Pairwise Attribute Normalization

Peter Landry† Ryan Webb‡

January 13, 2017

PRELIMINARY

Abstract

We propose a theory of multi-attribute choice with foundation in the neuroscience of sensory perception. The theory relies on a pairwise comparison of attributes by means of divisive normalization, a neural computation widely observed in cortex across many sensory modalities and species. The theory captures and unifies a number of previously disparate phenomena observed in the empirical literature, including the decoy effect, compromise effect, similarity effect, “majority rules” intransitive preference cycles, attribute-splitting effects, and comparability biases. A general formulation of the theory contains standard binary utility representations as a special case. The relation between pairwise normalization and the previously-established concept of attribute salience found in the economics literature is also explored.

*We thank P. Bordalo, D. Freeman, A.A.J. Marley, and C. Stewart for helpful feedback, and F. Wever for excellent research assistance.
†Rotman School of Management, University of Toronto, and Department of Management, University of Toronto, Mississauga. E-mail: peter.landry@rotman.utoronto.ca.
‡Rotman School of Management & Dept. of Economics, University of Toronto. E-mail: ryan.webb@rotman.utoronto.ca.
1 Introduction

Given a choice among alternatives that vary on multiple dimensions (i.e. attributes), numerous empirical studies have documented how the composition of the choice set can influence preferences. For instance, a revealed preference for one alternative in binary choice may reverse when a third alternative is added to the choice set. A well-known example is the (asymmetric) dominance effect (also called the decoy effect or attraction effect), in which introducing an alternative that is dominated by one alternative (on all attribute dimensions) can shift preference towards the dominant alternative (Huber, Payne, and Puto, 1982).\(^1\) A related behaviour, the compromise effect, describes how adding an extreme alternative to the choice set can shift preference towards the alternative that now lies intermediate on each attribute (Simonson, 1989).

Both the dominance and compromise effects are striking because — in the canonical model of rational choice — a preference ranking between two alternatives is immutable and should not depend on the presence of other alternatives in the choice set. For this reason, these two behavioural phenomena have received an abundance of attention in the theoretical literature, with particular focus paid to the manner in which alternatives are compared within and across attribute dimensions. For instance, Tversky and Simonson (1993) originally proposed that the dominance effect resulted from an asymmetric “gain/loss” weighting of attributes relative to an exogenously specified reference point. More recent studies have formalized how the dominance effect can arise by comparing attributes to a reference point which depends on the properties of the choice set itself, including the range of alternatives in the set (Koszegi and Szeidl, 2013; Cunningham, 2013; Bushong, Rabin, and Schwartzstein, 2016). Of this latter class of models, Bordalo, Gennaioli, and Shleifer (2013) introduce the concept of attribute salience — which operationalizes basic principles of sensory perception — to determine how attention is allocated to various attributes in order to generate both the dominance and compromise effects.\(^2\) More generally, Ok, Ortoleva, and Riella (2014) have characterized how the dominance effect can arise within a model which allows the reference point to be revealed endogenously from choice data, without the need to even specify attribute dimensions.\(^3\)

\(^1\)Empirical demonstrations of the dominance effect occur both in the lab and in the field, ranging from incentivized lotteries (Soltani, De Martino, and Camerer, 2012), to consumer goods (Doyle, O’Connor, Reynolds, and Bottomley, 1999), to medical decision-making (Schwartz and Chapman, 1999), to mate choice in animals (Lea and Ryan, 2015). Results have been demonstrated for both numerical and perceptual attributes, across and within-subject (Huber, Payne, and Puto, 2014; Trueblood, Brown, Heathcote, and Busemeyer, 2013, though see Frederick, Lee, and Baskin (2014) for potential limitations on robustness in perception). Additional relevant citations can be found in Ok, Ortoleva, and Riella (2014).

\(^2\)They also address a broader range of choice phenomena than just the dominance and compromise effects.

\(^3\)An additional line of literature demonstrates that the compromise effect can arise from information asymmetries between
However the range of multi-attribute choice anomalies is broader than just the dominance and compromise effects. Reviewed with more detail in Section 2, the wider class of multi-attribute choice anomalies includes: majority-rules transitivity violations (i.e. Condorcet cycles along attribute dimensions); subadditive attribute weighting (more commonly known as “attribute-splitting” effects); and comparability biases (underweighting of attributes that are not present for some alternatives). In this article we propose a simple, parsimonious, model of multi-attribute choice that provides a unified explanation for this broader range of phenomena. Of primary interest, the mechanism for how attributes are evaluated in relation to each other is grounded in an established normative principle of neurobiological computation.

Over 50 years of research in neuroscience has documented that the brain encodes information in relative, not absolute terms. For example, neurons in the visual cortex monotonically encode the luminance (brightness) of a particular region of visual space, much like pixels on a computer monitor (Hubel and Wiesel, 1968). However the activity of these neurons is also suppressed by the activity of nearby neurons encoding nearby regions of visual space via a computation called Divisive Normalization (Heeger, 1992; Carandini, Heeger, and Movshon, 1997). Similar results have been observed in other sensory systems including olfaction and audition, and across species ranging from primates to the fruit fly (Carandini and Heeger, 2012, for a review). More recently, the evidence for divisive normalization has been extended to the areas of the brain involved in value-based choice. The neural activity encoding the reward associated with a choice option is independently suppressed (normalized) by the value of the reward for other choice options (Louie, Grattan, and Glimcher, 2011).

The brain encodes information in a relative form because it is capacity constrained. Since the brain has a limited number of neurons, each with a bounded response range, sensory/value information must be efficiently compressed within this finite range. To see how divisive normalization implements a bounded representation, consider comparing two alternatives that vary on only one attribute dimension with positive values $x$ and $y$. In the simplest conceivable version of the

---

4Theoretical explanations for such transitivity violations include a bounded rational agent who is unable to construct complete rankings (Gans, 1996), and regret of unchosen options (Loomes and Sugden, 1982)

5The normative criterion in sensory neuroscience is the Efficient Coding Hypothesis (Barlow, 1961), which states that a neural encoding should minimize the redundancy of information encoded between representational units. Divisive normalization accomplishes this criterion by reducing the mutual information of a sensory stimulus (Schwartz and Simoncelli, 2001; Ballé, Laparra, and Simoncelli, 2016a). Ballé, Laparra, and Simoncelli (2016b) demonstrate that Divisive Normalization compresses sensory information more efficiently than standard computational algorithms, such as JPEG compression.
divisive normalization computation, the value of $x$ is normalized in relation to the value of $y$ according to

$$v(x; y) \equiv \frac{x}{x + y}. \quad (1)$$

If we presume that a consumer’s preference between the two alternatives is determined by which of their normalized valuations $v(x; y)$ or $v(y; x)$ is larger, this bounded value function maintains monotonicity since $v(x; y) > v(x; y)$ if and only if $x > y$.

While existing choice models examining the role of divisive normalization have allowed for more general forms of normalization than (1), these models have only considered choice alternatives which vary on a single dimension. Such models predict patterns of violations of the axiom of Independence of Irrelevant Alternatives that are consistent with observed behaviour (Louie, Khaw, and Glimcher, 2013), and capture behaviour better than standard econometric methods for handling such violations (Webb, Glimcher, and Louie, 2016). From a normative perspective, normalization preserves relative valuations and minimizes choice errors subject to a neural resource constraint (Webb, Glimcher, and Louie, 2016). More broadly, Steverson, Brandenburger, and Glimcher (2016) provide an axiomatization for a general form of normalization and demonstrate that it optimizes the costs of choice errors (modelled via Shannon entropy).

However, the uni-dimensional implementation of divisive normalization cannot address the dominance effect, nor the range of multi-attribute choice anomalies observed in the empirical literature (Louie, Glimcher, and Webb, 2015). Here, we extend divisive normalization into the realm of multiple attributes and characterize its implications for a wide range of choice behaviour.

In particular, we consider the following simple specification of attribute normalization. Letting $x = (x_1, \ldots, x_n) \in \mathbb{R}^n_+$ denote an alternative as defined by the vector of its $n$ observable attribute values, **pairwise attribute normalization** is defined by

$$v(x; C_x) \equiv \sum_{y \in C_x} \sum_{i=1}^n \frac{x_i}{x_i + y_i}, \quad (2)$$

where $C_x \equiv X \setminus \{x\}$ is the set of all alternatives in the choice set $X$ besides $x$.

There are two key assumptions which embed how normalization assimilates information in a multi-attribute setting. First, attribute information is normalized within attribute. This is the defining property of normalization from the sensory perception literature and allows the normalization function to be separable across attributes. Second, normalization is pairwise in the sense that each pair of alter-
natives is compared separately within each attribute dimension.\textsuperscript{6} Substantial empirical evidence supports a pairwise comparison of attributes in decision-making.\textsuperscript{7}

We demonstrate that this simple relative comparison can capture a wide range of behavioural phenomena in multi-attribute choice (reviewed in detail below), including the dominance effect, compromise effect, majority-rules transitivity violations, attribute-splitting effects, and comparability biases. The model makes predictions not only about where (in attribute space) phenomena like the dominance effect will arise, but also makes predictions about the strength of the effect that are consistent with empirical results.\textsuperscript{8} In doing so, we demonstrate that pairwise normalization implements fundamental properties of perception — principally, the property of \textit{diminishing sensitivity} — thus providing a neurobiological mechanism for the concept of attribute salience proposed in Bordalo, Gennaioli, and Shleifer (2013). In addition, the effective salience function implied by (2) is \textit{subadditive}, a property that serves a crucial role in generating the other multi-attribute choice anomalies (majority-rules intransitivities, attribute-splitting effects, and comparability biases) addressed by our framework.

Though the model presumes a specific functional form, it is a functional form that has both positive and normative justification. In Section 7, we consider a more general parametric form of normalization studied in the neuroscience literature. The inclusion of two parameters alleviates some extreme properties of the model and admits a link from pairwise normalization to standard utility theory, including hybrids of the two approaches. It also generates new empirical implications, including a similarity effect commonly reported in the empirical literature (e.g. Tversky, 1972).

\textsuperscript{6}Marley (1991) first introduced a pairwise comparison model in a stochastic setting under a general comparison process. Tversky and Simonsen (1993) then restricted the comparison of attributes via an asymmetric linear weighting of gain/loss differences, not normalization. This formulation not only required an exogenous specification of the reference point (i.e. what determines a gain vs. a loss), but also requires the degree of asymmetry to be specified ex-ante. Similar approaches aimed at capturing dominance and compromise effects can be found in dynamic process models in the psychology literature via asymmetric weighting (Busemeyer and Townsend, 1993; Trueblood, Brown, and Heathcote, 2014) or via loss aversion (Usher and McClelland, 2001). An earlier approach to incorporating the bounded activity of neurons in a choice model also relies on a linear asymmetric weighting by the range of attributes (Soltani, De Martino, and Camerer, 2012). In our model, no reference point or asymmetry is imposed exogenously, instead context dependence arises naturally from the relative comparison of the attributes themselves via normalization. In this sense, our approach is related to the endogenous reference point described in Ok, Ortolaeva, and Riella (2014), though we do still require the attribute space to be defined.

\textsuperscript{7}For instance, eye-tracking studies reveal that individuals repeatedly compare pairs of alternatives along a single attribute dimension, as opposed to within-alternative comparisons across attributes (Russo and Dosher, 1983; Noguchi and Stewart, 2014). Support for within-attribute comparisons is also evident in lottery comparisons (Ariely et al., 2011) and the neuroscience literature on valuation. Multiple meta-studies, covering over 200 separate studies, associate activity in the pre-frontal cortex with valuation before and during decision-making (Levy and Glimcher, 2012; Bartra, McGuire, and Kable, 2013; Clithero and Rangel, 2013). Lesions in this region yield highly inconsistent choice behaviour in human subjects, including violations of the General Axiom of Revealed Preference (Camille, Griffiths, Vo, Fellows, and Kable, 2011). In a clinical study, Fellows (2006) found that healthy control subjects (with an intact medial pre-frontal cortex) exhibit pairwise attribute comparisons, however subjects with a lesion of the medial pre-frontal cortex do not. This suggests that pair-wise attribute comparisons may be fundamentally related to the choice process of consistent individuals.

\textsuperscript{8}To our knowledge, the experiment conducted by (Soltani, De Martino, and Camerer, 2012) is the only empirical study to examine the strength of decoy effects. We review this study in detail in section 3.2.
Of course, further parametric generalizations can easily be added to increase the flexibility of normalized preferences for empirical application. For instance, a formulation which imposes linear weights on attributes — commonly used in the discrete choice econometric literature — can be easily incorporated to address subjective preferences over attribute dimensions.\(^9\) More generally, \( x_i \) can be re-defined as a subjective “partworth” component of utility, yielding pairwise normalization in the subjective evaluation of each attribute. However we will work primarily with the parameter-free form given in (2) for two reasons. First, to emphasize the role of normalization in early stages of sensory processing in the human brain, and second, to emphasize the ability of pairwise normalization, in its most basic form, to capture a wide range of anomalous and previously-unconnected empirical regularities in multi-attribute choice.

The rest of this article proceeds as follows. Section 2 briefly reviews the relevant literature on multi-attribute choice, noting in particular the behavioural phenomena that pairwise normalization captures. Section 3 formally presents the model and details how pairwise normalization shifts preferences when a third alternative is added to the choice set (e.g. the dominance and compromise effect). Section 4 addresses majority-rules preference cycles in binary choice sets with three attribute dimensions. Section 5 addresses subadditivity and Section 6 addresses comparability biases. Section 7 demonstrates that a generalization of the theory, based on the general parameterization of divisive normalization, yields the above results and contains standard utility theory as a special case.

2 Related Literature on Multi-Attribute Choice Anomalies

We now elaborate on other multi-attribute choice anomalies (besides the dominance and compromise effects) that are captured by pairwise normalization.

2.1 Majority-Rules Intransitivities

Suppose three alternatives, \( x, y, \) and \( z \) differ on three dimensions such that \( x \) is better than \( y \) on two of three dimensions, \( y \) is better than \( z \) on two of three dimensions, and \( z \) is better than \( x \) on two of three dimensions. A number of studies have shown that binary choices in such environments exhibit a particular form of intransitivity, dubbed a “majority-rules” preference cycle, where \( x \) is chosen over \( y, y \) over \( z, \) and \( z \) over \( x \) (May, 1954; Russo and Dosher, 1983; Zhang, Hsee, and

\[ v(x) \equiv \sum_{y \neq x} \sum_i \beta_i \cdot \frac{x_i}{x_i + y_i} \] would capture both subjective attribute weights and issues with defining the units of the attributes. Also see footnote 16.
Table 1: Three hypothetical marriages used by May (1954).

<table>
<thead>
<tr>
<th>Spouse</th>
<th>Intelligence</th>
<th>Looks</th>
<th>Wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>very intelligent</td>
<td>plain looking</td>
<td>well off</td>
</tr>
<tr>
<td>y</td>
<td>intelligent</td>
<td>very good looking</td>
<td>poor</td>
</tr>
<tr>
<td>z</td>
<td>fairly intelligent</td>
<td>good looking</td>
<td>rich</td>
</tr>
</tbody>
</table>

For instance, May (1954) presented 62 college students with three prospective spouses, with each potential spouse characterized by intelligence, looks, and wealth (Table 1). Although only 17 of the 62 subjects ordered the candidates cyclically, all 17 of these subjects exhibited the majority rules cycle (x preferred to y, y to z, and z to x) with none choosing the reverse cycle. Obviously, the occurrence of such cycles depends on the preferences of subjects over attributes matching the implied ordering. As noted by Bar-Hillel and Margalit (1988), “45 other subjects managed to order the three candidates from most to least desirable without cyclicality. Apparently, the subjects either considered one single dimension over-ridingly important (for instance, 21 of the 45 transitive orderings were from most to least intelligent), or could trade off advantages and disadvantages across dimensions. Indeed, the difficulty in obtaining systematic violations of transitivity in such contexts arises because choice patterns that transparently violate normative desiderata are labile.”

A recent study by Tsetsos, Moran, Moreland, Chater, Usher, and Summerfield (2016a) avoids this issue by matching monetary incentives directly to objectively measurable attributes. In the experiment, subjects chose between two alternatives (of a total set of three), each characterized by nine sequentially occurring bars of different heights, presented in a simultaneous stream. The height of the two bars in each presentation correspond to the values of the “attribute.” At the end of each trial, they were asked to choose the stream with the larger average height, receiving monetary reward proportional to their choice accuracy.

The key manipulation of the experiment is that the streams are judiciously chosen by the experimenter. While the average height of the two bars over nine attributes are equal, one of the bars is larger on six of the nine attributes. They find that 17 of 28 subjects yielded a significant majority-rules preference cycle. In a follow-up experiment specifically designed to assess cycles within-subject, they find a statistically significant majority-rules cycle in 15 of 21 subjects, with 11 exhibiting violations of (the stricter) Weak Stochastic Transitivity condition.10

---

10 Also see Davis-Stober, Park, Brown, and Regenwetter (2016); Tsetsos, Moran, Moreland, Chater, Usher, and Summerfield
Only 1 subject exhibited the opposite cycle.

### 2.2 Attribute Splitting (Subadditive Attribute Weighting)

In some choice contexts, the number of attributes can depend on the resolution with which attributes are defined. For example, two relevant attributes of a decision over sports cars could be the performance and running costs of each car. Suppose the faster car had high running costs and the slower one had lower costs. A decision between the cars would thus hinge on the tradeoff between performance and costs.

However a “cost” attribute can be decomposed into separate sub-attributes, such as fuel costs and maintenance costs. Empirical research documents that attribute-splitting of this sort (and the way in which a product is described more generally) can affect consumers’ valuations of a product. Decision-makers tend to place more cumulative weight on the value of two sub-attributes than on a single, composite attribute (Weber, Eisenfuhr, and von Winterfeldt, 1988; Weber and Borcherding, 1993, for a short review). Returning to our car example, since the slower of the two cars has both lower fuel costs and maintenance costs, it appears better in comparison to the faster car if it is compared on all three dimensions — speed, fuel costs, maintenance costs. Hämäläinen and Alaja (2008) demonstrate the effect can be attenuated, but not eliminated, via direct instruction and training of decision-makers.

The attribute-splitting effect is related to the more general property of subadditive weighting of attributes.\(^{11}\) Therefore, the attribute-splitting effect can also be related to a range of empirical findings that demonstrate subadditivity in other contexts besides multi-attribute consumer choice. For example, Kahneman and Knetsch (1992) document subadditivity in subjects’ valuations of a public good (as measured by willingness-to-pay). Bateman, Munro, Rhodes, Starmer, and Sugden (1997) demonstrate that valuations of meal vouchers are subadditive. Many studies have demonstrated subadditivity in subjects judgments (or weighting) of probability\(^{12}\), and subadditivity has also been used to characterize the tendency for subjects to display aversion to ambiguous lottery choices (i.e., the Ellsberg Paradox; Schmeidler, 1989). Read (2001) demonstrates subadditivity also arises in intertemporal choice settings, as subjects discount a future reward at a fixed time to a greater extent if the delay is divided into subintervals than if it remains undivided.

---

\(^{11}\)More formally, for two sets \(A\) and \(B\) closed under addition, a subadditive function \(f : A \to B\) has the property \(f(x + y) \leq f(x) + f(y)\), \(\forall x, y, \in A\).

\(^{12}\)See, for example, Fischhoff, Slovic, and Lichtenstein (1978), Starmer and Sugden (1993), Tversky and Koehler (1994), and Rottenstreich and Tversky (1997).
2.3 Comparability

A long line of experimental research addresses the “comparability” of attributes (also called “alignable differences”), where an attribute dimension is comparable if it is present for all alternatives and not comparable if it is only present in one alternative. To illustrate, consider the experiment by Slovic and MacPhillamy (1974) in which subjects are asked to assess the ability of students based on a variety of test scores (Table 2). For one of the tests, a score is present for both students, however for each of the remaining two tests only one student has a score.\textsuperscript{13} They find that subjects overweight the comparable attribute dimensions in their assessment relative to non-comparable attributes, a result which has now been replicated in numerous studies (Markman and Medin, 1995; Nowlis and Simonson, 1997; Zhang and Markman, 1998; Zhang and Fitzsimons, 1999; Herrmann, Heitmann, and Morgan, 2009).\textsuperscript{14} Interestingly, the effect strengthens as non-comparable attribute dimensions are added (Gourville and Soman, 2005), yielding a systematic violation of the regularity axiom.

Table 2: Example of comparable and non-comparable attributes from Slovic and MacPhillamy (1974).

<table>
<thead>
<tr>
<th></th>
<th>Student A</th>
<th>Student B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement</td>
<td>618</td>
<td>561</td>
</tr>
<tr>
<td>English</td>
<td>-</td>
<td>572</td>
</tr>
<tr>
<td>Quantitative</td>
<td>382</td>
<td>-</td>
</tr>
</tbody>
</table>

The comparability bias implies that an alternative can be preferred to another simply because it is dominant on the comparable dimension, even if the non-comparable attribute dimensions are not valued symmetrically. For instance, an individual considering two automobiles might value the presence of a sunroof more than a leather interior, but opt for the latter simply because it has a few more horsepower than the former. A closely-related finding from this literature involves valuation when options are evaluated jointly versus separately (Hsee, Loewenstein, Blount, and Bazerman, 1999; Hsee and Zhang, 2010). For example, the higher horsepower car with the leather interior could be preferred to the lower horsepower car with a sunroof when evaluated jointly as a choice, but the valuations of each attribute elicited separately would favour the latter.

\textsuperscript{13}Other examples include alternatives in which features are either present or not. For example, consider a set of microwave ovens which vary on internal capacity (comparable attribute), where each has a unique additional feature (moisture sensor, programmable menu, online help, etc...). (Gourville and Soman, 2005).

\textsuperscript{14}Of particular note, Slovic and MacPhillamy (1974) considered a treatment in which direct instruction was given to avoid the comparability bias (“Try not to pay more attention to a test simply because both students have a score on it”). Instead of reducing the bias, subjects appeared to simply use the given information less consistently.
3 Model

We now present a model of pairwise attribute normalization. A consumer chooses from a finite choice set, $X$. Each alternative $x \in X$ is defined with respect to $n > 0$ attributes, where $x_i \geq 0$ is the unnormalized input value for alternative $x$ on attribute $i \leq n$. To restrict to non-trivial choice problems between any two alternatives $x$ and $y$, let $x_1 > y_1$ and $x_2 > y_2$ without loss of generality, and denote $C_x \equiv X \setminus \{x\}$ as the set of all alternatives in $X$ besides $x$.

We now consider the consumer’s overall normalized valuation of $x \in X$ as given by pairwise attribute normalization:

$$v(x; C_x) = \sum_{y \in C_x} \sum_{i=1}^{n} \frac{x_i}{x_i + y_i}.$$  

(3)

Intuitively, we can interpret (3) as arising from a series of pairwise, attribute-level comparisons in which the elementary normalization computation is applied. When evaluating $x$, each comparison expresses $x$’s value in relation to some other alternative $y \in X$ on a given attribute $i$, yielding a normalized valuation $v_i(x_i; y_i) = \frac{x_i}{x_i + y_i}$ between 0 and 1. The consumer’s overall evaluation of $x$ is then the aggregation of all $v_i(x_i; y_i)$ terms, reflecting every possible comparison — to each alternative $y \in C_x$ and on each attribute dimension $i = 1, \ldots, n$.

We use (3) as the basis of our descriptive choice model in that the consumer is presumed to choose $x \in X$ if $v(x; C_x) > v(y; C_y)$ for all $y \in C_x$. If $v(x; C_x) > v(y; C_y)$, we will say that $x$ is “preferred” to $y$ given $X$. In cases where $v(x; C_x) = v(y; C_y)$ for $y \neq x$, we will say that the consumer is “indifferent” between $x$ and $y$ given $X$. Whenever we refer to a preference between two alternatives, it will therefore be implicit that we are referring to a strict preference.

To place our results within a common interpretive framework, we introduce the direct contrast function, $\Delta$, which quantifies the effective contrast between two alternatives on a single attribute dimension:

$$\Delta(x_i, y_i) \equiv |v_i(x_i; y_i) - v_i(y_i; x_i)| = \left| \frac{x_i - y_i}{x_i + y_i} \right|.$$  

Specifically, $\Delta(x_i, y_i)$ reflects the perceived magnitude of the difference between the attribute-$i$ valuations of $x$ and of $y$, after each has been normalized in relation to the other, also expressed on a scale from 0 to 1. Note that this definition parallels the definition of contrast in the visual perception literature (Carandini and Heeger, 2012). We refer to this as a “direct” measure of contrast because it arises from comparisons between $x_i$ and $y_i$ (later we will compare this to a notion of “indirect”
contrast based on indirect comparisons through a third value \( z_i \).

To illuminate key features of the direct contrast function and its role in shaping preferences under pairwise attribute normalization, it is useful to consider the concept of salience introduced by Bordalo, Gennaioli, and Shleifer (2012, 2013). A salience function operationalizes two fundamental properties of sensory perception.

**Definition 1.** A salience function \( s(a, b) \) is symmetric, continuous, and satisfies:

1. **Ordering:** if \( [a, b] \subset [a', b'] \), then \( s(a', b') > s(a, b) \).
2. **Diminishing Sensitivity:** for any \( \epsilon > 0 \), \( s(a + \epsilon, b + \epsilon) < s(a, b) \).

In general, the first argument of a salience function \( s(\cdot, \cdot) \) corresponds to the level or intensity of a stimulus, while the second argument represents the intensity of a “reference” stimulus to which the first is compared. **Ordering** states that a larger difference in intensity between two stimuli is easier to discern, therefore is more salient. **Diminishing Sensitivity** states that, for a given difference in intensities between two stimuli, increasing the magnitudes of those intensities reduces salience. This latter condition is a pervasive property of perception, more commonly known as Weber’s Law (Weber, 1834).\(^{15}\)

We now show that pairwise attribute normalization implements the properties of salience, as applied in comparisons between two alternatives on a single attribute dimension. All proofs are in the appendix.

**Proposition 1.** \( \Delta(x_i, y_i) \) is a salience function.

The effect of normalization on preferences can be understood through the familiar properties of a salience function. As a simple illustration, consider the case of two attribute binary choice. The consumer’s choice problem is equivalent to determining whether normalization renders \( x \)'s advantage on the first attribute more or less salient than \( y \)'s advantage on the second. That is to say, the consumer chooses \( x \) if \( \Delta(x_1, y_1) > \Delta(x_2, y_2) \) and chooses \( y \) otherwise, with both ordering and diminishing sensitivity implemented by the contrast function \( \Delta(x_i, y_i) \).\(^{16}\)

In addition to this relationship between direct contrast and preferences under pairwise attribute normalization, two other formulations equivalently represent consumer preferences in the case of two-attribute binary choice.

**Proposition 2.** Let \( X = \{x, y\} \) and \( n = 2 \). The following are equivalent:

\(^{15}\)As an illustration, the weight difference between a 20-gram rock and a 30-gram rock is less salient than the difference between a 10-gram rock and a 40-gram rock (ordering), but more salient than the difference between a 1020-gram rock and a 1030-gram rock (diminishing sensitivity).

\(^{16}\)In our base normalization model, it is easy to see that \( \Delta(x_i, y_i) \) also satisfies degree-zero homogeneity, \( \Delta(kx_i, ky_i) = \Delta(x_i, y_i) \) for any \( k > 0 \) — a convenient mathematical property shared by the salience functions forms used by Bordalo, Gennaioli, and Shleifer (2012, 2013). In particular, this ensures that, if need be, normalization can be defined over the subjective evaluation of attributes (utils) provided they are scalar transformations of the attribute.
(I) **Pairwise Attribute Normalization:** \(v(x; y) > v(y; x)\)

(II) **Direct Contrast:** \(\Delta(x_1, y_1) > \Delta(x_2, y_2)\).

(III) **Cobb-Douglas:** \(x_1 x_2 > y_1 y_2\).

(IV) **Weighted-Additive Utility:** \(\theta_1 x_1 + \theta_2 x_2 > \theta_1 y_1 + \theta_2 y_2\), where
\[
\theta_1 \equiv \frac{x_2 + y_2}{\sum_i x_i + y_i},
\quad \text{and} \quad
\theta_2 \equiv \frac{x_1 + y_1}{\sum_i x_i + y_i} \quad \text{(so that} \quad \theta_1 + \theta_2 = 1)\).

From (III), pairwise attribute normalization generates preferences equivalent to those arising from a symmetric Cobb-Douglas utility function, \(u(x) = x_1 x_2\). This equivalence is reassuring since it ensures that our base model maintains some well-understood and well-behaved features inherent in Cobb-Douglas utility. For instance, preferences will be transitive (in the two attribute case) because \(x_1 x_2 > y_1 y_2\) and \(y_1 y_2 > z_1 z_2\) necessarily imply \(x_1 x_2 > z_1 z_2\). Moreover, preferences over attributes are strictly convex, with the marginal rate of substitution between \(x_1\) and \(x_2\) given by \(x_2 / x_1\).\(^{17}\) However we will see in later sections that the capacity of our base model to generate Cobb-Douglas utility (along with some of its more convenient features) does not extend to trinary choice nor does it extend to choice problems in which alternatives are defined on three attribute dimensions.

The weighted-additive utility model (IV) is a common formulation for examining multi-attribute choice, both in theoretic and econometric applications. A number of previous models have examined how distortions of the attributes, via unequal weights on each attribute that sum to one, give rise to dominance and compromise effects (Bordalo, Gennaioli, and Shleifer, 2013).\(^{18}\) From (IV), we observe that normalization can be re-formulated in this manner, and can readily assess whether the consumer overweights or underweights a particular attribute by observing which among \(\theta_1\) and \(\theta_2\) is larger. Moreover, the precise forms of these attribute weights \(\theta_1\) and \(\theta_2\) reveal an intrinsic element of inter-attribute competition that exists in our model. For example, the weight \(\theta_1 = \frac{x_2 + y_2}{\sum_i x_i + y_i}\) describes how, exactly, \(x_2\) and \(y_2\) affect the weighting of attribute 1, in addition to their (opposite) effects on \(\theta_2\).\(^{19}\)

Finally, the relation between pairwise attribute normalization and direct contrast (I and III) emphasizes a key aspect of our model. The salience of the difference

---

\(^{17}\)Given (III), the link to direct contrast (II) also suggests a simple rule to assess whether \(x\)'s advantage is more or less salient than \(y\)'s advantage. Since \(\Delta(x_1, y_1)\) is a salience function, we can simply compare the ratios of each attribute dimension. That is \(x\)'s advantage is more salient than \(y\)'s advantage if \(\frac{x_1}{y_1} > \frac{y_2}{x_2}\). Though here we abstract from prices to focus on goods with multiple, non-negative quality dimensions, this rule resembles the role of price-to-quality ratios as a determinant of salience in Bordalo, Gennaioli, and Shleifer (2013), and conforms with their observation that consumers are drawn to such alternatives.

\(^{18}\)As noted earlier, Tversky and Simonson (1993) specify the weights \(\theta_i\) asymmetrically depending on whether they are a gain or loss relative to an exogenously specified reference point. In Koszegi and Szeidl (2013), \(\theta_i\) is an increasing function of the range of attributes on that dimensions, while in Bushong, Rabin, and Schwartzstein (2016) it is a decreasing function of this range.

\(^{19}\)The direct contrast function representation in (III) portrays preferences as arising from distortions inherent in two separate intra-attribute comparisons, where perceptions of \(x\)'s value in relation to \(y\) on attribute 1 are unaffected by perceptions on attribute 2 (and vice versa).
between $x_i$ and $y_i$ does not merely play an intermediate role in determining the effective value difference on which preferences are based. Instead, salience and the normalized value difference are equivalent. In this sense, attributes do not exist independently within the model (or, indeed, in neural activity), instead they are always represented in a relative form. We should also note that the form of salience arising in our base model embeds a very strong form of diminishing sensitivity. Holding the ratio between $x_i$ and $y_i$ fixed, the absolute difference between $x_i$ and $y_i$ has no bearing on the salience of their difference, and by extension, no bearing on choice behavior. As we will see in subsequent sections, the precise features of salience as generated by pairwise attribute normalization are more subtle in more complex choice environments, and can be moderated by the general parameterization of normalization we study in Section 7.

3.1 Basic Context-Dependence

We now consider trinary choice problems (with alternatives still defined on two attributes), focusing in particular on how preferences between two alternatives, $x$ and $y$, can be swayed by a third alternative, $z$. For ease of exposition, it will be useful to consider the binary indifference benchmark $v(x; y) = v(y; x)$. If $x$ and $y$ lie on the same binary-choice indifference curve, a strict preference between $x$ and $y$ with $X = \{x, y, z\}$ is necessarily attributable to the presence of $z$ in the choice set.

We now identify special cases in which preferences will not depend on context.

**Lemma 1.** Suppose $z$ is on the same binary-choice indifference curve as either $x$ or $y$. Then $x$ is preferred to $y$ given $X = \{x, y\}$ if and only if $x$ is preferred to $y$ given $X = \{x, y, z\}$. If, in addition, $v(x; y) = v(y; x)$, the consumer will be indifferent between all three alternatives given $X = \{x, y, z\}$.

Hence, the preference between $x$ and $y$ in binary choice is preserved with $z$ in the choice set (i.e., context-independent), in the event that the consumer would be indifferent in a hypothetical binary choice between $z$ and one of either $x$ or $y$. Therefore, $z$ can only sway preferences between $x$ and $y$ if it rests on its own binary-choice indifference curve. For the case of indifference between $x$ and $y$ in binary choice, this means $z$ must either be superior or inferior to both $x$ and $y$.

---

20 Here, we will continue to equate preferences with normalized valuations in the sense that “$x$ is preferred to $y$” will be taken to mean $v(x; y, z) > v(y; x, z)$. However, a preference for $x$ over $y$ no longer guarantees $x$ will be chosen because $z$ may be preferred to both $x$ and $y$ (and hence $z$ would be chosen from the full choice set).

21 It is implicit (and straightforward to show) that the mechanisms that drive context effects in our model can capture “complete” preference reversals too, in the sense that if the consumer is indifferent between $x$ and $y$ in binary choice but prefers $x$ to $y$ in trinary choice with $X = \{x, y, z\}$, then there exists a $x'$ such that $y$ is preferred to $x'$ in binary choice but $x'$ is preferred to $y$ in trinary choice given $X = \{x', y, z\}$.
in that it resides on a higher or lower binary-choice indifference curve. The next lemma demonstrates that superiority or inferiority of $z$ is maintained in trinary choice:

**Lemma 2.** Given $x$ is indifferent to $y$ in binary choice, the following are equivalent:

i. $z$ is preferred to $x$ given $X = \{x, z\}$,

ii. $z$ is preferred to $y$ given $X = \{y, z\}$,

iii. $z$ is preferred to $x$ given $X = \{x, y, z\}$,

iv. $z$ is preferred to $y$ given $X = \{x, y, z\}$.

Therefore, given indifference in binary choice, $z$ is preferred to both $x$ and $y$ in trinary choice if and only if it would be preferred to each among $x$ and $y$ in binary choice. For this reason, we will simply refer to $z$ as being “superior” (or “inferior”) without differentiating between binary and trinary choice sets.

We now consider situations in which preferences are indeed context-dependent. To guide analysis, we define the indirect contrast function as follows:

$$\tilde{\Delta}(x_i, y_i; z_i) \equiv \left| v_i(x_i; z_i) - v_i(y_i; z_i) \right| = \left| \frac{x_i}{x_i + z_i} - \frac{y_i}{y_i + z_i} \right|.$$ 

The indirect contrast function measures the difference between the attribute values $x_i$ and $y_i$, after each has been normalized in relation to $z_i$. As with $\Delta(x_i, y_i)$, $\tilde{\Delta}(x_i, y_i; z_i)$ captures a form of perceptual contrast between $x_i$ and $y_i$, except it is generated by the separate comparisons of $x_i$ and $y_i$ to $z_i$ by means of pairwise attribute normalization.

**Proposition 3.** For any $z_i > 0$, $\tilde{\Delta}(x_i, y_i; z_i)$ is a salience function.

Thus the indirect contrast function $\tilde{\Delta}$, like the direct contrast function $\Delta$, also implements the defining properties of salience: ordering and diminishing sensitivity.

To highlight the dual roles of direct and indirect contrast in shaping context-dependent preferences, we can use (3) to re-write the condition for $x$ to be preferred to $y$ in trinary choice (i.e. $v(x; y, z) > v(y; x, z)$) as follows:

$$\Delta(x_1, y_1) + \tilde{\Delta}(x_1, y_1; z_1) > \Delta(x_2, y_2) + \tilde{\Delta}(x_2, y_2; z_2).$$

(4)

Thus, the consumer’s preferences can be equivalently captured by comparing the cumulative — direct plus indirect — contrast of $x$’s advantage over $y$ on $x$’s strong attribute to that of $y$’s advantage over $x$ on the other attribute.

Under our benchmark of binary indifference, $\Delta(x_1, y_1) = \Delta(x_2, y_2)$, therefore (4) reduces to a simple comparison of indirect contrasts. Adding $z$ to the choice set
will create a strict preference for \( x \) if and only if the indirect contrast on \( x \)'s strong attribute is greater than on \( y \)'s strong attribute. The next result establishes this feature, while offering another equivalent intuitive condition.

**Proposition 4.** Given \( v(x; y) = v(y; x) \), the following are equivalent:

i. \( v(x; y, z) > v(y; x, z) \) (i.e., \( x \) is preferred to \( y \) given \( X = \{x, y, z\} \)).

ii. \( \bar{\Delta}(x_1, y_1; z_1) > \bar{\Delta}(x_2, y_2; z_2) \).

iii. \( (x_1x_2 - z_1z_2)\left(\frac{z_1}{z_2} - \frac{x_1}{y_2}\right) > 0 \).

While the first two representations in Proposition 4 capture the link between context-dependent preferences and indirect salience discussed above, the third embeds two criteria that determine how the inclusion of \( z \) in the choice set sways preferences between \( x \) and \( y \): (a) whether \( z \) is inferior or superior to \( x \) and \( y \), as captured by the sign of \((x_1x_2 - z_1z_2)\); and (b) whether \( \frac{z_1}{z_2} \) is larger or smaller than \( \frac{x_1}{y_2} \). Figure 1 presents intuition for these conditions graphically. In particular, \( \frac{z_1}{z_2} - \frac{x_1}{y_2} = 0 \) defines a line from the origin. If \( \frac{z_1}{z_2} - \frac{x_1}{y_2} > 0 \), then \( z \) is more “similar” to \( x \) than to \( y \) in the sense that \( z \)'s attributes are tilted more towards \( x \)'s strong attribute than to \( y \)'s strong attribute, while \( \frac{z_1}{z_2} - \frac{x_1}{y_2} < 0 \) implies the opposite.\(^{22}\)

![Figure 1: The addition of \( z \) to the choice set induces a strict preference for \( x \) over \( y \) if \( z \) is located in either of the gray regions. It has the opposite effect if \( z \) is located in either of the white regions. The line \( z_2 = \frac{y_2}{x_2} z_1 \), derived from representation (III) in Proposition 4, serves as a boundary between these regions. If \( z \) is in the hatched region, it is superior to both \( x \) and \( y \).](image)

Therefore the presence of \( z \) in the choice set shifts preferences towards \( x \) and away from \( y \) if and only if either of the following conditions hold:

\(^{22}\)As a simple illustration, suppose \( x \) and \( y \) are symmetric (\( y_2 = x_1 \) and \( y_1 = x_2 \)). Then \( \frac{z_1}{z_2} - \frac{x_1}{y_2} > 0 \) holds if and only if \( z_1 > z_2 \), meaning \( z \)'s strength is the same as \( x \)'s strength. In Figure 1, this boundary is depicted by the 45-degree line \( z_1 = \frac{y_2}{x_2} z_2 \).
i. z is inferior and is more similar to x than to y,

ii. z is superior and is more similar to y than to x.

The first condition applies in the cases of the dominance and compromise effects, which are established by the following two corollaries.23

**Corollary 1 (Dominance Effect).** Given x is indifferent to y in binary choice, suppose x asymmetrically dominates z in that $x_1 \geq z_1 > y_1$ and $y_2 > x_2 \geq z_2$ with $x \neq z$. Then x will be preferred to y given $X = \{x, y, z\}$.

**Corollary 2 (Compromise Effect).** Given x is indifferent to y in binary choice, suppose x is a compromise between y and z in that $z_1 > x_1 > y_1$ and $y_2 > x_2 > z_2$. Then x will be preferred to y given $X = \{x, y, z\}$, provided x is also preferred to z.

Pairwise attribute normalization therefore provides a single mechanism through which these context effects can arise. In doing so, the theory imposes no artificial boundary on the decoy alternative z depending on whether its attribute $z_1$ is slightly lower or slightly higher than $x_1$. Instead, the dominance and compromise effects arise because z disproportionately (but continuously) amplifies the salience of x’s advantage relative to the salience of y’s advantage within the subregion of attribute space for which z is inferior and more similar to x than to y.

These effects can be traced to the manner in which the indirect contrast function $\tilde{\Delta}(x_i, y_i; z_i)$ depends on $z_i$. In particular, the extent to which $z_i$ amplifies the salience between $x_i$ and $y_i$ depends on its proximity to $x_i$ and $y_i$.

**Lemma 3.** $\tilde{\Delta}(x_i, y_i; z_i) > \tilde{\Delta}(x_i, y_i; z'_i)$ if and only if

$$\left| \frac{\ln(x_i) + \ln(y_i)}{2} - \ln(z_i) \right| < \left| \frac{\ln(x_i) + \ln(y_i)}{2} - \ln(z'_i) \right|.$$

Lemma 3 states that the indirect contrast between $x_i$ and $y_i$ is greater if $z_i$ is more “central”—in the sense of being closer to the mean of $x_i$ and $y_i$ in logarithmic space—than if $z_i$ is more “extreme.” From this, the intuition for the dominance and compromise effects arises naturally. If x asymmetrically dominates z, then z is more central on x’s strong attribute and more extreme on x’s weak attribute since $x_1 > z_1 > y_1$ and $z_2 < x_2 < y_2$. As a result, z generates greater salience for x’s advantage over y than for y’s advantage over x, which shifts the preference towards x.

The compromise effect can likewise be understood as a consequence of the fact that z is more central on x’s strong attribute and more extreme on x’s weak

---

23We will address the second condition shortly in Section 3.2.
attribute. However, since \( z_i \) lies extreme to \( x_i \) and \( y_i \) on both dimensions (i.e., \( z_1 > x_1 > y_1 \) and \( z_2 < x_2 < y_2 \)), the “centrality” of \( z_1 \) compared to \( z_2 \) is understood in terms of \( z \)'s inferiority to \( x \) and \( y \). The advantage of \( z_1 \) relative to \( x_1 \) and \( y_1 \) is not as large as the disadvantage of \( z_2 \) relative to \( x_2 \) and \( y_2 \), making \( z_1 \) relatively central and \( z_2 \) relatively extreme in comparison.\(^{24}\)

### 3.2 Strength of Context Effects

The magnitudes of the direct and indirect contrast functions, together, determine the preference in a trinary choice set by means of equation (4). However, they provide more than a simple description of the regions in which \( z \) will shift preference; they also provide a measure of where \( z \) must be located to shift any preference relation. This result can be equivalently interpreted as the strength of a decoy effect.\(^{25}\)

![Figure 2: The strength of decoy effects under pairwise normalization.](image)

Figure 2 depicts the difference in indirect contrast \( \tilde{\Delta}(x_1, y_1; z_1) - \tilde{\Delta}(x_2, y_2; z_2) \),

---

\(^{24}\)To illustrate this more precisely, suppose \( z \) instead exists on the same binary-choice indifference curve as \( x \) and \( y \) (so that it is neither inferior nor superior). In this case, we can use \( z_1, z_2 = x_1, x_2 = y_1, y_2 \) from the Cobb-Douglas representation (II) to derive \( \ln(z_1) - \frac{1}{2}(\ln(x_1) + \ln(y_1)) = \frac{1}{2}(\ln(x_2) + \ln(y_2)) - \ln(z_2) \), implying \( z_1 \) and \( z_2 \) are equally central in the sense of Proposition 3. From this point, \( z \) can be made inferior by decreasing \( z_1 \) or \( z_2 \) (or both). Since \( z_1 > x_1 > y_1 \) and \( z_2 < x_2 < y_2 \), decreasing \( z_1 \) brings it closer to \( x_1 \) and \( y_1 \) while decreasing \( z_2 \) causes it to be farther from \( x_2 \) and \( y_2 \). In both cases, \( z_1 \) becomes more central in relation to \( z_2 \).

\(^{25}\)The strength of decoy effects has an additional interpretation in the domain of stochastic choice. A common formulation of a stochastic choice model consists of some value function plus a stochastic error term (e.g., Loomes and Sugden, 1995). Therefore the likelihood of choosing an alternative from a choice set is a monotonic transformation of the difference in evaluations between alternatives, or in this case, \( v(x; y, z) - v(y; x, z) \).
for a range of decoy locations. Given $v(x; y) = v(y; x)$, this is equivalent to the difference in normalized valuations, $v(x; y, z) - v(y; x, z)$. Two results are immediately apparent. First, the difference in normalized valuations is larger in magnitude when $z$ is located further from the indifference curve. Second, this difference is larger when $z_1$ is located between $x_1$ and $y_1$. Proposition 5 provides a means to compare the strength of decoy effects by describing the ability of an alternative $z$ to reverse a strict binary preference. Specifically, if the presence of $z$ in a choice set alters a strict preference of $y$ over $x$ so that indifference is reached, then the context effect can be strengthened — creating a strict preference for $x$ over $y$ — either by changing $z_1$ to a more central attribute value $z'_1$ or by changing $z_2$ to a more extreme attribute value $z'_2$. 

**Proposition 5.** Suppose the consumer prefers $y$ to $x$ given $X = \{x, y\}$ but is indifferent between $y$ and $x$ given $X = \{x, y, z\}$ with $z = (z_1, z_2)$.

i. For $z' = (z'_1, z_2)$, $x$ is preferred to $y$ given $X = \{x, y, z'\}$ if and only if $z'_1$ is more central (less extreme) than $z_1$:

$$\left| \frac{\ln(x_1) + \ln(y_1)}{2} - \ln(z'_1) \right| < \left| \frac{\ln(x_1) + \ln(y_1)}{2} - \ln(z_1) \right|.$$ 

ii. For $z' = (z_1, z'_2)$, $x$ is preferred to $y$ given $X = \{x, y, z'\}$ if and only if $z'_2$ is more extreme (less central) than $z_2$:

$$\left| \frac{\ln(x_2) + \ln(y_2)}{2} - \ln(z'_2) \right| > \left| \frac{\ln(x_2) + \ln(y_2)}{2} - \ln(z_2) \right|.$$ 

A corollary to this result is that the dominance effect is stronger than the compromise effect.

**Corollary 3.** Suppose $y$ is preferred to $x$ given $X = \{x, y\}$. Let $z^d \equiv (z_1^d, z_2)$ be a decoy, and $z^c \equiv (z_1^c, z_2)$ be a compromise (i.e. $z_1^c > x_1 > z_1^d > y_1$ and $y_2 > x_2 > z_2$). Then, if the consumer is indifferent between $x$ and $y$ given $X = \{x, y, z^c\}$, the consumer will prefer $x$ to $y$ given $X = \{x, y, z^d\}$.

These predictions are consistent with existing results on the strength of decoy effects. Soltani, De Martino, and Camerer (2012) present subjects with a series of choices over lotteries (in which the two relevant attributes are probability and reward). Each subject first makes a series of binary choices and binary indifference curves are established, thereby controlling for risk preferences. Then each subject is presented with a series of three-lottery choice sets in which two of the lotteries lie near the binary indifference curve, with the location of the third lottery varied. This within-subject design demonstrates that the dominance effect is strongest...
when the decoy lottery \( z \) is shifted further from the binary indifference curve (Figure 3). Moreover, this result holds both when \( z \) is an inferior decoy lottery to \( x \), when \( z \) is an inferior decoy lottery to \( y \), and when \( z \) is superior to \( y \).

This final result is demonstrated by means of a phantom design. Just before the choice is made, the superior alternative \( z \) is removed from the choice set and the subject reveals a preference between the remaining two alternatives (\( x \) and \( y \)).

Recall that pairwise normalization predicts that a superior decoy alternative will alter the relative preference between the lower ranking alternatives in this manner (see Figure 2).\(^{26}\)

Figure 3: The strength of decoy effects, reproduced from Soltani, De Martino, and Camerer (2012). For each quadrant, decoy lotteries are median split into those lotteries lying “close” to the indifference curve and those lying “far”. The red and blue markers denote the decoy locations predicted to have the strongest effects, as described in Proposition 5 and Figure 2.

4 Preference Cycles

We now return to binary choice settings while allowing alternatives to vary on three dimensions. In contrast to preferences in two-attribute binary choice, we will show that preferences in three-attribute binary choice may no longer be transitive under pairwise attribute normalization. We should emphasize that our results do not imply that preferences must be intransitive. Instead, pairwise normalization predicts a particular form of intransitivity when the choice sets, or the subjective evaluation of attributes, meet certain conditions. We then describe empirical results which support these predictions.

To start, the following example demonstrates that pairwise normalization can generate preference cycles when the choice set has the property of “cyclical majority-

\(^{26}\)Note also that Soltani, De Martino, and Camerer (2012) find evidence that a decoy \( z \) located in region D5 decreases the likelihood that \( x \) is chosen, commonly referred to as a similarity effect. We will address the ability of pairwise normalization to capture this phenomenon in Section 7.3.
dominance”.

**Example 1.** Suppose \( x = (a, b, c), \ y = (b, c, a), \) and \( z = (c, a, b) \) with \( a > b > c \). Then we have a “majority rules” preference cycle whereby each option is preferred to that for which it is better on two out of three dimensions. That is, \( x \) is preferred to \( y \) given \( X = \{ x, y \} \), \( y \) is preferred to \( z \) given \( X = \{ y, z \} \), and \( z \) is preferred to \( x \) given \( X = \{ x, z \} \). To see this, we can compute

\[
v(x; y) - v(y; x) = v(y; z) - v(z; y) = v(z; x) - v(x; z)
\]

\[
= \frac{c - a}{a + c} + \frac{a - b}{a + b} + \frac{b - c}{b + c} = \frac{(a - b)(b - c)(a - c)}{(a + b)(b + c)(a + c)} > 0.
\]

When there is a cyclical majority-dominance relationship among three alternatives — i.e., without loss of generality, \( x \) is better than \( y \) on two of three attributes, \( y \) is better than \( z \) on two of three attributes, and \( z \) is better than \( x \) on two of three attributes — then preferences could, in principle, be intransitive in one of two ways: (i) a majority rules cycle (as in the example above); or (ii) a “minority rules” cycle that runs opposite to the majority rule (in our example, this would entail preferences for \( y \) over \( x \), \( z \) over \( y \), and \( x \) over \( z \)). The next result demonstrates that pairwise normalization cannot yield preference cycles that go against the majority rule.\(^{27}\)

**Proposition 6.** Consider three potential alternatives in which each is the best on exactly one of three attributes, second best on another, and worst on the remaining attribute. Then, if preferences among each pair of alternatives in binary choice are intransitive, it must be the case that for any two alternatives \( x \) and \( x' \), \( x \) will be preferred to \( x' \) given \( X = \{ x, x' \} \) if and only if \( x_i > x'_i \) for all but one \( i \in \{1, 2, 3\} \).

In the domain of subjective preferences, an empirical demonstration of Proposition 6 is challenging because it does not guarantee preferences among any three alternatives satisfying the majority-dominance criterion will in fact be intransitive. With heterogeneity in the subjective weighting of attributes, we would therefore only expect a subset of subjects to exhibit this pattern. However a clear hypothesis is still maintained. For those subjects who do exhibit intransitive preferences, we should expect to see a majority-rules cycle instead of a minority-rules cycle. The experiment by conducted by May (1954) detailed in Section 2 exhibits such a pattern.

However if attributes are constructed such that they are equally-valued across

\(^{27}\)Note, it is possible that pairwise normalization can yield intransitive preferences other than a cyclical majority-dominance relationship, provided that the choice set does not satisfy the premise of majority-dominance. Try, for example, \( x = (20, 2, 0), \ y = (2, 21, 0), \) and \( z = (2, 1, 10) \). In this case, pairwise attribute normalization implies \( z \) is preferred to \( y \), \( y \) is preferred to \( x \), and \( x \) is preferred to \( z \).
subjects, this would maximize the incidence of preference cycles. This is the approach taken in the experiment by Tsetsos, Moran, Moreland, Chater, Usher, and Summerfield (2016a), also detailed in Section 2. The choice behaviour of subjects was consistent with a systematic majority-rules preference cycle as predicted by pairwise normalization.

In an effort to replicate these results, we conducted a simple laboratory experiment in which 173 subjects were asked to make a sequence of three binary choices between holiday packages to Niagara Falls. Each alternative varied on three attribute dimensions (the hotel room, the dinner option, and the associated tour). The choices were fully incentivized, with each subject entered into a lottery, and the winning subject received the holiday package corresponding to one of their binary choices selected at random.\footnote{Examples of such lottery designs for incentive compatibility can be found in XX.}

The choice alternatives were constructed so that each alternative in a binary choice should have been dominant on two attributes (see Table 3). The total cost of each option was equivalent to C$350.\footnote{The monetary value of the dining and touring options were roughly equivalent, however the lodging component was higher. Therefore we could not completely control the degree to which different attributes were weighted.}

<table>
<thead>
<tr>
<th>Package A</th>
<th>Hotel</th>
<th>Dinner</th>
<th>Tour</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>King Room w/ view</td>
<td>$5 McDonald’s voucher</td>
<td>bus tour</td>
</tr>
<tr>
<td>Package B</td>
<td>Queen Room w/o view</td>
<td>dinner cruise for two</td>
<td>walking map</td>
</tr>
<tr>
<td>Package C</td>
<td>Motel w/in driving distance</td>
<td>$40 at local restaurant</td>
<td>helicopter tour</td>
</tr>
</tbody>
</table>

Of the 173 subjects, we found that 17, or 10% of subjects, displayed an intransitive cycle in their binary choices. This is a proportion of intransitivity typically observed in experimental data, though slightly lower than observed by May (1954). However of the 17 transitivity violations we observed, 13 of them (76.5%) were of the majority-rules form. This proportion is significantly greater than the proportion displaying the alternative minority-rules cycle ($p = 0.044$, two-tailed).

### 4.1 Subadditivity of $\Delta$ and Its Role in Preference Cycles

To provide some intuition for how majority-rules preference cycles arise in our model, we will now highlight a relevant feature of the direct contrast function $\Delta$, \textit{subadditivity}. Specifically, for any $a$, $b$, and $c$ such that $a > b > c \geq 0$, (strict) subadditivity of $\Delta$ implies that

\[
\Delta(a, b) + \Delta(b, c) > \Delta(a, c).
\]
If, as in Example 1, a, b, and c are (in some order) the attribute values for each of three potential alternatives, then the inequality shown above fully accounts for the preference cycle predicted by the model. That is, for any two of three alternatives in Example 1, one alternative will have a normalized advantage of \( \Delta(a, c) \) on one attribute while the other will have normalized advantages of \( \Delta(a, b) \) and \( \Delta(b, c) \). Even though the total unnormalized advantage(s) for each alternative are equal, \( a - c = (a - b) + (b - c) \), subadditivity ensures the two smaller advantages outweigh the larger advantage.

Subadditivity of within-attribute value comparisons (or of salience, depending on your interpretation) has not, to our knowledge, been previously implicated as a potential basis for majority-rules preference cycles in deterministic multi-attribute choice. However, Bar-Hillel and Margalit (1988) entertain the notion of subadditive sensitivity to losses as a potential basis for majority-rules preference cycles among lotteries.

5 Splitting Effects

We now examine the effect of attribute-splitting under pairwise attribute normalization, beginning from our benchmark of binary indifference between \( \mathbf{x} \) and \( \mathbf{y} \). We then consider decomposing an attribute into two sub-attributes, 1a and 1b, where for simplicity we assume that \( \mathbf{x} \)'s advantage on the original attribute is preserved on both sub-attributes.

**Proposition 7.** Given \( \mathbf{x} \) is indifferent to \( \mathbf{y} \) in binary choice, define \( \mathbf{x}' \equiv (x'_a, x'_b, x_2) \) and \( \mathbf{y}' \equiv (y_1, y_1, y_2) \) 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 \( \mathbf{x}' \) will be preferred to \( \mathbf{y}' \) given \( X = \{ \mathbf{x}', \mathbf{y}' \} \).

Thus, splitting an attribute into two sub-attributes strengthens preferences for the alternative that was stronger on that initial attribute, provided the advantage is maintained on each of the two sub-attributes. This prediction can be understood as the joint byproduct of diminishing sensitivity and subadditivity of \( \Delta \). To see this, let \( \delta^a = x_{1a} - y_{1a} \) and \( \delta^b = x_{1b} - y_{1b} \) represent the unnormalized advantages for \( \mathbf{x} \) on each subattribute. Noting \( x_1 - y_1 = \delta^a + \delta^b \), subadditivity of \( \Delta \) implies

\[
\Delta(y_1 + \delta^a, y_1) + \Delta(x_1, y_1 + \delta^a + \delta^b, y_1 + \delta^a) > \Delta(x_1, y_1).
\]

Since \( y_1 \geq y_{1a} \) and \( y_1 + \delta^a > y_{1b} \), diminishing sensitivity implies

\[
\Delta(y_{1a} + \delta^a, y_{1a}) > \Delta(y_1 + \delta^a, y_1), \quad \text{and} \quad \Delta(y_{1b} + \delta^b, y_{1b}) > \Delta(y_1 + \delta^a + \delta^b, y_1 + \delta^a).
\]
Combining these conditions with the previous inequality, we see the subattributes must generate greater total contrast than the higher-level attribute from which they are split:

$$\Delta(x_{1a}, y_{1a}) + \Delta(x_{1b}, y_{1b}) > \Delta(x_1, y_1).$$

Therefore subadditivity and diminishing sensitivity along each attribute dimension combine to generate attribute splitting effects.

### 6 Comparability Biases

Next, we consider how preferences might depend on whether or not alternatives are comparable on each dimension. In particular, we consider a binary choice between \(x\) and \(y\) in which an attribute \(i\) is comparable if it belongs to both alternatives (i.e. \(x_i > 0\) and \(y_i > 0\)). If, instead, \(x_i > 0\) and \(y_i = 0\) (or the reverse), \(i\) is regarded as a non-comparable attribute because it is only present for one alternative.

To isolate the effect of comparability on preferences, we will again work from our benchmark of binary indifference, while presuming \(x\) and \(y\) are comparable on both dimensions in that \(x_1, x_2, y_1, y_2\) are all nonzero. We then compare this benchmark to preferences among two modified alternatives that have the same attribute-values as \(x\) and \(y\) (respectively), but are no longer comparable on both attributes. That is, without loss of generality, \(x' = (x_1, x_2, 0)\) and \(y' = (y_1, 0, y_2)\):

**Proposition 8.** Given \(x\) is indifferent to \(y\) in binary choice, define \(x' = (x_1, x_2, 0)\) and \(y' = (y_1, 0, y_2)\). Then \(x'\) is preferred to \(y'\) given \(X = \{x', y'\}\).

Thus, preferences are stronger for an alternative if it has an advantage (in this case, \(y_2 > x_2\)) that is comparable than it would be if that same advantage was non-comparable, ceteris paribus. Or put differently, normalization effectively generates a greater weighting of comparable attributes relative to non-comparable attributes.

This result can be understood as another consequence of the subadditivity of \(\Delta\). To see this, first note that subadditivity implies \(\Delta(0, x_2) + \Delta(x_2, y_2) > \Delta(0, y_2)\), which we re-write as

$$\Delta(x_2, y_2) > \Delta(0, y_2) - \Delta(0, x_2).$$

The left-side of this inequality reflects the salience of \(y\)'s advantage over \(x\), \(y_2 > x_2\), in the event that \(x_1\) and \(x_2\) are comparable. The right-side captures the salience of \(y\)'s advantage in the event that \(x_1\) and \(x_2\) are non-comparable. Thus, a lack of comparability between \(y_2\) and \(x_2\) diminishes their effective difference under pairwise attribute normalization, which weakens the consumer’s preference for \(y\) relative to \(x\).
7 Parametric Generalizations

In this section, we consider a two-parameter version of the pairwise attribute normalization model. This generalization is based on a parametric form of the single-dimension normalization equation studied in the neuroscience literature (Carandini and Heeger, 2012), recently adapted to the study of risk preferences by Tymula and Glimcher (2016). Similar to our construction of the basic model, we adapt this parametric normalization computation to a multi-attribute choice environment by means of pairwise, attribute-level comparisons.

Formally, the component of the normalized value of \( x \in X \) arising from a single comparison to some \( y \in C_x \) on attribute \( i \) is now given by
\[
v_i^*(x_i; y_i) = \frac{x_i^\alpha}{\sigma^\alpha + x_i^\alpha + y_i^\alpha}.
\]

This formulation introduces two new parameters: the semi-saturation constant \( \sigma \geq 0 \) and the rectification constant \( \alpha > 0 \) (intuition for the behavioral relevance of each parameter will be provided shortly). Our basic normalization computation
\[
v_i(x_i, y_i) = x_i^\alpha x_i^\alpha + y_i^\alpha
\]
is therefore subsumed as a special case of \( v_i^*(x_i; y_i) \) with \( \alpha = 1 \) and \( \sigma = 0 \).

The overall valuation of an alternative in the two-parameter model is then (as before) an aggregation of the normalized valuations corresponding to each possible comparison:
\[
v^*(x; C_x) \equiv \sum_{y \in C_x} \sum_{i=1}^n \frac{x_i^\alpha}{\sigma^\alpha + x_i^\alpha + y_i^\alpha}.
\] (5)

Our first lemma establishes the relevant condition for which \( x \) will be preferred to \( y \) in two-attribute binary choice:

**Lemma 4.** In the two-parameter model given by (5) with \( X = \{x, y\} \) and \( n = 2 \), \( x \) is preferred to \( y \) if and only if
\[
\sigma^\alpha (x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha > \sigma^\alpha (y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha.
\] (6)

This condition implies that the generalized model nests, as special cases, several classic preference representations.

**Proposition 9.** Let \( X = \{x, y\} \) and \( n = 2 \). In each of the following cases and given \( u(x) \neq u(y) \), \( x \) is preferred to \( y \) if and only if \( u(x) > u(y) \):

(I) **Cobb-Douglas:** \( \sigma = 0 \) and any \( \alpha > 0 \), with \( u(a, b) \equiv ab \).

(II) **CES:** \( \sigma > 0 \) sufficiently large and any \( \alpha > 0 \), with \( u(a, b) \equiv (a^\alpha + b^\alpha)^{1/\alpha} \).

(III) **Linear:** \( \sigma > 0 \) sufficiently large and \( \alpha = 1 \), with \( u(a, b) \equiv a + b \).

(IV) **Rank-Based Lexicographic:** \( \sigma > 0 \) and \( \alpha > 0 \) both sufficiently large, with \( u(a, b) \equiv \max\{a, b\} \). In the event that \( u(x) = u(y) \), \( x \) is then preferred to
Proposition 9.I demonstrates that the Cobb-Douglas equivalence in the basic model with \( \sigma = 0 \) and \( \alpha = 1 \) (Proposition 2) extends to any \( \alpha > 0 \). From 9.II, we see that preferences effectively converge to those given by a CES (constant elasticity of substitution) utility function in the large-\( \sigma \) limit, where \((1 - \alpha)^{-1}\) represents the elasticity of substitution between \( x_1 \) and \( x_2 \). As a special case of CES preferences, 9.III demonstrates that preferences converge to linear additive utility for \( \sigma \to \infty \) in the event that \( \alpha = 1 \), in which case \( x_1 \) and \( x_2 \) are perfect substitutes.

In general, binary-choice preferences in our two-parameter model (with \( n = 2 \)) are a composite of preferences under the Cobb-Douglas and CES representations. If \( u(x) \geq u(y) \) holds for both the Cobb-Douglas and CES utility functions (with at least one inequality non-binding), then \( v^*(x; y) > v^*(y; x) \), meaning \( x \) is necessarily preferred to \( y \). In the event that the Cobb-Douglas and CES utility representations disagree as to which among \( x \) and \( y \) brings greater utility, increasing \( \sigma \) increases the effective weight of CES utility (relative to Cobb-Douglas) in determining preferences in our model. That is, in such cases, there is a threshold \( \sigma' \) such that preferences will coincide with Cobb-Douglas if \( \sigma < \sigma' \) and with CES if \( \sigma > \sigma' \).

In the case of \( \alpha = 1 \), CES utility is additive and separable, which means \( \sigma \) then effectively parameterizes the separability of attributes in the consumer’s valuation. That is, increasing \( \sigma \) makes \( x_1 \) and \( x_2 \) are more separable in the consumer’s valuation of \( x \), while decreasing \( \sigma \) has the opposite effect.

Representation (IV) in Proposition 9 demonstrates how preferences in two-attribute binary choice effectively converge with those given by a rank-based lexicographic model for the case in which both parameters are arbitrarily large. That is to say, in this case, preferences can be described as follows:

- The consumer’s preference among \( x \) and \( y \) is determined on the basis of which alternative’s highest attribute value is greater. That is, \( x \) will be preferred if \( \max\{x_1, x_2\} > \max\{y_1, y_2\} \) while \( y \) will be preferred if \( \max\{x_1, x_2\} < \max\{y_1, y_2\} \).
- In the event that \( x \) and \( y \)’s highest attribute values are the same, the consumer’s preference is then determined on the basis of which alternative is better on its weaker attribute. That is, given \( \max\{x_1, x_2\} = \max\{y_1, y_2\} \), \( x \) will be preferred if \( \min\{x_1, x_2\} > \min\{y_1, y_2\} \) while \( y \) will be preferred if \( \min\{x_1, x_2\} < \min\{y_1, y_2\} \).

Comparing the CES and rank-based lexicographic preferences helps illustrate a
more general role for \( \alpha \), whereby increasing \( \alpha \) leads the consumer to base binary preferences more on a comparison of their strengths and less on a comparison of their weaknesses (provided \( \sigma > 0 \), since \( \alpha \) has no effect on preferences with \( \sigma = 0 \) in light of 9.1.

### 7.1 Contrast and Salience in the Two-Parameter Model

Next, we explore how \( \sigma \) and \( \alpha \) can affect the relationship between pairwise attribute normalization and notions of salience. To do this, we define the relevant direct contrast function under (5) as

\[
\Delta^*(x_i, y_i) \equiv |v_i^*(x_i; y_i) - v_i^*(y_i; x_i)| = \left| \frac{x_i^\alpha - y_i^\alpha}{\sigma^\alpha + x_i^\alpha + y_i^\alpha} \right|.
\]

**Proposition 10.** \( \Delta^*(x_i, y_i) \) is a salience function if and only if \( \sigma = 0 \) or \( \alpha \leq 1 \) (or both).

Thus, the direct contrast function \( \Delta^* \) preserves its status as a salience function when generalizing one of the two implicit parametric restrictions (\( \sigma = 0 \) or \( \alpha = 1 \)) from the basic model, or when generalizing both with \( \alpha < 1 \). If \( \sigma > 0 \) and \( \alpha > 1 \), however, \( \Delta^* \) is no longer a salience function because it does not exhibit diminishing sensitivity over its full domain. Proposition 11 examines this case in greater detail, while highlighting the importance of \( \sigma \) in determining the effective bounds of diminishing sensitivity.

**Proposition 11.** Given \( \sigma > 0 \), \( \alpha > 1 \), and, without loss of generality, \( x_i \geq y_i \).

(i) Holding \( y_i \) fixed, the following conditions are equivalent:

a. \( \Delta^*(x_i, y_i) \) satisfies diminishing sensitivity for all \( x_i \geq y_i \).

b. \( \Delta^*(x_i, y_i) \) is concave in \( x_i \) for all \( x_i \geq y_i \).

c. \( \sigma \leq \hat{\sigma}(y_i) \), where

\[
\hat{\sigma}(y_i) \equiv \left( \frac{2}{\alpha - 1} \right)^{1/\alpha} y_i.
\]

(ii) For \( \sigma > \hat{\sigma}(y_i) \), the following statements are true:

a. \( \Delta^*(x_i, y_i) \) exhibits increasing sensitivity if \( y_i < x_i < \bar{x}(\sigma, y_i) \) and decreasing sensitivity if \( x_i > \bar{x}(\sigma, y_i) \), where

\[
\bar{x}(\sigma, y_i) \equiv \left\{ x_i : \sigma^\alpha = \frac{2(x_i - y_i)}{x_i^{1-\alpha} - y_i^{1-\alpha}} \right\} > y_i.
\]

b. \( \Delta^*(x_i, y_i) \) is strictly convex in \( x_i \) for all \( x_i \) satisfying \( y_i < x_i < \bar{x}(\sigma, y_i) \) and
strictly concave in \( x_i \) for all \( x_i > \bar{x}(\sigma, y_i) \), where

\[
\bar{x}(\sigma, y_i) \equiv \left( \frac{(\alpha - 1)(y^{\alpha} + \sigma^{\alpha})}{\alpha + 1} \right)^{1/\alpha} > y_i.
\]

c. \( \bar{x}(\sigma, y_i) \) and \( \hat{x}(\sigma, y_i) \) are increasing in both arguments.

To convey key features of \( \Delta^*(x_i, y_i) \) with \( \sigma > 0 \) and \( \alpha > 1 \), Proposition 11 effectively fixes the smaller attribute value, taken here to be \( y_i \) without loss of generality, while implicitly allowing the larger attribute value \( x_i \) to vary. Of particular relevance, part (i) shows that \( \Delta^*(x_i, y_i) \) satisfies diminishing sensitivity for all \( x_i \geq y_i \) if and only if \( \sigma \) is sufficiently small (in relation to \( y_i \)), as described by \( \sigma < \hat{\sigma}(y_i) \). The proposition also reveals an equivalence between the preservation of diminishing sensitivity and the shape of \( \Delta^*(x_i, y_i) \) as it relates to \( x_i \) — and a dual role for \( \hat{\sigma}(y_i) \) in determining both properties — in the sense that \( \Delta^*(x_i, y_i) \) is strictly concave in \( x_i \) (for all \( x_i \geq y_i \)) if and only if \( \sigma < \hat{\sigma}(y_i) \).

![Figure 4: Parametrization of the direct contrast function for \( x_i > y_i \): varying \( \sigma \) while holding \( \alpha > 1 \) (Left), and varying \( \alpha \) while holding \( \sigma > 1 \) (Right).](image)

Part (ii) of Proposition 11 describes the properties of \( \Delta^*(x_i, y_i) \) when \( \sigma > \hat{\sigma}(y_i) \). In particular, \( \Delta^*(x_i, y_i) \) exhibits increasing sensitivity in the region \( y_i < x_i < \bar{x}(\sigma, y_i) \); for \( x_i > \bar{x}(\sigma, y_i) \), diminishing sensitivity is restored. Similarly, \( \Delta^*(x_i, y_i) \) is convex in \( x_i \) when \( y_i < x_i < \hat{x}(\sigma, y_i) \) and concave in \( x_i \) for \( x_i > \hat{x}(\sigma, y_i) \). Hence, the value \( \bar{x}(\sigma, y_i) \) represents the threshold between increasing sensitivity and diminishing sensitivity while \( \hat{x}(\sigma, y_i) \) represents the cutoff between the convex region and the concave region of \( \Delta^*(x_i, y_i) \). In light of the fact that convexity implies increasing responsiveness to \( x_i \) and concavity implies the opposite, the threshold
\(\hat{x}(\sigma, y_i)\) can also be understood as the point at which \(\Delta^*(x_i, y_i)\) is most responsive to changes in \(x_i\) (as measured by \(\partial\Delta^*(x_i, y_i)/\partial x_i > 0\)).

We can now interpret the role of the parameters in shaping relative attribute-level valuations in our framework. The magnitude of the semi-saturation constant \(\sigma\) determines the effective size of the increasing sensitivity and convexity regions. In particular, \(\hat{x}(\sigma, y_i)\) and \(\hat{x}(\sigma, y_i)\) are both increasing in \(\sigma\) (item c, part ii of Proposition 11), which means for larger values of \(\sigma\), \(x_i\) must be farther from \(y_i\) before diminishing sensitivity kicks in, and the point at which \(\Delta^*(x_i, y_i)\) is most responsive to changes in \(x_i\) will likewise shift to the right.  

In addition to its role in permitting the increasing sensitivity and convexity regions, the rectification constant \(\alpha > 1\) determines the extent to which the responsiveness of \(\Delta^*(x_i, y_i)\) is concentrated over a small range of \(x_i\), as opposed to being dispersed over a large range. That is, as \(\alpha > 1\) increases, \(\Delta^*(x_i, y_i)\) becomes more responsive to changes in \(x_i\) when \(x_i\) is near its point of maximum responsiveness \(\hat{x}(\sigma, y_i)\), but becomes less responsive when \(x_i\) is far from \(\hat{x}(\sigma, y_i)\). For example, with \(\alpha = 1\), the responsiveness of \(\Delta^*(x_i, y_i)\) to \(x_i\) is relatively dispersed over a wide range of values; in the limit as \(\alpha \to \infty\), however, \(\Delta^*(x_i, y_i)\) assumes the shape of a step-function in \(x_i\), in which case \(\Delta^*(x_i, y_i)\) is infinitely responsive at its point of maximum responsiveness but is unresponsive to changes in \(x_i\) everywhere else.

### 7.2 Robustness

We now establish conditions under which our main behavioral effects, in their exact forms as originally-presented in the parameter-free model, are robust to our parametric generalizations:

**Proposition 12.** In the two-parameter model given by (5):

(i) The dominance and compromise effects (Corollaries 1-2) hold for any \(\alpha > 0\) given \(\sigma = 0\).

(ii) The attribute-splitting effect (Proposition 7) holds as long as \(\sigma = 0\) or \(0 < \alpha \leq 1\) (or both).

(iii) The majority-rules intransitivity (Proposition 6) and comparability bias (Proposition 8) hold for all \(\sigma \geq 0\) and \(\alpha > 0\).

Thus, Proposition 12 establishes the robustness of our behavioral predictions for all but the following two cases:

---

\(^{30}\)While \(\sigma\) has been typically treated as a constant in the neuroscience literature (e.g. Shevell, 1977; Heeger, 1992; Louie, Grattan, and Glimcher, 2011), recent work suggests that \(\sigma\) arises dynamically in neural systems from the history of stimuli (LoFaro, Louie, Webb, and Glimcher, 2014; Louie, LoFaro, Webb, and Glimcher, 2014). See Tymula and Glimcher (2016) for an application of this formulation in the domain of risk preferences, in which \(\sigma\) serves as a dynamic reference point of past lotteries.
1. The attribute-splitting effect with $\alpha > 1$ and $\sigma > 0$.

2. The dominance and compromise effects with $\sigma > 0$.

The lack of an attribute-splitting effect with $\alpha > 1$ and $\sigma > 0$ can be understood as a consequence of the fact that $\Delta^*$ does not universally satisfy diminishing sensitivity under this parameterization. Intuitively, the presence of an increasing sensitivity region permits a higher normalized valuation when two attributes are combined into one, with greater gains from combining larger as opposed to smaller attributes. Thus, combining attributes will expand an alternative’s normalized advantage with respect to those attributes, or equivalently, reducing the advantage on the composite attribute by splitting it into its subattributes — the opposite of our effect in the basic model with diminishing sensitivity.

The robustness of the dominance and compromise effects with $\sigma > 0$, and the ability of this general model to capture an additional behavioral phenomenon known as a “similarity effect,” is addressed in the next subsection.

7.3 Context Effects with $\sigma > 0$

To analyze context effects in the two-parameter normalization model, we will again work from a benchmark assumption that $x$ and $y$ lie on the same binary-choice indifference curve, so that a strict preference between $x$ and $y$ in trinary choice can equivalently be understood as a context effect induced by the presence of the third alternative $z$ in the choice set. Our analysis will, to a large degree, focus on the critical role of $\sigma$ in determining the nature of such context effects. Of particular interest, we will see how the inclusion of a given $z$ in a choice set with $x$ and $y$ may generate a strict preference for $x$ over $y$ for some values of $\sigma$ yet generate the reverse preference for other values of $\sigma$. In expressing such results, however, there is an added complication with $\sigma \geq 0$. Namely, if $x$ and $y$ are on the same binary-choice indifference curve given $\sigma \geq 0$, changing $\sigma$ will in general break this indifference, making it unclear as to whether an accompanying change in the preference between $x$ and $y$ in trinary choice is due to $\sigma$’s potential role in shaping context-dependent preferences, or due to its effect on binary-choice preferences. Therefore, for ease of exposition, we adopt a simpler benchmark in which the consumer’s indifference between $x$ and $y$ in binary choice is preserved even as $\sigma$ varies.

**Assumption 1.** $v^*(x; y|\sigma) = v^*(y; x|\sigma)$ for all $\sigma \geq 0$, or equivalently, $x$ and $y$ are symmetric in that $x = (a, b)$ and $y = (b, a)$ for some $a > 0$ and $b > 0$.

Given this benchmark binary-choice indifference assumption, the next lemma addresses the the degree to which a preference among $x$ and $y$ in trinary choice
may vary with $\sigma$ (and, to a lesser extent, $\alpha$).

**Lemma 5.** Given Assumption 1, $X = \{x, y, z\}$, and any $\alpha > 0$. If $x$ is preferred to $y$ with $\sigma = 0$, then $x$ is preferred to $y$ for $\sigma > 0$ unless

$$z_1^\alpha z_2^\alpha < x_1^\alpha x_2^\alpha < (z_1^\alpha + \sigma^\alpha)(z_2^\alpha + \sigma^\alpha). \quad (7)$$

From (7), we can see that introducing $\sigma > 0$ into the model can only undo a strict, context-induced preference for $x$ over $y$ in the event that $x$ is modestly superior to $z$ in the sense that, taking $\alpha = 1$ for illustrative purposes, $x = (x_1, x_2)$ is preferred to $z = (z_1, z_2)$, but would not preferred to a hypothetical $z' = (z_1 + \sigma, z_2 + \sigma)$ reflecting an improvement to $z$ of magnitude $\sigma > 0$ on each dimension.

The following corollary relates this discussion to our main context effects of interest with $\sigma > 0$.

**Corollary 4.** Given $X = \{x, y, z\}$ and Assumption 1, suppose $z$ either satisfies the conditions required for a dominance effect (Corollary 1) or for a compromise effect (Corollary 2) with $\sigma = 0$. Then for $\sigma > 0$, $x$ will be preferred to $y$ (i.e., the corresponding context effect will still hold) as long as $z$ is sufficiently “inferior” in that $(z_1^\alpha + \sigma^\alpha)(z_2^\alpha + \sigma^\alpha) < x_1^\alpha x_2^\alpha$. If $z$ is not sufficiently inferior in that $(z_1^\alpha + \sigma^\alpha)(z_2^\alpha + \sigma^\alpha) > x_1^\alpha x_2^\alpha$, then $y$ will be preferred to $x$.

![Figure 5: A similarity effect predicted by pairwise normalization. When $\sigma > 0$, nearby locations of a decoy $z$ (in all directions) shift preference towards $y$ from $x$.](image)

With $\sigma = 0$, the dominance and compromise effects were assured as long as $z$ was inferior to $x$ and $y$ (i.e., $z_1 z_2 < x_1 x_2 = y_1 y_2$). With $\sigma > 0$, however, Corollary 4 states that if $z$ is only modestly inferior to $x$ and $y$, its inclusion in the choice set can have the opposite effect, creating a preference for $y$ instead of $x$ in trinary.
choice (see Figure 5). In effect, $\sigma$ shrinks the region in which the dominance and compromise effects arises (Figure 6).

Also note, with $\sigma > 0$, $x$ no longer borders the region for which the inclusion of $z$ in the choice set shifts preferences towards $x$. Consequently, any $z$ within a sufficiently small neighborhood around $x$ will now induce a strict preference for $y$ over $x$. Empirical results displaying this property have been previously termed “similarity effects” whereby adding a third alternative to a choice set that is similar to one of the existing alternatives, in the sense that it is located nearby in attribute space, can shift preferences towards the dissimilar alternative (e.g. Tversky, 1972).

The nature of context effects under pairwise attribute normalization with $\sigma > 0$ provides a lens through which we can understand findings from the Soltani, De Martino, and Camerer (2012) experiment (Figure 3) first discussed in Section 3.2. In this experiment, a set of decoy lotteries lay within a region nearby, but not necessarily preferred to, the target lottery (Region D5). The authors find that decoys in this region shift preference to the opposite lottery, and this effect is stronger when the decoy lottery is further away. Pairwise attribute normalization is consistent with this result when $\sigma > 0$, in addition to maintaining the predictions for Regions D1-D6 (see Figure 6).

Finally, we address the ability of the more general formulation of pairwise attribute normalization to address the magnitude of context effects. In the baseline model ($\sigma = 0$, $\alpha = 1$), the dominance effect is stronger for a $z$ which is inferior (dominated) than for a $z$ which is superior (dominant). For example, compare the regions above the binary indifference curve in Figure 2 to the regions below. Soltani, De Martino, and Camerer (2012) find the magnitude of these effects to be roughly equivalent between regions D1 and D6. The general formulation can capture this result by varying the strength of these regions in two ways: via the semi-saturation constant $\sigma$ (Corollary 5), and via the rectification constant $\alpha$ (Figure 7).

**Corollary 5.** Under Assumption 1, suppose $z_2 < x_2 < y_2 < z'_2$ and $y_1 < z_1 < x_1$ so that $z_L \equiv (z_1, z_2)$ is asymmetrically dominated by $x$ while $z_H \equiv (z_1, z'_2)$ strictly dominates $y$ (but not $x$). For $\tilde{\sigma}_0 \equiv \left[ \sqrt{\left( \frac{z_1 - z_2}{2} \right)^2 + x_1^\alpha x_2^\alpha - \frac{z_1^2 + z_2^2}{2}} \right]^{1/\alpha}$:

- If $\sigma < \tilde{\sigma}_0$, $x$ is preferred to $y$ for both $X = \{x, y, z_L\}$ and $X = \{x, y, z_H\}$.
- If $\sigma > \tilde{\sigma}_0$, $y$ is preferred to $x$ for $X = \{x, y, z_L\}$, while $x$ is (still) preferred to $y$ for $X = \{x, y, z_H\}$.

Hence, increasing $\sigma > 0$ simultaneously weakens the dominance effect for dominated $z$ and strengthens it for $z$ which are dominant to $y$. In fact, the former may
vanish while the latter remains. In contrast, the rectification parameter $\alpha$ can vary the magnitude of these effects without requiring a change in the size of the regions inducing the dominance, compromise, and similarity effects.

Figure 6: Changing $\sigma$.

Figure 7: $\alpha > 1, \sigma = 0$ (Left). $\alpha < 1, \sigma > 0$ (Centre). $\alpha > 1, \sigma > 0$ (Right).
8 Conclusions

We propose a parsimonious model of multi-attribute choice which accounts for a wide variety of choice phenomena observed in the empirical literature, including the dominance and compromise effects, majority-rules intransitivity cycles, sub-additivity of attributes, and comparability biases. The model is grounded in the neurobiological principles of perception, and takes as a primitive the neurobiological mechanism for how sensory information is compared in a relative formulation. It incorporates a measure for the magnitude of perceived attributes, which we term a contrast function, and we demonstrate that this function implements a neurobiological mechanism for the salience concept introduced by Bordalo, Gennaioli, and Shleifer (2012, 2013). In addition to the range of multi-attribute choice phenomena captured, the model also provides predictions for the strength of these effects that are borne out in empirical studies. Finally, we demonstrate that a general formulation of the model, derived from the functional form typically studied in the neuroscience literature, robustly captures the multi-attribute choice phenomena.
References


CARANDINI, M., D. J. HEEGER, AND J. A. MOVSHON (1997): “Linearity and
Normalization in Simple Cells of the Macaque Primary Visual Cortex,” The
Journal of Neuroscience, 17(21), 8621–8644.

involved in the computation of subjective value,” Social cognitive and affective
neuroscience, pp. 1–14.


“Reported violations of rationality may be aggregation artifacts,” Proceedings
of the National Academy of Sciences, 113(33), E4761–3.

rationale for the attraction and compromise effects,” Theoretical Economics, 7(1),
125–162.

DOYLE, J. R., D. J. O’CONNOR, G. M. REYNOLDS, AND P. A. BOTTOMLEY
(1999): “The robustness of the asymmetrically dominated effect: Buying frames,
phantom alternatives, and in-store purchases,” Psychology & Marketing, 16(3),
225–243.

FELLOWS, L. K. (2006): “Deciding how to decide: ventromedial frontal lobe dam-

FISCHHOFF, B., P. SLOVIC, AND S. LICHTENSTEIN (1978): “Fault trees: Sensi-
tivity of estimated failure probabilities to problem representation.,” Journal of


GOURVILLE, J. T., AND D. SOMAN (2005): “Overchoice and Assortment Type:

environmental decision analysis,” Ecological Economics, 68(1-2), 556–569.


Noguchi, T., and N. Stewart (2014): “In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions,” *Cognition*, 132(1), 44–56.


40
A Appendix

A.1 Proof of Proposition 1

Ordering. Without loss of generality, suppose $x_i > y_i$. Then $\partial \Delta(x_i, y_i) / \partial x_i = 2y_i/(x_i + y_i)^2 > 0$ and $\partial \Delta(x_i, y_i) / \partial y_i = -2x_i/(x_i + y_i)^2 < 0$. Therefore, $x'_i \geq x_i$ and $y'_i \leq y_i$, with at most one of these inequalities binding, imply $\Delta(x'_i, y'_i) > \Delta(x_i, y_i)$.

Diminishing sensitivity. With $x_i > y_i$ and $\epsilon > 0$, $\Delta(x_i + \epsilon, y_i + \epsilon) = (x_i - y_i)/(x_i + y_i + 2\epsilon) < (x_i - y_i)/(x_i + y_i) = \Delta(x_i, y_i)$.

\[\Box\]

A.2 Proof of Proposition 2

By definition, $v(x; y) > v(y; x)$ holds if and only if

\[
\frac{x_1}{x_1 + y_1} + \frac{x_2}{x_2 + y_2} > \frac{y_1}{x_1 + y_1} \frac{y_2}{x_2 + y_2}.
\]  

(8)

Rearranging (8), we can see that it is equivalent to $\frac{x_1 - y_1}{x_1 + y_1} > \frac{y_2 - x_2}{x_2 + y_2}$, which itself is equivalent to $\Delta(x_1, y_1) > \Delta(x_2, y_2)$.

Multiplying both sides of (8) by $(x_1 + y_1)(x_2 + y_2) > 0$ and rearranging, we can see that it is also equivalent to $x_1x_2 > y_1y_2$.

Multiplying both sides of (8) by $\frac{(x_1 + y_1)(x_2 + y_2)}{x_1 + y_1 + x_2 + y_2} > 0$, we see that it is also equivalent to $\theta_1x_1 + \theta_2x_2 > \theta_1y_1 + \theta_2y_2$ with $\theta_1 = \frac{x_1 + y_1}{\sum_i x_i + y_i}$ and $\theta_2 = \frac{x_2 + y_2}{\sum_i x_i + y_i}$.

\[\Box\]

A.3 Proof of Lemma 1

We first show that $v(x; y) > v(y; x)$ implies $v(x; y, z) > v(y; x, z)$. If $x$ and $z$ are on the same binary-choice indifference curve, $v(x; z) = v(z; x) = 1$ and $v(y; z) < 1 < v(z; y)$; if $y$ and $z$ are on the same binary-choice indifference curve, $v(y; z) = v(z; y) = 1$ and $v(x; z) > 1 > v(z; x)$. In either case, $v(x; z) > v(y; z)$. Given $v(x; y) v(y > x)$, this implies $v(x; y, z) = v(x; y) + v(x; z) > v(y; x) + v(y; z) = v(y; x, z)$, as desired.

We now show the converse, $v(x; y, z) > v(y; x, z)$ implies $v(x; y) > v(y; x)$, and proceed by contradiction. Now $v(x; y, z) > v(y; x, z)$ implies $v(x; y) + v(x; z) > v(y; x) + v(y; z)$. It must then be the case that $v(x; z) > v(y; z)$ since $v(y; x) \geq v(x; y)$ (by assumption).

If $x$ and $z$ are on the same binary-choice indifference curve, $v(x; z) = v(z; x) = 1$ and $v(y; z) < 1 < v(z; y)$; if $y$ and $z$ are on the same binary-choice indifference curve, $v(y; z) = v(z; y) = 1$ and $v(x; z) > 1 > v(z; x)$. In either case, we can
use (C-D) from Proposition 2 to obtain \( x_1 x_2 > y_1 y_2 \), which contradicts \( v(y; x) \geq v(x; y) \).

If Assumption ?? also applies, \( x_1 x_2 = y_1 y_2 \) and \( z_1 z_2 \in \{ x_1 x_2, y_1 y_2 \} \). Thus, \( x_1 x_2 = y_1 y_2 = z_1 z_2 \). This means \( v(x; y) = v(y; x) = v(x; z) = v(y; z) = v(z; y) = v(z; x) = v(z; y) \) = 1. In turn, this means \( v(x; y, z) = v(x; y) + v(x; z) = v(y; x, z) = v(y; x) + v(y; z) = v(z; x, y) = v(z; x) + v(z; y) \), as desired. ■

A.4 Proof of Lemma 2

From Proposition 2, we know that \( v(z; x) > v(x; z) \) (i-a) is equivalent to \( z_1 z_2 > x_1 x_2 \) and with Assumption ?? we know \( x_1 x_2 = y_1 y_2 \). Taken together, it follows that (i-a) is equivalent to \( v(z; y) > v(y; z) \) (i-b).

Next, let \( x', y', \text{ and } z' \) denote the corresponding alternatives as labeled in Lemma 1. Taking \( z = x', x = z' \), and \( y = y' \), it then follows from Lemma 1 that (i-b) is equivalent to (ii-b). If we instead take \( z = x' \), \( x = y' \), and \( y = z' \), Lemma 1 establishes equivalence of (i-a) and (ii-a). This completes the proof. ■

A.5 Proof of Proposition 3

Ordering: Without loss of generality, take \( x_i \geq y_i \) and consider any \( x'_i \) and \( y'_i \) such that \([y_i, x_i] \in [y'_i, x'_i] \). Since \( v(a; b) = \frac{a}{a+b} \) is increasing in \( a \), \( v_i(x'_i; z_i) \geq v_i(x_i; z_i) \) and \( v_i(y'_i; z_i) \leq v_i(y_i; z_i) \), with at most one inequality binding. This implies \( \hat{\Delta}(x_i, y_i; z_i) = v(x_i; z_i) - v(y_i; z_i) < v(x'_i; z_i) - v(y'_i; z_i) = \hat{\Delta}(x'_i, y'_i; z_i) \), as desired.

Diminishing sensitivity: Using \( \hat{\Delta}(x_i + \epsilon, y_i + \epsilon; z_i) = \frac{x_i + \epsilon}{x_i + y_i + \epsilon z_i} - \frac{y_i + \epsilon}{y_i + x_i + \epsilon z_i} \) given \( x_i > y_i \), \( \frac{d}{d\epsilon} \hat{\Delta}(x_i + \epsilon, y_i + \epsilon; z_i) = \frac{\frac{1}{z_i}}{(x_i + y_i + \epsilon z_i)^2} - \frac{\frac{1}{y_i} x_i}{(y_i + x_i + \epsilon z_i)^2} \leq 0 \) (strict with \( x_i > y_i \), as desired. ■

A.6 Proof of Proposition 4

Under Assumption ??, \( v(x; y) = v(y; x) \), so that \( v(x; y, z) > v(y; x, z) \) if and only if \( v(x; z) > v(y; z) \), i.e.,

\[
\frac{x_1}{x_1 + z_1} + \frac{x_2}{x_2 + z_2} \geq \frac{y_1}{y_1 + z_1} + \frac{y_2}{y_2 + z_2}.
\]  (9)

Rearranging (9), we can see that this condition is equivalent to \( \hat{\Delta}(x_1, y_1; z_1) > \hat{\Delta}(x_2, y_2; z_2) \), as desired.

Next, if we define \( \tilde{w}_i \equiv \frac{w_i}{\sqrt{x_i x_2}} \) for \( i \in \{1, 2\} \), and each \( w \in \{ x, y, z \} \), (9) can be equivalently expressed as \( \frac{x_1}{x_1 + z_1} + \frac{x_2}{x_2 + z_2} \geq \frac{y_1}{y_1 + z_1} + \frac{y_2}{y_2 + z_2} \). Using \( x_1 x_2 = \tilde{y}_1 \tilde{y}_2 = 1 \) to substitute out \( x_2 \) and \( \tilde{y}_2 \), this condition becomes \( \frac{x_1}{\tilde{x}_1 + z_1} + \frac{1}{1 + \tilde{x}_1 z_2} > \frac{\tilde{y}_1}{\tilde{y}_1 + z_1} + \frac{1}{1 + \tilde{y}_1 z_2} \).
After cross-multiplying and collecting terms, then factoring out $\tilde{x}_1 - \tilde{y}_1 > 0$, we find that (9) holds if and only if $(1 - \tilde{z}_1 \tilde{z}_2)(\tilde{z}_1 - \tilde{x}_1 \tilde{y}_1 \tilde{z}_2) > 0$. If we multiply the first term of this inequality by $x_1 x_2 > 0$ and the second term by $\sqrt{x_1 x_2} > 0$, the inequality is preserved. After substituting out $\frac{y_1}{x_2} = \frac{x_1 y_1 y_2}{x_1 y_2} = \frac{x_1}{y_2}$ along with each $\tilde{w}_i = \frac{w_i}{\sqrt{x_1 x_2}}$, we find (9) holds if and only if $(x_1 x_2 - z_1 z_2)(\tilde{z}_1 - \tilde{x}_1 \tilde{y}_1 \tilde{z}_2) > 0$, as desired. 

A.7 Proof of Corollary 1

We have $x_1 > y_1$ by assumption, $x_1 x_2 > z_1 z_2$ follows from $x_1 \geq z_1$ and $x_2 \geq z_2$ with $x \neq z$, and $\frac{x_1}{z_2} > \frac{y_1}{x_2} = \frac{x_1 y_1 y_2}{x_1 y_2 x_2} = \frac{x_1}{y_2}$ since $z_1 > y_1$, $x_2 \geq z_2$, and $x_1 x_2 = y_1 y_2$. Hence, $(x_1 x_2 - z_1 z_2)(\tilde{z}_1 - \tilde{x}_1 \tilde{y}_1 \tilde{z}_2) > 0$ (since both terms are positive), so from Proposition 4 we see that $v(x; y, z) > v(y; x, z)$, as desired. 

A.8 Proof of Corollary 2

We have $x_1 > y_1$ by assumption, $x_1 x_2 > z_1 z_2$ follows from Proposition 2 given $z$ is inferior, and $\frac{x_1}{z_2} > \frac{y_1}{x_2} = \frac{x_1 y_1 y_2}{x_1 y_2 x_2} = \frac{x_1}{y_2}$ since $z_1 > y_1$, $x_2 > z_2$, and $x_1 x_2 = y_1 y_2$. Again, $(x_1 x_2 - z_1 z_2)(\tilde{z}_1 - \tilde{x}_1 \tilde{y}_1 \tilde{z}_2) > 0$ (since both terms are positive), so from Proposition 4 we see that $v(x; y, z) > v(y; x, z)$, as desired. 

A.9 Proof of Lemma 3

Without loss of generality, take $x_i > y_i$. In this case, $\tilde{\Delta}(x_i; y_i; z_i) > \tilde{\Delta}(x_i; y_i; z'_i)$ is equivalent to $\frac{x_i}{x_i + z_i} - \frac{y_i}{y_i + z_i} > \frac{x_i}{x_i + z_i'} - \frac{y_i}{y_i + z_i'}$. Multiplying both sides by $\frac{(x_i + z_i)(y_i + z_i')(x_i + z_i')(y_i + z_i)}{x_i y_i}$

$> 0$, subtracting common terms, then rearranging, we see that the inequality is equivalent to $x_i y_i (z_i - z'_i) > z_i z'_i (z_i - z'_i)$.

Case (a) If $z_i > z'_i, x_i y_i (z_i - z'_i) > z_i z'_i (z_i - z'_i)$ holds if and only if $x_i y_i > z_i z'_i$, which can be expressed as $\frac{\sqrt{y_i}}{z_i} > \frac{z_i}{\sqrt{x_i y_i}}$. Taking the natural logarithm of both sides, we see this is equivalent to $\frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z_i) > \ln(z_i) - \frac{1}{2} (\ln(x_i) + \ln(y_i))$. Since $z_i > z'_i$, we also know $\frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z'_i) > \frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z_i).$ Therefore, the condition is equivalent to $\frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z'_i) > \frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z_i)$. Since $z_i z'_i < x_i y_i$, taking $z_i > z'_i$ means $z'_i < \sqrt{x_i y_i}$, so that $\frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z'_i) > 0.$ Thus, for $z_i > z'_i, \tilde{\Delta}(x_i; y_i; z_i) > \tilde{\Delta}(x_i; y_i; z'_i)$ holds if and only if $\frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z_i') > \frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z_i)$.

Case (b) If $z_i < z'_i, x_i y_i (z_i - z'_i) > z_i z'_i (z_i - z'_i)$ holds if and only if $x_i y_i < z_i z'_i$, which can be expressed as $\frac{\sqrt{y_i}}{z_i} < \frac{z_i}{\sqrt{x_i y_i}}$. Taking the natural logarithm of both sides, we see this is equivalent to $\ln(z'_i) - \frac{1}{2} (\ln(x_i) + \ln(y_i)) > \frac{1}{2} (\ln(x_i) + \ln(y_i)) - \ln(z_i)$. 

43
Since \( z_i < z'_i \), we also know \( \ln(z'_i) - \frac{1}{2}(\ln(x_i) + \ln(y_i)) > \ln(z_i) - \frac{1}{2}(\ln(x_i) + \ln(y_i)) \). Therefore, the condition is equivalent to \( \ln(z'_i) - \frac{1}{2}(\ln(x_i) + \ln(y_i)) > \frac{1}{2}(\ln(x_i) + \ln(y_i)) - \ln(z_i) \). Since \( z_i z'_i > x_i y_i \), taking \( z_i < z'_i \) means \( z'_i > \sqrt{x_i y_i} \), so that \( \ln(z'_i) - \frac{1}{2}(\ln(x_i) + \ln(y_i)) > 0 \). Thus, for \( z'_i > z_i \), \( \tilde{\Delta}(x_i, y_i; z_i) > \tilde{\Delta}(x_i, y_i; z'_i) \) also holds if and only if \( \frac{1}{2}(\ln(x_i) + \ln(y_i)) - \ln(z'_i) > \frac{1}{2}(\ln(x_i) + \ln(y_i)) - \ln(z_i) \). ■

A.10 Proof of Proposition 5

For \( w \in \{z^1, z^2\} \), we can re-write \( v(x; y, w) > v(y; x, w) \) as \( \Delta(x_1, y_1) + \tilde{\Delta}(x_1, y_1; w_2) > \Delta(x_2, y_2) + \tilde{\Delta}(x_2, y_2; w_2) \).

For part (i) and given \( v(x; y, z) = v(y; x, z) \), the condition reduces to \( \tilde{\Delta}(x_1, y_1; z'_1) > \tilde{\Delta}(x_1, y_1; z_1) \), which from Lemma 3 holds if and only if \( \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(z'_1) > \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(z_1) \).

For part (ii) and given \( v(x; y, z) = v(y; x, z) \), the condition reduces to \( \tilde{\Delta}(x_2, y_2; z'_2) < \tilde{\Delta}(x_2, y_2; z_2) \), which from Lemma 3 holds if and only if \( \frac{1}{2}(\ln(x_2) + \ln(y_2)) - \ln(z'_2) < \frac{1}{2}(\ln(x_2) + \ln(y_2)) - \ln(z_2) \). ■

A.11 Proof of Corollary 3

Given \( x_1 > z^d_1 > y_1 \), \( \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(y_1) > \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(z^d_1) > \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(x_1) \). Next, observe \( \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(y_1) \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(x_1) = \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(x_1) \). Together, these observations imply \( \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(z^d_1) \). Given \( z^c_1 > x_1 > y_1 \), we also have \( \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(z^c_1) \). Thus, \( \frac{1}{2}(\ln(x_1) + \ln(y_1)) - \ln(z^c_1) \). Taking \( z = z^c \) and \( z' = z^d \), part (i) of Proposition 5 establishes \( v(x; y, z^d) > v(y; x, z^d) \) given \( v(x; y, z^c) = v(y; x, z^c) \). Taking \( z = z^d \) and \( z' = z^c \), part (i) of Proposition 5 also establishes \( v(x; y, z^c) < v(y; x, z^c) \) given \( v(x; y, z^d) = v(y; x, z^d) \). ■

A.12 Proof of Proposition 6

Without loss of generality, assume \( x_1 > z_1 > y_1, y_2 > x_2 > z_2, \) and \( z_3 > y_3 > x_3 \). We will now proceed by contradiction. If preferences in binary-choice are intransitive, but do not follow the “majority rule” ordering, it must be that \( x \) is preferred to \( y, y \) to \( z \), and \( z \) to \( x \). Noting that normalized valuations are invariant to a scaling of all attribute-i values by a common factor \( \lambda_i > 0 \), we take \( \lambda_1 \equiv z_1^{-1}, \lambda_2 \equiv x_2^{-1}, \lambda_3 \equiv y_3^{-1} \) and define \( \tilde{w}_i \equiv \lambda_i w_i \) for all \( w \in \{x, y, z\} \) and \( i = 1, 2, 3 \). Also define \( k_i \equiv \max\{\tilde{w}_i\} - 1 > 0 \) and \( q_i \equiv 1 - \min\{\tilde{w}_i\} > 0 \) so that the ordered, rescaled
attribute values are \((1 + k_i, 1, 1 - q_i)\) for each \(i\). Our three preference relations now imply:

\[
\begin{align*}
v(x; y) &> v(y; x) \quad \frac{k_2}{2 + k_2} + \frac{q_3}{2 - q_3} < \frac{k_1 + q_1}{2 + k_1 - q_1}, \\
v(y; z) &> v(z; y) \quad \frac{k_3}{2 + k_3} + \frac{q_1}{2 - q_1} < \frac{k_2 + q_2}{2 + k_2 - q_2}, \\
v(z; x) &> v(x; z) \quad \frac{k_1}{2 + k_1} + \frac{q_2}{2 - q_2} < \frac{k_3 + q_3}{2 + k_3 - q_3}.
\end{align*}
\]

Summing these conditions, we get:

\[
\sum_i \left( \frac{k_i}{2 + k_i} + \frac{q_i}{2 - q_i} \right) < \sum_i \left( \frac{k_i + q_i}{2 + k_i - q_i} \right).
\]

Therefore, \(\frac{k_i}{2 + k_i} + \frac{q_i}{2 - q_i} < \frac{k_i + q_i}{2 + k_i - q_i}\) for at least one \(i \in \{1, 2, 3\}\). Combining the fractions on the left-side of this inequality, we get \(\frac{2(k_i + q_i)}{(2 + k_i)(2 - q_i)} < \frac{k_i + q_i}{2 + k_i - q_i}\), which holds if and only if \(2(2 + k_i - q_i) < (2 + k_i)(2 - q_i)\), i.e. if and only if \(-q_i k_i > 0\), a contradiction. ■

### A.13 Proof of Proposition 7

Given \(v_1(x_1; y_1) + v_2(x_2; y_2) = v_1(y_1; x_1) + v_1(y_2; x_2)\), \(x'\) is preferred to \(y'\) if

\[
\begin{align*}
& (v_1a(x_1a; y_1a) + v_1b(x_{1b}; y_{1b}) - v_1(x_1; y_1)) \\
& - (v_1a(y_1a; x_1a) + v_1b(y_{1b}; x_{1b}) - v_1(y_1; x_1)) \\
= & \Delta(x_{1a}; y_{1a}) + \Delta(x_{1b}; y_{1b}) - \Delta(x_1; y_1) \\
= & (x_{1a} - y_{1a})(x_{1b} + y_{1b})^2 + (x_{1b} - y_{1b})(x_{1a} + y_{1a})^2 \\
& (x_{1a} + y_{1a})(x_{1b} + y_{1b})(x_{1a} + y_{1a} + x_{1b} + y_{1b}) > 0.
\end{align*}
\]

Given \(0 < x_{1a} \geq y_{1a}\) and \(0 < x_{1b} \geq y_{1b}\) (with at most one of these two inequalities binding), both terms in the numerator of the last expression are non-negative and at least one term is strictly positive. Since the denominator is positive too, we get the desired result. ■

### A.14 Proof of Proposition 8

\(x'\) is preferred to \(y'\) if and only if \(v_1(x_1; y_1) + v_2(x_2; 0) > v_1(y_1; x_1) + v_2(y_2; 0)\), i.e.,

\[
\frac{x_1}{x_1 + y_1} + \frac{x_2}{x_2 + 0} > \frac{y_1}{x_1 + y_1} + \frac{y_2}{y_2 + 0},
\]

which reduces to \(x_1 > y_1\). ■
A.15 Proof of Lemma 4

Given $X = \{x, y\}$, $x$ is preferred to $y$ if and only if $v^*(x; y) - v^*(y; x) = \sum_i \frac{x_i^\alpha - y_i^\alpha}{\sigma^\alpha + x_i^\alpha + y_i^\alpha} > 0$. Combining terms and factoring out the denominator, we see that this condition is equivalent to:

$$\sum_i (x_i^\alpha - y_i^\alpha) \prod_{j \neq i} (\sigma^\alpha + x_j^\alpha + y_j^\alpha) > 0.$$ 

For the two-attribute case, this reduces to (6). $\blacksquare$

A.16 Proof of Proposition 9

(I) Using (6) given $\sigma = 0$, Lemma 4 indicates that $x$ is preferred to $y$ if and only if $x_1^\alpha x_2^\alpha > y_1^\alpha y_2^\alpha$, which is equivalent to $x_1 x_2 > y_1 y_2$.

(II) Given $u(a, b) = (a^\alpha + b^\alpha)^{1/\alpha}$, we see that $u(x) > u(y)$ holds if and only if $u(x)^\alpha > u(y)^\alpha$, i.e., $x_1^\alpha + x_2^\alpha > y_1^\alpha + y_2^\alpha$. Let $\sigma_0 = (\frac{2(y_1^\alpha y_2^\alpha - x_1^\alpha x_2^\alpha)}{x_1^\alpha + x_2^\alpha + y_1^\alpha + y_2^\alpha})^{1/\alpha} < \infty$. Observe $\sigma_0(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha = \sigma_0(y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha$. Thus, $\sigma^\alpha(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha > \sigma^\alpha(y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha$ for all $\sigma > \sigma_0$, implying $x$ is preferred to $y$ from Lemma 4.

The converse is established by contradiction. Namely, suppose $x$ is preferred to $y$ but $u(x) < u(y)$, or equivalently, $x_1^\alpha + x_2^\alpha < y_1^\alpha + y_2^\alpha$. From Lemma 4, we see that, together, these conditions require $x_1^\alpha x_2^\alpha > y_1^\alpha y_2^\alpha$, so that $x_1^\alpha x_2^\alpha - y_1^\alpha y_2^\alpha > 0$. By inspection, we can now see $\sigma > \sigma_0$ with $\sigma_0 > 0$ as defined above implies $\sigma^\alpha(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha < \sigma^\alpha(y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha$, which from Lemma 4 implies $y$ is preferred to $x$. Hence, we have a contradiction, so that a preference for $x$ over $y$ necessarily requires $u(x) > u(y)$ for sufficiently large $\sigma > 0$, as desired.

(III) Follows as a special case of (II), taking $\alpha = 1$.

(IV) Given $u(a, b) = \max\{a, b\}$, letting $x_1 = \max\{x_1, x_2\}$ and $y_2 = \max\{y_1, y_2\}$, without loss of generality, we see $u(x) > u(y)$ holds if and only if $x_1 > y_2$. Observe, $\sigma^\alpha(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha \geq \sigma^\alpha x_1^\alpha$. Given any $\sigma > y_2$, we also see $\sigma^\alpha(y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha \leq 2\sigma^\alpha y_1^\alpha + 2y_1^\alpha y_2^\alpha < 4\sigma^\alpha y_2^\alpha$. From Lemma 4, we can then see that a sufficient condition for $x$ to be preferred to $y$ given any $\sigma > y_2$ is $\sigma^\alpha x_1^\alpha > 4\sigma^\alpha y_2^\alpha$. Factoring out $\sigma^\alpha > 0$ then taking the natural logarithm, we see this condition is equivalent to $\alpha \ln(x_1) > \alpha \ln(y_2) + \ln(4)$. Taking $\alpha_0 \equiv \frac{\ln(4)}{\ln(x_1) - \ln(y_2)} > 0$, we see $\alpha \ln(x_1) > \alpha \ln(y_2) + \ln(4)$ holds for any $\alpha > \alpha_0$ and $\sigma > y_2$, so that $x$ must be preferred to $y$ for sufficiently large $\alpha$ and $\sigma$, as desired.

The converse is established by contradiction. Namely, suppose $x$ is preferred to $y$ but $u(x) < u(y)$, or equivalently, $y_2 > x_1$. Using Lemma 4 while applying the same logic outlined above (except switching the roles of $x$ and $y$), we see it must be the case that, for any $\sigma > x_1$, $\alpha \ln(x_1) + \ln(4) > \alpha \ln(y_2)$ by virtue of the preference for $y$. Hence, we have a contradiction, so that a preference for $y$ over $x$ necessarily requires $u(x) < u(y)$ for sufficiently large $\alpha$ and $\sigma$, as desired.

46
for \( x \) over \( y \). Defining \( \alpha' \equiv \frac{\ln(4)}{\ln(y_2) - \ln(x_1)} > 0 \) (positive because \( y_2 > x_1 \)), we can see \( \alpha > \alpha' \) implies \( \alpha \ln(x_1) + \ln(4) < \alpha \ln(y_2) \). Hence, we have a contradiction, so that a preference for \( x \) over \( y \) (with \( u(x) \neq u(y) \)) necessarily requires \( u(x) > u(y) \) for sufficiently large \( \sigma > 0 \) and \( \alpha > 0 \), as desired.

In the case of \( u(x) = u(y) \), i.e., \( x_1 = y_2 \), we see from Lemma 4 that, in this case, \( x \) will be preferred to \( y \) if and only if \( \sigma^a(x_1^a + x_2^a) > \sigma^a(x_1^a + x_2^a) + 2x_1^ax_2^a > \sigma^a(x_1^a + y_2^a) + 2x_1^ay_1^a \). Subtracting \( \sigma^a x_1^a \) from both sides, then dividing both sides by \( \sigma^a + 2x_1^a > 0 \), we see this condition is equivalent to \( x_2 > y_1 \). Given \( u_0(a, b) \equiv \min\{a, b\} \) with \( x_1 \geq x_2 \) and \( y_2 \geq y_1 \), we see that \( x_2 > y_1 \) is equivalent to \( u_0(x) > u_0(y) \), as desired. ■

### A.17 Proof of Proposition 10

Without loss of generality, suppose \( x_1 \geq y_1 \). Then \( \partial \Delta^*(x_i, y_i) / \partial x_i = \frac{\alpha x_i^a(y_i^a + 2 \epsilon)}{x_i^a(y_i^a + y_i^a)^2} > 0 \) and \( \partial \Delta^*(x_i, y_i) / \partial y_i = -\frac{\alpha y_i^a(y_i^a + 2 \epsilon)}{y_i^a(y_i^a + x_i^a)^2} < 0 \). Therefore, \( x_i^a \geq x_i \) and \( y_i^a \leq y_i \), with at most one of these inequalities binding, imply \( \Delta^*(x_i', y_i') > \Delta^*(x_i, y_i) \). Thus, ordering is satisfied for all \( \sigma > 0 \) and \( \alpha > 0 \).

In light of the above work, must demonstrate \( \Delta^*(x_i, y_i) \) exhibits diminishing sensitivity for all \( x_i \geq 0 \) and \( y_i \geq 0 \) (with \( x_i \geq y_i \), without loss of generality) if and only if \( \sigma = 0 \) or \( \alpha \leq 1 \) (or both). Observe

\[
\frac{d[\Delta^*(x_i + \epsilon, y_i + \epsilon)]}{d\epsilon} = -\frac{\alpha x_i^a y_i^a(2(x_i^a - y_i^a) + \sigma^a(x_i^{1-a} - y_i^{1-a}))}{x_i y_i(x_i^a + y_i^a + \sigma^a)^2}.
\]

We can see \( \frac{d[\Delta^*(x_i + \epsilon, y_i + \epsilon)]}{d\epsilon} < 0 \) if and only if \( 2(x_i - y_i) + \sigma^a(x_i^{1-a} - y_i^{1-a}) > 0 \). Given \( x_i \geq y_i \), this clearly holds for \( \sigma = 0 \) and also for \( \alpha \leq 1 \) because, together, \( \alpha \leq 1 \) and \( x_i \geq y_i \) guarantee \( x_i^{1-a} - y_i^{1-a} \geq 0 \). Thus, \( \Delta^*(x_i + \epsilon, y_i + \epsilon) < \Delta^*(x_i, y_i) \) for all \( \epsilon > 0 \) given \( \sigma = 0 \) or \( \alpha \leq 1 \) (or both).

Thus, to complete the proof, we only need to show that given any \( \sigma > 0 \) and \( \alpha > 1 \), there exist \( x_i \geq 0 \) and \( y_i \geq 0 \) with \( x_i \geq y_i \) such that \( 2(x_i - y_i) + \sigma^a(x_i^{1-a} - y_i^{1-a}) < 0 \). Take \( y_i = \frac{\sigma^{(a-1)/a}}{2} \) and let \( x_i = y_i + \delta \). Substituting these into \( 2(x_i - y_i) + \sigma^a(x_i^{1-a} - y_i^{1-a}) \) then differentiating with respect to \( \delta \), we get \( 2 - 2^a < 0 \) for \( \alpha > 1 \). Also note \( 2(x_i - y_i) + \sigma^a(x_i^{1-a} - y_i^{1-a}) = 0 \) given \( x_i = y_i \), i.e., given \( \delta = 0 \). Together, these imply \( 2(x_i - y_i) + \sigma^a(x_i^{1-a} - y_i^{1-a}) < 0 \) for \( y_i = \frac{\sigma^{(a-1)/a}}{2} \) and \( x_i = y_i + \delta \), provided \( \delta > 0 \) is sufficiently small, as desired. ■

### A.18 Proof of Proposition 11

**Part (i).** Let \( h(x_i, y_i, \sigma, \alpha) \equiv 2(x_i - y_i) + \sigma^a(x_i^{1-a} - y_i^{1-a}) \). From our work in the proof of Proposition 10, we can then see that \( \Delta^*(x_i, y_i) \) satisfies diminishing
sensitivity for all $x_i \geq y_i$ as in condition (a) if and only if $h(x_i | y_i, \sigma, \alpha) > 0$ for all $x_i \geq y_i$. Observe $h'(x_i | y_i, \sigma, \alpha) = 2 - \frac{(\alpha - 1)\sigma^\alpha}{x_i^{\alpha-1}}$, so that $h'(x_i | y_i, \sigma, \alpha) = 0$ if and only if $x_i = x_i^* \equiv \sigma \left( \frac{\alpha - 2}{\alpha - 1} \right)^{1/\alpha}$. Also observe, $h''(x_i | y_i, \sigma, \alpha) = \frac{\alpha(\alpha - 1)^{\alpha-1}}{x_i^{\alpha-1}} > 0$. Thus, $h(x_i | y_i, \sigma, \alpha)$ is uniquely minimized by $x_i = x_i^*$. Given $h'(x_i | y_i, \sigma, \alpha) > 0$ for all $x_i > x_i^*$ and $h(y_i | y_i, \sigma, \alpha) = 0$, $\Delta^*(x_i, y_i)$ satisfies diminishing sensitivity for all $x_i \geq y_i$ if and only if $x_i^* \leq y_i$ or $h(x_i^* | y_i, \sigma, \alpha) > 0$ (or both). Given $h'(x_i | y_i, \sigma, \alpha) < 0$ for all $x_i < x_i^*$ and $h(y_i | y_i, \sigma, \alpha) = 0$, $x_i^* > y_i$ implies $h(x_i^* | y_i, \sigma, \alpha) < 0$. Taken together, these last two observations imply $\Delta^*(x_i, y_i)$ satisfies diminishing sensitivity for all $x_i \geq y_i$ if and only if $x_i^* \leq y_i$, which, using the definitions of $x_i^*$ and of $\hat{\sigma}(y_i)$, we can see this is equivalent to $\hat{\sigma}(y_i) \equiv \left( \frac{2}{\alpha - 1} \right)^{1/\alpha} y_i$ as in condition (c).

Taking $\partial^2 \Delta^*(x_i, y_i) / \partial x_i^2$, multiplying through by $x_i^2 (x_i^\alpha + y_i^\alpha + \sigma^\alpha) > 0$, diving by $\alpha x_i^\alpha (2y_i^\alpha + \sigma^\alpha) > 0$, and rearranging, we see $\Delta^*(x_i, y_i)$ is concave in $x_i$ for all $x_i \geq y_i$ if and only if $x_i^\alpha (1 + \alpha) \geq (\alpha - 1)(y_i^\alpha + \sigma^\alpha)$ for all $x_i \geq y_i$. Since the left-side of this inequality is clearly increasing in $x_i$, $\Delta^*(x_i, y_i)$ is concave in $x_i$ for all $x_i \geq y_i$ if and only if the inequality holds at $x_i = y_i$, i.e., if and only if $y_i^\alpha (1 + \alpha) \geq (\alpha - 1)(y_i^\alpha + \sigma^\alpha)$. Solving for $\sigma$, we see this condition is equivalent to $\sigma \leq \hat{\sigma}(y_i) \equiv \left( \frac{2}{\alpha - 1} \right)^{1/\alpha} y_i$, as desired.

Part (ii). Using our definition of $h(x_i | y_i, \sigma, \alpha)$ and $\tilde{x}(\sigma, y_i)$, it is readily verifiable that $h(\tilde{x}(\sigma, y_i) | y_i, \sigma, \alpha) = 0$. Given $x_i^* > y_i$ for $\sigma \geq \hat{\sigma}(y_i)$ from part (i), $h(y_i | y_i, \sigma, \alpha) = 0$, $h'(x_i | y_i, \sigma, \alpha) < 0$ for all $x_i < x_i^*$, and $h'(x_i | y_i, \sigma, \alpha) > 0$ for all $x_i > x_i^*$, it follows that $\tilde{x}(\sigma, y_i) > x_i^*$, implying $h(x_i | y_i, \sigma, \alpha) > 0$ for $y_i < x_i < \tilde{x}_i$ and $h(x_i | y_i, \sigma, \alpha) > 0$ for $x_i > \tilde{x}_i$, as desired for statement (a).

Recalling from part (i) that $\Delta^*(x_i, y_i)$ is concave in $x_i$ if and only if $x_i^\alpha (1 + \alpha) \geq (\alpha - 1)(y_i^\alpha + \sigma^\alpha)$, we can rearrange this inequality to see that it binds at $\hat{x}(\sigma, y_i)$. By inspection, we can then see that $x_i < \hat{x}(\sigma, y_i)$ implies $x_i^\alpha (1 + \alpha) < (\alpha - 1)(y_i^\alpha + \sigma^\alpha)$ and $x_i > \hat{x}(\sigma, y_i)$ implies $x_i^\alpha (1 + \alpha) > (\alpha - 1)(y_i^\alpha + \sigma^\alpha)$, implying the desired result.

Expressing $h(\tilde{x}, \sigma, y) \equiv h(\tilde{x}(\sigma, y_i) | y_i, \sigma, \alpha) = 2(\tilde{x} - y_i) + \sigma^\alpha (\tilde{x}^{1 - \alpha} - y_i^{1 - \alpha}) = 0$, we see $\partial h(\tilde{x}, \sigma, y) / \partial \tilde{x} = 2 - (\alpha - 1)\tilde{x}^{-\alpha}\sigma^\alpha$, $\partial h(\tilde{x}, \sigma, y) / \partial y_i = -2 + (\alpha - 1)y_i^{-\alpha}\sigma^\alpha$, and $\partial h(\tilde{x}, \sigma, y) / \partial \sigma = \sigma \alpha^{-1}(\frac{1}{\tilde{x}^{\alpha-1}} - \frac{1}{\tilde{x}^{\alpha-1}}) < 0$.

[Next step: establish $\partial h(\tilde{x}, \sigma, y) / \partial \tilde{x} > 0$ and $\partial h(\tilde{x}, \sigma, y) / \partial y_i < 0$ to get desired result]
A.19  Proof of Proposition 12
A.20  Proof of Lemma 5
A.21  Proof of Corollary 4
A.22  Proof of Corollary 5
A.23 Proof of Proposition 12

Parts (i) and (ii) follow in light of the fact that \( \sigma = 0 \) implies that \( x \sim^* y \) and \( x \sim y \) are equivalent. In particular, \( x \sim^* y \) given \( \sigma = 0 \) reduces to \( x_1^\alpha x_2^\alpha = y_1^\alpha y_2^\alpha \) in the same way \( x \sim y \) reduced to \( x_1 x_2 = y_1 y_2 \), and in turn, we see that \( x_1^\alpha x_2^\alpha = y_1^\alpha y_2^\alpha \) holds if and only if our original condition \( x_1 x_2 = y_1 y_2 \) also holds. Next, from Proposition 2, we know \( x \) is preferred to \( y \) given \( X = \{x, y, z\} \) (given \( x \sim y \)) if and only if \( (x_1 x_2 - z_1 z_2) \left( \frac{a_1}{z_2} - \frac{a_1}{y_2} \right) > 0 \). We can similarly derive \( x \succ^*_z y \) (given \( x \sim^* y \) and \( \sigma = 0 \)) if and only if \( (x_1^\alpha x_2^\alpha - z_1^\alpha z_2^\alpha) \left( \frac{x_1^\alpha}{z_2^\alpha} - \frac{x_1^\alpha}{y_2^\alpha} \right) > 0 \), which again holds if and only if the analogous condition from the original, parameter-free model also holds.

The robustness of part (iii) with general \( \alpha > 0 \) and \( \sigma = 0 \) is straightforward in light of the equivalence of \( x \sim^* y \) and \( x \sim y \).

For part (iv) let \( x_1^\alpha = px_1 \) and \( y_1^\alpha = qy_1 \) so that \( x_1^\alpha = (1 - p)x_1 \) and \( y_1^\alpha = (1 - q)y_1 \). We want to show
\[
(v^*(px_1; qy_1)) - v^*(qy_1; px_1)) + (v^*((1 - p)x_1; (1 - q)y_1)) - v^*((1 - q)y_1; (1 - p)x_1)) > (v^*(x_1; y_1)) - v^*(y_1; x_1)).
\]
That is,
\[
\frac{(px_1)^\alpha - (qy_1)^\alpha}{(px_1)^\alpha + (qy_1)^\alpha} + \frac{((1 - p)x_1)^\alpha - ((1 - q)y_1)^\alpha}{((1 - p)x_1)^\alpha + ((1 - q)y_1)^\alpha} > \frac{x_1^\alpha - y_1^\alpha}{x_1^\alpha + y_1^\alpha}.
\]
Since \( x_1^\alpha \geq y_1^\alpha \) and \( x_1^\alpha \geq y_1^\alpha \), both terms on the left-side are non-negative (with at least one term strictly positive). Letting \( \gamma_a \equiv \left( \frac{q}{p} \right)^\alpha \) and \( \gamma_b \equiv \left( \frac{1 - q}{1 - p} \right)^\alpha \), this condition simplifies to:
\[
\frac{x_1^\alpha - \gamma_a y_1^\alpha}{x_1^\alpha + \gamma_a y_1^\alpha} + \frac{x_1^\alpha - \gamma_b y_1^\alpha}{x_1^\alpha + \gamma_b y_1^\alpha} > \frac{x_1^\alpha - y_1^\alpha}{x_1^\alpha + y_1^\alpha}.
\]
Since \( \min\{\gamma_a, \gamma_b\} < 1 \) except when \( \gamma_a = \gamma_b = 1 \), \( \gamma_a \neq \gamma_b \) implies
\[
\frac{x_1^\alpha - \gamma_a y_1^\alpha}{x_1^\alpha + \gamma_a y_1^\alpha} + \frac{x_1^\alpha - \gamma_b y_1^\alpha}{x_1^\alpha + \gamma_b y_1^\alpha} \geq 2 \cdot \frac{x_1^\alpha - \min\{\gamma_a, \gamma_b\} y_1^\alpha}{x_1^\alpha + \min\{\gamma_a, \gamma_b\} y_1^\alpha} > \frac{x_1^\alpha - y_1^\alpha}{x_1^\alpha + y_1^\alpha},
\]
For \( \gamma_a = \gamma_b \), we see
\[
\frac{x_1^\alpha - \gamma_a y_1^\alpha}{x_1^\alpha + \gamma_a y_1^\alpha} + \frac{x_1^\alpha - \gamma_b y_1^\alpha}{x_1^\alpha + \gamma_b y_1^\alpha} = 2 \cdot \frac{x_1^\alpha - y_1^\alpha}{x_1^\alpha + y_1^\alpha} > \frac{x_1^\alpha - y_1^\alpha}{x_1^\alpha + y_1^\alpha},
\]
as desired.

For part (v), given \( x \sim^* y \) with \( \sigma = 0 \), \( x' \succ^* y' \) with \( x' = (x_1, x_2, 0) \) and \( y' = (y_1, 0, y_2) \) if and only if \( \frac{x_1^\alpha}{x_1^\alpha + y_1^\alpha} + \frac{x_2^\alpha}{x_2^\alpha + 0^\alpha} > \frac{y_1^\alpha}{y_1^\alpha + y_2^\alpha} + \frac{y_2^\alpha}{y_2^\alpha + 0^\alpha} \), which reduces to
\( x_1 > y_1 \), as desired.

For part (vi), \( x \succ^* y \) with \( \sigma = 0 \) holds if and only if \( x_1x_2 = y_1y_2 \). Next, \( x' \succ^* y' \) with \( x' = (x_1 + d_1, x_2 + d_2) \) and \( y' = (y_1 + d_1, y_2 + d_2) \) if and only if \( (x_1 + d_1)(x_2 + d_2) > (y_1 + d_1)(y_2 + d_2) \), which is equivalent to \( x_1d_2 + x_2d_1 > y_1d_2 + y_2d_1 \). With \( x_1 > y_1 \) and \( y_2 > y_1 \), we can rearrange this condition to give \( d_2 \frac{x_1}{x_1 - y_1} > \frac{y_2 - x_2}{x_1 - y_1} \).

\[ \text{A.24 Proof of Lemma ??} \]

\( x \succ^* y \) if and only if \( v^*(x; y) - v^*(y; x) = \sum_i \frac{x_i^\alpha - y_i^\alpha}{\sigma + x_i + y_i} > 0 \). Combining terms and factoring out the denominator, we see that the condition for which \( x \succ^* y \) becomes:

\[ \sum_i (x_i^\alpha - y_i^\alpha) \prod_{j \neq i} (\sigma + x_j^\alpha + y_j^\alpha) > 0. \]

For the two-attribute case, this reduces to

\[ \sigma^\alpha(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha > \sigma^\alpha(y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha. \] (10)

Given \( x_1x_2 > y_1y_2 \) (i.e., \( x \succ y \)), we see that this condition is immediately satisfied since \( x_1 + x_2 > y_1 + y_2 \). If \( x_1x_2 \leq y_1y_2 \) (\( y \succeq x \)), we can solve for \( \sigma \) using \( \sigma^\alpha(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha = \sigma^\alpha(y_1^\alpha + y_2^\alpha) + 2y_1^\alpha y_2^\alpha \). Since \( v^*(x; y) - v^*(y; x) = \sigma^\alpha(x_1^\alpha + x_2^\alpha) + 2x_1^\alpha x_2^\alpha - \sigma^\alpha(y_1^\alpha + y_2^\alpha) - 2y_1^\alpha y_2^\alpha \) is increasing in \( \sigma \) given \( x_1^\alpha + x_2^\alpha > y_1^\alpha + y_2^\alpha \), \( \sigma > \bar{\sigma} \) ensures the inequality in (10) is satisfied.

\[ \text{A.25 Proof of Proposition ??} \]

Since \( x \sim^* y \), \( v^*(y; x, z) - v^*(x; y, z) = v^*(y; z) - v^*(x; z) = \frac{y_1}{\sigma + y_1 + x_1} + \frac{y_2}{\sigma + y_2 + x_2} - \frac{x_1}{\sigma + x_1 + z_1} - \frac{x_2}{\sigma + x_2 + z_2} \). Substituting out each \( z_i = x_i \), we see that \( y \succ^*_z x \) if and only if

\[ v^*(y; z) - v^*(x; z) = \frac{y_1}{\sigma + y_1 + x_1} + \frac{y_2}{\sigma + y_2 + x_2} - \frac{x_1}{\sigma + 2x_1} - \frac{x_2}{\sigma + 2x_2} > 0. \] (11)

Given \( x \sim^* y \), we know from equation (10) that \( \sigma = \frac{2(y_1y_2 - x_1x_2)}{(x_1 + x_2)(y_1 + y_2)} \). Multiplying both sides of equation (11) by \( (\sigma + x_1 + y_1)(\sigma + x_2 + y_2)(\sigma + 2x_1)(\sigma + 2x_2) \), then substituting out \( \sigma \) using the previous expression, and combining terms we get that \( y \succ^*_z x \) if and only if

\[ \frac{2(x_1 - x_2)(x_1 - y_1)(x_2 - y_2)(x_1 - x_2 + y_1 - y_2)(x_1x_2 - y_1y_2)}{(x_1 + x_2 - y_1 - y_2)^2} > 0 \] (12)

Again from \( x \sim^* y \) and \( x \neq y \), we know (a) \( x_i > y_i \) for some \( i \in \{1, 2\} \) and \( y_j > x_j \) for \( j = 3-i \), so that \( (x_1 - y_1)(x_2 - y_2) < 0 \); (b) \( \sigma(x_1 + x_2) + 2x_1x_2 = \sigma(y_1 + y_2) + 2y_1y_2 \), so that \( x_1 + x_2 > y_1 + y_2 \) if and only if \( x_1x_2 < y_1y_2 \). Given these facts, and
because the denominator in (12) is necessarily positive, this condition reduces to:

$$(x_1 - x_2)(x_1 - x_2 + y_1 - y_2)(x_1 + x_2) - (y_1 + y_2) > 0.$$ If \(i = 1\) (so that \(x_1 - x_2 > 0\), we see that \(y >^*_x x\) if and only if one of the following two conditions holds: (i) \(x_i + y_1 > x_j + y_j\) and \(x_1 + x_2 > y_1 + y_2\); (ii) \(x_i + y_1 < x_j + y_j\) and \(x_1 + x_2 < y_1 + y_2\).

If \(i = 2\) (so that \(x_1 - x_2 < 0\), we can again see that \(y >^*_x x\) if and only if either (i) or (ii) holds. Thus, we need to show that either (i) or (ii) always holds. To do this, we will show that \(x_1 + x_2 \geq y_1 + y_2\) implies \(x_i + y_1 \geq x_j + y_j\) so that conditions (i) and (ii) reduce to \(x_1 + x_2 > y_1 + y_2\) and \(x_1 + x_2 < y_1 + y_2\), respectively.

Without loss of generality, suppose \(x_1 + x_2 > y_1 + y_2\) but \(x_i + y_1 < x_j + y_j\). Combining these inequalities gives \(y_i - x_j < x_i - y_j < x_j - y_i\). Thus, \(x_j > y_i\). Since \(x \sim^* y\) with \(x_1 + x_2 > y_1 + y_2\), equation (10) implies \(y_i y_j > x_i x_j\), i.e. \(y_j > \frac{x_i x_j}{y_i}\). Therefore \(y_i + \frac{x_i x_j}{y_i} < y_j + y_i < x_i + x_j\). Multiplying by \(y_i\) gives \(y_i^2 + x_i x_j < y_i (x_i + x_j)\), which is equivalent to \((x_i - y_i)(x_j - y_i) < 0\). This is a contradiction because \(x_i > y_i\) by assumption and \(x_j > y_i\) from above. We can similarly verify that \(x_1 + x_2 < y_1 + y_2\) but \(x_i + y_1 > x_j + y_j\) is also a contradiction, so that equation (12) must hold as long as \(x_1 + x_2 \neq y_1 + y_2\).

If \(x_1 + x_2 = y_1 + y_2\), this with equation (10) implies \(x_1 = y_2\) and \(x_2 = y_1\). In this case, \(y >^*_x x\) if and only if \(v^*(y; x, z) - v^*(x; y, z) = \frac{x_2}{\sigma + x_1 + x_2} + \frac{x_1}{\sigma + x_1 + x_2} - \frac{x_1}{\sigma + 2x_1} + \frac{x_2}{\sigma + 2x_2} > 0\). This condition is equivalent to \(\frac{\Delta(x_1 - x_2)}{\sigma(x_1 + x_2)(\sigma + 2x_1)(\sigma + 2x_2)} > 0\), which clearly holds.

**A.26 Proof of Lemma ??**

Observe \(\Delta^*_i(x_i, y_i; z_i) = \frac{x_i}{\sigma + x_i + z_i} - \frac{y_i}{\sigma + y_i + z_i} = \frac{x_i}{x_i + (z_i + \sigma)} - \frac{y_i}{y_i + (z_i + \sigma)} = \tilde{\Delta}(x_i, y_i; z_i + \sigma)\).

**A.27 Proof of Corollary 4**

Since \(x \sim^* y\) for all \(\sigma \geq 0\) with \(x \neq y\), we must have \(x_1 = y_2\) and \(x_2 = y_1\). In light of Lemma ??, \(x >^*_z y\) if and only if \(\tilde{\Delta}(x_1, y_1; z_1 + \sigma) - \tilde{\Delta}(x_2, y_2; z_2 + \sigma) = \frac{x_1}{x_1 + z_1} + \frac{x_2}{x_2 + z_2} - \frac{x_1}{x_1 + z_1} - \frac{x_2}{x_2 + z_2} = \frac{z_1 - x_2}{x_1 + z_1} + \frac{x_2 - x_1}{x_2 + z_2} > 0\). Cross-multiplying and factoring out the denominator, we get \(x >^*_z y\) if and only if \((x_1 - x_2)(z_1 - z_2)(x_1 x_2 - z_1 z_2') > 0\). Since \((x_1 - x_2)(z_1 - z_2) > 0\) if and only if \(z_i > z_j\), and \((x_1 x_2 - z_1 z_2') > 0\) if and only if \(z' < x, y\), parts (i) and (ii) follow. Since \(z_1 > z_2\) if \(z\) is a decoy for \(x\), and also if \(x\) is a compromise between \(y\) and \(z\), \(z_1 z_2' = (z_1 + \sigma)(z_2 + \sigma) < x_1 x_2\) must hold for \(x\) is preferred to \(y\) given \(X = \{x, y, z\}\) in either case.
A.28 Proof of Proposition 6*

We again proceed by contradiction, scaling all values as in the proof of Proposition 6 while maintaining the definitions of $k_i$ and $q_i$. Our three preference relations now imply:

\[
\begin{align*}
(x & \succ^* y) \quad \frac{k_2}{2 + \lambda_2 \sigma + k_2} + \frac{q_3}{2 + \lambda_3 \sigma - q_3} < \frac{k_1 + q_1}{2 + \lambda_1 \sigma + k_1 - q_1}, \\
(y & \succ^* z) \quad \frac{k_3}{2 + \lambda_3 \sigma + k_3} + \frac{q_1}{2 + \lambda_1 \sigma - q_1} < \frac{k_2 + q_2}{2 + \lambda_2 \sigma + k_2 - q_2}, \\
(z & \succ^* x) \quad \frac{k_1}{2 + \lambda_1 \sigma + k_1} + \frac{q_2}{2 + \lambda_2 \sigma - q_2} < \frac{k_3 + q_3}{2 + \lambda_3 \sigma + k_3 - q_3}.
\end{align*}
\]

Summing these conditions, we get:

\[
\sum_i \left( \frac{k_i}{2 + \lambda_i \sigma + k_i} + \frac{q_i}{2 + \lambda_i \sigma - q_i} \right) < \sum_i \left( \frac{k_i + q_i}{2 + \lambda_i \sigma + k_i - q_i} \right).
\]

Therefore, \(\frac{k_i}{2 + \lambda_i \sigma + k_i} + \frac{q_i}{2 + \lambda_i \sigma - q_i} < \frac{k_i + q_i}{2 + \lambda_i \sigma + k_i - q_i}\) for at least one \(i \in \{1, 2, 3\}\). Combining the fractions on the left-side of this inequality, we get \(\frac{(2 + \lambda_i \sigma)(k_i + q_i)}{(2 + \lambda_i \sigma + k_i)(2 + \lambda_i \sigma - q_i)} < \frac{k_i + q_i}{2 + \lambda_i \sigma + k_i - q_i}\), which holds if and only if \((2 + \lambda_i \sigma)(2 + \lambda_i \sigma + k_i - q_i) < (2 + \lambda_i \sigma + k_i)(2 + \lambda_i \sigma - q_i)\), i.e. if and only if \(-q_i k_i > 0\), a contradiction. ■

A.29 Proof of Proposition 7*

We want to show

\[
\begin{align*}
\left( v^*(x^a; y^a) + v^*(x^b; y^b) - v^*(x^a; y^a) \right) - \left( v^*(y^a; x^a) + v^*(y^b; x^b) - v^*(y^a; x^a) \right) \\
= \frac{(x^a - y^a)(x^b + y^b)^2 + (x^b - y^b)(x^a + y^a)^2 + 2\sigma(x^a x^b - y^a y^b)}{(\sigma + x^a + y^a)(\sigma + x^b + y^b)(\sigma + x^a + y^a + x^b + y^b)} > 0,
\end{align*}
\]

given \(x_i > y_i\), \(0 < x^a \geq y^a\), \(0 < x^b \geq y^b\), \(x_i = x^a + x^b\), and \(y_i = y^a + y^b\). Since \(0 < x^a \geq y^a\) and \(0 < x^b \geq y^b\) (with at most one inequality binding), each of the three terms in the numerator of this expression is positive, as desired. ■

A.30 Proof of Proposition ??

\[
v(x; y) - v(y; x) = \frac{a - b}{\sigma + a + b} + \frac{b}{\sigma + b} - \frac{a}{\sigma + a} = \frac{ab(a - b)}{(\sigma + a)(\sigma + b)(\sigma + a + b)} > 0 \text{ if and only if } a > b.
\]

■
A.31  Proof of Lemma ??

For all $w \in X$, define

$$\tilde{w}_i = \left(\frac{w_i}{\sigma}\right)^\alpha.$$  \hfill (13)

Now $x \succeq y$ if and only if $v(x; y) - v(y; x) = \tilde{x}_1 - \tilde{y}_1 + \frac{\tilde{x}_2 - \tilde{y}_2}{1 + \tilde{x}_1 + \tilde{y}_1} \geq 0$. Multiplying through by the denominators, this condition is equivalent to $(\tilde{x}_1 - \tilde{y}_1)(1 + \tilde{x}_2 + \tilde{y}_2) + (\tilde{x}_2 - \tilde{y}_2)(1 + \tilde{x}_1 + \tilde{y}_1) \geq 0$. Expanding the products and rearranging, this condition becomes $\tilde{x}_1 + \tilde{x}_2 + 2\tilde{x}_1\tilde{x}_2 \geq \tilde{y}_1 + \tilde{y}_2 + 2\tilde{y}_1\tilde{y}_2$. Multiplying through by $\sigma^{2\alpha}$ and reverting to our original notation gives (??). ■