American Psychological Association

Decision

2325-9965/18/$12.00

http://dx.doi.org/10.1037/dec0000094

Attention and Reference Dependence

Sudeep Bhatia
University of Pennsylvania

Russell Golman
Carnegie Mellon University

We present a model of reference dependence for riskless choice, in which we assume that reference points affect choice by directing the decision maker's attention toward the particular attributes associated with the reference object. This model makes no assumptions about the curvature of utility and does not assume a built-in asymmetry in gains and losses. Nonetheless it is able to generate a type of loss aversion and can explain behavioral anomalies related to reference dependence. In addition, the model makes a number of novel predictions that differentiate it from existing accounts of reference dependence and can easily be extended to predict instances of choice set dependence and anchoring effects. Sensibly, for decisions about economic goods, the model implies acyclic sequential choices and any sequence of short-sighted reference-dependent decisions is guaranteed to converge to a stable personal equilibrium.

Keywords: attention, reference dependence, endowment effect, loss aversion, behavioral economics

Choice behavior is influenced by salient objects, such as reference points. These objects affect the decision maker's valuations, leading to choice reversals and other deviations from rationality. Reference points, for example, are often preferred over competing objects, generating the endowment effect and the status quo bias. In addition, objects that are clear improvements over the reference point and objects that involve only small tradeoffs from the reference point, are often chosen over objects involving large tradeoffs from the reference point (Kahneman, Knetsch, & Thaler, 1991; Tversky & Kahneman, 1991).

Standard economic models of reference dependence involve some deviations from traditional consumer theory. Prospect theory (Tversky & Kahneman, 1991) and related models (e.g., Köszegi & Rabin, 2006) assume nonconvex preferences and explicit asymmetries in gains and losses. In addition to being relatively intractable (e.g., not guaranteeing continuous demand functions), such preferences can lead a decision maker into cyclic choices when making sequential decisions among economic goods. These decision makers could be vulnerable to money pumps, and, more distressingly for economists wishing to apply these models in the field, a stable choice function (under Tversky and Kahneman's model) or a personal equilibrium consistent with one's expectations (under Köszegi and Rabin's model) does not always exist (see Gul & Pesendorfer, 2006).

Recent work on the cognitive processes underlying choice suggests an alternate approach to studying reference dependence. A number of researchers (Ashby, Dickert, & Glöckner, 2012; Ashby, Walasek, & Glöckner, 2015; Bhatia, 2013, 2017b; Carmon & Ariely, 2000; Johnson, Häubl, & Keinan, 2007; Nayakankuppam & Mishra, 2005; Pachur & Scheibehenne, 2012; Willemsen, Backenholt, & Johnson, 2011) find that endowments direct the decision maker's attention toward their most prominent attributes. This is subsequently shown to generate an increased weight on these attributes, altering

Sudeep Bhatia, Department of Psychology, University of Pennsylvania; Russell Golman, Department of Social and Decision Sciences, Carnegie Mellon University.

Correspondence concerning this article should be addressed to Sudeep Bhatia, Department of Psychology, University of Pennsylvania, Solomon Labs, 3701 Walnut Street, Philadelphia, PA 19104. E-mail: bhatiasu@sas.upenn.edu

preferences in favor of the endowment and producing the endowment effect.

These findings suggest that reference-dependent behaviors are driven by attentional processes. Reference points do not act as frames, altering perceptions of gains and losses. Rather they act as primes, directing attention toward information that they are strongly associated with. As choice objects are most strongly associated with their most prominent attributes, it is these attributes that receive a higher weight in the decision task. Changing reference points can thus affect the weights on these attributes and, consequently, the choice between competing objects.

This article presents a model of reference dependence motivated by this intuition. The attention-biased utility of an object is a linear combination of valuations for each attribute of the object. Although attribute valuation is stable and reference independent, the weight on each attribute in attention-biased utility depends on the attention devoted to that attribute, which in turn depends on the amount of that attribute in the reference object. Changing the reference point alters the attentional weights in the attention-biased utility function. This affects the decision maker's choice and can generate preference reversals.

This model requires minimal deviations from standard consumer theory. It does not make any assumptions about the curvature of valuation functions: If valuation functions are strictly concave then so is attention-biased utility. In addition, the model does not assume a built-in asymmetry in gains and losses. As a result, desirable theoretical properties, such the existence of continuous demand functions (given a reference point), are guaranteed with standard assumptions. Nonetheless, reductions in consumption have a stronger impact on attention-biased utility than corresponding gains, generating the well-known phenomenon of loss aversion. Indeed, the proposed attention-biased utility model can explain many of the observed anomalies in the domain of reference-dependent riskless choice.

Although attention-biased utility can parsimoniously capture many of the empirical phenomena explained by prospect theory and other behavioral economic models, it also generates novel predictions that distinguish it from these earlier theories. For example, attention-biased utility predicts reversals of reference-dependent effects for choices involving attributes with negative and strictly decreasing valuations. Brenner, Rottenstreich, Sood, and Bilgin (2007) have documented reversals of the endowment effect for undesirable objects, suggesting that reference points do indeed operate differently in the negative domain.

Sequential choice is of particular interest in a reference-dependent context because of the possibility of preference reversals. This article shows that maximization of attention-biased utility for sequential decisions between economic goods generates long-term choice behavior that satisfies basic consistency requirements. Decision makers do not cycle through available goods indefinitely (and are not vulnerable to money pumps); rather, choices are guaranteed to stabilize, and a long-run choice function, or an expectations-consistent personal equilibrium, is guaranteed to exist.

In a related vein, this article also considers the choice of reference points that maximize utility. We establish that the impact of a reference point on utility from a particular object is nonmonotonic, and high reference points are often optimal. This result helps us understand the common preference for attainable but nontrivial outcomes as aspirational reference points, or goals (Diener, Suh, Lucas, & Smith, 1999; Locke, Shaw, Saari, & Latham, 1981). With a more general conception of reference points, the model also accommodates other behavioral anomalies, such as choice set effects (Huber, Payne, & Puto, 1982; Simonson, 1989) and anchoring effects (Ariely, Loewenstein, & Prelec, 2003; Johnson & Schkade, 1989), which have recently been attributed to attentional biases.

The next section reviews research in economics, psychology and neuroscience on the role of attention in choice. The Model of Attention-Biased Utility section provides a formal theory of attention-biased utility. The Implications section explores the implications of attention-biased utility for loss aversion and reference dependent choice. The Novel Predictions section offers predictions regarding undesirable choice objects along with other novel predictions. The Endogenous Reference Points section analyzes sequential choices as well as optimal reference points with attention-biased utility. The Multiple Reference Points section extends our theory to explain observed instances of anchoring and choice set dependence and is followed by the conclusion.

Attention and Choice

Attention is one of the most important psychological variables in behaviorally motivated economic theories of choice. Herbert Simon's early approach to understanding deviations from rationality was entirely driven by attention-based constraints on the decision maker's ability to process information and explore the choice set (Simon, 1955). Sims (2003), recognizing costs associated with information processing, proposed that people rationally direct their attention to valuable information sources. Gabaix, Laibson, Moloche, and Weinberg (2006) have proposed a model that captures boundedly rational attentional allocation in costly information acquisition tasks. This model outperforms fully rational attention allocation and makes a number of powerful predictions regarding information acquisition in markets and societies. Caplin, Dean, and Martin (2011) have further explored attentional allocation in large and complex choice sets and have found that decision makers frequently use Simon's (1955) satisficing heuristic to make choices. In the same vein, Masatlioglu, Nakajima, and Ozbay (2012) have provided a theoretical framework for inferring attentional allocation from choice behavior, and Matejka and McKay (2014) have studied the implications of rational attention allocation for probabilistic choice.

Although people can often choose to voluntarily direct their attention as they see fit, attention can also be drawn spontaneously and involuntarily to salient or surprising stimuli. In some cases, choices may be motivated in part by the desire to exert some control over the stimuli that spontaneously attract attention (Golman & Loewenstein, 2016). In other cases, biased attention, or accessibility, plays a key role in shaping choice at a subconscious level (Kahneman, 2003). Fehr and Rangel (2011) argued that attention can affect how attributes are weighted in a decision value computation.

Using this approach, behavioral economists have proposed various attentional biases to explain a range of observed behavioral anomalies. Bordalo, Gennaioli, and Shleifer (2012a), for example, provided an attentional explanation for many of the classical experimental findings on risky choice, including violations of the independence assumption and risk seeking behavior in losses. Bordalo, Gennaioli, and Shleifer (2012b) use a similar approach to explain the endowment effect as well as its reversal, and Bordalo, Gennaioli, & Shleifer, (2015) do so for biases in consumer choice. Köszegi and Szeidl (2013) studied intertemporal choice using an attentional model of attribute weighting. Their model generates present-biased behavior and, in addition, provides a range of predictions about the settings in which this behavior is most pronounced. Golman, Loewenstein, Molnar, and Saccardo (2017) proposed that attention to the presence or absence of information leads to demand for good news and avoidance of bad news, and Golman, Loewenstein, and Gurney (2017) used the same attentional model to account for source preference under risk and ambiguity.

Understandably, attention is also of considerable interest to scholars of decision making outside of economics. As a well understood and easily observable cognitive variable, attention may naturally play a role in many of the key mechanisms underlying choice. Recently, psychologists and neuroscientists interested in these mechanisms have started to explore the various determinants and consequences of attention in decision making. Carmon and Ariely (2000), for example, find that decision makers direct their attention toward the attribute of the objects they possess: Owners focus on attributes of the traded object, whereas nonowners are more likely to focus on the expenditure involved in the trade. Because increased attention to a particular attribute increases that attribute's weight in the decision, this attentional bias leads to a discrepancy in buying and selling prices, generating the endowment effect.

Similar results are noted by Nayakankuppam and Mishra (2005), who found that owners are more likely to attend to the attributes that the endowed object is strongest on, and less likely to attend to the attributes that the endowed object is weakest on, relative to nonowners. Johnson et al. (2007) replicated these findings and, in addition, discovered differences in the order that owners and nonowners attend to the various attributes, in the decision task. Decision makers generally attend to the strongest attributes of the object that they possess before focusing on the weakest attributes of their possessed object, or attributes of objects that they do not own. Willemsen et al. (2011) extended these findings beyond the endowment effect. Ashby et al. (2012, 2015) used eye-tracking and response time restrictions, and Bhatia (2017b) used quantitative model fits, to further demonstrate the role of biased attention in reference dependence. Finally, Pachur and Scheibehenne (2012) demonstrated the existence of these attentional biases for risky choice. These articles note that the attentional biases displayed by the decision makers can predict buying and selling prices, and choice probabilities. In addition, altering where these decision makers focus their attention can eliminate the endowment effect.

The endowment effect can also be generated without explicit endowments. Dhar and Simonson (1992) and Dhar, Nowlis, and Sherman (1999) found that increasing the salience of a desirable item increases its share in the choice set, relative to its competitors. In contrast, increasing the salience of an undesirable item decreases its share relative to its competitors. Importantly, this behavior stems from biased attention toward the focal item's attributes. Reducing this attentional bias can eliminate differences in choice shares.

Bushong, King, Camerer, and Rangel (2010) presented related findings. They noted that the physical presence of a good increases that good's desirability and subsequently increases the decision maker's willingness-to-pay for the good. Other research by Krajbich, Armel, and Rangel (2010) and Reutskaja, Camerer, and Rangel (2011) suggested that visual attention toward an item has a direct relationship with the decision maker's preference toward that item. This has been documented for both actual choice and the neural mechanisms that determine choice (Lim, O'Doherty, & Rangel, 2011).

These results suggest that endowments may not act as frames, as assumed in prospect theory and related models of reference dependence. Rather they act primes, directing the decision maker's attention toward relevant attributes. Although this intuition is sufficient to explain some of the above-mentioned results, a formal model is necessary to explore its implications for other reference-dependent anomalies, as well as for anomalies in other domains. A formal model can also highlight similarities and differences between an attention-based model of reference dependence and prospect theory, as well as standard theories of rational choice.

Model of Attention-Biased Utility

Consider an N-attribute choice space consisting of objects $x \in \mathbb{R}_{+}^{N}$. The decision maker is assumed to choose from a choice set $X \subset \mathbb{R}_{+}^{N}$, given a reference point $r \in \mathbb{R}_{+}^{N}$. We assume that there exist N strictly monotonic valuation functions $V_i = V_i(x_i)$ corresponding to the decision maker's reference-independent valuation of attribute i in x. For simplicity, we will set $V_i(0) = 0$ for all i.

Under the assumption that attributes can be valued independently, rational choice of $x \in X$ is consistent with maximization of utility $U^*(x) = \sum_{i=1}^{N} V_i(x_i)$. We assume, however, that the decision maker is subject to reference-dependent attentional biases, according to which her weight on each attribute depends on the reference object's amount of that attribute. Specifically, we assume that there exist N non-negative and strictly increasing attention functions $\alpha_i = \alpha_i(r_i)$ representing the decision maker's attentional weight on attribute i given a reference point r. The decision maker chooses according to $\tilde{U}(x \mid r) = \sum_{i=1}^{N} \alpha_i (r_i) \cdot V_i(x_i)$. Note that we can add a constant to overall utility without changing the underlying preferences. Thus, we can normalize $\tilde{U}(x \mid r)$ so that the overall choice utility of the reference point is zero. We obtain the utility function $U(x \mid r)$ with:

$$U(x|r) = \sum_{i=1}^{N} \alpha_i(r_i) \cdot V_i(x_i) - \alpha_i(r_i) \cdot V_i(r_i). \quad (1)$$

We will refer to any function of the above form[1] as an *attention-biased utility function*. If

[1] In general, preferences may not always permit an additively separable representation based on attributes that can be valued independently. In these settings we might assume that there exists a function $f: \mathbb{R}^N \rightarrow \mathbb{R}^M$, mapping physical (nonseparable) attributes, to separable mental attributes. Thus, for any x and r we would obtain $f(x) = y$ and $f(r) = s$. Attention and valuation functions could subsequently be defined on mental attributes instead of physical attributes, and Equation 1 would become $U(x \mid r) = \sum_{j=1}^{M} \alpha_j(s_j(r)) \cdot V_j(y_j(x)) - \alpha_j(s_j(r)) \cdot V_j(s_j(r))$. Of course, the freedom to define an unobservable set of mental attributes would make the model so flexible as to be practically unfalsifiable. To make testable predictions, we should specify the function mapping physical attributes into mental attributes. Taking this function to be the identity mapping, as we do, is a convenient simplification when appropriate.

$\alpha_i(r_i) = c$ for a constant c, for all i, then the decision maker does not display any reference-dependent attentional bias and simply maximizes total value, U^*.

Attribute Priming and Attention-Biased Utility

In this article we apply attention-biased utility to study choices between multiattribute objects that are offered to the decision maker without delay or uncertainty (as in e.g., Keeney & Raiffa, 1993). Our use of attention-biased utility in this domain is motivated primarily by the experimental results of Carmon and Ariely (2000), Johnson et al. (2007), and others. This work has consistently found that the attributes present in the reference object are more salient and thus have higher weights in the decision. This is reflected in the strictly increasing attention functions $\alpha_i(r_i)$. Although this experimental finding is fairly robust and conclusive, it may be useful to briefly outline a theoretical justification for why reference points should have such an effect on attention.

People are sensitive to their context, so that salient but irrelevant items in their immediate environment can nonetheless be attended to and processed, and subsequently influence cognition and behavior. For example, presenting individuals with a word such as *dog*, activates the mental representation of a dog, which then activates the various attributes associated with dogs. Individuals who have been presented with this word are more likely to think about dog-related attributes and use the concept *dog* in subsequent (unrelated) cognitive tasks. These individuals are also more sensitive to dog-related attributes in other items, and are thus more likely to process concepts like *cat*, which share a lot of their salient attributes with dogs. These effects, known as priming effects, are a well-documented feature of human cognition across multiple domains, including perception, language, memory, and categorization (see Neely, 1991 or Wiggs & Martin, 1998, for reviews). The findings documented by Carmon and Ariely (2000) and Johnson et al. (2007), and others can be seen as an example of priming effects in the domain of preferential choice. Salient (but possibly irrelevant) choice objects, such as reference points, influence choice in roughly the same way that salient words influence language processing or salient images influence object perception and recognition: They increase the activation of their associated attributes, making these attributes more important for the task at hand.

Constraints on Attention-Biased Utility

Strictly speaking, our attention-biased utility model in Equation 1 is a class of reference-dependent utility functions with freedom to choose appropriate attention functions for each attribute. In principle, the attention functions could be specified by collecting data about attentional capture during the choice process. In practice, the attention-biased utility model is useful for making qualitative predictions that do not depend on the specific functional form of the attention functions. In the sections below, we describe behavioral predictions that hold generally for any specification of the attention-biased utility model. The simple assumption that attention to an attribute is increasing in the amount of that attribute in the reference point suffices to explain the full set of reference-dependent behavioral patterns observed in prior experimental work. Still, if we were to augment our model with additional assumptions about the determinants of attention, we would be able to derive more specific predictions.

Additional constraints on the attention functions limit the flexibility of the model, permitting formal model testing. When attributes are symmetric, we might make the simplifying assumption that the attention functions should take on the same functional form for all attributes. Of course, attention toward attributes could be different, reflecting, for example, attentional biases independent of the reference point. However, it is likely that the relationship between reference points and attribute attention is similar across attributes, and that a common functional form suffices across different attributes. For additional parsimony, the attention functions could simply be specified as a linear function of the attributes in the reference point, as in Bhatia's (2017b) recent model fitting work. Different attributes may still have differences in their baseline activation but all attributes would then described by the same linear functional form.

This constraint is incorporated by an existing cognitive model of dynamic attention in multiattribute choice (Bhatia, 2013). Bhatia shows that a simple class of associative networks embedded in a sequential sampling framework generates the reference-dependent attentional biases discussed here. Static attention functions for the attention-biased utility model could be derived from this dynamic model of attention following the approach laid out by Johnson and Busemeyer (2016). Thus, Bhatia (2013) allows us to interpret the ratio $\frac{\alpha_i}{\alpha_j}$ as the odds of attending to attribute i relative to attribute j at any point in time during the decision process (a property of the model drawn from prior dynamic theories of multiattribute choice—see Roe, Busemeyer, & Townsend, 2001, as well as Bhatia, 2014; Busemeyer & Townsend, 1993; Johnson & Busemeyer, 2016 for related work for risky choice).

Bhatia's (2013) model also allows us to specify what a reference point is in psychological terms. According to the proposed theory, any choice option that is more salient than the other options in the choice set, will exert a stronger influence on attribute attention. This will alter the weighting of these attributes in the decision task, and bias final preferences. In this view, reference points are merely options that are salient to the decision maker. This intuition will help us extend the theory of attention-biased utility to domains beyond reference dependence in later sections.

There are still many other influences on attention that are not directly incorporated into the structure of attention-biased utility. Extrapolating from research about visual attention in multiattribute choice (see Orquin & Loose, 2013, for a review), attention may be attracted toward attributes with high contrast between choice objects (Duncan & Humphreys, 1989), toward attributes with abrupt onset (Yantis & Jonides, 1984), or toward surprising attributes (Itti & Baldi, 2009). Attention may also be drawn to attributes on which choice objects take on extreme values within the range or distribution of feasible attribute values. And, of course, attention can then be shaped by other forms of priming (Neely, 1991) or distraction (Lavie, 2005). All of these influences on attention are compatible with attention-biased utility, as long as attention to an attribute remains increasing in the amount of that attribute in the reference point.

Dimensional Reference Dependence

Attention-biased utility makes no assumptions about the shape of the valuation functions V_i. If V_i are strictly increasing and everywhere concave then so is U, regardless of the reference point. There are also no asymmetries in gains and losses; rather the impact of the reference point is dimensional: Objects that are strongest on the reference point's primary attributes are the ones that are chosen. This is in contrast to models of reference dependence based on prospect theory (Köszegi & Rabin, 2006; Tversky & Kahneman, 1991) in which built-in gain–loss asymmetries generate preferences for objects that are unambiguous gains over the reference point, relative to objects that involve tradeoffs from the reference point.

Our attention-biased utility model resembles Bordalo et al.'s (2012b) recent attention-based explanation of the endowment effect, which also relies on dimensional attention weights rather than an explicit asymmetry in gains and losses. According to that theory, however, the attention toward a particular attribute in an object depends (discontinuously) on the dispersion of the amount of that attribute across the choice set, and the endowment effect only arises if people seemingly sometimes forget to consider the entire choice set when valuing the endowed good. Moreover, Bordalo et al.'s (2012b) model fails to account for a reference dependent choice anomaly documented by Herne (1998), Effect 5 in the Choices section,[2] as well as a number of other effects discussed in this article.

Dimensional reference dependence is also a property of an earlier model proposed by Munro and Sugden (2003). Although their constant elasticity of substitution based utility differs from attention-biased utility in many important ways, both forms create a preference for objects based on their dimensional overlap with the reference point, rather than their position as gains or losses relative to the reference point.

This dimensional bias allows for a particularly convenient geometric interpretation. Consider the simple case where $\alpha_i(r_i) = V_i(r_i)$ for all i, a setting in which attention is proportional

[2] This is the finding that decision makers' preference for objects that dominate the reference point are stronger for extreme valued reference points even when the reference point is dominated on its secondary (nonextreme) attribute.

to valuation (as would be the case if the decision maker attends to attributes based on their desirability). Here, reference-dependent utility can simply be captured as a dot product of the vector of valuations of the reference point, $V_r = \langle V_1(r_1), V_2(r_2), \ldots, V_N(r_N)\rangle$, with the vector of valuations of the evaluated object, $V_x = \langle V_1(x_1), V_2(x_2), \ldots, V_N(x_N)\rangle$. The decision maker in turn chooses the object whose vector of valuations has the highest projection onto V_r. Changing x so as to increase V_x in the direction of V_r will lead to the highest increase in utility from x, whereas changing x to increase V_x in a direction orthogonal to V_r will lead to absolutely no increase in utility for x. One intuitive implication of this is that if the attributes of the reference point do not overlap with those of the object in consideration, adding more of the reference point's attributes to the object has no effect on utility.

Figure 1 demonstrates indifference curves generated by attention-biased utility in a two-attribute choice space. I_x and I_y are indifference curves for settings where x and y are reference points, respectively. These curves intersect indicating the possibility of preference reversals.

Figure 2 presents the same scenario in valuation space. V_x and V_y are the valuation vectors of x and y, I_{V_x} and I_{Vy} are indifference curves for settings where x and y are reference points, respectively. I_{V_x}, for example, consists of all valuation vectors V_z such that object z is indifferent to x when x is the reference point. Valuation vectors lying above I_{V_x} correspond to choice alternatives that are preferred over x when x is the reference point. The opposite is true for vectors lying below I_{V_x}. When x is the reference point, all valuation vectors with the same projection onto V_x will lie on the same indifference curve. Also note that the projection of V_y onto V_x is smaller than the projection of V_x onto itself (i.e., its magnitude), demonstrating that x is preferred over y when x is the reference point. The opposite holds for the projection of V_x and V_y onto V_y, indicating that y is preferred to x when y is the reference point.

Although this example is valid only for the setting where $\alpha_i(r_i) = V_i(r_i)$ for all i, the intuition behind it holds for more general cases as well. For strictly increasing V_i, objects for which V projects maximally onto α are the ones that are chosen. α itself is a function of r. Of course α_i can depend on r_i differently for different i, implying that attentional biases may vary across attributes.

This intuition allows us to derive the first result of this article. Proposition 1 describes the general setting in which reference points

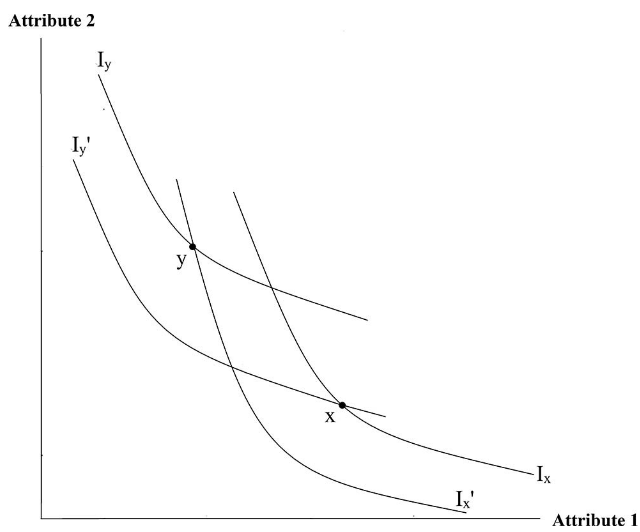

Figure 1. Indifference curves I_x and I_y when x and y, respectively, are the reference points. These curves intersect indicating the possibility of preference reversals.

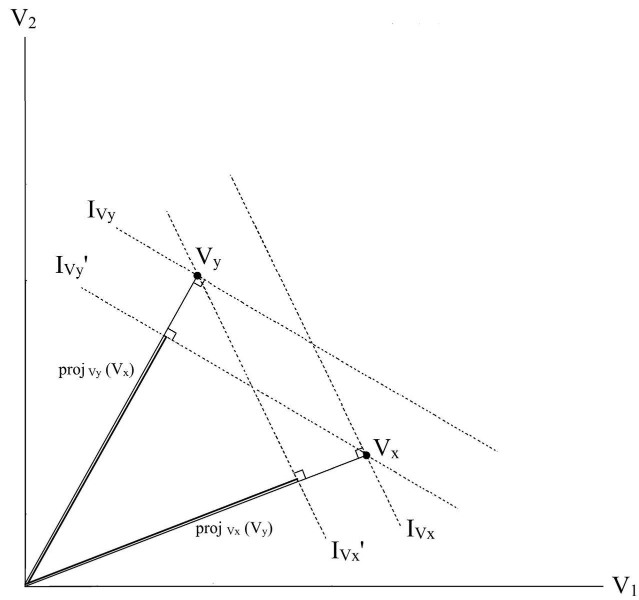

Figure 2. Indifference curves I_{V_x} and I_{V_y} in valuation space, when x and y respectively are the reference points. The dotted lines correspond to the projection of V_x and V_y onto themselves and each other.

necessarily change utility in favor of a particular object relative to another, for strictly increasing V_i. It states that the attention-biased utility function exhibits increasing differences in the consumption object and the reference point on each attribute, that is, a reference point r will necessarily bias preferences in favor of x over y, relative to another reference point s, if r is not smaller than s on the attributes that favor x, and s is not smaller than r on the attributes that favor y, and if r and s differ on at least one attribute on which x and y differ. Put simply, this means that having an attribute in the reference object complements consumption of that attribute.

> **Proposition 1.** If all V_i are strictly increasing, then the attention-biased utility function $U(x \mid r)$ exhibits strictly increasing differences in (x_i, r_i) for all i.

Proposition 1 holds because reference points that are strongest on the attributes that favor x relative to y are the ones that direct the most attention toward these attributes. Increased attention leads to higher weighting which then amplifies any differences in valuation on these attributes.

Implications

This part of the article studies how reference points affect preferences between objects. It first shows that attention-biased utility implicitly generates loss averse preferences without requiring any nonstandard assumptions on utility except for reference-dependent attentional weighting. The subsequent section outlines how attention-biased utility can be used to explain a range of reference-dependent anomalies documented in the behavioral literature.

Loss Aversion

A number of researchers have noted that decision makers are loss averse. Moving from an inferior consumption state y to a superior consumption state x affects the decision maker's utility less than the equivalent move from x to y. Thus, changes to consumption that are perceived as losses loom larger than changes perceived as gains (Thaler, 1980; Tversky & Kahneman, 1991).

Although loss aversion has been established indirectly through revealed preference, it is assumed to be a psychologically realistic property as well. As discussed in Novemsky and Kahneman (2005), most researchers accept loss aversion as both a description and an explanation of the phenomenon being studied. This psychological fact motivates the standard account of reference dependence, incorporated in Tversky and Kahneman's (1991) prospect theory as well as in Köszegi and Rabin's (2006) more recent expectations based framework. These models assume that negative changes relative to the reference point affect utility more than positive changes relative to the reference point. Particularly, there is a kink in utility at reference point r, such that a change in consumption from r to $r + \delta$ (for $\delta \in \mathbb{R}_+^N$) is less desirable than the change from r to $r - \delta$ is undesirable.

Note, however, that the explicit asymmetry assumed in these models is between a gain and a loss that are not quite comparable. The difference between a gain and a loss is conflated with a wealth effect. In principle, loss aversion should be identified by comparing a gain and a loss involving the same levels of consumption. Although prospect theory's kink in utility at the reference point can induce loss aversion, that is, can make the negative change from x to y matter more than the equivalent positive change from y to x, it is not the only mechanism to do so. Indeed Proposition 2 shows that loss aversion also emerges from attention-biased utility with strictly increasing V_i. This is an implication of the model, rather than a built-in assumption.

Proposition 2. If V_i are strictly increasing, and x and y are any two objects such that $x_i \geq y_i$ for all $i \in \{1, 2, \ldots, N\}$, and $x_i > y_i$ for some $i \in \{1, 2, \ldots, N\}$, we have $U(x|x) - U(y|x) > U(x|y) - U(y|y)$.

The intuition for this result is the following: for increasing V_i, a superior reference point directs more attention toward at least one underlying attribute, relative to a dominated reference point. The attributes that receive a higher attentional weight are precisely those that the superior reference point dominates on. This amplifies any difference in the consumption utility of these attributes, leading to a greater effect on utility for negative deviations from the reference point (moving from a superior to inferior consumption state) relative to equivalent positive deviations (moving from an inferior to a superior consumption state). In other words, losses involve more attention to the attributes being lost than gains do to attributes being gained. This leads to a greater decrease in utility for losses, than an increase in utility for gains, corresponding to a type of loss aversion.[3]

Behavior

The mechanism responsible for the emergence of loss aversion from attention-biased utility can also explain findings regarding endowments and other reference points. These findings have generally been documented by using either monetary measures of preference, such as willingness-to-pay or willingness-to-accept, or through explicit choices between two or more items. The next section explores the implications of attention-biased utility with regards to monetary measures of preference. The Choices section explores reference-dependent choices.

Measures of preference. Let us consider four different measures of preference, as formalized in Bateman, Munro, Rhodes, Starmer, and Sugden (1997): willingness-to-pay (WTP), willingness-to-accept (WTA), equivalent-loss (EL), and equivalent-gain (EG). The endowment of a particular item can be represented by the superior reference state, x, whereas not being endowed with the item can be represented by the dominated reference state, y. For our present analysis, we can limit ourselves to two attributes, i and j, and use these measures to study how changes to attribute i impact the decision maker's preferences in units of attribute j. We will hold, $x_i > y_i$ and $x_j = y_j$. Because of the independence between attributes for consumption utility and attention, this two-attribute setting easily generalizes to more complex cases.

Because we are only considering two attributes, we can write preferences for any choice z given any reference point r as $U(z_i, z_j \mid r)$, and

[3] The built-in gain–loss asymmetry in prospect theory provides one account for anomalous low-stakes risk aversion, which Rabin (2000) showed cannot be accounted for with utility function curvature. The attention-biased utility function presented here, while generating loss aversion, does not by itself generate this kind of risk aversion. However, a related model of utility derived from beliefs, which similarly relies on attention weights, can provide an alternative account for risk aversion over small gambles (Golman, Loewenstein, & Gurney, 2017).

the four measures used to measure preference can be defined as:

1. $U(x_i, x_j - \text{WTP} \mid y) = U(y_i, x_j \mid y)$
2. $U(y_i, x_j + \text{WTA} \mid x) = U(x_i, x_j \mid x)$
3. $U(x_i, x_j - \text{EL} \mid x) = U(y_i, x_j \mid x)$
4. $U(y_i, x_j + \text{EG} \mid y) = U(x_i, x_j \mid y)$

Willingness-to-pay by this definition is the largest loss on attribute j that the decision maker is willing to incur to increase consumption from y_i to x_i. Likewise, willingness-to-accept is the lowest gain on attribute j that the decision maker is willing to accept to reduce consumption from x_i to y_i. Equivalent-loss is the largest reduction of attribute j that the decision maker is willing to incur to avoid reducing consumption from x_i to y_i. Finally, equivalent-gain is the smallest increase in attribute j that the decision maker is willing to accept to avoid increasing consumption from y_i to x_i.

Standard Hicksian theory predicts that EL = WTP and EG = WTA (see Bateman et al., 1997, for a discussion). However, a number of researchers find that both EG and EL ratings deviate from WTP and WTA respectively (Bateman et al., 1997; Kahneman, Knetsch, & Thaler, 1990; Knetsch, 1989; Loewenstein & Adler, 1995). Individuals are willing to pay more to avoid losing an item than they are willing to pay to gain the item, generating EL > WTP. Likewise, individuals require more money to give up an item that they own than they do to forego acquiring an item that they do not own, generating EG < WTA. These results are often seen to be direct implications of the built-in gain–loss asymmetry assumed in prospect theory and its generalizations. However, Proposition 3 shows that these results are also generated by attention-biased utility.

Proposition 3. For strictly increasing V_i, and for any two objects x and y, with $x_i > y_i$ and $x_j = y_j$, attention-biased utility generates EL > WTP and EG < WTA.

The intuition for this result is related to that for Propositions 1 and 2. Because $x_i > y_i$, reference point x leads to increased attention toward attribute i relative to reference point y. On the other hand, because $x_j = y_j$ attention toward attribute j is constant regardless of the reference point. This leads to a bias in favor of attribute i when x is the reference point relative to when y is the reference point, generating the observed inequalities between EL and WTP, and EG and WTA. In other words, the higher attention to the attributes of a possessed item make it so that individuals are willing to pay more to avoid losing the item than they are willing to pay to gain the item. Likewise, individuals require more money to give up an item that they own than they do to forego acquiring an item that they do not own. Of course, these inequalities also imply that WTP < WTA, so that the willingness to pay to obtain an item is lower than the willingness to accept to give up the item.

Choices. Reference dependence is also associated with a number of behavioral tendencies involving explicit choices between objects. For example, Knetsch and Sinden (1984), Samuelson and Zeckhauser (1988), and Knetsch (1989) found that the endowment or status quo is more likely to be chosen from the choice set, relative to competing objects. Knetsch (1989) also found that decision makers are more likely to select one object over another if they are endowed with the first object, compared to the setting in which they are endowed with neither of the two objects (see also Morewedge & Giblin, 2015).

In addition, reference dependence can be observed when the most frequently chosen object is not the reference point itself. For example, Tversky and Kahneman (1991) noted that decision makers generally prefer choice objects that are strict improvements over their reference point, relative to objects that involve tradeoffs with their reference point. Similar findings have been replicated by Herne (1998), who also finds that extreme reference points that are very weak on some attributes and very strong on other attributes lead to stronger biases in preference, relative to more evenly distributed reference points. Finally, Tversky and Kahneman (1991) noted that decision makers tend to prefer objects that involve small tradeoffs from the reference point relative to objects that involve larger tradeoffs from the reference point. These biases can generate preference reversals as reference points are varied.

These results have been experimentally documented with two-attribute choice sets consisting of two desirable objects x and y that do not dominate each other and two or more reference points, r and s, that may or may not be dominated by x or y. If we write $P_{x,y}(r) = U(x \mid r) - U(y \mid r)$, as the relative preference for x over y

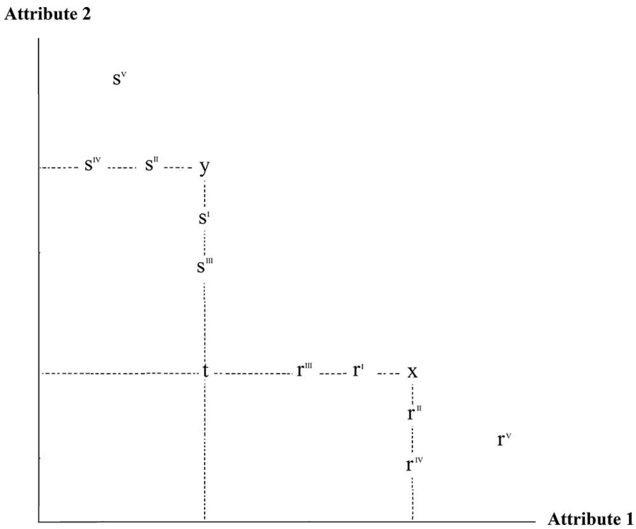

Figure 3. Reference-dependent anomalies in choice.

when r is the reference point, and consider any x and y such that $x_1 > y_1$ and $y_2 > x_2$, then based on the choice objects in Figure 3, the above effects can be written as:

1. Conservatism (Knetsch, 1989; Knetsch & Sinden, 1984; Samuelson & Zeckhauser, 1988): If t is such that $t_1 = y_1$ and $t_2 = x_2$, then $P_{x,y}(x) > P_{x,y}(t) > P_{x,y}(y)$.
2. Inner improvements versus tradeoffs (Herne, 1998; Tversky & Kahneman, 1991): If r^I and s^I are such that $x_1 > r_1^I > y_1 = s_1^I$ and $y_2 > s_2^I > x_2 = r_2^I$, then $P_{x,y}(r^I) > P_{x,y}(S^I)$.
3. Outer improvements versus tradeoffs (Herne, 1998): If r^{II} and s^{II} are such that $x_1 = r_1^{II} > y_1 > s_1^{II}$ and $y_2 = s_2^{II} > x_2 > r_2^{II}$, then $P_{x,y}(r^{II}) > P_{x,y}(S^{II})$.
4. Inner extreme versus balanced reference points (Herne, 1998): If r^I, s^I, r^{III} and s^{III} are such that $x_1 > r_1^I > r_1^{III} > y_1 = s_1^I = s_1^{III}$ and $y_2 > s_2^I > s_2^{III} > x_2 = r_2^I = r_2^{III}$, then $P_{x,y}(r^I) - P_{x,y}(S^I) > P_{x,y}(r^{III}) - P_{x,y}(S^{III})$.
5. Outer extreme versus balanced reference points (Herne, 1998): If r^{II}, s^{II}, r^{IV} and s^{IV} are such that $x_1 = r_1^{II} = r_1^{IV} > y_1 > s_1^{II} > s_1^{IV}$ and $y_2 = s_2^{II} = s_2^{IV} > x_2 > r_2^{II} > r_2^{IV}$, then $P_{x,y}(r^{IV}) - P_{x,y}(S^{IV}) > P_{x,y}(r^{II}) - P_{x,y}(S^{II})$.
6. Small versus large tradeoffs (Tversky & Kahneman, 1991): If r^V and s^V are such that $r_1^V > x_1 > y_1 > s_1^V$ and $s_2^V > y_2 > x_2 > r_2^V$, then $P_{x,y}(r^V) > P_{x,y}(S^V)$.

Conservatism (Effect 1) is the finding that decision makers tend to stay with their reference point instead of choosing competing objects. The improvements versus tradeoffs effects pertain to the finding that decision makers prefer objects that dominate the reference point compared to objects that involve tradeoffs from the reference point. These have been documented both when the reference point is dominated on its primary attribute (Effect 2) and when it is dominated on its secondary attribute (Effect 3). The extreme versus balanced reference point effects pertain to the finding that these improvements-versus-tradeoffs effects are stronger for extreme valued reference points compared to moderate valued reference points, both when the reference point is dominated on its primary attribute (Effect 4) and when it is dominated on its secondary attribute (Effect 5). Finally, the small versus large tradeoffs effect (Effect 6) corresponds to the finding that objects that involve small tradeoffs from the reference point are preferred over objects that involve large tradeoffs from the reference point.

Tversky and Kahneman (1991) used Effects 1, 2, and 6 to justify the application of the prospect theory valuation function to capture reference dependence in riskless choice. Effects 1 and 2 are explained through the gain–loss asymmetry, whereas Effect 6 requires both the gain–loss asymmetry and convex utility in losses. Herne (1998) provided further evidence for the descriptive power of the prospect theory valuation function. Effect 3 is explained by the gain–loss asymmetry, whereas Effects 4 and 5 are explained by both the gain–loss asymmetry and convex utility in losses.

Effects 1–6 were initially compiled by Munro and Sugden (2003), who also noted that one general condition implied Effects 1–6. This condition is as follows:

Condition 1. For all attributes i, j, and for all objects x and y, such that $x_i > y_i$, $y_j > x_j$, and $x_k = y_k$ for all $k \neq i, j$, and for all reference points r and s, such that $r_i > s_i$ and $r_k = s_k$ for $k \neq i$, we have $P_{x,y}(r) > P_{x,y}(s)$.

Munro and Sugden were reluctant to propose Condition 1 as a fundamental property of reference-dependent choice, as they did not have any explanation for why reference points may bias preferences in this way. Note however that Proposition 1 implies that any attention-biased utility function with strictly increasing V_i satisfies Condition 1. Hence Condition 1 is a natural implication of attentional processes underlying reference-dependent choice. According to the proposed model, the move from s to r favors x relative to y because it leads to increased attention toward, and subsequently a higher weight on attribute i. Because x is more valuable on attribute i relative to y, this biases choice in favor of x instead of y. The fact that attention-biased utility satisfies Condition 1, also implies that it can capture Effects 1–6. This gives us Proposition 4.

Proposition 4. Attention-biased utility with strictly increasing V_i generates reference-dependent Effects 1–6.

Note that Proposition 4, like Propositions 1–3, holds regardless of the choice of consumption utility functions. If V_i are assumed to be globally concave, then a globally concave U is able to generate effects 1–6. Linear V_i can also generate these effects. Built-in gain–loss asymmetries in valuation along with convex utility in losses, as assumed in prospect theory or other standard models of reference dependence, are not necessary. Effects 1–6 all follow from the dimensional weighting mechanism at play in attention-biased choice.

Also note that our explanation of Effects 1–6 is not limited to reference points generated by endowments, the status-quo, or expectations. According to our theory, any focal object will act as a reference point and will bias choice in the manner predicted by Effects 1–6. When applied to Effect 1 (conservatism), this can explain the findings of Bushong et al. (2010), Dhar and Simonson (1992), Dhar et al. (1999), Krajbich et al. (2010), Lim et al. (2011), and Reutskaja et al. (2011).

Novel Predictions

The previous section provides an analysis of attention-biased utility for desirable goods with strictly increasing V_i. These goods, like mugs, chocolates, or pens, are by far the most commonly used stimuli in experiments on reference dependence. Recent research has, however, begun to examine reference dependence with regards to undesirable objects. Although reference points like endowments are more likely to be chosen relative to their competitors in choices among desirable goods, this work finds that endowments are less likely to be chosen in choices among undesirable goods (Brenner et al., 2007).

The proposed framework allows for a simple interpretation of these findings. According to attention-biased utility, endowments and other reference points increase the salience of their attributes. This generally leads to an increased preference for desirable endowments. Undesirable reference points, however, have undesirable attributes. Increased attention toward these attributes should reduce the desirability of the endowment. This is formalized in the following proposition.

Proposition 5. If V_i are strictly decreasing, and if x and y are any two nonidentical objects, then we have $U(x \mid x) - U(y \mid x) < U(x \mid y) - U(y \mid y)$.

Though the attention-biased utility form for desirable objects is identical to that for undesirable objects, many of its properties differ. Applying Equation 1 with strictly decreasing valuation functions, we find a reversal of the endowment effect for undesirable goods. Attention-biased utility suggests that in the domain of undesirable goods, a decision maker continually wants what she does not have. As conveyed by the popular adage, the grass is always greener on the other side.

Note that the result presented here also holds for salient choice options that are not endowments. Indeed, Dhar et al. (1999) have found that directing attention to undesirable objects, makes them less likely to be chosen. In contrast, prospect theory and associated models of reference dependence are unable to capture either reversals of the endowment effect in then negative domain or the preference for nonfocal undesirable objects. A gain–loss asymmetry is assumed to apply regardless of the underlying valence of the choice object. Hence biases in favor of the reference point are expected to emerge for both desirable and undesirable reference points. This is not borne out in the data.

The reversal of the endowment effect for undesirable goods, entailed by attention-biased utility, has significant economic consequences. Whereas the endowment effect would lead to undertrading relative to an efficient market, its reversal would lead to overtrading in markets for undesirable goods. In the presence of transaction costs, too much trade would be economically inefficient. We would observe individuals continually trying to pass along the undesirable good rather than allowing it to sit with whoever can best tolerate it. Indeed, in the labor market, we observe just that. Adverse workplace conditions generate higher than average voluntary labor turnover (Cottini, Kato, & Westergaard-Nielsen, 2011). Although it may seem intuitive that workers want to leave hazardous or unpleasant jobs, standard economic theory suggests that a wage premium should perfectly compensate for poorer working conditions (Rosen, 1974). Indeed, Herzog and Schlottmann (1990) found a wage premium for manufacturing jobs that expose workers to fatal injury risk but found that workers are nevertheless more likely to leave more hazardous jobs.

Attention-biased utility also makes a number of other unique predictions. For example, as illustrated by Proposition 1, more valuable goods are predicted to display a stronger endowment effect. This prediction, which is independent of the functional forms of V_i and α_i, finds confirmatory evidence in the neuroscience literature (Knutson et al., 2008). Although prospect theory and associated models of reference dependence can generate this type of behavior in certain settings, it is not always guaranteed.

Relatedly, attention-biased utility predicts that increasing the amount of an object in the reference point can increase the relative preference for the object. Thus, for example, a decision maker offered a choice between a mug and a pen would be more likely to choose the mug if the reference point (such as her expectations) contained three mugs relative to if it contained only two mugs. Interestingly, models based on prospect theory make the opposite prediction. According to these models, increasing the amount of an object in a reference point reduces the relative preference for the object if the reference point dominates the object. (This is due to the assumption of diminishing sensitivity in losses, which leads to a larger drop in the gain–loss utility of an object relative to a competitor, as the amount of that object in the reference point is increased). These theories would thus predict that a decision maker would be *less* likely to choose a mug if the reference point contained three mugs relative to if it contained only two mugs. Indeed, Schurr and Ritov (2013) documented the absence of an endowment effect for partial losses of one's endowment, which is consistent with attention-biased utility, but not prospect theory. Bhatia (2017b) considered a number of related predictions differentiating attention-biased utility from the standard accounts of reference dependence proposed by Tversky and Kahneman (1991) and Köszegi and Rabin (2006) and found that behavior largely conforms to the predictions of attention-biased utility.

Endogenous Reference Points

Thus far we have explored the implications of reference dependence, holding reference points as being fixed and exogenous. Yet there a number of settings where the reference points are determined, either directly or indirectly, by the decision maker. The following sections explore two such settings. The first relates to sequential choice under the assumption that current endowments serve as reference points. Because choice determines endowment, which in turn determines preference, reference dependence involves particularly interesting dynamics. The second section relates to settings where the decision maker or the policymaker are able to exert some control over the reference point. This section establishes a nonmonotonicity in the impact of reference points on utility, which has implications for theories of goals and aspirations as reference points.

Equilibrium Choice

Reference points may evolve as a person makes choices. As a person adapts to a new reference point, he may then change his mind about the choices he just made. It is quite possible that moving from one object to another, and thus changing the endowment, alters the reference point, subsequently changing the most desirable option in the choice set, and leading to revised choice (see, e.g., Barkan & Busemeyer, 2003). Choice-acclimatization and, subsequently, choice-revision is especially likely for an unsophisticated decision maker who does not anticipate that the reference point will change or that such a change would affect her preference.

As an example consider a two-attribute choice set $X = \{x, y, z\}$ with $x = (6, 4)$, $y = (7, 3)$, and $z = (8, 1)$. The decision maker has valuation and attention functions such that $V_i(x_i) = x_i$ and $\alpha_i(r_i) = r_i$, for $i = 1, 2$. Assume that the decision maker's initial endowment is x. Because $U(y \mid x) > U(z \mid x) = U(x \mid x)$, the decision maker first selects y, After accepting y, the decision maker's endowment changes to y, and she now finds herself desiring z, as $U(z \mid y) > U(y \mid y) = U(x \mid y)$. This leads to a second choice, which moves the decision maker from y to z. At z she is content, as $U(z \mid z) > U(y \mid z) = U(x \mid z)$, and makes no more choices.

Analysis that explores only solitary choices will ignore long-run behavior, or equilibrium choices, like the one presented above. Studying these equilibrium choices, as well as the ways in which decision makers may obtain these choices, is important to extend theories of individual decision making to more complex settings involving firms, markets, and societies. Such an analysis will also clarify the behavior of sophisticated decision makers who use rational expectations of their own choices as reference points (see, e.g., Köszegi and Rabin, 2006). This section will outline the properties of equilibrium choice under attention-biased utility. For simplicity, it will assume that the decision maker's current endowment serves as her reference point.

Consider the set of all objects that can constitute an equilibrium choice for the decision maker. Once endowed with one of these options, the decision maker finds her endowment to be at least as desirable as any other alternative available in the choice set. For an attention-biased utility function U, and any choice set X, this set can be defined as $C^E(X, U) = \{x \in X : U(x \mid x) \geq U(y \mid x), \forall y \in X\}$. This set is analogous to set of personal equilibria proposed by Köszegi and Rabin (2006).

Consider, also, the strict relation $>_U$, defined as $x >_U y$ if and only if $U(x \mid y) > U(y \mid y)$. We say $x >_U y$ if the decision maker would strongly prefer giving up y for x when endowed with y. C^E can now be expressed as $C^E(X, U) = \{x \in X : \sim y >_U x, \forall y \in X\}$, making C^E the set of maximal elements in X with respect to $>_U$.

We can also define a decision maker's sequential choice behavior. Particularly, we say that for a choice set X, and for any $x, y \in X$, the decision maker trades y for x if $U(x \mid y) = max\{U(z \mid y), \forall z \in X\} > U(y \mid y)$. Note the presumption that if $U(x \mid y) = \{U(y \mid y)$, the decision maker does not engage in the trade, and instead chooses to remain with her endowment, y. Also note that multiple trades are possible for any endowment, and that trading y for x implies $x >_U y$. In addition, for any x, y such that $x >_U y$, there exists some choice set X such that the decision maker trades y for x (the simplest example being $X = \{x, y\}$).

Finally, define a *trading sequence* as a sequence $\langle x(0), x(0), \ldots x(t), \ldots \rangle$ such that for all t, either the decision maker trades $x(t)$ for $x(t + 1)$ or $x(t + 1) = x(t)$. For a finite choice set, we say that a trading sequence terminates at time T if no additional trades are possible once the decision maker is endowed with $x(T)$, that is, $x(T + 1) = x(T)$. If a trading sequence does not terminate, then the decision maker would continue to exchange her endowments for other choice options indefinitely.

Proposition 6 and Corollary 1 describe the equilibrium implications of attention-biased utility. They answer two related questions. 1. Is $\succ_U$ acyclic? 2. Does C^E capture all possible outcomes of a choice task? That is, do all trading sequences necessarily terminate in C^E? In the domain of desirable goods, the answers are yes and yes.

Proposition 6. For strictly increasing V_i, and for any x, y, and z such that $y \succ_U x$ and $z \succ_U y$, it cannot be that $x \succ_U z$.

Corollary 1. For strictly increasing V_i, and for any finite, nonempty choice set X, all possible trading sequences with $x(0) \in X$ terminate, with $x(T) \in C^E(X, U)$.

Proposition 6 shows that $\succ_U$ is acyclic, when valuations are increasing in each attribute. This implies the corollary that the set of equilibrium choices is nonempty and is reached in a finite sequence of trades for a finite choice set. When V_i are strictly concave, we can generalize this claim to infinite compact and convex choice sets (as proved in Munro & Sugden, 2003). Unsophisticated decision makers who maximize only immediate reference-dependent utility, engage in a sequence of trades that necessarily lead them to an equilibrium choice option.

The acyclicity established in Proposition 6 demonstrates that revealed preferences, as described by $\succ_U$, meet one of the main assumptions of rational choice theory. Note, however, that Proposition 6 does not establish negative transitivity, so these revealed preferences cannot necessarily be represented with a (reference-independent) utility function. Nevertheless, convergence of trading sequences to an equilibrium choice set implies that it is possible to perform traditional economic analysis on the equilibrium outcomes in markets with traders characterized by reference-dependent attention-biased utility. Munro and Sugden (2003), for example, undertake this analysis. They lay out a set of minimal conditions for which trading behavior with reference-dependent, but acyclic, endowment-based orderings leads to Pareto optimal outcomes in an economy. Sequential choices are consistent, decision makers eventually stabilize, and money pumps do not exist, even though solitary choices may display reference-dependent inconsistencies.

The consistency of long-run behavior under attention-biased utility allows for a reconciliation of the real world with the laboratory. The experimental work discussed in this paper has established quite clearly that contextual factors such as reference points affect choices. However, experienced decision makers typically display stable preference in economically relevant settings (List, 2004, but see also Loomes, Starmer, & Sugden, 2010). Our framework illustrates that reference dependence in individual choices can be consistent with stable choice behavior in settings in which people get used to their choices.

This insight should inspire caution in trying to make welfare judgments based on revealed preferences, even when they appear to be stable. Coherent sequential choices belie an arguably irrational psychological arbitrariness (Ariely et al., 2003). It is not necessary for optimal choices, in terms of total value, to lie in C^E: Choice options that maximize total value U^* may be rejected in favor of objects that do not. Recall the previous example involving choice set $X = \{x, y, z\}$ with $x = (6, 4)$, $y = (7, 3)$, and $z = (8, 1)$, and valuation and attention functions such that $V_i(x_i) = x_i$ and $\alpha_i(r_i) = r_i$, for $i = 1, 2$. In this setting, both x and y generate more value for the decision maker than z. However, we have $C^E(X, U) = \{z\}$. Settings like this point to a particularly troubling disconnect between welfare measures involving choice and welfare measures involving the maximization of underlying valuations.

To the extent that people can anticipate their eventual choices, their expectations may serve as reference points. A model of rational-expectations-based reference dependence incorporating the prospect theory valuation function has been proposed by Köszegi and Rabin (2006). This model provides a number of important insights regarding the reference-dependent behavior of sophisticated decision makers (see also Abeler, Falk, Goette, & Huffman, 2011; Ericson & Fuster, 2011; Gill & Prowse, 2012; Köszegi & Rabin, 2007, 2009; Loomes, Orr, & Sugdenm, 2009; Loomes & Sugden, 1986; Sugden, 2003). However, as demonstrated by Gul and Pesendorfer (2006), Köszegi and Rabin's model can generate cyclic sequential choices, implying that there may be no choice consistent with rational expectations (and thus no rational expectations at all). Proposition 6 here guarantees that with attention-biased utility a personal equilibrium consistent with expectations as reference points always exists.

Optimal Reference Points

Reference points determine not only choice but also the amount of hedonic utility that the decision maker receives by selecting a particular choice. Hence, keeping actual attained outcomes constant, it is possible to increase or decrease the decision maker's wellbeing by changing the reference point that the outcome is compared to.

To determine the properties of an optimal reference point, we can once again appeal to empirical research on the interplay of reference points, goals, and utility. Considerable research has shown that outcomes higher than the reference point are generally considered successes and are associated with higher levels of subjective wellbeing relative to outcomes lower than the reference point (Heath, Larrick, & Wu, 1999). Importantly, however, the relationship between utility and the level of a reference point is not monotonic. Adopting an extremely low reference point does not necessarily increase utility. Rather, the reference points that most enhance utility from a consumption object tend to be dominated by that object, but also nontrivial, that is, not at the minimum possible levels (see, e.g., Diener et al., 1999, for a review).

Prospect theory predicts, counter to this evidence, that the optimal reference point in $\mathbb{R}_+^N$ (given a fixed actual consumption object) is always $\mathbf{0}$. However, as Proposition 7 demonstrates, attention-biased utility accords with the empirical research on goal setting. Given a particular consumption object the optimal reference point is inferior to the object on every attribute, but, more interestingly, is often strictly positive on each attribute.

Proposition 7. For any attention-biased utility with strictly increasing V_i and any consumption object x, the optimal reference object (maximizing $U(x \mid r)$) is some $r^*(x, U)$ such that $r_i^*(x, U) \in [0, x_i)$ when $x_i > 0$ and $r_i^*(x, U) = 0$ when $x_i = 0$. A sufficient, but not necessary, condition for $r_i^*(x, U) > 0$ when $x_i > 0$ is $\alpha_i(0) = 0$. In addition, for any α_i and unbounded (strictly increasing) V_i, there exists a threshold $\bar{x}_i$ such that if consumption exceeds this threshold $x_i > \bar{x}_i$, we then have $r_i^*(x, U) > 0$.

This result stems from the fact that reference points that are highly valued on a object's strongest attributes direct more attention toward these attributes, thereby increasing the object's overall valuation. Reference points that are too good, however, reduce the object's valuation because the utility function contrasts the actual object against them. Determining the utility maximizing reference point involves optimizing these tradeoffs. For a wide range of attention-biased utility functions, the optimal reference point for a sufficiently desirable object necessarily has intermediate values: values that are greater than zero, but of course smaller than those of the object itself. Moreover, it is straightforward to observe that more ambitious goals (higher reference points) are more motivating because they increase marginal utility.

Multiple Reference Points

Using a utility model derived from a formal cognitive theory of attention gives us insight about the determination of a reference point. There has been much debate within economics about what can serve as a reference point. Many economists consider reference points to be best defined as being (internally consistent) expectations over future outcomes (Abeler et al., 2011; Ericson & Fuster, 2011; Gill & Prowse, 2012; Köszegi & Rabin, 2006, 2007, 2009). Although some choice reversals are indeed driven by changes to expectations, restricting the reference point exclusively to expectations fails to account for a range of other related anomalies observed in preferential choice. Past endowments (Strahilevitz & Loewenstein, 1998), the endowments of neighbors and close others (Clark & Oswald, 1996; McBride, 2001), goals (Heath et al., 1999), and focal outcomes in the choice task (Bushong et al., 2010; Dhar & Simonson, 1992; Dhar et al., 1999; Krajbich et al., 2010) have all been shown to act as reference points in generating similar reference-dependent choice patterns. These are all outcomes that the decision maker does not usually expect to obtain.

The premise borrowed from Bhatia's (2013) cognitive model and much psychological research generally is that any choice option that is particularly salient will serve as a prime and thus exert an influence on attribute attention. This will alter the weighting of these attributes in the decision task and influence constructed preferences. In this view, reference points are merely options that are especially salient to the decision maker. The decision maker need not expect to receive these options, but expectations certainly can be (and often are) more salient than their alternatives.

This definition can be used to extend the proposed model to study phenomena such as anchoring and choice set dependence (Ariely et al., 2003; Huber et al., 1982; Simonson, 1989), effects that we show can be predicted by the attention-biased utility model if anchors and elements of the choice set are assumed to influence choice in a manner similar to endowments and other (more standard) reference points. Indeed, considerable psychological evidence suggests that both choice set effects and anchoring effects emerge from biased attribute attention (Chapman & Johnson, 1994, 1999; Strack & Mussweiler, 1997).

These extensions of our model require that we specify a theory permitting multiple reference points, as every object in a choice set may be salient and simultaneously serve as a reference point. As above, we consider an N-attribute choice space consisting of objects $x \in \mathbb{R}^N_+$. Once again, there exist N strictly monotonic valuation functions $V_i = V_i(x_i)$, though we may no longer be able to identify them as reference-independent valuations of attributes as we cannot create a choice environment without any salient options.

Now, the decision maker is assumed to choose from a choice set $X \subset \mathbb{R}^N_+$, given reference points $r_1, r_2, \ldots, r_K$ (with each $r_k \in \mathbb{R}^N_+$) having salience $s_1, s_2, \ldots s_K$ respectively (with each $s_k \in \mathbb{R}_+$). There exist N non-negative attention functions $\alpha_i = \alpha_i(r_{1,i}, \ldots, r_{K,i}; s_1, \ldots, s_K)$ representing the decision maker's attentional weight on attribute i due to the reference points. Each attention function α_i is strictly increasing in each $r_{K,i}$, exhibits strictly increasing differences in $r_{K,i}$ and s_k, and is symmetric across all k. The attention-biased utility function becomes

$$U(x|r_1, \ldots, r_K; s_1, \ldots, s_K) = \sum_{i=1}^{N} \alpha_i(r_{1,i}, \ldots, r_{K,i}; s_1, \ldots, s_K) \cdot \left(V_i(x_i) - \bar{V}_i\right), \quad (2)$$

where the normalization constants $\bar{V}_i$ do not affect choice and parsimoniously may be taken to be the weighted average valuations of the reference points $\bar{V}_i = \frac{\sum s_k \cdot V_i(r_{k,i})}{\sum s_k}$.

To connect with the simple case of a single reference point, we assume that α_i is independent of $r_{k,i}$ when $s_k = 0$, and we recover Equation 1 with reference point r_{k^*} when $s_k = 1$ for $k = k^*$ and $s_k = 0$ otherwise. We may actually believe that all of the objects in the choice set are salient and serve as reference points, but in some cases one particular object (say an endowment or the expected choice) is more salient and described as the reference point. All of our earlier results continue to hold in this more nuanced setting as well.

Choice Set Effects

Preferences are strongly influenced by available, yet irrelevant, options. Adding, removing, or otherwise changing these options can alter the decision maker's choices and lead to a range of preference reversals. The best known of these reversals is the asymmetric dominance effect, according to which a dominated decoy option to a choice set increases the choice share of the dominating option (Huber et al., 1982; Pettibone & Wedell, 2000). A related reversal is termed the *compromise effect*. This refers to the increase in the preference for a choice option due to the presence of an extreme decoy that makes the option appear as a compromise (Simonson, 1989; Simonson & Tversky, 1992). If we assume that available options are more salient than those not in the choice set, so that the addition of dominated and extreme decoys to a choice set alters attribute attention, then we can consider choice set dependence to be a special case of the theory of attention-biased utility with multiple reference points.

We now assume that the objects that are part of a choice set are all (equally) salient, and thus all serve as reference points. Hence, for a choice set X containing K elements, the decision maker maximizes $U(x \mid X; s_1, \ldots, s_K)$, where $s_1 = s_2 = \ldots = s_K$. Keeping the salience s_k fixed, we may drop it from our notation and refer to the attention-biased utility in this context as $U(x \mid X)$.

For any x and $y \in X$, we let $P_{x,y}(X) = U(x \mid X) - U(y \mid X)$ refer to the relative preference for x over y

given the choice set X. For the case of two objects presenting a tradeoff in a two-attribute space, consider without loss of generality any x and y such that $x_1 > y_1$ and $y_2 > x_2$. The asymmetric dominance and compromise effects can be written as follows:

1. Asymmetric dominance (Huber et al., 1982): If z^I and z^{II} are such that $x_1 > z_1^I > y_1 > z_1^{II}$ and $y_2 > z_2^{II} > x_2 > z_2^I$, then $P_{x,y}(x, y, z^I) > P_{x,y}(x, y, z^{II})$.
2. Compromise (Simonson, 1989): If z^{III} and z^{IV} are such that $z_1^{III} > x_1 > y_1 > z_1^{IV}$ and $z_2^{IV} > y_2 > x_2 > z_2^{III}$, then $P_{x,y}(x, y, z^{III}) > P_{x,y}(x, y, z^{IV})$.

In the definitions above, z^I and z^{II} are decoys that are dominated by x and y respectively, whereas z^{III} and z^{IV} are decoys that make x appear as a compromise and make y appear as a compromise respectively. Replacing z^I with z^{II}, z^{III} with z^{IV}, or vice versa, can bias preference. These effects closely resemble the improvements versus tradeoffs and the small versus large tradeoffs effects introduced earlier in the paper. Thus, not surprisingly, our model is able to predict the emergence of these effects if we assume that the salience weights for z^I and z^{II}, and for z^{III} and z^{IV}, are constant. This is illustrated in Proposition 8.

Proposition 8. Attention-biased utility with strictly increasing V_i generates the asymmetric dominance and compromise effects.

Proposition 8 stems from the same properties of the model as Proposition 1, as decoys in the above formulation act as reference points. For example, changing the decoy from z^I to z^{II} shifts attention away from Attribute 1 and toward Attribute 2 because z^I has relatively more of attribute 1 and z^{II} has relatively more of attribute 2. The same shift occurs if the decoy changes from z^{III} to z^{IV}.

Unlike the traditional reference-dependent biases or the anchoring effects discussed next, the attentional explanation for the asymmetric dominance and compromise effects has not been formally tested in the domain of preferential choice. However, Trueblood, Brown, Heathcote, and Busemeyer (2013) have found these effects in perceptual judgments, and Bhatia (2014) has found that asymmetrically dominated decoys bias attention in judgment tasks in a manner consistent with these results, leading to a type of asymmetric dominance effect for the formation of belief. Additional details about these effects, as well as the psychological theory justifying our specification of attention-biased utility with multiple reference points can be found in Bhatia (2013).

Anchoring Effects

Stated measures of preference (like many other judgments) exhibit the anchoring effect. The typical anchoring study involves the valuation of a choice item in terms of either WTP or WTA. Prior to the valuation, however, an arbitrary high or low number (the anchor) is generated, and decision makers are asked whether they would be willing to pay (or accept) that amount of money to obtain (or give up) the choice item. Typically, high anchors generate high WTP and WTA responses, whereas low anchors generate low WTP and WTA responses (see, e.g., Ariely et al., 2003; Beggs & Graddy, 2009; Johnson & Schkade, 1989). Research by Ariely et al., (2003) also found that highly desirable choice items generate a higher WTP and WTA than less desirable items, regardless of the anchor, suggesting that although absolute valuation may be arbitrary (in the sense that it is manipulated by anchoring), relative preference is nonetheless coherent.

As with the endowment effect, a number of scholars have discovered that attention plays a critical role in predicting the emergence of anchoring. Notably, Chapman and Johnson (1994, 1999), Strack and Mussweiler (1997), and Mussweiler and Strack (1999) found that anchors bias attention toward information that is consistent with the anchor. High anchors in a willingness-to-pay task focus the decision maker on highly desirable attributes of the item in consideration. This leads to high valuations, close to the anchor. The analogous finding holds for low anchors. As with the endowment effect, the extent of this attentional bias predicts the strength of the anchoring effect. In addition, the anchoring effect can be removed by refocusing the decision maker on the other attributes in the choice task (Mussweiler, Strack, & Pfeiffer, 2000). These results suggest that anchoring—like the endowment effect, other reference-dependent effects, and the choice set effects already discussed—could be explained by an attentional mechanism. The key assumption here would be that anchors, which are particularly salient outcomes in the choice task, serve as secondary reference points (in addition to endowments), with judgments of WTP and WTA in the presence of anchors, being made using attention-biased utility applied to these reference points.

As in Measures of Preference section, we can consider a two-attribute space, where i represents the attribute at hand and j represents money or other attributes that are used to represent the decision maker's preferences for i. Owning the item is represented by x_i, whereas not owning the item is captured by $y_i < x_i$. The decision maker's wealth prior to the evaluation task is y_j. During the evaluation task the decision maker is asked to list an amount WTP such that $U(x_i, y_j - \text{WTP} \mid y, r_2) = U(y_i, y_j \mid y, r_2)$.

High anchor settings initially ask the decision maker whether she wants to buy the item for a_H, whereas the low anchor settings initially ask the decision maker whether she wants to buy the item for a_L. Here $a_H > a_L$. A high anchor can thus be formalized as a object $A_H = (x_i, y_j - a_H)$, which represents ownership of the item for the price a_H. Likewise, a low anchor can be formalized as a object $A_L = (x_i, y_j - a_L)$, which represents ownership of the item for low price of a_L.

We assume that asking decision makers to evaluate the anchor increases the salience of the anchor. The anchor thus serves as a secondary reference point, r_2. In effect, the decision maker is asked to select either WTP_{A_H} or WTP_{A_L} such that $U(x_i, y_j - \text{WTP}_{A_H} \mid y, A_H) = U(y_i, y_j \mid y, A_H)$ or $U(x_i, y_j - \text{WTP}_{A_L} \mid y, A_L) = U(y_i, y_j \mid y, A_L)$.

Proposition 9 shows that any attention-biased utility function with strictly increasing V_i will generate $\text{WTP}_{A_H} > \text{WTP}_{A_L}$. This is a natural implication of the attentional mechanisms that are responsible for the reference-dependent anomalies discussed in the previous sections. Anchoring with $a_H > a_L$ means that the decision maker considers the object A_H with lower final wealth than A_L. Thus, A_H focuses less attention on the expenditure required to obtain the item relative to the attention on the item itself. A lower weight on attribute j relative to attribute i implies that the decision maker is more willing to give up attribute j to obtain attribute i. This leads to $\text{WTP}_{A_H} > \text{WTP}_{A_L}$.

Corollary 2 establishes that this mechanism does not eradicate the observed coherence of anchored choice. For any $z_i > x_i > y_i$, the willingness-to-pay for z_i is always greater than the willingness-to-pay for x_i, regardless of the anchor or reference point involved in the choice task (see Ariely et al., 2003).

Proposition 9. If V_i are strictly increasing, then for any anchors A_H and A_L such that $a_H > a_L$, we have $\text{WTP}_{A_H} > \text{WTP}_{A_L}$.

Corollary 2. If V_i are strictly increasing, then for any anchor A, and any $z_i > x_i > y_i > y_i$ with $U(z_i, y_j - \text{WTP}_z \mid y, A) = U(y_i, y_j \mid y, A)$ and $U(x_i, y_j - \text{WTP}_x \mid y, A) = U(y_i, y_j \mid y, A)$, we have $\text{WTP}_z > \text{WTP}_x$.

Anchoring effects emerge from the same assumptions that generate traditional reference-dependent effects and context effects. These anomalies arise from attention-biased utility with strictly increasing valuation functions V_i.

Concluding Comments

Individual choice displays a systematic and pervasive relationship with a large range of normatively irrelevant contextual factors. A number of these contextual factors involve salient, or otherwise exceptional, choice options, such as reference points. Current approaches to understanding the impact of these options on choice are based on prospect theory, and thus require assumptions such as asymmetries in gains and losses and convex utility in losses. These lead to notable deviations from standard economic theory (e.g., unique choices with liner budget sets are no longer guaranteed), making many economic applications of these models technically challenging.

Recent work in psychology and neuroscience suggests a novel approach. Salient choice options have been shown to alter the decision maker's attention toward the attributes involved in choice task (Ashby et al., 2012, 2015; Bhatia, 2013, 2017b; Carmon & Ariely, 2000; Johnson et al., 2007; Nayakankuppam & Mishra, 2005; Pachur & Scheibehenne, 2012; Willemsen et al., 2011). This can affect the weighting of attributes and subsequently influence choice. This article presents a formal, psychologically grounded model of this priming-based mechanism and shows that it explains a number of behavioral anomalies associated with reference dependence, including findings such as the endowment effect, and related biases traditionally assumed to support prospect theory accounts of reference dependence (Herne, 1998; Tversky & Kahneman, 1991). This model also features a number of properties that make it desirable from an economic perspective, including concave utility and the guaranteed existence of a stable equilibrium for sequential choices. The model also generates predictions that distinguish it from prospect theory, such as the reversal of the endowment effect for undesirable choice objects as well as the strengthening of the endowment effect for more valuable goods. Many of the model's unique predictions are supported by recent experimental evidence (e.g., Brenner et al., 2007), suggesting that it may provide a better behavioral account of reference dependence than existing approaches, in some settings. Finally, extensions to the model allow it to capture other phenomena related to reference dependence, such as anchoring effects and context effects (Ariely et al., 2003; Huber et al., 1982; Simonson, 1989), which have also been attributed to biased attention in the psychology literature (Chapman & Johnson, 1994, 1999; Strack & Mussweiler, 1997).

References

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review, 101,* 470–492.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *Quarterly Journal of Economics, 112,* 1937–1996.

Ashby, N. J. S., Dickert, S., & Glöckner, A. (2012). Focusing on what you own: Biased information uptake due to ownership. *Judgment and Decision Making, 7,* 254–267.

Ashby, N. J., Walasek, L., & Glöckner, A. (2015). The effect of consumer ratings and attentional allocation on product valuations. *Judgment and Decision Making, 10,* 172.

Barkan, R., & Busemeyer, J. R. (2003). Modeling dynamic inconsistency with a changing reference point. *Journal of Behavioral Decision Making, 16,* 235–255.

Bateman, I., Munro, A., Rhodes, B., Starmer, C., & Sugden, R. (1997). A test of the theory of reference-dependent preferences. *Quarterly Journal of Economics, 118,* 73–105.

Beggs, A., & Graddy, K. (2009). Anchoring effect: Evidence from art auctions. *American Economic Review, 99,* 1027–1039.

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review, 120,* 522–543.

Bhatia, S. (2014). Confirmatory search and asymmetric dominance. *Journal of Behavioral Decision Making, 27,* 468–476.

Bhatia, S. (2017a). Attention and attribute overlap in preferential choice. *Quarterly Journal of Experimental Psychology, 70,* 1174–1196.

Bhatia, S. (2017b). Comparing theories of reference-dependent choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 1490–1507.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012a). Salience theory of choice under risk. *Quarterly Journal of Economics, 127,* 1243–1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012b). Salience in experimental tests of the endowment effect. *American Economic Review, 102,* 47–52.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2015). Salience theory of judicial decisions. *The Journal of Legal Studies, 44*(S1), S7–S33.

Brenner, L., Rottenstreich, Y., Sood, S., & Bilgin, B. (2007). On the psychology of loss aversion: Possession, valence, and reversals of the endowment effect. *Journal of Consumer Research, 34,* 369–376.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432.

Bushong, B. B., King, L. M., Camerer, C. F., & Rangel, A. (2010). Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. *American Economic Review, 100,* 1556–1571.

Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *American Economic Review, 101,* 2899–2922.

Carmon, Z., & Ariely, D. (2000). Focusing on the forgone: How value can appear so different to buyers and sellers. *Journal of Consumer Research, 27,* 360–370.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making, 7,* 223–242.

Chapman, G. B., & Johnson, E. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes, 79,* 115–153.

Clark, A. E., & Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics, 61,* 359–381.

Cottini, E., Kato, T., & Westergaard-Nielsen, N. (2011). Adverse workplace conditions, highinvolvement work practices and labor turnover: Evidence from Danish linked employer–employee data. *Labour Economics, 18,* 872–880.

Dhar, R., Nowlis, S., & Sherman, S. J. (1999). Comparison effects on preference construction. *Journal of Consumer Research, 26,* 293–306.

Dhar, R., & Simonson, I. (1992). The effect of the focus of comparison on consumer preferences. *Journal of Marketing Research, 29,* 430–440.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125,* 276–302.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Ericson, K. M. M., & Fuster, A. (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *Quarterly Journal of Economics, 126,* 1879–1907.

Fehr, E., & Rangel, A. (2011). Neuroeconomic foundations of economic choice—recent advances. *Journal of Economic Perspectives, 25,* 3–30.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96,* 1043–1068.

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review, 102,* 469–503.

Golman, R., & Loewenstein, G. (2016). Information gaps: A theory of preferences regarding the presence and absence of information. *Decision*. Advance online publication. http://dx.doi.org/10.1037/dec0000068

Golman, R., Loewenstein, G., & Gurney, N. (2017). *Information gaps for risk and ambiguity*. Working paper.

Golman, R., Loewenstein, G., Molnar, A., & Saccardo, S. (2017). *The demand for, and avoidance of, information*. Manuscript in preparation.

Gul, F., & Pesendorfer, W. (2006). *The revealed preference implications of reference dependent preferences*. Working Paper.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38,* 79–109.

Herne, K. (1998). Testing the reference-dependent model: An experiment on asymmetrically dominated reference points. *Journal of Behavioral Decision Making, 11,* 181–192.

Herzog, H., & Schlottmann, A. (1990). Valuing risk in the workplace: Market price, willingness to pay, and the optimal provision of safety. *Review of Economics and Statistics, 72,* 463–470.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research, 9,* 90–98.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research, 49,* 1295–1306.

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Attributes of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 461–474.

Johnson, E. J., & Schkade, D. A. (1989). Bias in utility assessments: Further evidence and explanations. *Management Science, 35,* 406–424.

Johnson, J. G., & Busemeyer, J. R. (2016). A computational model of the attention process in risky choice. *Decision, 3,* 254.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review, 93,* 1449–1475.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy, 98,* 1325–1348.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives, 5,* 193–206.

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: preferences and value tradeoffs*. Cambridge, UK: Cambridge University Press.

Knetsch, J. (1989). The endowment effect and evidence of nonreversible indifference curves. *American Economic Review, 79,* 1277–1284.

Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics, 99,* 507–521.

Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., & Loewenstein, G. (2008). Neural antecedents of the endowment effect. *Neuron, 58,* 814–822.

Köszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics, 121,* 1133–1165.

Köszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review, 97,* 1047–1073.

Köszegi, B., & Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review, 99,* 909–936.

Köszegi, B., & Szeidl, A. (2013). A model of focusing in economic choice. *Quarterly Journal of Economics, 128,* 53–104.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13,* 1292–1298.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences, 9,* 75–82.

Lim, S. L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience, 31,* 13214–13223.

List, J. A. (2004). Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica, 72,* 615–625.

Locke, E., Shaw, K., Saari, L., & Latham, G. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin, 90,* 125–152.

Loewenstein, G., & Adler, D. (1995). A bias in the prediction of tastes. *The Economic Journal, 105,* 929–937.

Loomes, G., Orr, S., & Sugdenm, R. (2009). Taste uncertainty and status quo effects in consumer choice. *Journal of Risk and Uncertainty, 39,* 113–135.

Loomes, G., Starmer, C., & Sugden, R. (2010). Preference reversals and disparities between willingness to pay and willingness to accept in repeated markets. *Journal of Economic Psychology, 31,* 374–387.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies, 53,* 271–282.

Masatlioglu, Y., Nakajima, D., & Ozbay, E. (2012). Revealed attention. *American Economic Review, 102,* 2183–2205.

Matejka, F., & McKay, A. (2014). *Rational inattention to discrete choices: A new foundation for the multinomial logit model*. Working paper.

McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. *Journal of Economic Behavior & Organization, 45,* 251–278.

Morewedge, C. K., & Giblin, C. E. (2015). Explanations of the endowment effect: An integrative review. *Trends in Cognitive Sciences, 19,* 339–348.

Munro, A., & Sugden, R. (2003). On the theory of reference-dependent preferences. *Journal of Economic Behavior & Organization, 50,* 407–428.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 80,* 136–164.

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin, 26,* 1142–1150.

Nayakankuppam, D., & Mishra, H. (2005). The endowment effect: Rose-tinted and dark-tinted glasses. *Journal of Consumer Research, 32,* 390–395.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition, 11,* 264–336.

Novemsky, N., & Kahneman, D. (2005). The boundaries of loss aversion. *Journal of Marketing Research, 43,* 119–128.

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144,* 190–206.

Pachur, T., & Scheibehenne, B. (2012). Constructing preference from experience: The endowment effect reflected in external information search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1108–1116.

Pettibone, J. C., & Wedell, D. H. (2000). Examining models of nondominated decoy effects across judgment and choice. *Organizational Behavior and Human Decision Processes, 81,* 300–328.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica, 68,* 1281–1292.

Reutskaja, E., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure. *American Economic Review, 101,* 900–926.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological review, 108,* 370.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82,* 34–55.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1,* 7–59.

Schurr, A., & Ritov, I. (2013). The effect of giving it all up on valuation: A new look at the endowment effect. *Management Science, 60,* 628–637.

Simon, H. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158–174.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 281.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics, 50,* 665–690.

Strack, F., & Mussweiler, T. (1997). Explaining the Enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73,* 437–446.

Strahilevitz, M. A., & Loewenstein, G. (1998). The effect of ownership history on the valuation of objects. *Journal of Consumer Research, 25,* 276–289.

Sugden, R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory, 111,* 172–191.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1,* 39–60.

Trueblood, J., Brown, S., Heathcote, A., & Busemeyer, J. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science, 24,* 901–908.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106,* 1039–1061.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology, 8,* 227–233.

Willemsen, M. C., Backenholt, U., & Johnson, E. J. (2011). Choice by value encoding and value construction: Processes of loss aversion. *Journal of Experimental Psychology: General, 140,* 303–324.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601–621.

(*Appendix follows*)

Appendix

Proofs

Proof of Proposition 1

Consider any x and y and reference points r and s such that $x_i \geq y_i$ if and only if $r_i \geq s_i$ with at least one pair of inequalities strict. We can write:

$$[U(x|r) - U(y|r)] - [U(x|s) - U(y|s)]$$
$$= \sum_{i=1}^{N} [\alpha_i(r_i) - \alpha_i(s_i)] \cdot [V_i(x_i) - V_i(y_i)]. \quad (3)$$

Because V_i and α_i are strictly increasing for all i, we have $[\alpha_i(r_i) - \alpha_i(s_i)] \cdot [V_i(x_i) - V_i(y_i)] \geq 0$ with the inequality strict when $x_i \neq y_i$and $r_i \neq s_i$. Hence the sum over all i, $[U(x \mid r) - U(y \mid r)] - [U(x \mid s) - U(y \mid s)]$, is positive, giving us $[U(x \mid r) - U(y \mid r)] > [U(x \mid s) - U(y \mid s)]$. □

Proof of Proposition 2

Proposition 2 is easily obtained by substituting r with x and s with y, in Proposition 1.

Proof of Proposition 3

Let us show that WTP < EL. the same steps can be used to show that eg < WTA.

Taking WTP and EL as defined in the paper, and noting that $x_j = y_j$, we obtain:

$$\alpha_i(y_i) \cdot V_i(x_i) + \alpha_j(x_j) \cdot V_j(x_j - \text{WTP})$$
$$= \alpha_i(y_i) \cdot V_i(y_i) + \alpha_j(x_j) \cdot V_j(x_j) \quad (4)$$

$$\alpha_i(x_i) \cdot V_i(x_i) + \alpha_j(x_j) \cdot V_j(x_j - \text{EL})$$
$$= \alpha_i(x_i) \cdot V_i(y_i) + \alpha_j(x_j) \cdot V_j(x_j). \quad (5)$$

Subtracting Equation 5 from 4, and simplifying, gives us:

$$\alpha_j(x_j) \cdot [V_j(x_j - \text{WTP}) - V_j(x_j - \text{EL})]$$
$$= [\alpha_i(x_i) - \alpha_i(y_i)] \cdot [V_i(x_i) - V_i(y_i)]. \quad (6)$$

Because V_i and α_i are both strictly increasing in their arguments, the right-hand side of Equation 6 is positive. Then, because α_i is always positive, we get $V_j(x_j - \text{WTP}) > V_j(x_j - \text{EL})$. Once again, since V_i is strictly increasing, this implies WTA < EL. □

Proof of Proposition 4

Proposition 1 implies Condition 1, which is sufficient to generate effects 1–6.

Proof of Proposition 5

Consider any undesirable x and y. we can write:

$$[U(x|x) - U(y|x)] - [U(x|y) - U(y|y)]$$
$$= \sum_{i=1}^{K} [\alpha_i(x_i) - \alpha_i(y_i)] \cdot [V_i(x_i) - V_i(y_i)]. \quad (7)$$

Note that V_i is nonpositive and decreasing for all i. Hence if $V_i(x_i) \geq V_i(y_i)$, we have $x_i \leq y_i$, which implies $\alpha_i(x_i) \leq \alpha_i(y_i)$. This means that $[\alpha_i(x_i) - \alpha_i(y_i)] \cdot [V_i(x_i) - V_i(y_i)] \leq 0$ for all i. Because x and y are nonidentical, we have some i such that $V_i(x_i) \neq V_i(y_i)$, which implies that $[\alpha_i(x_i) - \alpha_i(y_i)] \cdot [V_i(x_i) - V_i(y_i)] < 0$ for some i. Hence the sum over all i, $[U(x \mid x) - U(y \mid x)] - [U(x \mid y) - U(y \mid y)]$ is negative, giving us $[U(x \mid x) - U(y \mid x)] < [U(x \mid y) - U(y \mid y)]$. □

(*Appendix continues*)

Proof of Proposition 6

Assume, for a contradiction, that $y \succ_U x$, $z \succ_U y$, and $x \succ_U z$. By this assumption,

$$[U(z|z) - U(x|z)] + [U(y|y) - U(z|y)] + [U(x|x) - U(y|x)] < 0. \quad (8)$$

We can expand this sum of utility differences (in Equation 8) as

$$\begin{aligned}&\sum_{i=1}^{N} \alpha_i(z_i) \cdot [V_i(z_i) - V_i(x_i)] + \alpha_i(y_i) \cdot [V_i(y_i) - V_i(z_i)] + \alpha_i(x_i) \cdot [V_i(x_i) - V_i(y_i)] \\ &= \sum_{i=1}^{N} [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)] \\ &+ [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] \\ &+ [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]. \quad (9)\end{aligned}$$

The last two terms must be strictly positive for all i because both α_i and V_i are monotonically increasing functions, so each term is either the product of two positives (if $z_i > x_i$ or $y_i > z_i$, respectively) or the product of two negatives (if $z_i < x_i$ or $y_i < z_i$ resp.).

Observe that if we add $\sum_{i=1}^{N} [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(z_i) - V_i(x_i)]$ to the sum in Equation 9, we obtain

$$\begin{aligned}&\sum_{i=1}^{N} [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(z_i) - V_i(x_i)] \\ &+ [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)] \\ &+ [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] \\ &+ [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)] \\ &= \sum_{i=1}^{N} [\alpha_i(y_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(x_i)], \quad (10)\end{aligned}$$

which, of course, is strictly positive by the same argument as above. We can separate out the contributions on each attribute and isolate the term we just added in Equation 10, obtaining for any i,

$$\begin{aligned}&[\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(z_i) - V_i(x_i)] \\ &= [\alpha_i(y_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(x_i)] \\ &- [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)] \\ &- [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] \\ &- [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)] \\ &> - [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)] \\ &- [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] \\ &- [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]. \quad (11)\end{aligned}$$

Now observe that $[\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(z_i) - V_i(x_i)]$ has the same sign as $[\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)]$. (They are both positive if and only if $y_i > z_i > x_i$ or $y_i < z_i < x_i$.) Thus, we can put a lower bound on this latter term, $[\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)] > - [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] - [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]$.

We plug this lower bound into Equation 11 to get

$$\sum_{i=1}^{N} \alpha_i(z_i) \cdot [V_i(z_i) - V_i(x_i)] + \alpha_i(y_i) \cdot [V_i(y_i) - V_i(z_i)] + \alpha_i(x_i) \cdot [V_i(x_i) - V_i(y_i)] > 0.$$

This contradicts Equation 8.

Proof of Corollary 1

First let us note that if a trading sequence terminates at T, then by the definition of a trade, and of C^E, we have $x(T) \in C^E(X, U)$. Conversely, for any t, if $x(t) \in C^E(X, U)$ then the trading sequence has terminated at $T \leq t$.

(*Appendix continues*)

Now consider any X and any U such that for some time t we have some $x(t) \in X$, but $x(t) \notin C^E(X, U)$. Because $x(t) \notin C \in (X, U)$, there exists some $x(t+1) \in X \backslash \{\cup_{t' \le t} x(t')\}$ such that the decision maker trades $x(t)$ for $x(t+1)$. (We can rule out $x(t+1) = x(t')$ for some $t' \le t$ by the acyclicity of the preference relation $>_U$.) Because X is finite, it is impossible that this holds for arbitrarily large t—the premise that $x(t) \notin C^E(X, U)$ must be false for some large enough t. □

Proof of Proposition 7

The contribution to attention-biased utility on attribute i, as a function of r_i, goes from positive to negative at $r_i = x_i$. Attention-biased utility is continuous, so it must achieve a maximum with $0 \le r_i < x_i$ whenever $x_i > 0$. If $\alpha_i(0) = 0$, then the optimal r_i^* must be strictly positive whenever $x_i > 0$ because there would be no contribution to utility with $r_i = 0$. Moreover, the derivative $\frac{\partial U(x \mid r)}{\partial r_i}|_{r_i=0} = \alpha_i'(0) \cdot (V_i(x_i) - V_i(0)) - \alpha_i(0) \cdot V_i'(0)$ is increasing without bound in x_i as long as V_i is unbounded, so for sufficiently high x_i, the optimal r_i^* is again strictly positive. □

Proof of Proposition 8

For simplicity we write $\alpha_i(X; s_1, \ldots, s_K)$ as $\alpha_i(X)$, where X is the choice set. For the asymmetric dominance effect, $P_{x,y}(x, y, z^I) - P_{x,y}(x, y, z^{II})$ can be simplified and written as

$$[\alpha_1(x, y, z^I) - \alpha_1(x, y, z^{II})] \cdot [V_1(x_1) - V_1(y_1)] + [\alpha_2(x, y, z^I) - \alpha_2(x, y, z^{II})] \cdot [V_2(x_2) - V_2(y_2)].$$

This is guaranteed to be positive as $x_1 > z_1^I > y_1 > z_1^{II}$ and $y_2 > z_2^{II} > x_2 > z_2^I$ implies that $\alpha_1(x, y, z^I) > \alpha_1(x, y, z^{II})$, $V_1(x_1) > V_1(y_1)$, $\alpha_2(x, y, z^I) < \alpha_2(x, y, z^{II})$, and $V_2(x_2) < V_2(y_2)$.

The same argument holds for the compromise effect, where $z_1^{III} > x_1 > y_1 > z_1^{IV}$ and $z_2^{IV} > y_2 > x_2 > z_2^{III}$.

Proof of Proposition 9

Taking the definitions provided in the paper, and writing $x_j^H = y_j - a_H$ and $x_j^L = y_j - a_L$ we get

$$\alpha_i(y_i, x_i; s_1, s_2) \cdot V_i(x_i) + \alpha_j(y_j, x_j^H; s_1, s_2) \cdot V_j(y_j - \text{WTP}_{A_H}) = \alpha_i(y_i, x_i; s_1, s_2) \cdot V_i(y_i) + \alpha_j(y_j, x_j^H; s_1, s_2) \cdot V_j(y_j) \quad (12)$$

and

$$\alpha_i(y_i, x_i; s_1, s_2) \cdot V_i(x_i) + \alpha_j(y_j, x_j^L; s_1, s_2) \cdot V_j(y_j - \text{WTP}_{A_L}) = \alpha_i(y_i, x_i; s_1, s_2) \cdot V_i(y_i) + \alpha_j(y_j, x_j^L; s_1, s_2) \cdot V_j(y_j). \quad (13)$$

Putting together Equations 12 and 13 gives us:

$$\alpha_j(y_j, x_j^H; s_1, s_2) \cdot [V_j(y_j) - V_j(y_j - \text{WTP}_{A_H})] = \alpha_j(y_j, x_j^L; s_1, s_2) \cdot [V_j(y_j) - V_j(y_j - \text{WTP}_{A_L})] > 0. \quad (14)$$

We know that $\alpha_j(y_j, x_j^L; s_1, s_2) > \alpha_j(y_j, x_j^H; s_1, s_2) > 0$, as α_j is strictly increasing in r_2 and $x_j^L > x_j^H$ because $a_H > a_L$. This implies that $V_j(y_j) - V_j(y_j - \text{WTP}_{A_H}) > V_j(y_j) - V_j(y_j - \text{WTP}_{A_L}) > 0$. Again, because V_j is strictly increasing, this implies that $\text{WTP}_{A_H} > \text{WTP}_{A_L}$. □

Proof of Corollary 2

Because the reference points y and A are fixed, U can be seen as a standard (reference independent) utility function, that is strictly increasing in its arguments. For such functions we know $z_i > x_i$ guarantees that $\text{WTP}_z > \text{WTP}_x$.

Received January 29, 2017
Revision received September 28, 2017
Accepted December 12, 2017 ■