{\mathbf{x}, \mathbf{z}\})$ for any $\mathbf{x}$ and $\mathbf{z}$ satisfying $x_1 > z_1$, $x_2 > z_2$, and $x_1 + x_2 + x_3 = z_1 + z_2 + z_3$ with $N = 3$. Noting these conditions imply $z_3 > x_3$ and $z_3 - x_3 = x_1 + x_2 - z_1 - z_2$ while letting $\delta_n \equiv x_n - z_n > 0$ for $n = 1, 2$, under (17) we have $V(\mathbf{x}, \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{z}, \{\mathbf{x}, \mathbf{z}\}) = (h(\delta_1) - h(\delta_1 + \delta_2)) \cdot \delta_1 + (h(\delta_2) - h(\delta_1 + \delta_2)) \cdot \delta_2$. Thus, $V(\mathbf{x}, \{\mathbf{x}, \mathbf{z}\}) - V(\mathbf{z}, \{\mathbf{x}, \mathbf{z}\}) > 0$ must hold since $h(\delta_1 + \delta_2) < h(\delta_n)$ for $n = 1, 2$ given h is decreasing, implying majority-rule preference cycles are robustly captured.

Splitting Bias (**Y**). Given $\delta^a = x_{1a} - y_{1a}$ and $\delta^b = x_{1b} - y_{1b}$, the splitting bias is captured if $(h(\delta^a) - h(\delta^a + \delta^b)) \cdot \delta^a + (h(\delta^b) - h(\delta^a + \delta^b)) \cdot \delta^b > 0$, which must hold since $h(\delta^a + \delta^b) < \min\{h(\delta^a), h(\delta^b)\}$ with h decreasing.

Alignability Effect (**Y**). Since binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$ implies $x_1 + x_2 = y_1 + y_2$ under (17), $x_1 - y_1 = y_2 - x_2$. Using this relation, we can then see that the alignability effect is captured if $(h(y_2 - x_2) - h(y_2)) \cdot (y_2 - x_2) + (h(x_2) - h(y_2)) \cdot x_2 > 0$, which must hold since $h(y_2) < \min\{h(x_2), h(y_2 - x_2)\}$ with h decreasing.

Diversification Bias (N). Same as Koszegi and Szeidl (2013) — see above.

Feature Bias (N). Same as Koszegi and Szeidl (2013) — see above.

C.6. Soltani et al. (2012)

For Soltani et al.'s (2012) model, which assumes that alternatives are defined on two attribute dimensions, we use the following value function:

$$V(\mathbf{x}; X) = w_1 \cdot \frac{x_1 - \min_{\mathbf{x}' \in X}\{x_1'\}}{\max_{\mathbf{x}' \in X}\{x_1'\} - \min_{\mathbf{x}' \in X}\{x_1'\}} + w_2 \cdot \frac{x_2 - \min_{\mathbf{x}' \in X}\{x_2'\}}{\max_{\mathbf{x}' \in X}\{x_2'\} - \min_{\mathbf{x}' \in X}\{x_2'\}}. \tag{18}$$

For simplicity, this formulation assumes that the "representation factors" are zero (i.e. $f_s = f_t = 0$, in their notation). It also assumes that the attribute weights $w_1, w_2 > 0$ are the same for all alternatives, but it still allows $w_1 \neq w_2$. The reason we do not impose $w_1 = w_2$ is that (18) would then imply that the DM *must* be indifferent between $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ in binary choice given $x_1 > y_1$ and $y_2 > x_2$.

Compromise Effect (S). Applied to (18), and given $x_1 > y_1$ and $y_2 > x_2$, the DM is indifferent between $\mathbf{x}$ and $\mathbf{y}$ in binary choice if and only if $w_1 = w_2$. Taking $w_1 = w_2 = 1$ with $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, and given $\mathbf{z}$ makes $\mathbf{x}$ a compromise, we see $V(\mathbf{x}; X) = \frac{x_1 - y_1}{z_1 - y_1} + \frac{x_2 - z_2}{y_2 - z_2}$, $V(\mathbf{y}; X) = \frac{y_1 - y_1}{z_1 - y_1} + \frac{y_2 - z_2}{y_2 - z_2} = 1$, and $V(\mathbf{z}; X) = \frac{z_1 - y_1}{z_1 - y_1} + \frac{z_2 - z_2}{y_2 - z_2} = 1$. Since $V(\mathbf{y}; X) = V(\mathbf{z}; X)$, $\mathbf{z}$ is not chosen, while $\mathbf{x}$ is chosen if and only if $\frac{x_1 - y_1}{z_1 - y_1} + \frac{x_2 - z_2}{y_2 - z_2} > 1$. Letting $\mathbf{x} = (1, 1)$, $\mathbf{y} = (0, 2)$, and $\mathbf{z} = (z_1, 0)$, it is then readily verifiable that this condition holds for $z_1 < 2$, in which case (18) predicts a compromise effect, but is violated for $z_1 > 2$, in which case the opposite effect is predicted.

Dominance Effect, Weak (N) *and Strict* (**Y**). Again taking $w_1 = w_2 = 1$ (without loss of generality given binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$), if $\mathbf{x}$ asymmetrically dominates $\mathbf{z}$, $\mathbf{x}$ will be chosen from $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ if and only if $V(\mathbf{x}; X) > V(\mathbf{y}; X)$, where $V(\mathbf{x}; X) = \frac{x_1 - y_1}{x_1 - y_1} + \frac{x_2 - z_2}{y_2 - z_2} = 1 + \frac{x_2 - z_2}{y_2 - z_2}$ and $V(\mathbf{y}; X) = \frac{y_1 - y_1}{x_1 - y_1} + \frac{y_2 - z_2}{y_2 - z_2} = 1$. Thus, $V(\mathbf{x}; X) > V(\mathbf{y}; X)$ if $x_2 > z_2$, but not if $x_2 = z_2$.

Decoy-Range Effect (N). In the decoy-range effect, the DM chooses $\mathbf{y}$ over $\mathbf{x}$ in binary choice, which (it is readily verifiable) holds under (18) if and only if $w_2 > w_1$ (since $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) =$

$w_1 \cdot \frac{x_1-y_1}{x_1-y_1} + w_2 \cdot \frac{x_2-x_2}{y_2-x_2} = w_1$ and $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = w_1 \cdot \frac{y_1-y_1}{x_1-y_1} + w_2 \cdot \frac{y_2-x_2}{y_2-x_2} = w_2$). The DM is also indifferent between $\mathbf{x}$ and $\mathbf{y}$ given $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, with $x_1 > z_1 > y_1$ and $y_2 > x_2 \geq z_2$. Noting that, in this case, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = w_1\big(\frac{x_1-y_1}{x_1-y_1} - \frac{y_1-y_1}{x_1-y_1}\big) + w_2\big(\frac{x_2-z_2}{y_2-z_2} - \frac{y_2-z_2}{y_2-z_2}\big) = w_1 + w_2\big(\frac{x_2-z_2}{y_2-z_2} - 1\big)$, the decoy range effect is captured if and only if $\frac{\partial}{\partial z_2}\big(w_1 + w_2\big(\frac{x_2-z_2}{y_2-z_2} - 1\big)\big) = w_2 \cdot \frac{x_2-y_2}{(y_2-z_2)^2} < 0$, which must hold since $y_2 > x_2$.

Relative Difference Effect (N). Using (18) and with $\mathbf{x}'$ and $\mathbf{y}'$ as defined in Proposition 1, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ implies $w_1 = w_2$, which implies $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) = V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\})$. Thus, the relative difference effect is not captured.

Majority-Rule Preference Cycles (-), *Splitting Bias* (-), *Alignability Effect* (-). As formalized here, majority-rule preference cycles, the splitting bias, and the alignability effect can only be expressed with a model that allows more than two attributes. Since Soltani et al.'s (2012) model assumes that alternatives are defined on two dimensions, it does not make predictions regarding these behavioral effects.[35]

Diversification Bias (N). Since $N = 2$ under (18), the balanced allocation in Proposition 7 is $\mathbf{x} = \big(\frac{A}{2}, \frac{A}{2}\big)$. To show that the diversification bias is never captured, it suffices to show that there is always a $\mathbf{x}' \neq \mathbf{x}$ with $x_1' + x_2' = A$ such that $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) \leq V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ under (18). If $w_1 > w_2$, we can take $\mathbf{x}' = (A, 0)$, in which case $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = w_2 < V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = w_1$. If $w_1 < w_2$, we can take $\mathbf{x}' = (0, A)$, in which case $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = w_1 < V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = w_2$. Lastly, if $w_1 = w_2$, we can take any $\mathbf{x}' = (x_1', x_2')$ with $x_1' + x_2' = A$, in which case $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = w_1 = w_2$.

Feature Bias (S). Since $N = 2$ under (18), the alternatives in Proposition 8 are $\mathbf{x} = (x_1, q)$ and $\mathbf{x}' = (x_1 + q, 0)$ with $q > 0$. We can then use (18) to compute $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) = w_2$ and $V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = w_1$. Thus, the feature bias is captured if $w_2 > w_1$, but not if $w_1 > w_2$.

C.7. Joint normalization

We now consider the "joint normalization" model given in (5) with $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x_n'$.

Compromise Effect (N). Given $y_2 = \frac{x_1 x_2}{y_1}$ (Proposition 16) and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x_n'$ implies $V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{(x_1-y_1)(y_1 z_2 - z_1 x_2)}{(x_1+y_1+z_1)(x_1 x_2 + y_1 x_2 + y_1 z_2)}$. With $z_1 > x_1 > y_1$ and $z_2 < x_2$, $y_1 z_2 - z_1 x_2 < 0$ must hold. This, along with $x_1 > y_1$, ensures that the compromise effect cannot be captured with $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x_n'$.

[35] That said, it may be reasonable to consider the following extension of the model in (18) to three attributes: $V(\mathbf{x}; X) = \sum_{n=1}^{3} w_n \cdot \frac{x_n - \min_{\mathbf{x}' \in X}\{x_n'\}}{\max_{\mathbf{x}' \in X}\{x_n'\} - \min_{\mathbf{x}' \in X}\{x_n'\}}$. It is then straightforward to show that this model robustly captures majority-rule preference cycles. However, this model still cannot make concrete predictions regarding the splitting bias because the valuations of $\mathbf{x}'$ and $\mathbf{y}'$ in Proposition 5 would depend on the weighting parameters w_{1a} and w_{1b}, while binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$ does not restrict w_{1a} and w_{1b}. Similarly, the implied valuations of $\mathbf{x}'$ and $\mathbf{y}'$ in Proposition 6 would depend on w_3, but this parameter is not restricted by binary-choice indifference between $\mathbf{x}$ and $\mathbf{y}$, implying this extension of Soltani et al.'s model also does not make concrete predictions regarding the alignability effect.

Dominance Effect, Weak (N) *and Strict* (N). Given $y_2 = \frac{x_1 x_2}{y_1}$ (Proposition 16) and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x'_n$ implies $V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{(x_1 - y_1)(y_1 z_2 - z_1 x_2)}{(x_1 + y_1 + z_1)(x_1 x_2 + y_1 x_2 + y_1 z_2)}$. With $z_1 > y_1$ and $z_2 \leq x_2$, $y_1 z_2 - z_1 x_2 < 0$ holds. This, with $x_1 > y_1$, ensures the dominance effect with $z_2 = x_2$ or $z_2 < x_2$ cannot be captured with $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x'_n$.

Decoy-Range Effect (N). Given $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x'_n$ and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, we can compute $\frac{\partial [V(\mathbf{x}; X) - V(\mathbf{y}; X)]}{\partial z_2} = \frac{y_2 - x_2}{(x_2 + y_2 + z_2)^2} > 0$, which ensures that the decoy-range effect cannot be captured with $r(x_n) = \sum_{\mathbf{x}' \in X \setminus \mathbf{x}} x'_n$.

Relative Difference Effect (**Y**), *Majority-Rule Preference Cycles* (**Y**), *Splitting Bias* (**Y**), *Alignability Effect* (**Y**), *Diversification Bias* (**Y**), and *Feature Bias* (**Y**). As established in the proof of Proposition 16, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ holds (for any N) under joint normalization if and only if $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ holds under (1). Therefore, all of these effects (which are formalized in terms of binary choices) must be captured by maximum normalization since they are captured by pairwise normalization.

C.8. Average normalization

We now consider "average normalization" as given in (5) with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$.

Compromise Effect (S). To show that average normalization sometimes predicts the compromise effect and sometimes predicts the opposite, suppose $\mathbf{x} = (2, 1)$, $\mathbf{y} = (1, 2)$, and $\mathbf{z} = (2.1, z_2)$. Then, if $z_2 = 0$, $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx .0035 > 0$, consistent with the compromise effect. If $z_2 = .6$, $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx -.00028 < 0$, which is the opposite of the compromise effect (we can also verify $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx .11 > 0$ in this case, ensuring $\mathbf{z}$ is not chosen in trinary choice).

Dominance Effect, Weak (S) *and Strict* (S). To show that average normalization sometimes predicts the dominance effect and sometimes predicts the opposite, suppose $\mathbf{x} = (2, 1)$, $\mathbf{y} = (1, 2)$. Then $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (2, .5)) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (2, .5)) \approx .00041 > 0$ and $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (1.9, .5)) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (1.9, .5)) \approx .000671 > 0$, consistent with the weak and strict dominance effects. However, $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$ also implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (2, .7)) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (2, .7)) \approx -.000341 < 0$ and $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (1.9, .7)) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} | \mathbf{z} = (1.9, .7)) \approx -.000084 < 0$, in which case the average normalization model predicts the opposite of the weak and strict dominance effects.

Decoy-Range Effect (**Y**). Since $V(\mathbf{x}; X)$ is invariant to re-scaling all x'_n with $\mathbf{x}' \in X$ on a given dimension by the same positive constant, it suffices to show that the result holds with $\mathbf{x} = (1 + 2\mu\gamma, 1 - 2\gamma)$ and $\mathbf{y} = (1 - 2\mu\gamma, 1 + 2\gamma)$ (which effectively sets the average of x_n and y_n to 1 on each dimension), where $\gamma \in \left(0, \frac{1}{2}\right]$ and $\mu \in (0, 1)$ since $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) < V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$. We can then let $\mathbf{z} = ((1 - \beta)x_1 + \beta y_1, \lambda x_2)$ with $\beta \in [0, 1)$ and $\lambda \in [0, 1]$ with either $\beta > 0$ or $\lambda < 1$ (or both), so that $\mathbf{x}$ asymmetrically dominates $\mathbf{z}$. Using these definitions, we can then compute $\frac{\partial [V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}]}{\partial z_2} = \frac{12\gamma(1 - 2\gamma)G(\lambda, \gamma)}{((5 + (1 - 2\gamma)\lambda)^2 - 36\gamma^2)^2}$ with $G(\lambda, \gamma) \equiv 4 - 9(1 - 4\gamma^2) + 4\lambda(1 - 2\gamma) + \lambda^2(1 - 2\gamma)^2$. It then suffices to show $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \geq V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ implies $\frac{\partial [V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}]}{\partial z_2} < 0$, or equivalently, $G(\lambda, \gamma) < 0$. Now $\frac{\partial G(\lambda, \gamma)}{\partial \lambda} \geq 0$ since

$\gamma \leq \frac{1}{2}$. Therefore, $G(\lambda, \gamma)$ is maximized at $\lambda = 1$, implying that it suffices to show that either $G(\lambda, \gamma) < 0$ or $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) < V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ must hold with $\lambda = 1$. Now $G(1, \gamma) = 4\gamma(10\gamma - 3)$, implying $G(1, \gamma) < 0$ for all $\gamma < \frac{3}{10}$. Next, $\frac{V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\}|\lambda=1) - V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\}|\lambda=1)}{\partial\mu} = \frac{27\gamma H(\mu,\beta,\gamma)}{(3+2\mu(2-\beta)\gamma)^2(3-2\mu(1+\beta)\gamma)^2}$, where $H(\mu, \beta, \gamma) \equiv 3 + 4\mu(1 - 2\beta)\gamma + 4\mu^2(1 - \beta + \beta^2)\gamma^2$. Thus, $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\}|\lambda=1) - V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\}|\lambda=1)]}{\partial\mu} > 0$ if and only if $H(\mu, \beta, \gamma) > 0$. Next, observe $\frac{\partial H(\mu,\beta,\gamma)}{\partial\beta} = -4\mu\gamma(2 + \mu(1 - 2\beta)\gamma) < 0$ since $\mu, \beta, \gamma < 1$. Thus, $H(\mu, \beta, \gamma)$ is minimized at $H(\mu, 1, \gamma) = 3 - 4\mu\gamma + 4\mu^2\gamma^2$. Noting $\frac{\partial H(\mu,1,\gamma)}{\partial\mu} = -4\gamma(1 - 2\mu\gamma) < 0$ since $\mu < 1$ and $\gamma \leq \frac{1}{2}$, $H(\mu, 1, \gamma)$ must be minimized at $H(1, 1, \gamma) = 2 + (1 - 2\gamma)^2 > 0$. Thus, $H(\mu, \beta, \gamma) > 0$ for all μ, β, γ, which implies $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\}|\lambda=1) - V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\}|\lambda=1)]}{\partial\mu} > 0$. Thus, it suffices to show that $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}|\mu = \lambda = 1) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}|\mu = \lambda = 1) = -\frac{12(1-\beta)\gamma^3(8\gamma-3\beta+2\beta\gamma)}{(3+2\gamma)(3-4\gamma)(3+2(2-\beta)\gamma)(3-2(1+\beta)\gamma)} < 0$ for all $\gamma > \frac{3}{10}$. Since $\beta < 1$ and $\gamma \leq \frac{1}{2}$, we can see this condition reduces to $-\frac{8\gamma-3\beta+2\beta\gamma}{3-2(1+\beta)\gamma} < 0$, which then holds if $\min\{8\gamma - 3\beta + 2\beta\gamma, 3 - 2(1 + \beta)\gamma\} > 0$. Since $8\gamma - 3\beta + 2\beta\gamma > 0$ for all $\beta < \frac{8\gamma}{3-2\gamma}$, $3 - 2(1 + \beta)\gamma > 0$ for all $\beta < \frac{3-2\gamma}{2\gamma}$, and $\frac{8\gamma}{3-2\gamma} \leq \frac{3-2\gamma}{2\gamma}$ (since $16\gamma^2 \leq 4 \leq (3 - 2\gamma)^2$ with $\gamma \leq \frac{1}{2}$), the condition further reduces to $\beta < \frac{8\gamma}{3-2\gamma}$. Since $\frac{8\gamma}{3-2\gamma}$ is increasing in γ (by inspection) and $\frac{8\gamma}{3-2\gamma} = 1 > \beta$ given $\gamma = \frac{3}{10}$, $\beta < \frac{8\gamma}{3-2\gamma}$ must hold for all $\gamma \geq \frac{3}{10}$, as desired.

Relative Difference Effect (**Y**). The relative difference effect must be captured under average normalization because (a) it is captured under (1) (Proposition 1), and (b) binary choices are equivalent under average normalization and (1) (Proposition 16).

Majority-Rule Preference Cycles (**Y**). First it is readily verifiable that a majority-rule preference cycle is possible under average normalization by taking, for example, $\mathbf{x} = (2, 1, 0)$, $\mathbf{x}' = (1, 0, 2)$, and $\mathbf{x}'' = (0, 2, 1)$. To show that the opposite, minority-rule preference cycle cannot hold, we proceed by contradiction.

Next, suppose $y_n = 0$ for some n and $\mathbf{y} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\}$, and $\mathbf{z} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\} \setminus \mathbf{y}$ is the alternative that is better than $\mathbf{y}$ on two of three dimensions. Without loss of generality, suppose $y_1 = 0 < z_1$, $0 \leq y_2 < z_2$, and $y_3 > z_3$. Then $V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}\}) - V(\mathbf{y}; \{\mathbf{z}, \mathbf{y}\}) = \Delta(z_1, 0) + \Delta(z_2, y_2) - \Delta(y_3, z_3)$. However, this must be positive since $\Delta(z_1, 0) = \frac{2}{3} \geq \Delta(y_3, z_3)$ with $\Delta(z_2, y_2) > 0$ given $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$. This implies that a minority rule preference cycle cannot exist if $y_n = 0$ for some n and $\mathbf{y} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\}$.

Now, assuming from here on that all attribute values are nonzero, if a minority-rule preference cycle exists, there must be at least one alternative that is chosen even though the product of its attribute values is weakly less than the product of the other alternative's attribute values. That is, there must be some $\mathbf{y} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\}$ and $\mathbf{z} \in \{\mathbf{x}, \mathbf{x}', \mathbf{x}''\} \setminus \mathbf{y}$ for which: (a) $\mathbf{z}$ is better than $\mathbf{y}$ on two of three dimensions; (b) $z_1z_2z_3 \geq y_1y_2y_3$, and (c) $V(\mathbf{y}; \{\mathbf{z}, \mathbf{y}\}) > V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}\})$. Without loss of generality, suppose $z_1 > y_1$, $z_2 > y_2$, and $y_3 > z_3$. Then define $\mathbf{y}'$ such that $y'_n = y_n$ for $n = 1, 2$ and $y'_3 = \frac{z_1z_2z_3}{y_1y_2} \geq y_3$. Then, a necessary condition for $V(\mathbf{y}; \{\mathbf{z}, \mathbf{y}\}) > V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}\})$ to hold is that $V(\mathbf{y}'; \{\mathbf{z}, \mathbf{y}'\}) > V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}'\})$. Next, let $a = \frac{y'_1}{z_1} < 1$ and $b = \frac{y'_2}{z_2} < 1$, implying $\frac{z_3}{y_3} = ab$. We can then use (5) with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$ to compute $V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{z}, \mathbf{y}'\}) = \frac{2(1-a)(1-b)(1-ab)(9(1+b)+9a^2b(1+b)+a(9-22b+9b^2))}{(3+a)(1+3a)(3+b)(1+3b)(3+ab)(1+3ab)}$. Thus, $V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{z}, \mathbf{y}'\}) > 0$ if and only if $9(1 + b) + 9a^2b(1 + b) + a(9 - 22b + 9b^2) > 0$. Without loss of generality, suppose $b \leq a$, and let $b = \lambda a$ with $\lambda \in [0, 1]$. With this substitution, we then see that $V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}'\}) -$

$V(\mathbf{y}'; \{\mathbf{z}, \mathbf{y}'\}) > 0$ is equivalent to $Z(a, \lambda) \equiv 9 - 22a^2\lambda + 9a^4\lambda^2 + 9a(1+\lambda) + 9a^3\lambda(1+\lambda) > 0$. Next, we can see that $\lambda = \lambda^*(a) \equiv \frac{-9+22a-9a^2}{18a^2(1+a)}$ uniquely solves the first-order condition, $\frac{\partial Z(a,\lambda)}{\partial \lambda} = 0$. We can then observe $Z(a, \lambda^*(a)) = \frac{(9-a)(9a-1)(9+2a+9a^2)}{36a(1+a)} < 0$ if and only if $a < \frac{1}{9}$. We can also observe $\lambda^*(a) > 0$ can only hold if $-9 + 22a - 9a^2 > 0$, which cannot be the case with $a < \frac{1}{9}$ since $-9 + 22a - 9a^2 < -9 + 22a < -9 + \frac{22}{9} < 0$ given $a < \frac{1}{9}$. Thus, since the solution to λ^* does not apply, for $Z(a, \lambda) < 0$ to hold, it must hold at one of the bounds, $\lambda \in \{0, 1\}$. However, $Z(a, 0) = 9(1 + a) > 0$ and $Z(a, 1) = 5 + 4(1 - a^2) + 18a(1 - a) + 18a^3 + 9a^4 > 0$. Therefore, $Z(a, \lambda) > 0$ must hold for all applicable a, λ, implying $V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{z}, \mathbf{y}'\}) > 0$ must hold, thus guaranteeing $V(\mathbf{z}; \{\mathbf{z}, \mathbf{y}\}) - V(\mathbf{y}; \{\mathbf{z}, \mathbf{y}\}) > 0$ and contradiction the presumed minority-rule preference cycle. This establishes that only majority-rule preference cycles can be captured.

Splitting Bias (**Y**). Let $\lambda_a \equiv \frac{y_{1a}}{x_{1a}}$, $\lambda_b \equiv \frac{y_{1b}}{x_{1b}}$, and $\lambda \equiv \frac{y_1}{x_1}$. We can then use (5) with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x_n'$ to compute $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{x_{1a}}{x_{1a}+.5(x_{1a}+y_{1a})} - \frac{y_{1a}}{y_{1a}+.5(x_{1a}+y_{1a})} + \frac{x_{1b}}{x_{1b}+.5(x_{1b}+y_{1b})} - \frac{y_{1b}}{y_{1b}+.5(x_{1b}+y_{1b})} - \frac{x_1}{x_1+.5(x_1+x_2)} + \frac{y_1}{y_1+.5(x_1+x_2)}$, which reduces to $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{2+3\lambda_a-\lambda_a^2}{3+10\lambda_a+3\lambda_a^2} + \frac{2+3\lambda_b-\lambda_b^2}{3+10\lambda_b+3\lambda_b^2} - \frac{2+3\lambda-\lambda^2}{3+10\lambda+3\lambda^2} = f(\lambda_a) + f(\lambda_b) - f(\lambda)$, where $f(s) = \frac{2+3s-s^2}{3+10s+3s^2} \geq 0$ for $s \in [0, 1]$. Now $f'(s) = -\frac{11+18s+19s^2}{(3+10s+3s^2)^2} < 0$. Since the definitions of λ_a, λ_b, and λ ensure $\min\{\lambda_a, \lambda_b\} \geq \lambda$ (binding if and only if $\lambda_a = \lambda_b = \lambda$), it follows that $f(\lambda_a) + f(\lambda_b) > \max\{f(\lambda_a), f(\lambda_b)\} \geq f(\lambda)$, ensuring $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) > 0$, implying the splitting bias is captured under average normalization.

Alignability Effect (S). Using $y_2 = \frac{x_1 x_2}{y_1}$ given $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\})$ (Proposition 16), the alignability effect holds with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x_n'$ since $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{x_1}{x_1+.5(x_1+y_1)} + \frac{x_2}{x_2+.5x_2} - \frac{y_1}{y_1+.5(x_1+y_1)} - \frac{y_2}{y_2+.5y_2} = \frac{2(x_1-y_1)(x_1+y_1)}{(3x_1+y_1)(x_1+3y_1)} > 0$.

Diversification Bias (**Y**). Given $x_n = \frac{A}{N}$ for all n, suppose $x_n' > x_n$ and let $\Delta_n^+ = \frac{x_n'-x_n}{x_n}$. Using (5) with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x_n'$, we can compute $\mathbf{x}'$'s advantage over $\mathbf{x}$ on n as $A_n^+ \equiv \frac{x_n'}{x_n'+.5(x_n+x_n')} - \frac{x_n}{x_n+.5(x_n+x_n')} = \frac{2\Delta_n^+(2+\Delta_n^+)}{(4+\Delta_n^+)(4+3\Delta_n^+)}$, implying $\frac{\partial[A_n^+/\Delta_n^+]}{\partial \Delta_n^+} = -\frac{1}{2(4+\Delta_n^+)^2} - \frac{3}{2(4+3\Delta_n^+)^2} < 0$. Noting $\lim_{x_n' \to x_n} \left\{\frac{A_n^+}{\Delta_n^+}\right\} = \frac{1}{4}$, it then follows that $A_n^+ < \frac{\Delta_n^+}{4}$ given $x_n' > x_n$.

Now suppose $x_n' < x_n$ and let $\Delta_n^- = \frac{x_n-x_n'}{x_n}$. Using (5) with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x_n'$, we can then compute $\mathbf{x}$'s advantage over $\mathbf{x}'$ on n as $A_n^- \equiv \frac{x_n}{x_n+.5(x_n+x_n')} - \frac{x_n'}{x_n'+.5(x_n+x_n')} = \frac{2\Delta_n^-(2-\Delta_n^-)}{(4-\Delta_n^-)(4-3\Delta_n^-)}$. Next, we can compute $\frac{\partial[A_n^-/\Delta_n^-]}{\partial \Delta_n^-} = \frac{1}{2(4-\Delta_n^-)^2} + \frac{3}{2(4-3\Delta_n^-)^2} > 0$. Noting $\lim_{x_n' \to x_n} \left\{\frac{A_n^-}{\Delta_n^-}\right\} = \frac{1}{4}$, it then follows that $A_n^- > \frac{\Delta_n^-}{4}$ given $x_n' < x_n$.

Now let $\Delta^+ = \sum_{n=1}^N \Delta_n^+ \cdot \mathrm{I}[x_n' > x_n]$, $\Delta^- = \sum_{n=1}^N \Delta_n^- \cdot \mathrm{I}[x_n' < x_n]$, $A^+ = \sum_{n=1}^N A_n^+ \cdot \mathrm{I}[x_n' > x_n]$, and $A^- = \sum_{n=1}^N A_n^- \cdot \mathrm{I}[x_n' < x_n]$. We can then see that $\mathbf{x}$ is chosen over $\mathbf{x}' \neq \mathbf{x}$ as long as $A^- > A^+$. From our above work, we then see $A^- > \sum_{n=1}^N \frac{\Delta_n^-}{4} \cdot \mathrm{I}[x_n' < x_n] = \frac{\Delta^-}{4}$, and $A^+ < \sum_{n=1}^N \frac{\Delta_n^+}{4} \cdot \mathrm{I}[x_n' > x_n] = \frac{\Delta^+}{4}$. Furthermore, given $\sum_{n=1}^N x_n = \sum_{n=1}^N x_n' = A$ and $x_n = \frac{A}{N}$ for all n, it must be the case that $\Delta^+ = \Delta^-$. This ensures $A^- > A^+$, implying the diversification bias is captured under average normalization.

Feature Bias (**Y**). Given $\mathbf{x}$ and $\mathbf{x}'$ are as defined in Proposition 8, let $\mathbf{y} = (x_{n'}, q)$ and $\mathbf{y}' = (x_{n'} + q, 0)$ where n' is the dimension with $x'_{n'} = x_{n'} + q$. Using (5) with $r(x_n) = ||X||^{-1} \sum_{\mathbf{x}' \in X} x'_n$, we can verify $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{y}; \{\mathbf{y}, \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{y}, \mathbf{y}'\})$, which must be positive from Proposition 16 since $y_1 y_2 = x_{n'} q > 0 = (x_{n'} + q) \cdot 0 = y'_1 y'_2$, which ensures the feature bias holds under average normalization.

C.9. Maximum normalization

We now consider "maximum normalization" as given in (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\}$.

Compromise Effect (N). Given $y_2 = \frac{x_1 x_2}{y_1}$ (Proposition 16), $z_1 > x_1$, and $z_2 < x_2$, (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = \frac{x_1}{x_1+z_1} + \frac{x_2}{x_2+y_2} - \frac{y_1}{y_1+z_1} + \frac{y_2}{y_2+y_2} = \frac{x_1}{x_1+z_1} - \frac{x_1}{x_1+x_1} + \frac{y_1}{y_1+x_1} - \frac{y_1}{y_1+z_1} = -\frac{(x_1-y_1)(z_1-y_1)(z_1-x_1)}{(2(x_1+z_1)(y_1+z_1)(y_1+x_1))} < 0$, implying the compromise effect is not captured under maximum normalization.

Dominance Effect, Weak (N) *and Strict* (N). Given $z_1 < x_1$, and $z_2 \leq x_2$, (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = \frac{x_1}{x_1+x_1} + \frac{x_2}{x_2+y_2} - \frac{y_1}{y_1+x_1} + \frac{y_2}{y_2+y_2} = V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) = 0$, implying the weak and strict dominance effects are not captured.

Decoy-Range Effect (N). Given $z_1 < x_1$, and $z_2 \leq x_2$, $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\}) - V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\})]}{\partial z_2} = 0$ under (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\}$, implying the decoy-range effect is not captured.

Relative Difference Effect (**Y**), *Majority-Rule Preference Cycles* (**Y**), *Splitting Bias* (**Y**), *Alignability Effect* (**Y**), *Diversification Bias* (**Y**), and *Feature Bias* (**Y**). As established in the proof of Proposition 16, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ holds (for any N) under maximum normalization if and only if $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ holds under (1). Therefore, all of these effects (which are formalized in terms of binary choices) are captured by maximum normalization since they are captured by pairwise normalization.

C.10. Minimum normalization

We now consider "minimum normalization" as given in (5) with $r(x_n) = \min_{\mathbf{x}' \in X}\{x'_n\}$.

Compromise Effect (**Y**). Given $y_2 = \frac{x_1 x_2}{y_1}$ (Proposition 16), $z_1 > x_1$, and $z_2 < x_2$, (5) with $r(x_n) = \min_{\mathbf{x}' \in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = \frac{x_1}{x_1+y_1} + \frac{x_2}{x_2+z_2} - \frac{y_1}{y_1+y_1} + \frac{y_2}{y_2+z_2} = \frac{x_2}{x_2+z_2} - \frac{x_2}{x_2+x_2} + \frac{y_2}{y_2+x_2} - \frac{y_2}{y_2+z_2} = \frac{(x_1-y_1)(x_2-z_2)(x_1x_2-y_1z_2)}{2(x_1+y_1)(x_2+z_2)(x_1x_2+y_1z_2)}$. This last expression is positive since $x_1 > y_1$ and $x_2 > z_2$ (noting these inequalities also ensure $x_1 x_2 > y_1 z_2$), implying the compromise effect is captured under minimum normalization.

Dominance Effect, Weak (N) *and Strict* (**Y**). Given $y_2 = \frac{x_1 x_2}{y_1}$ (Proposition 16), $x_1 > z_1 > y_1$, and $z_2 \leq x_2$, (5) with $r(x_n) = \min_{\mathbf{x}' \in X}\{x'_n\}$ again implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = \frac{x_1}{x_1+y_1} + \frac{x_2}{x_2+z_2} - \frac{y_1}{y_1+y_1} + \frac{y_2}{y_2+z_2} = \frac{x_2}{x_2+z_2} - \frac{x_2}{x_2+x_2} + \frac{y_2}{y_2+x_2} - \frac{y_2}{y_2+z_2} = \frac{(x_1-y_1)(x_2-z_2)(x_1x_2-y_1z_2)}{2(x_1+y_1)(x_2+z_2)(x_1x_2+y_1z_2)}$. We can then see that this last expression equals zero if $z_2 = x_2$, implying the weak dominance effect is not captured under (5) with $r(x_n) = \min_{\mathbf{x}' \in X}\{x'_n\}$. We can also see that this expression is otherwise positive with $x_1 > y_1$ and $x_2 > z_2$ (again noting these inequalities also ensure $x_1 x_2 > y_1 z_2$), implying the strict dominance effect is captured under minimum normalization.

Decoy-Range Effect (**Y**). Given $z_1 < x_1$, and $z_2 \leq x_2$, $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\})-V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\})]}{\partial z_2} = \frac{(y_2-x_2)(z_2^2-x_2y_2)}{(x_2+z_2)^2(y_2+z_2)^2}$ under (5) with $r(x_n) = \min_{\mathbf{x}'\in X}\{x'_n\}$. This expression must be negative since $y_2 > x_2 > z_2$, implying the decoy-range effect is captured.

Relative Difference Effect (**Y**), *Majority-Rule Preference Cycles* (**Y**), *Splitting Bias* (**Y**), *Alignability Effect* (**Y**), *Diversification Bias* (**Y**), and *Feature Bias* (**Y**). As established in the proof of Proposition 16, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\})$ holds (for any N) under minimum normalization if and only if $V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) > V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\})$ holds under (1). Therefore, all of these effects (which are formalized in terms of binary choices) are captured by minimum normalization since they are captured by pairwise normalization.

C.11. Max-min normalization

We now consider "max-min normalization" as given in (5) with $r(x_n) = \max_{\mathbf{x}'\in X}\{x'_n\} - \min_{\mathbf{x}'\in X}\{x'_n\}$.

Compromise Effect (S). To show that max-min normalization sometimes predicts the compromise effect and sometimes predicts the opposite, suppose $\mathbf{x} = (2, 1)$, $\mathbf{y} = (1, 2)$, and $\mathbf{z} = (2.1, z_2)$. Then, if $z_2 = 0$, (5) with $r(x_n) = \max_{\mathbf{x}'\in X}\{x'_n\} - \min_{\mathbf{x}'\in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx .0023 > 0$, consistent with the compromise effect. If $z_2 = .2$, (5) with $r(x_n) = \max_{\mathbf{x}'\in X}\{x'_n\} - \min_{\mathbf{x}'\in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx -.0002 < 0$, which is the opposite of the compromise effect (we can also verify $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{z}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx .25 > 0$, ensuring $\mathbf{z}$ is not chosen).

Dominance Effect, Weak (N) *and Strict* (S). With $y_2 > x_2 = z_2$ and $x_1 > z_1 > y_1$, $r(x_n) = \max_{\mathbf{x}'\in X}\{x'_n\} - \min_{\mathbf{x}'\in X}\{x'_n\}$ for $n = 1, 2$ is not affected by the inclusion of $\mathbf{z}$ in X. Thus, $V(\mathbf{x}'; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{x}'; \{\mathbf{x}, \mathbf{y}\}) = 0$ for $\mathbf{x}' \in \{\mathbf{x}, \mathbf{y}\}$, implying that max-min normalization does not predict a dominance effect in this case. To show that max-min normalization sometimes predicts the dominance effect and sometimes predicts the opposite with $x_1 > z_1 > y_1$ and $y_2 > x_2 > z_2$, first verify $\mathbf{x} = (3, 1)$, $\mathbf{y} = (1, 3)$, and $\mathbf{z} = (2, .5)$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx .0069 > 0$, consistent with the dominance effect. However, $\mathbf{x} = (2, 1)$, $\mathbf{y} = (1, 2)$, and $\mathbf{z} = (1.5, .5)$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) \approx -.0048 > 0$, which is the opposite of the dominance effect.

Decoy-Range Effect (**Y**). To show that the decoy-range effect is captured under max-min normalization, first note that $x_1 > z_1 > y_1$ and $y_2 > x_2 = z_2$ imply $\max\{x'_n : \mathbf{x}' \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}\} \in \{x_n, y_n\}$ and $\min\{x'_n : \mathbf{x}' \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}\} \in \{x_n, y_n\}$ for each $n = 2$. It then follows that, with $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\})$, $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) > V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ must hold given $x_1 > z_1 > y_1$ and $y_2 > x_2 = z_2$. Next, given $z_2 = \min\{x'_2 : \mathbf{x}' \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}\}$, we can compute $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\})-V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\})]}{\partial z_2} = \frac{x_2}{(x_2+y_2-z_2)^2} - \frac{y_2}{(2y_2-z_2)^2}$ under (5) with $r(x_n) = \max_{\mathbf{x}'\in X}\{x'_n\} - \min_{\mathbf{x}'\in X}\{x'_n\}$. We can then see that $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\})-V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\})]}{\partial z_2} < 0$ if and only if $z_2 < y_2 - \sqrt{x_2y_2}$. This implies $z_2 < y_2 - \sqrt{x_2y_2}$ for $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\})$ and $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$ to both hold with $\mathbf{z}$ asymmetrically dominated by $\mathbf{x}$ under (5) with $r(x_n) = \max_{\mathbf{x}'\in X}\{x'_n\} - \min_{\mathbf{x}'\in X}\{x'_n\}$. Next, given $V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}\}) > V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}\})$, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}\})$, $z'_1 = z_1$, and $z'_2 < z_2$, we must have $V(\mathbf{x}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\}) > V(\mathbf{y}; \{\mathbf{x}, \mathbf{y}, \mathbf{z}'\})$ since $z'_2 < z_2 < y_2 - \sqrt{x_2y_2}$ and $\frac{\partial[V(\mathbf{x};\{\mathbf{x},\mathbf{y},\mathbf{z}\})-V(\mathbf{y};\{\mathbf{x},\mathbf{y},\mathbf{z}\})]}{\partial z_2} < 0$ for all $z_2 < y_2 - \sqrt{x_2y_2}$. Thus, the decoy-range effect is captured with max-min normalization.

Relative Difference Effect (**Y**). The relative difference effect must be captured because (a) it is captured under (1) (Proposition 1), and (b) binary choices are equivalent under (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$ and (1) (Proposition 16).

Majority-Rule Preference Cycles (S). To show that max-min normalization sometimes predicts majority-rule preference cycles and sometimes predicts the opposite, suppose $\mathbf{x} = (a, b, c)$, $\mathbf{x}' = (b, c, a)$, and $\mathbf{x}'' = (c, a, b)$ with $a = 1$ and $b = .5$. Then, if $c = .2$, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) - V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\}) = V(\mathbf{x}''; \{\mathbf{x}, \mathbf{x}''\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}''\}) = \frac{13}{360} > 0$ under (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$, thus generating a majority-rule preference cycle. If $c = .4$, (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{x}'; \{\mathbf{x}', \mathbf{x}''\}) - V(\mathbf{x}''; \{\mathbf{x}', \mathbf{x}''\}) = V(\mathbf{x}''; \{\mathbf{x}, \mathbf{x}''\}) - V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}''\}) = -\frac{1}{40} < 0$, in which case there is a minority-rule preference cycle instead.

Splitting Bias (**Y**). Using (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$, $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{(\delta^a)^2}{x_{1a}(\delta^a + x_{1a})} + \frac{(\delta^b)^2}{x_{1b}(\delta^b + x_{1b})} - \frac{(\delta^a + \delta^b)^2}{(x_{1a} + x_{1b})(\delta^a \delta^b + x_{1a} + x_{1b})}$ with $\delta^a = x_{1a} - y_{1a}$ and $\delta^b = x_{1b} - y_{1b}$. Combining the right-side fractions, we see $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{(x_{1a}\delta^b)^2(x_{1a}(2\delta^a + \delta^b + 2x_{1b}) + \phi) + (x_{1b}\delta^a)^2(\phi' + x_{1b}(\delta^a + 2\delta^b + 2x_{1a})) + (x_{1b}^2 \delta^a - x_{1a}^2 \delta^b)^2}{x_{1a}(\delta^a + x_{1a})x_{1b}(\delta^b + x_{1b})(x_{1a} + x_{1b})(\delta^a \delta^b + x_{1a} + x_{1b})}$, where $\phi \equiv \delta^a(\delta^a + \delta^b + x_{1b})$ and $\phi' \equiv \delta^b(\delta^a + \delta^b + x_{1a})$. Since all terms in both the numerator and the denominator are positive, this final expression must be positive, implying the splitting bias is captured under max-min normalization.

Alignability Effect (**Y**). Using (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$ and $y_2 = \frac{x_1 x_2}{y_1}$ (as implied by Proposition 16) with $\mathbf{x}'$ and $\mathbf{y}'$ as defined in Proposition 6, we can compute $V(\mathbf{x}'; \{\mathbf{x}', \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{x}', \mathbf{y}'\}) = \frac{(x_1 - y_1)^2}{x_1(2x_1 - y_1)} > 0$, which ensures the alignability effect holds under max-min normalization.

Diversification Bias (S). Here, we show that the diversification bias is always captured for $N = 2$ but may not be captured with $N > 2$. Given $X = \{\mathbf{x}, \mathbf{x}'\}$, with $N = 2$ it is readily verifiable that (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$ implies $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = \frac{2\delta^3}{(1+\delta)(1+2\delta)}$, where $\delta = \frac{|x_1 - x'_1|}{x_1}$. Thus, $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) > 0$ for all $\mathbf{x}' \neq \mathbf{x}$ with $N = 2$, consistent with the diversification bias. However, suppose $N = A = 5$ (implying $\mathbf{x} = (1, 1, 1, 1, 1)$) and $\mathbf{x}' = \left(\frac{10}{6}, \frac{5}{6}, \frac{5}{6}, \frac{5}{6}, \frac{5}{6}\right)$. We can then compute $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = -\frac{2}{105} < 0$ under (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$, implying the diversification bias is not captured in this case.

Feature Bias (**Y**). Given $\mathbf{x}$ $\mathbf{x}'$ are as defined in Proposition 8, let $\mathbf{y} = (x_{n'}, q)$ and $\mathbf{y}' = (x_{n'} + q, 0)$ where n' is the dimension for which $x'_{n'} = x_{n'} + q$. Using (5) with $r(x_n) = \max_{\mathbf{x}' \in X}\{x'_n\} - \min_{\mathbf{x}' \in X}\{x'_n\}$, we can verify $V(\mathbf{x}; \{\mathbf{x}, \mathbf{x}'\}) - V(\mathbf{x}'; \{\mathbf{x}, \mathbf{x}'\}) = V(\mathbf{y}; \{\mathbf{y}, \mathbf{y}'\}) - V(\mathbf{y}'; \{\mathbf{y}, \mathbf{y}'\})$, which must be positive from Proposition 16 since $y_1 y_2 = x_{n'} q > 0 = (x_{n'} + q)0 = y'_1 y'_2$, ensuring the feature bias holds under max-min normalization.

C.12. Discussion of other models

As mentioned in the footnote of Table 1, there are other notable models addressing context-dependent choice — in particular, the compromise and/or dominance effects — besides those whose predictions were explicitly classified in Table 1. Here, we briefly describe other notable theories, and address the relevant differences in structure and/or domain that make them less amenable to direct behavioral comparisons with our model.

To begin, in Kamenica's (2008) model, a compromise effect can arise as a rational response to information asymmetries between consumers and a monopolist firm. Specifically, consumers who are initially uninformed of their preferences can infer from a trinary choice set — representing the firm's (endogenously-chosen) product line — that their preferences correspond to a "type" for which the intermediate option is optimal, while such inferences cannot be made from a binary choice set. Among other key differences with our single decision-maker model, Kamenica's model features a continuum of consumers who are heterogeneous in their preferences and in whether or not they are initially informed of their preferences, as well as a monopolist firm that endogenously chooses its product line. In addition, consumers who exhibit a compromise effect in Kamenica's model infer from the presence of the third alternative that, in the current state, some consumers prefer that third alternative even though those who exhibit a compromise effect do not. It is therefore less clear how a dominance effect could arise through such "contextual inference" unless consumers can prefer a dominated alternative over an alternative that dominates it. Furthermore, alternatives in Kamenica's model are only defined on two dimensions (quality and price), so behavioral patterns requiring the consideration of more than two attributes (e.g. majority-rule preference cycles) are outside its scope.

Next, De Clippel and Eliaz (2012) model two-attribute choice as an intra-personal bargaining problem, where two "selves" disagree on the relative importance of the two attributes. In trinary choice, the two selves ultimately settle on the alternative that is not the worst on any dimension, giving rise to a compromise effect as well as a dominance effect. Behaviors requiring the consideration of more than two attributes are not addressed by this model, and it is not clear how it would be adapted to a N-attribute setting (presumably requiring N-selves). Perhaps, however, the key behavioral feature — a "fallback" bargaining outcome that favors an alternative that is not worst on any dimension — could be generalized. Still, this mechanism would not explain binary-choice phenomena such as majority-rule preference cycles, the diversification bias, and the feature bias (as formalized in Propositions 4, 7, and 8) in which both alternatives under consideration are worst on at least one dimension.

In another approach, Ok et al. (2015) axiomatically characterize a reference-dependent model capturing a form of the attraction effect that, unlike most other theories, does not require exogenously-defined attributes. Since our model and the behaviors it addresses are — as with the large majority of theories addressing such context-dependent behaviors — formalized in terms of exogenously-defined attributes, its predictions are not amenable to direct comparisons with the predictions of Ok et al.'s model.

Guo (2016) proposes a model in which the presented choice set endogenously determines the extent to which a rational consumer engages in costly deliberation as a means to learn their own preferences. As Guo demonstrates, this "contextual deliberation" mechanism can produce a compromise effect. As in Kamenica's (2008) model — and unlike our approach — Guo's model features consumers who are imperfectly informed of their own preferences, while context-dependent behavior arises through context-dependent learning. In addition, Guo's model and our model address somewhat different behaviors. For instance, Guo's model captures choice overload effects, which are not typically addressed by multi-attribute choice models (including ours), but does not consider the dominance effect.

Lastly, in Natenzon's (2019) model, an imperfectly-informed decision-maker can exhibit probabilistic forms of the compromise and dominance effects. Since context-dependence in this model is driven by correlations among stochastic perceptual errors in valuation, its predictions are not amenable to classification according to the deterministic criteria used to classify behavior in Table 1.[36] Even if that were not an issue, the usefulness of classifying the predictions of Natenzon's model with respect to the behaviors in Table 1 may still be questionable since its predictions are highly dependent on the values of its parameters, but unlike the theoretical models considered, Natenzon's model is primarily empirical, and its parameters are meant to be estimated.

Appendix D. Model restrictions in Fig. 2

This appendix describes parametric and functional form assumptions used to create the graphs in Fig. 2 (as well as Fig. 4) for the Tversky and Simonson (1993), Kivetz et al. (2004a), Soltani et al. (2012), and Bordalo et al. (2013) models. As noted in the text, these graphs depicted each model's predicted effect of adding a third alternative **z** on the DM's relative valuation of two alternatives, **x** and **y**, which are equally-valued in binary choice. With one exception (addressed below), we used $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$ to generate the graphs in Fig. 2. In turn, the parametric and functional form assumptions described below were selected due to their simplicity and adherence to the more general restrictions of the model to which they were applied.

Tversky and Simonson (1993). To generate the graph for Tversky and Simonson's (1993) model, we used the value function (13) described in Appendix C. For any $\theta > 0$, it is then readily verifiable under (13) that, with $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$ and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$,

$$V(\mathbf{x}; X) - V(\mathbf{y}; X) \propto \frac{\max\{2-z_1,0\}+\max\{1-z_2,0\}}{|2-z_1|+|1-z_2|} - \frac{\max\{1-z_1,0\}+\max\{2-z_2,0\}}{|1-z_1|+|2-z_2|},$$

which generates the regions shown in Fig. 2 for Tversky and Simonson's (1993) model.

Kivetz et al. (2004a). To generate the graph for Kivetz et al.'s (2004a) model, we used the value function (14) described in Appendix C. It is then readily verifiable under (14) that, with $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$ and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$,

$$V(\mathbf{x}; X) - V(\mathbf{y}; X) = \sum_{n=1}^{2} (3 - n - \min\{1, z_n\})^c - (n - \min\{1, z_n\})^c,$$

which, taking any $c \in (0, 1)$, generates the regions shown in Fig. 2 for the Kivetz et al. (2004a) model.

Soltani et al. (2012). To generate the graph for Soltani et al.'s (2012) model, we used the value function (18), along with the additional restriction $w_n = w > 0$ for $n = 1, 2$. It is then readily verifiable under (18) that, with $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$ and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$,

$$V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{w}{\max\{z_1,2\}-\min\{z_1,1\}} - \frac{w}{\max\{z_2,2\}-\min\{z_2,1\}},$$

which generates the regions shown in Fig. 2 for the Soltani et al. (2012) model.

[36] That is to say, if stochastic perceptual errors were excluded from Natenzon's model, choice behavior would no longer be context-dependent, making the classification exercise trivial. By contrast, the stochastic components of Kivetz et al.'s (2004a) model could be omitted (without loss of insight) when classifying its predictions because context-dependence arises through its deterministic components.

Bordalo et al. (2013) — Two Quality Attributes. To generate the graph for Bordalo et al.'s (2013) model with alternatives defined on two quality dimensions, we used the value function (15) described in Appendix C. It is then readily verifiable under (15) that, with $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 2)$ and $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$,

$$V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{2\delta^{g_x(\mathbf{z})} + \delta^{1-g_x(\mathbf{z})}}{\delta^{g_x(\mathbf{z})} + \delta^{1-g_x(\mathbf{z})}} - \frac{\delta^{g_y(\mathbf{z})} + 2\delta^{1-g_y(\mathbf{z})}}{\delta^{g_y(\mathbf{z})} + \delta^{1-g_y(\mathbf{z})}},$$

where $g_x(\mathbf{z}) \equiv \mathrm{I}\left[\frac{|3-z_1|}{9+z_1} > \frac{z_2}{6+z_2}\right]$ and $g_y(\mathbf{z}) \equiv \mathrm{I}\left[\frac{z_1}{6+z_1} > \frac{|3-z_2|}{9+z_2}\right]$. We can then see that, for any $\delta \in (0, 1)$, these expressions generate the regions shown in Fig. 2.

Bordalo et al. (2013) — Price and Quality. To generate the graph for Bordalo et al.'s (2013) model with alternatives defined by its price and a single quality attribute, we used $\mathbf{x} = (p_x, q_x) = (1, 1)$ and $\mathbf{y} = (p_y, q_y) = (2, 2)$ while otherwise applying the same restrictions in (15), which can still be applied with price as an attribute simply by treating the price of $\mathbf{z} \in \{\mathbf{x}, \mathbf{y}\}$ as a negative quality attribute with value $-p_z$. Given $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, the predicted value difference between $\mathbf{x}$ and $\mathbf{y}$ in trinary choice is then

$$V(\mathbf{x}; X) - V(\mathbf{y}; X) = \frac{\delta^{1-g_x(\mathbf{z})} - \delta^{g_x(\mathbf{z})}}{\delta^{g_x(\mathbf{z})} + \delta^{1-g_x(\mathbf{z})}} - 2 \cdot \frac{\delta^{1-g_y(\mathbf{z})} - \delta^{g_y(\mathbf{z})}}{\delta^{g_y(\mathbf{z})} + \delta^{1-g_y(\mathbf{z})}},$$

where now $g_x(\mathbf{z}) \equiv \mathrm{I}[p_z > q_z]$ and $g_y(\mathbf{z}) \equiv \mathrm{I}\left[\frac{|3-p_z|}{9+p_z} < \frac{|3-q_z|}{9+q_z}\right]$. In turn, these expressions (again, with any $\delta \in (0, 1)$) generate the regions shown in Fig. 2.

References

Arieli, Amos, Ben-Ami, Yaniv, Rubinstein, Ariel, 2011. Tracking decision makers under uncertainty. Am. Econ. J. Microecon. 3, 68–76.

Azar, Ofer, 2008. The effect of relative thinking on firm strategy and market outcomes: a location differentiation model with endogenous transportation costs. J. Econ. Psychol. 29, 684–697.

Ballé, J., Laparra, V., Simoncelli, E., 2017. End-to-end optimized image compression. In: International Conference on Learning Representations. https://arxiv.org/abs/1611.01704.

Bardolet, David, Fox, Craig, Lovallo, Dan, 2007. Naive Diversification and Partition Dependence in Capital Allocation Decisions. Unpublished manuscript.

Barlow, Horace, 1961. Possible principles underlying the transformation of sensory messages. Sens. Commun. 1, 217–234.

Bateman, Ian, Munro, Alistair, Rhodes, Bruce, Starmer, Chris, 1997. Does part–whole bias exist? An experimental investigation. Econ. J. 107, 322–332.

Benartzi, Shlomo, Thaler, Richard, 2001. Naive diversification strategies in defined contribution saving plans. Am. Econ. Rev. 91, 79–98.

Bernheim, Douglas, 2009. On the potential of neuroeconomics: a critical (but hopeful) appraisal. Am. Econ. J. Microecon. 1, 1–41.

Birnbaum, Michael, Bahra, Jeffrey, 2007. Gain-loss separability and coalescing in risky decision making. Manag. Sci. 53, 1016–1028.

Bordalo, Pedro, Gennaioli, Nicola, Shleifer, Andrei, 2013. Salience and consumer choice. J. Polit. Econ. 121, 803–843.

Bushong, Benjamin, Rabin, Matthew, Schwartzstein, Joshua, 2019. A Model of Relative Thinking. Working paper.

Carandini, Matteo, Heeger, David, 1994. Summation and division by neurons in primate visual cortex. Science 264, 1333–1335.

Carandini, Matteo, Heeger, David, 2012. Normalization as a canonical neural computation. Nat. Rev., Neurosci. 13, 51–62.

Carandini, Matteo, Heeger, David, Movshon, Anthony, 1997. Linearity and normalization in simple cells of the macaque primary visual cortex. J. Neurosci. 17, 8621–8644.

Carpenter, Gregory, Glazer, Rashi, Nakamoto, Kent, 1994. Meaningful brands from meaningless differentiation: the dependence on irrelevant attributes. J. Mark. Res. 31, 339–350.

Castillo, Geoffrey, 2020. The attraction effect and its explanations. Games Econ. Behav. 119, 123–147.

Daviet, Remi, 2018. Methods for statistical analysis and prediction of discrete choices. University of Toronto PhD Thesis.
Daviet, Remi, Webb, Ryan, 2020. A Double Decoy Experiment to Distinguish Theories of Dominance Effects. SSRN working paper.
De Clippel, Geoffroy, Eliaz, Kfir, 2012. Reason-based choice: a bargaining rationale for the attraction and compromise effects. Theor. Econ. 7, 125–162.
Doyle, John, O'Connor, David J., Reynolds, Gareth, Bottomley, Paul, 1999. The robustness of the asymmetrically dominated effect: buying frames, phantom alternatives, and in-store purchases. Psychol. Mark. 16, 225–243.
Dumbalska, S., Li, V., Tsetsos, K., Summerfield, C., 2020. A map of decoy influence in human multi-alternative choice. Proc. Natl. Acad. Sci. 117 (40), 25169–25178.
Frederick, Shane, Lee, Leonard, Baskin, Ernest, 2014. The limits of attraction. J. Mark. Res. 51, 487–507.
Frisch, Deborah, 1993. Reasons for framing effects. Organ. Behav. Hum. Decis. Process. 54, 399–429.
Frydman, Cary, Jin, Lawrence, 2020. Efficient Coding and Risky Choice. Working paper.
Furl, Nicholas, 2016. Facial-attractiveness choices are predicted by divisive normalization. Psychol. Sci. 27, 1379–1387.
Geyskens, Inge, Gielens, Katrijn, Gijsbrechts, Els, 2010. Proliferating private-label portfolios: how introducing economy and premium private labels influences brand choice. J. Mark. Res. 47, 791–807.
Gourville, John, Soman, Dilip, 2005. Overchoice and assortment type: when and why variety backfires. Mark. Sci. 24, 382–395.
Guo, Liang, 2016. Contextual deliberation and preference construction. Manag. Sci. 62, 2977–2993.
Guo, Julie, Tymula, Agniezka, 2020. Waterfall Illusion in Risky Choice. University of Sydney Working Paper.
Hamalainen, Raimo, Alaja, Susanna, 2008. The threat of weighting biases in environmental decision analysis. Ecol. Econ. 68, 556–569.
Hartline, Keffer, Wagner, Henry, 1952. The peripheral origin of nervous activity in the visual system. Cold Spring Harbor Symp. Quant. Biol. 17, 125–141.
Heeger, David, 1992. Normalization of cell responses in cat striate cortex. Vis. Neurosci. 9, 181–197.
Holper, Lisa, van Brussel, Leo, Schmidt, Luzia, Schulthess, Sven, 2017. Adaptive value normalization in the prefrontal cortex is reduced by memory load. eNeuro: Cogn. Behav. 4, 1–20.
Hsee, Chrisopher, Loewenstein, George, Blount, Sally, Bazerman, Max, 1999. Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. Psychol. Bull. 25, 576–590.
Huber, Joel, Payne, John, Puto, Christopher, 1982. Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. J. Consum. Res. 9, 90–98.
Huber, Joel, Payne, John, Puto, Christopher, 2014. Let's be honest about the attraction effect. J. Mark. Res. 51, 520–525.
Humphrey, Steven, 1995. Regret aversion or event-splitting effects? More evidence under risk and uncertainty. J. Risk Uncertain. 11, 263–274.
Hunt, Laurence, Dolan, Raymond, Behrens, Timothy, 2014. Hierarchical competitions subserving multi-attribute choice. Nat. Neurosci. 17, 1613–1626.
Itthipuripat, Sirawaj, Cha, Kexin, Rangsipat, Napat, 2015. Value-based attentional capture influences context-dependent decision-making. J. Neurophysiol. 114, 560–569.
Jacobi, Sarah, Hobbs, Benjamin, 2007. Quantifying and mitigating the splitting bias and other value tree-induced weighting biases. Decis. Anal. 4, 194–210.
Kahneman, Daniel, Knetsch, Jack, 1992. Valuing public goods: the purchase of moral satisfaction. J. Environ. Econ. Manag. 22, 57–70.
Kahneman, Daniel, Tversky, Amos, 1984. Choices, values, frames. Am. Psychol. 39, 341–350.
Kamenica, Emir, 2008. Contextual inference in markets: on the informational content of product lines. Am. Econ. Rev. 98, 2127–2149.
Khaw, Mel, Glimcher, Paul, Louie, Kenway, 2017. Normalized value coding explains dynamic adaptation in the human valuation process. Proc. Natl. Acad. Sci. 114, 2127–2149.
Kivetz, Ran, Netzer, Oded, Srinivasan, V., 2004a. Alternative models for capturing the compromise effect. J. Mark. Res. 41, 237–257.
Kivetz, Ran, Netzer, Oded, 2004b, Srinivasan, V. Extending compromise effect models to complex buying situations and other context effects. J. Mark. Res. 41, 262–268.
Koszegi, Botond, Szeidl, Adam, 2013. A model of focusing in economic choice. Q. J. Econ. 128, 53–104.
Król, Michał, Król, Magdalena, 2019. Inferiority, not similarity of the decoy to target, is what drives the transfer of attention underlying the attraction effect: evidence from an eye-tracking study with real choice. J. Neurosci. Psychol. Econ. 12, 88–104.
Lea, Amanda, Ryan, Michael, 2015. Irrationality in mate choice revealed by tungara frogs. Science 349, 964–966.
Liew, Shi Xian, Howe, Piers, Little, Daniel, 2016. The appropriacy of averaging in the study of context effects. Psychon. Bull. Rev. 23, 1639–1646.

LoFaro, Thomas, Louie, Kenway, Webb, Ryan, 2014. The temporal dynamics of cortical normalization models of decision-making. Lett. Biomath. 1, 209–220.
Louie, Kenway, Glimcher, Paul, Webb, Ryan, 2015. Adaptive neural coding: from biological to behavioral decision-making. Curr. Opin. Behav. Sci. 5, 91–99.
Louie, Kenway, Grattan, Lauren, Glimcher, Paul, 2011. Reward value-based gain control: divisive normalization in parietal cortex. J. Neurosci. 31, 10627–10639.
Louie, Kenway, LoFaro, Thomas, Webb, Ryan, 2014. Dynamic divisive normalization predicts time-varying value coding in decision-related circuits. J. Neurosci. 34, 16046–16057.
Louie, Kenway, Khaw, Mel, Glimcher, Paul, 2013. Normalization is a general neural mechanism for context-dependent decision making. Proc. Natl. Acad. Sci. 110, 6139–6144.
Malkoc, Selin, Hedgcock, William, 2013, Hoeffler, Steve Between a rock and a hard place: the failure of the attraction effect among unattractive alternatives. J. Consum. Psychol. 23, 317–329.
Markman, Arthur, Medin, Douglas, 1995. Similarity and alignment in choice. Organ. Behav. Hum. Decis. Process. 63, 117–130.
Marley, Anthony, 1991. Context dependent probabilistic choice models based on measures of binary advantage. Math. Soc. Sci. 21, 201–231.
May, Kenneth, 1954. Intransitivity, utility, and the aggregation of preference patterns. Econometrica 22, 1–13.
Mellers, Barbara, Cooke, Alan, 1994. Trade-offs depend on attribute range. J. Exp. Psychol. Hum. Percept. Perform. 20, 1055.
Meyer, Robert, Zhao, Shenghui, Han, Jin, 2008. Biases in valuation vs. usage of innovative product features. Mark. Sci. 27, 1083–1096.
Mick, David, Fournier, Susan, 1998. Paradoxes of technology: consumer cognizance, emotions, and coping strategies. J. Consum. Res. 25, 123–143.
Mowen, Maryanne, Mowen, John, 1986. An empirical examination of the biasing effects of framing on business decisions. Decis. Sci. 17, 596–602.
Natenzon, Paulo, 2019. Random choice and learning. J. Polit. Econ. 127 (1), 419–457.
Netzer, Nick, 2009. Evolution of time preferences and attitudes toward risk. Am. Econ. Rev. 99, 937–955.
Noguchi, Takao, Stewart, Neil, 2014. In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. Cognition 132, 44–56.
Noguchi, Takao, Stewart, Neil, 2018. Multialternative decision by sampling: a model of decision making constrained by process data. Psychol. Rev. 125, 512–544.
Ohshiro, Tomokazu, Angelaki, Dora, DeAngelis, Gregory, 2011. A normalization model of multisensory integration. Nat. Neurosci. 14, 775–782.
Ok, Efe, Ortoleva, Pietro, Riella, Gil, 2015. Revealed (p)reference theory. Am. Econ. Rev. 105, 299–321.
Olsen, Shawn, Bhandawat, Vikas, Wilson, Rachel, 2010. Divisive normalization in olfactory population codes. Neuron 66, 287–299.
Padoa-Schioppa, Camillo, 2009. Range-adapting representation of economic value in the orbitofrontal cortex. J. Neurosci. 29, 14004–14014.
Payne, John, Bettman, James, Johnson, Eric, 1993. The Adaptive Decision Maker. Cambridge University Press, New York.
Qamar, Ahmad, Cotton, James, George, Ryan, Beck, Jeffrey, Prezhdo, Eugenia, 2013. Trial-to-trial, uncertainty-based adjustment of decision boundaries in visual categorization. Proc. Natl. Acad. Sci. 110, 20332–20337.
Rabinowitz, Neil, Willmore, Ben, Schnupp, Jan, 2011. Contrast gain control in auditory cortex. Neuron 70, 1178–1191.
Rangel, Antonio, Clithero, John, 2012. Value normalization in decision making: theory and evidence. Curr. Opin. Neurobiol. 22, 970–981.
Rayo, Luis, Becker, Gary, 2007. Evolutionary efficiency and happiness. J. Polit. Econ. 115, 302–337.
Read, Daniel, Loewenstein, George, 1995. Diversification bias: explaining the discrepancy in variety seeking between combined and separated choices. J. Exp. Psychol., Appl. 1, 34.
Reeck, Crystal, Wall, Daniel, Johnson, Eric, 2017. Search predicts and changes patience in intertemporal choice. Proc. Natl. Acad. Sci. 114, 11890–11895.
Rieskamp, Jorg, Busemeyer, Jerome, Mellers, Barbara, 2006. Extending the bounds of rationality: evidence and theories of preferential choice. J. Econ. Lit. 44, 631–661.
Robson, Arthur, 2001. The biological basis of economic behavior. J. Econ. Lit. 39, 11–33.
Robson, Arthur, Whitehead, Lorne, 2018. Rapidly Adaptive Hedonic Utility. Working paper.
Rubinstein, Ariel, 2002. Irrational diversification in multiple decision problems. Eur. Econ. Rev. 46, 1369–1378.
Russo, Edward, Dosher, Barbara, 1983. Strategies for multiattribute binary choice. J. Exp. Psychol. Learn. Mem. Cogn. 9, 676–696.

Russo, Edward, Rosen, Larry, 1975. An eye fixation analysis of multialternative choice. Mem. Cogn. 3, 267–276.
Rustichini, Aldo, Conen, Katherine, Cai, Xinying, Padoa-Schioppa, Camillo, 2017. Optimal coding and neuronal adaptation in economic decisions. Nat. Commun. 8, 1–14.
Schwartz, Janet, Chapman, Gretchen, 1999. Are more options always better? The attraction effect in physicians' decisions about medications. Med. Decis. Mak. 19, 315–323.
Schwartz, Odelia, Simoncelli, Eero, 2001. Natural signal statistics and sensory gain control. Nat. Neurosci. 4, 819–825.
Shevell, Steven, 1977. Saturation in human cones. Vis. Res. 17, 427–434.
Simonson, Itamar, 1989. Choice based on reasons: the case of attraction and compromise effects. J. Consum. Res. 16, 158–174.
Sinz, Fabian, Bethge, Matthias, 2013. Temporal adaptation enhances efficient contrast gain control on natural images. PLoS Comput. Biol. 9, e1002889.
Soltani, Alireza, De Martino, Benedetto, Camerer, Colin, 2012. A range-normalization model of context-dependent choice: a new model and evidence. PLoS Comput. Biol. 8, e1002607.
Starmer, Chris, Sugden, Robert, 1993. Testing for juxtaposition and event-splitting effects. J. Risk Uncertain. 6, 235–254.
Steverson, Kai, Brandenburger, Adam, Glimcher, Paul, 2019. Choice-theoretic foundations of the divisive normalization model. J. Econ. Behav. Organ.
Sullivan, Nicolette, Fitzsimons, Gavan, Platt, Michael, 2019. Indulgent foods can paradoxically promote disciplined dietary choices. Psychol. Sci. 30, 273–287.
Thompson, Debora, Hamilton, Rebecca, Rust, Roland, 2005. Feature fatigue: when product capabilities become too much of a good thing. J. Mark. Res. 42, 431–442.
Thompson, Debora, Norton, Michael, 2011. The social utility of feature creep. J. Mark. Res. 48, 555–565.
Trueblood, Jennifer, Brown, Scott, Heathcote, Andrew, 2013. Not just for consumers: context effects are fundamental to decision making. Psychol. Sci. 24, 901–908.
Trueblood, Jennifer, Brown, Scott, Heathcote, Andrew, 2015. The fragile nature of contextual preference reversals: reply to Tsetsos, Chater, and Usher (2015). Psychol. Rev. 122, 848–853.
Tserenjigmid, Gerelt, 2015. Theory of decisions by intra-dimensional comparisons. J. Econ. Theory 159, 326–338.
Tserenjigmid, Gerelt, 2019. Choosing with the worst in mind: a reference-dependent model. J. Econ. Behav. Organ. 157, 631–652.
Tsetsos, Konstantinos, Moran, Rani, Moreland, James, Chater, Nick, 2016. Economic irrationality is optimal during noisy decision making. Proc. Natl. Acad. Sci. 113, 3102–3107.
Turner, Brandon M., Schley, Dan R., Muller, Carly, Tsetsos, Konstantinos, 2017. Competing theories of multi-alternative, multi-attribute preferential choice. Psychol. Rev.
Tversky, Amos, 1972. Elimination by aspects: a theory of choice. Psychol. Rev. 79, 281–299.
Tversky, Amos, Russo, Edward, 1969. Substitutability and similarity in binary choices. J. Math. Psychol. 6, 1–12.
Tversky, Amos, Simonson, Itamar, 1993. Context-dependent preferences. Manag. Sci. 39, 1179–1189.
Tymula, Agnieszka, Glimcher, Paul, 2019. Expected Subjective Value Theory: a Representation of Decision Under Risk and Certainty. Working paper.
Wainwright, Martin, Schwartz, Odelia, Simoncelli, Eero, 2002. Natural image statistics and divisive normalization. In: Rao, Olsahausen, Lewicki (Eds.), Probabilistic Models of the Brain: Perception and Neural Function. MIT Press, London.
Webb, Ryan, Glimcher, Paul, Louie, Kenway, 2020a. The normalization of consumer valuations: context-dependent preferences from neurobiological constraints. Manag. Sci.
Webb, Ryan, Glimcher, Paul, Louie, Kenway, 2020b. Divisive normalization does influence decisions with multiple alternatives. Nat. Hum. Behav.
Weber, Ernst, 1834. On the Tactile Senses (with translation of De Tactu). Experimental Psychology Society, New York.
Weber, Martin, Borcherding, Katrin, 1993. Behavioral influences on weight judgments in multiattribute decision making. Eur. J. Oper. Res. 67, 1–12.
Weber, Martin, Eisenfuhr, Franz, Von Winterfeldt, Detlof, 1988. The effects of splitting attributes on weights in multiattribute utility measurement. Manag. Sci. 34, 431–445.
Wedell, Douglas, 1991. Distinguishing among models of contextually induced preference reversals. J. Exp. Psychol. Learn. Mem. Cogn. 17, 767–778.
Wilson, Nathan, Runyan, Caroline, Wang, Forea, 2012. Division and subtraction by distinct cortical inhibitory networks in vivo. Nature 488, 343–348.
Woodford, Michael, 2012. Inattentive Valuation and Reference-Dependent Choice. Working paper.
Xing, Jing, Heeger, David, 2001. Measurement and modeling of center-surround suppression and enhancement. Vis. Res. 41, 571–583.

Yamada, Hiroshi, Louie, Kenway, Tymula, Agnieszka, 2018. Free choice shapes normalized value signals in medial orbitofrontal cortex. Nat. Commun. 9, 162–171.
Yeung, Catherine, Soman, Dilip, 2005. Attribute evaluability and the range effect. J. Consum. Res. 32, 363–369.
Zhang, Jiao, Hsee, Christopher, Xiao, Zhixing, 2006. The majority rule in individual decision making. Organ. Behav. Hum. Decis. Process. 99, 102–111.
Zhang, Shi, Markman, Arthur, 1998. Overcoming the early entrant advantage: the role of alignable and nonalignable differences. J. Mark. Res. 35, 413–426.
Zimmermann, J., Glimcher, P.W., Louie, K., 2018. Multiple timescales of normalized value coding underlie adaptive choice behaviour. Nat. Commun. 9, 3206.