Attention and Reference Dependence

Sudeep Bhatia and Russell Golman *

Department of Social & Decision Sciences, Carnegie Mellon University

March 13, 2013

Abstract

We present a model of reference dependence, which assumes that reference points affect choice by directing the decision maker's attention towards the particular goods in the reference bundle. 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 loss aversion and can explain a large number of behavioral anomalies related to reference dependence. Additionally, this model produces acyclic sequential choices, and any sequence of short-sighted reference-dependent decisions is guaranteed to converge to a stable personal equilibrium. JEL Codes: D03, D11, D87.

*Address: 5000 Forbes Avenue, Pittsburgh, PA, 15213; e-mail: *sudeepb@andrew.cmu.edu*, *rgolman@andrew.cmu.edu*.

1 Introduction

Behavior is influenced by salient bundles, such as reference points. These bundles affect the decision maker's evaluation of available choice options, leading to choice reversals and other deviations from rationality. Reference points, for example, are often preferred over competing bundles, generating the endowment effect and the status quo bias. Additionally, bundles that are clear improvements over the reference point and bundles that involve only small tradeoffs from the reference point, are often chosen over bundles involving large tradeoffs from reference point (Tversky and Kahneman, 1991; Kahneman et al., 1991).

Standard models of reference dependence involve considerable deviations from traditional consumer theory. Prospect theory (Tversky and Kahneman, 1991) and related models (e.g., Kőszegi and Rabin, 2006) assume non-convex preferences and explicit asymmetries in gains and losses. Additionally, such preferences can lead a decision maker into cyclic sequential choices. These choices could be vulnerable to money pumps, and (more distressingly for economists wishing to apply these models in the field) the existence of a stable choice function (under Tversky and Kahneman's model) or a personal equilibrium consistent with one's expectations (under Kőszegi and Rabin's model) is not guaranteed (see Gul and Pesendorfer, 2006).

Recent work on the cognitive processes underlying choice suggests an alternate approach to studying reference dependence. A number of researchers (Carmon and Ariely, 2000; Nayakankuppam and Misra, 2005; Johnson et al., 2007; Willemsen et al., 2011; Pachur and Scheibehenne, 2012; Ashby et al., 2012) find that endowments direct the decision maker's attention towards their most prominent dimensions. This is subsequently shown to generate an increased weight on these dimensions, altering 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. Rather they act as *primes*, directing attention towards information that they are strongly associated with. As choice bundles are most strongly associated with their component goods, it is these goods that receive a higher weight in the decision task. Changing reference points can thus affect the weights on these goods and, consequently, the choice between competing bundles.

This paper presents a model of reference dependence, motivated by this intuition. Attention-biased utility of a consumption bundle is a linear combination of consumption utilities for each good in the bundle. While consumption utility is stable and reference independent, the weight on each good in attention-biased utility depends on the amount of that good in the reference bundle. Changing the reference point alters the attentional weights in the attention-biased utility function. This can affect the decision maker's choice and generate preference reversals.

This model requires minimal deviations from consumer theory. It does not make any assumptions about the curvature of consumption utility: if consumption utility is strictly concave then so is attention-biased utility. Additionally, this model does not assume a built-in asymmetry in gains and losses. Nonetheless, reductions in consumption have a stronger impact on attention-biased utility than corresponding gains, generating the well-known phenomenon of loss aversion. The proposed model can subsequently explain all the observed anomalies regarding reference-dependent choice.

Sequential choice is of particular interest in a reference-dependent context. This paper shows that maximization of attention-biased utility generates long-term choice behavior that satisfies basic consistency requirements. Decision makers do not cycle through available options indefinitely (and are not vulnerable to money pumps); rather, decisions 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 paper also considers the choice of reference points that maximize utility. It establishes that the impact of a reference point on utility from a particular bundle is non-monotonic, and that high reference points are often optimal. This result helps us understand the common preference for attainable but non-trivial outcomes as aspirational reference points, or goals.

While attention-biased utility can parsimoniously capture much 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. Particularly, attention-biased utility predicts reversals of reference-dependent effects for choices involving goods with negative and strictly decreasing consumption utility. In contrast, prospect theory makes no distinction between the negative and positive domains. Brenner et al. (2007) and Bhatia and Turan (2012) have documented reversals of the endowment effect for undesirable bundles, suggesting that reference points do indeed operate differently in the negative domain.

The next section reviews research in economics, psychology and neuroscience on the role of attention in choice. Section 3 provides a formal theory of attention-biased utility. Section 4 explores the implications of attention-biased utility for loss aversion and reference dependence. Section 5 explores sequential choices emerging from attention-biased utility, as well as optimal reference points under attention-biased utility. Section 6 studies undesirable choice objects. Section 7 concludes.

2 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 choice set (Simon, 1955). Daniel Kahneman (2003) has also argued that biased attention, or accessibility, plays a key role in shaping choice. More recently, Gabaix et al. (2006) have proposed a model that captures 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 et al. (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 et al. (2012) have provided a theoretical framework for inferring attentional allocation from choice behavior.

Attentional biases have also been used by economists to explain a range of observed behavioral anomalies. Bordalo et al. (2012a), for example, provide 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 et al. (2012b) use a similar approach to explain the endowment effect as well as its reversal. Finally, Kőszegi and Szeidl (2013) study 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.

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 can provide a rigorous account of 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 towards the dimensions of the bundles they possess: owners focus on aspects of the traded good, where as non-owners are more likely focus on aspects of the expenditure involved in the trade. Since increased attention to a particular dimension increases that dimension'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 Misra (2005), who find that owners are more likely to attend to the aspects that the endowed good is strongest on, and less likely to attend to the aspects that the endowed good is weakest on, relative to non-owners. Johnson et al. (2007) replicate these findings and, in addition, discover differences in the order that owners and non-owners attend to the various dimensions, in the decision task. Decision makers generally attend to the strongest dimensions of the bundle that they possess before focusing on the weakest dimensions of their possessed bundle, or dimensions of bundles that they do not own. Willemsen et al. (2011) extend these findings beyond the endowment effect. Ashby et al. (2012) use eye-tracking and response time restrictions to further demonstrate the role of biased attention in reference dependence. Finally, Pachur and Scheibehenne (2012) demonstrate the existence of these attentional biases for risky choice. These papers note that the attentional biases displayed by the decision makers can predict buying and selling prices, and choice probabilities. Additionally, altering where these decision makers focus their attention can eliminate the endowment effect.

While these attentional biases have typically been documented for desirable items such as mugs or pens (which are subject to the endowment effect (Kahneman et al., 1990)), Bhatia and Turan (2012) explore endowment-related attentional biases for undesirable goods. As with previous work they find that endowments direct the decision maker's attention towards their primary dimensions. Since the primary dimensions of undesirable items are themselves undesirable, this leads to stronger aversion to the reference point, and a reversal of the endowment effect. Reversals of this type have previously been documented by Brenner et al. (2007) and point to settings where standard models of reference dependence (e.g., prospect theory) may be inadequate descriptors of human behavior.

The endowment effect, as well as its reversal, can also be generated without explicit endowments. Dhar and Simonson (1992) and Dhar et al. (1999) find 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 towards the focal item's primary dimensions. Reducing this attentional bias can eliminate differences in choice shares.

Bushong et al. (2010) present related findings. They note 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 et al. (2010) and Reutskaja et al. (2011) suggests that visual attention towards an item has a direct relationship with the decision maker's preference towards that item. This has been documented for both actual choice and the neural mechanisms that determine choice (Lim et al., 2011).

A nearly identical set of results have been discovered for the anchoring effect. Chapman and Johnson (1994, 1999), Strack and Mussweiler (1997), and Mussweiler and Strack (1999), for example, find that anchors bias attention towards information that is consistent with the anchor. High anchors in a willingness-to-pay task focus the decision maker on highly desirable dimensions 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. Additionally, the anchoring effect can be removed by refocusing the decision maker on the other dimensions in the choice task (Mussweiler et al., 2000).

These results suggest that endowments do 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 towards relevant dimensions. While 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, such as those involving focal outcomes or anchors. 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.

3 Model

Consider an N-dimensional choice space consisting of bundles $x \in \mathbb{R}_+^N$. 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 consumption utility of good i in x. For simplicity, we will set $V_i(0) = 0$ for all i.

Traditionally rational choice (under the assumption that goods can be valued independently) selects $x \in X$, to maximize total consumption 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 good depends on the reference bundle's amount of that good. 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 good i given a reference point r. The decision maker chooses according to $\tilde{U}(x|r) = \sum_{i=1}^{N} \alpha_i(r_i) \cdot V_i(x_i)$. Note that we can add a constant to overall utility without changing the underlying preferences. Thus we can normalize $\tilde{U}(x|r)$ so that the overall choice utility of the reference point is zero. We obtain the utility function $U(x|r)$ with:

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

We will refer to any function of the above form[1] as an *attention-biased utility function*. If $\alpha_i(r_i) = c$ for a constant c, for all i, then the decision maker does not display any reference-dependent attentional bias, and instead maximizes total value, U^*.

Bhatia (2013a) provides a neuro-cognitive motivation of the above functional form. He shows that a simple class of neural networks can generate attention-biased utility if reference points are assumed to be more salient than competing choice objects. Activation from the nodes representing

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

these reference points spreads to associated nodes, representing their component goods. This in turn affects the inputs into the nodes representing choice object valuation. The activation of these valuation nodes can be described using equation 1, once the network settles. In general, attention-biased utility provides a formal representation of the attentional biases outlined in the previous section. Increasing the amount of a particular good in the reference bundle leads to more attention on the valuation of that good in each choice bundle, so that particular valuation has a greater impact on overall utility. Note also that when V_i are strictly increasing, the attentional weights can equivalently be seen as strictly increasing functions of $V_i(r_i)$.

Note that there are no restrictions on the shape of 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*: bundles that are strongest on the reference point's primary dimensions are the ones that are chosen. This is in contrast to models of reference dependence based on prospect theory (Tversky and Kahneman, 1991; Kőszegi and Rabin, 2006) in which built-in gain-loss asymmetries generate preferences for bundles that are unambiguous gains over the reference point, relative to bundles that involve tradeoffs from the reference point.

Our account 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 towards a particular good in a bundle depends (discontinuously) on the dispersion of the amount of that good 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 Section 4.2.2. In contrast, by adopting a very simple attention weighting function and positing that an endowment serves as a natural reference point, we capture the endowment effect as well as other reference-dependent choice anomalies without additional assumptions.

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

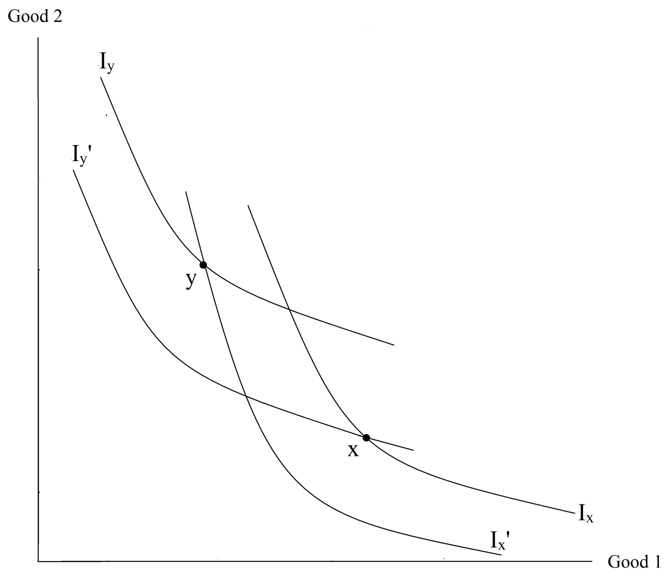

Figure 1: Indifference curves I_x and I_y when x and y respectively are the reference points

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. Here, reference-dependent utility can simply be captured as a dot product of the vector of valuations of the reference point, $V_r = \langle V_1(r_1), V_2(r_2), ..., V_N(r_N)\rangle$, with the vector of valuations of the evaluated bundle, $V_x = \langle V_1(x_1), V_2(x_2), ..., V_N(x_N)\rangle$. The decision maker in turn chooses the bundle 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.

Figure 1 demonstrates indifference curves generated by attention-biased utility in a two-good 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 bundle z is indifferent to x when x is the reference point. Valuation vectors lying above I_{V_x} correspond to choice alternatives

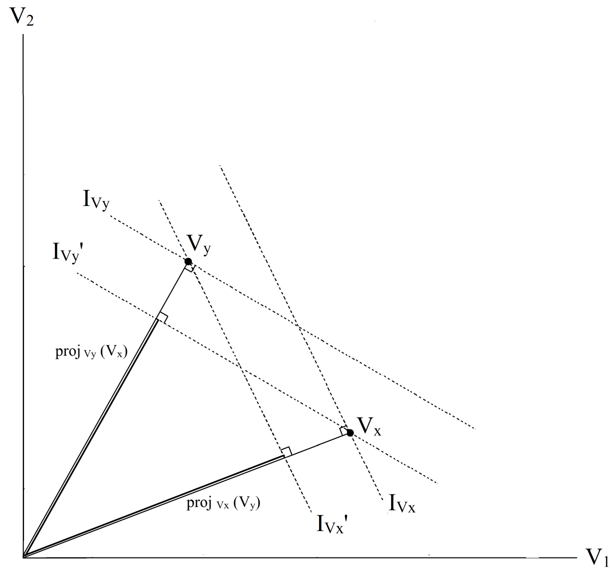

Figure 2: Indifference curves I_{V_x} and I_{V_y} in valuation space, when x and y respectively are the reference points

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, bundles 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 dimensions.

This intuition allows us to derive the first result of this paper. Proposition 1 outlines the general setting in which reference points necessarily change utility in favor of a particular bundle relative to another, for strictly increasing V_i. It states that 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 dimensions that favor x, and s is not smaller than r on the dimensions that favor y, and if r and s differ on at least one dimension on which x and y differ.

Proposition 1. *Let x and y be any non-identical bundles. Index goods such that $x_i \neq y_i$ for $i \in \{1, 2, ..., K\}$, and $x_i = y_i$ for $i \in \{K+1, K+2, ..., N\}$. If all V_i are strictly increasing, then for any reference points r and s such that*

- *if $x_i > y_i$ then $r_i \geq s_i$*
- *if $y_i > x_i$ then $s_i \geq r_i$*
- *$r_i \neq s_i$ for some $i \in \{1, 2, ..., K\}$*

we have $U(x|r) - U(y|r) > U(x|s) - U(y|s)$.

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

4 Implications

This part of the paper studies how reference points affect preferences between bundles. It first shows that attention-biased utility implicitly generates loss averse preferences without requiring any non-standard 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.

4.1 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 and 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 Kőszegi 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. While prospect theory's kink in utility at the reference point can induce loss aversion, i.e., 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 bundles such that $x_i \geq y_i$ for all $i \in \{1, 2, ..., N\}$, and $x_i > y_i$ for some $i \in \{1, 2, ..., N\}$, 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 towards at least one underlying good, relative to a dominated reference point. The goods 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 goods, generating loss aversion as the reference points are varied.[2]

4.2 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 valuation measures of preference, such as willingness-to-pay or

[2]The built-in gain-loss asymmetry in prospect theory provides one account for anomalous low-stakes risk aversion, which Rabin (2000) shows 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 and Loewenstein, 2012).

willingness-to-accept, or through explicit choices between two or more items. Section 4.2.1 explores the implications of attention-biased utility with regards to valuation measures of preference. Section 4.2.2 explores reference-dependent choices. 4.2.3 extends this reasoning to model anchoring effects.

4.2.1 Measures of Preference

Let us consider four different measures of preference, as formalized in Bateman et al. (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 goods i and j, and use these measures to study how changes to good i impact the decision maker's preferences in units of good j. We will hold, $x_i > y_i$ and $x_j = y_j$. Due to the independence between dimensions for consumption utility and attention, this two-good setting easily generalizes to more complex cases.

Since we are only considering two goods, we can write preferences for any choice z given any reference point r as $U(z_i, z_j|r)$, and the four measures used to measure preference can be defined as:

1. $U(x_i, x_j - WTP|y) = U(y_i, x_j|y)$
2. $U(y_i, x_j + WTA|x) = U(x_i, x_j|x)$
3. $U(x_i, x_j - EL|x) = U(y_i, x_j|x)$
4. $U(y_i, x_j + EG|y) = U(x_i, x_j|y)$

Willingness-to-pay by this definition is the largest loss on good 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 good 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 good 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 good 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 Batemen et al., 1997 for a discussion). However, a number of researchers find that both EG and EL ratings deviate from WTP and WTA respectively (Knetch, 1989; Kahneman et al., 1990; Loewenstein and Adler, 1995; Batemen et al., 1997). 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 bundles 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. Since $x_i > y_i$, reference point x leads to increased attention towards good i relative to reference point y. On the other hand, since $x_j = y_j$ attention towards good j is constant regardless of the reference point. This leads to a bias in favor of good 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.

4.2.2 Choices

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

Additionally, reference dependence can be observed when the most frequently chosen bundle is not the reference point itself. For example, Tversky and Kahneman (1991) note that decision makers generally prefer choice bundles that are strict improvements over their reference point, relative to bundles 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 goods and very strong on other goods lead to stronger biases in preference, relative to more evenly distributed reference points. Finally, Tversky and Kahneman (1991) note that decision

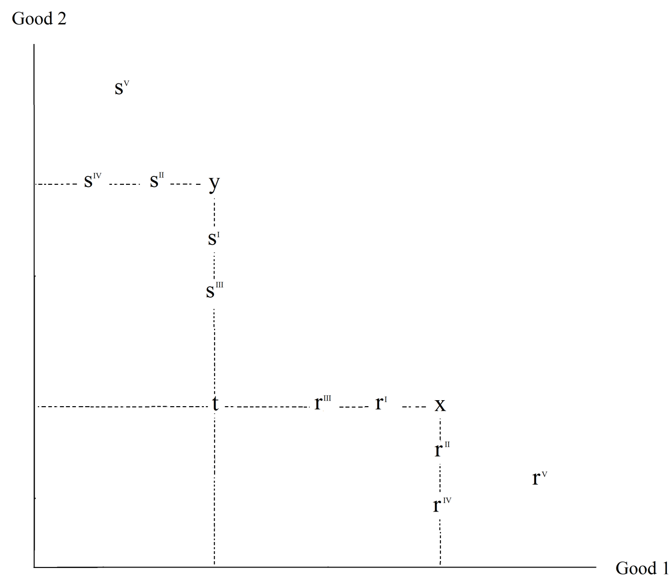

Figure 3: Reference-dependent anomalies in choice

makers tend to prefer bundles that involve small tradeoffs from the reference point relative to bundles 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-good choice sets consisting of two desirable bundles 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 $P_{x,y}(r) = U(x|r) - U(y|r)$, as the relative preference for x over y 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 and Sinden, 1984; Samuelson and Zeckhauser, 1988; Knetsch, 1989): 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 vs. tradeoffs (Tversky and Kahneman, 1991; Herne, 1998): 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 vs. 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 vs. 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 vs. 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 vs. large tradeoffs (Tversky and 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 $P_{x,y}(r^V) > P_{x,y}(s^V)$.

Tversky and Kahneman (1991) use 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, where as effect 6 requires both the gain-loss asymmetry and convex utility in losses. Herne (1998) provides 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 goods i, j, and for all bundles 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 $P_{x,y}(r) > P_{x,y}(s)$.*

Munroe 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 towards, and subsequently a higher weight on good i. Since x is more valuable on good 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. According to our theory, any focal bundle 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 Dhar and Simonson (1992), Dhar et al. (1999), Bushlong et al. (2010); Krajbich et al. (2010), Reutskaja et al. (2011) and Lim et al. (2011).

4.2.3 Anchoring Effects

The above two sections explore settings in which salient bundles, such as endowments and other reference points, affect the decision maker's choices between, and valuations of, available options. These sections show preferences are biased in favor of reference points and other options close to the reference point.

A similar type of bias has been shown to emerge with anchors. The typical anchoring study involves the valuation of a choice item in terms of either willingness-to-pay (WTP) or willingness-to-accept (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. Johnson and Schkade, 1989; Ariely et al., 2003; Beggs and Graddy, 2009). Research by Ariely et al., (2003) also finds 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.

This effect can be more formally represented in terms of the measures of preference outlined in section 4.2.1. For simplicity the discussion here will only explore willingness-to-pay. Analogous results hold for willingness-to-accept.

As in section 4.2.1., we can consider a two-good space, where i represents the good at hand and j represents money or other goods 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 x_j. During the evaluation task the decision maker is asked to list an amount WTP such that $U(x_i, x_j - WTP|r) = U(y_i, x_j|r)$.

High anchor settings ask the decision maker whether she wants to buy the item for a_H, whereas the low anchor settings 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 bundle $A_H = (x_i, x_j - a_H)$, which represents ownership of the item for the price a_H. Likewise a low anchor can be formalized as a bundle $A_L = (x_i, x_j - a_L)$, which represents ownership of the item for low price of a_L.

Recall that we are suggesting that reference points are bundles that are salient to the decision maker. We assume that asking decision makers to evaluate the anchor increases the salience of the anchor. The anchor thus serves as a reference point. In effect, the decision maker is asked to select either WTP_{A_H} or WTP_{A_L} such that $U(x_i, x_j - WTP_{A_H}|A_H) = U(y_i, x_j|A_H)$ or $U(x_i, x_j - WTP_{A_L}|A_L) = U(y_i, x_j|A_L)$.

Proposition 5 shows that any attention-biased utility function with strictly increasing V_i will generate $WTP_{A_H} > WTP_{A_L}$. 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 bundle 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 dimension j relative to dimension i implies that the decision maker is more willing to give up dimension j to obtain dimension i. This leads to $WTP_{A_H} > WTP_{A_L}$.

Corollary 1 establishes that the 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

(Ariely et al., 2003).

Proposition 5. *If V_i are strictly increasing, then for any anchors A_H and A_L such that $a_H > a_L$, we have $WTP_{A_H} > WTP_{A_L}$.*

Corollary 1. *If V_i are strictly increasing, then for any anchor A, and any $z_i > x_i > y_i$ with $U(z_i, x_j - WTP_z|A) = U(y_i, x_j|A)$ and $U(x_i, x_j - WTP_x|A) = U(y_i, x_j|A)$, we have $WTP_z > WTP_x$.*

Anchoring effects emerge from the same assumptions that generate reference dependence. Both these anomalies arise from attention-biased utility with any strictly increasing valuation functions V_i, including those that are universally concave without a built-in asymmetry in gains and losses.

5 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. Since 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 policy maker are able to exert some control over the reference point. This section establishes a non-monotonicity in the impact of reference points on utility, which has implications for theories of goals and aspirations as reference points.

5.1 Equilibrium Choice

Thus far we have studied attention-biased utility as it relates to solitary decisions, ones in which the decision maker is not allowed to change her mind after making her choice. Yet it is quite possible that moving from one bundle 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 and 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-good 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. Since $U(y|x) > U(z|x) = U(x|x)$, 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 $U(z|y) > U(y|y) > U(x|y)$. This leads to a second choice, which moves the decision maker from y to z. At z she is content, as $U(z|z) > U(y|z) > U(x|z)$, 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 in order 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. Kőszegi 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 bundles 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 $C^E(X, U) = \{x \in X : U(x|x) \geq U(y|x), \forall y \in X\}$. This set is analogous to set of personal equilibria proposed by Kőszegi and Rabin (2006).

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

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 $U(x|y) = max\{U(z|y), \forall z \in X\} > U(y|y)$. Note the presumption that if $U(x|y) = U(y|y)$, 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 $x \succ_U y$. Additionally, for any x, y such that $x \succ_U y$, 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 $\langle x(0), x(1), ..., x(t), ... \rangle$ 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 $x(T)$, i.e. $x(T+1) = x(T)$. 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 2 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? 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 2. *For strictly increasing V_i, and for any finite, non-empty 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 consumption utility is increasing in goods. This implies the corollary that the set of equilibrium choices is non-empty 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 and 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 (see e.g. List, 2003). 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 pure consumption utility, to lie in C^E: choice options that maximize total value U^* may be rejected in favor of bundles 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 consumption utility for the decision maker than z. However, we have $C^E(X, U) = \{z\}$. Settings like this points to a particularly troubling disconnect between welfare measures involving choice, and welfare measures involving the maximization of underlying valuations, or consumption utility.

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 Kőszegi and Rabin (2006). This model provides a number of important insights regarding the reference-dependent behavior of sophisticated decision makers (see also Loomes and Sugden, 1986; Kőszegi and Rabin, 2007 and 2009; Ericson and Fuster, 2011; Abeler et al., 2011; Gill and Prowse, 2012). However, as demonstrated by Gul and Pesendorfer (2006), Kőszegi 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.

5.2 Optimal Reference Points

Reference points determine not only choice but also the amount of 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.

In order 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 et al., 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 bundle tend to be dominated by that bundle, but also non-trivial, 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 bundle) is always $\mathbf{0}$. However, as Proposition 7 demonstrates, attention-biased utility accords with the empirical research on goal setting. Given a particular consumption bundle the optimal reference point is inferior to the bundle on every dimension, but, more interestingly, is often strictly positive on each dimension.

Proposition 7. *For any attention-biased utility with strictly increasing V_i and any consumption bundle x, the optimal reference bundle (maximizing $U(x|r)$) 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 $\alpha_i(0) = 0$. Additionally, 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 bundle's strongest dimensions direct more attention towards these dimensions, thereby increasing the bundle's overall valuation. Reference points that are too good, however, reduce the bundle's valuation because the utility function contrasts the actual bundle 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 bundle necessarily has intermediate values: values that are greater than zero, but of course smaller than those of the bundle itself. Moreover, it is straightforward to observe that more ambitious goals (higher reference points) are more motivating because they increase marginal utility.

6 Undesirable Goods

The above sections provide 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. While reference points like endowments are more likely to be chosen relative to their competitors in choices amongst desirable goods, this work finds that endowments are less likely to be chosen in choices amongst undesirable goods (Brenner et al., 2007; Bhatia and Turan, 2012).

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 primary dimensions. This generally leads to an increased preference for desirable endowments. Undesirable reference points, however, have undesirable dimensions. Increased attention towards these dimensions should reduce the desirability of the endowment. This is formalized in Proposition 8.

Proposition 8. *If V_i are strictly decreasing, and if x and y are any two non-identical bundles, then we have $U(x|x) - U(y|x) < U(x|y) - U(y|y)$.*

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. Moreover, sequential choice among undesirable goods with reference point adaptation does not converge to a stable equilibrium, as it would with desirable 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 bundles, 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 non-focal undesirable bundles. 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.

7 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 involve considerable deviations from economic theory, such as asymmetries in gains and losses, and convex utility in losses.

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 dimensions involved in choice task. This can affect the weighting of goods and subsequently influence choice. This paper presents a formal model of this mechanism and shows that it explains a range of behavioral anomalies associated with reference dependence. In particular, the model generates anomalies not captured by prospect theory, such as the reversal of the endowment effect for undesirable goods, as well as anomalies in other domains, such as anchoring. Another prediction of attention-biased utility – a stronger endowment effect for more valuable goods – also finds confirmatory evidence in the psychology literature (Knutson et al., 2008). Inspired by these successful predictions, Bhatia (2013b) systematically compares this model against prospect theory across settings in which they make contrasting predictions and finds that attention-biased utility offers a better quantitative fit to choice 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 et al., 2011). While 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) find a wage premium for manufacturing jobs that expose workers to fatal injury risk, but find that workers are nevertheless more likely to leave more hazardous jobs.

Attention is a relatively well understood cognitive mechanism. Within economics, attention has been used to make sense of a large range of individual behaviors, including response mode effects (Tversky et al., 1990), biased judgment (Kahneman, 2003), focalism (Kőszegi and Szeidl, 2013), suboptimal search (Gabaix et al., 2006; Caplin et al., 2011), and anomalies in risky choice (Bordalo et al., 2012a). New tools such as eye-tracking and brain-imaging allow for a rigorous test of attention-based economic models (see e.g. Fehr and Rangel, 2011 for a review). Explaining reference dependence in terms of attention thus not only provides a psychologically and neuroscientifically plausible account for these phenomena, but also promises that the underlying cognitive mechanism, as well as the patterns of choice, can be subjected to rigorous empirical tests. Moving beyond a collection of counterintutive observations and stylized facts to a coherent theory that withstands rigorous empirical testing has always been the way forward in studying deviations from economic rationality. The next step is then moving from a collection of disparate theories, each with their own domain of applicability, to an overarching theory that relies on a common (and empirically verifiable) mechanism to account for seemingly unrelated phenomena. We suggest that attention is that mechanism.

A Proofs

A.1 Proof of Proposition 1

Consider any x, y, r and s. Index goods such that $x_i \neq y_i$ for $i \in \{1, 2, ..., K\}$ and $x_i = y_i$ for $i \in \{K+1, K+2, ..., N\}$. We can thus write:

$$[U(x|r) - U(y|r)] - [U(x|s) - U(y|s)] = \sum_{i=1}^{K}[\alpha_i(r_i) - \alpha_i(s_i)] \cdot [V_i(x_i) - V_i(y_i)]. \tag{2}$$

Our conditions guarantee that $x_i > y_i$ implies $r_i \geq s_i$ and that $y_i > x_i$ implies $s_i \geq r_i$. Since V_i and α_i are strictly increasing for all i, and $r_i \neq s_i$ for some $i \in \{1, 2, ..., K\}$, we have $[\alpha_i(r_i) - \alpha_i(s_i)] \cdot [V_i(x_i) - V_i(y_i)] > 0$ for all $i \in \{1, 2, ..., K\}$, Hence the sum over all i, $[U(x|r) - U(y|r)] - [U(x|s) - U(y|s)]$, is positive, giving us $[U(x|r) - U(y|r)] > [U(x|s) - U(y|s)]$. □

A.2 Proof of Proposition 2

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

A.3 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:

$$\alpha_i(y_i) \cdot V_i(x_i) + \alpha_j(x_j) \cdot V_j(x_j - WTP) = \alpha_i(y_i) \cdot V_i(y_i) + \alpha_j(x_j) \cdot V_j(x_j) \tag{3}$$

$$\alpha_i(x_i) \cdot V_i(x_i) + \alpha_j(x_j) \cdot V_j(x_j - EL) = \alpha_i(x_i) \cdot V_i(y_i) + \alpha_j(x_j) \cdot V_j(x_j). \tag{4}$$

Subtracting Equation 4 from 3, and simplifying, gives us:

$$\alpha_j(x_j) \cdot [V_j(x_j - WTP) - V_j(x_j - EL)] = [\alpha_i(x_i) - \alpha_i(y_i)] \cdot [V_i(x_i) - V_i(y_i)]. \tag{5}$$

Since V_i and α_i are both strictly increasing in their arguments, the right hand side of Equation 5 is positive. Then, since α_i is always positive, we get $V_j(x_j - WTP) > V_j(x_j - EL)$. Once again, since V_i is strictly increasing, this implies $WTP < EL$. □

A.4 Proof of Proposition 4

Proposition 1 implies condition 1, which is sufficient to generate effects 1-6.

A.5 Proof of Proposition 5

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

$$\alpha_i(x_i) \cdot V_i(x_i) + \alpha_j(x_j^H) \cdot V_j(x_j - WTP_{A_H}) = \alpha_i(x_i) \cdot V_i(y_i) + \alpha_j(x_j^H) \cdot V_j(x_j) \tag{6}$$

$$\alpha_i(x_i) \cdot V_i(x_i) + \alpha_j(x_j^L) \cdot V_j(x_j - WTP_{A_L}) = \alpha_i(x_i) \cdot V_i(y_i) + \alpha_j(x_j^L) \cdot V_j(x_j). \tag{7}$$

Putting together equations (6) and (7) gives us:

$$\alpha_j(x_j^H) \cdot [V_j(x_j) - V_j(x_j - WTP_{A_H})] = \alpha_j(x_j^L) \cdot [V_j(x_j) - V_j(x_j - WTP_{A_L})] > 0. \tag{8}$$

We know that $\alpha_j(x_j^L) > \alpha_j(x_j^H) > 0$, as α_j is strictly increasing and $a_H > a_L$. This implies that $V_j(x_j) - V_j(x_j - WTP_{A_H}) > V_j(x_j) - V_j(x_j - WTP_{A_L}) > 0$. Again, since V_j is strictly increasing, this implies that $WTP_{A_H} > WTP_{A_L}$. □

A.5.1 Proof of Corollary 1

Since the reference point A is 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 $WTP_z > WTP_x$.

A.6 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,

$$[U(z|z) - U(x|z)] + [U(y|y) - U(z|y)] + [U(x|x) - U(y|x)] < 0. \tag{9}$$

We can expand this sum of utility differences (in Equation 9) as

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

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

$$[\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]. \quad (10)$$

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 $y_i > z_i$ resp.) 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 10, we obtain

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

$$+\; [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] \;+\; [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]$$

$$= \sum_{i=1}^{N} [\alpha_i(y_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(x_i)], \quad (11)$$

which of course is strictly positive by the same argument as above. We can separate out the contributions on each dimension and isolate the term we just added in Equation 11, obtaining for any i,

$$[\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(z_i) - V_i(x_i)]$$

$$= [\alpha_i(y_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(x_i)] \;-\; [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)]$$

$$-\; [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)] \;-\; [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]$$

$$> -\, [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(y_i) - V_i(z_i)] \;-\; [\alpha_i(z_i) - \alpha_i(x_i)] \cdot [V_i(z_i) - V_i(x_i)]$$

$$-\; [\alpha_i(y_i) - \alpha_i(z_i)] \cdot [V_i(y_i) - V_i(z_i)]. \quad (12)$$

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 10 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 9. □

A.6.1 Proof of Corollary 2

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 $T \leq t$.

Now consider any X and any U such that for some time t we have some $x(t) \in X$, but $x(t) \notin C^E(X,U)$. Since $x(t) \notin C^E(X,U)$ there exists some $x(t+1) \in X \backslash \{\bigcup_{t' \leq t} x(t')\}$ such that the decision maker trades $x(t)$ for $x(t+1)$. (We can rule out $x(t+1) = x(t')$ for some $t' \leq t$ by the acyclicity of the preference relation $\succ_U$.) 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. □

A.7 Proof of Proposition 7

The contribution to attention-biased utility on dimension 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 $0 \leq r_i < x_i$ 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 $\frac{\partial U(x|r)}{\partial r_i}\Big|_{r_i=0} = \alpha_i'(0) \cdot (V_i(x_i) - V_i(0)) \;-\; \alpha_i(0) \cdot V_i'(0)$ 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. □

A.8 Proof of Proposition 8

Consider any undesirable x and y. We can write:

$$[U(x|x) - U(y|x)] - [U(x|y) - U(y|y)] = \sum_{i=1}^{K}[\alpha_i(x_i) - \alpha_i(y_i)] \cdot [V_i(x_i) - V_i(y_i)]. \qquad (13)$$

Note that V_i is non-positive 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. Since x and y are non-identical, 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, $[U(x|x) - U(y|x)] - [U(x|y) - U(y|y)]$ is negative, giving us $[U(x|x) - U(y|x)] < [U(x|y) - U(y|y)]$. □

References

Abeler, J., Falk, A., Goette, L., & Huffman, D., (2011). Reference Points and Effort Provision. American Economic Review, 101(2), 470-92.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences. Quarterly Journal of Economics, 112(2), 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(3), 254-267.

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(1), 73-105.

Beggs, A. & Graddy, K. (2009). Anchoring Effect: Evidence from Art Auctions. American Economic Review, 99(3), 1027-1039.

Bhatia, S. (2013a). Associations and the Accumulation of Preference. Psychological Review. Forthcoming.

Bhatia, S. (2013b). Comparing Theories of Reference Dependence. Working Paper.

Bhatia, S., & Turan, N. (2012). Attention and Reversals of the Endowment Effect. Working Paper.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012a). Salience Theory of Choice Under Risk. Quarterly Journal of Economics, 127(3).

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012b). Salience in Experimental Tests of the Endowment Effect. American Economic Review, 102(3), 47-52.

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(3), 369-376.

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(4), 1-16.

Caplin, A., Dean, M., & Martin, D. (2011). Search and Satisficing. American Economic Review, 101(7), 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(3), 360-370.

Chapman, G., & Johnson, E. (1999). Anchoring, Activation, and the Construction of Values. Organizational behavior and human decision processes, 79(2), 115-153.

Cottini, E., Kato, T., & Westergaard-Nielsen, N. (2011). Adverse Workplace Conditions, High-Involvement Work Practices and Labor Turnover: Evidence from Danish Linked Employer-Employee Data. Labour Economics 18(6), 872-880.

Dhar, R., & Simonson, I. (1992). The Effect of the Focus of Comparison on Consumer Preferences. Journal of Marketing Research, 29(4), 430-440.

Dhar, R., Nowlis, S., & Sherman, S. J. (1999). Comparison Effects on Preference Construction. Journal of Consumer Research, 26(3), 293-306.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective Well-being: Three Decades of Progress. Psychological Bulletin, 125(2), 276-302.

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(4), 1879-1907.

Fehr, E., & Rangel,A. (2011). Neuroeconomic Foundations of Economic Choice - Recent Advances. Journal of Economic Perspectives, 25(4), 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(4), 1043-1068.

Gill, D., & Prowse, V. (2012). A Structural Analysis of Disappointment Aversion in a Real Effort Competition. American Economic Review, 102(1), 469-503.

Golman, R., & Loewenstein, G. (2012). Curiosity, Information Gaps, and the Utility of Knowledge. Available at SSRN: http://ssrn.com/abstract=2149362.

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(1), 79-109.

Herne, K. (1998). Testing the Reference-dependent Model: An Experiment on Asymmetrically Dominated Reference Points. Journal of Behavioral Decision Making, 11(3), 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(3), 463-470.

Johnson, E. J., & Schkade, D. A. (1989). Bias in Utility Assessments: Further Evidence and Explanations. Management Science, 35(4), 406-424.

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of Endowment: A Query Theory of Value Construction. Journal of experimental psychology: Learning, memory, and cognition, 33(3), 461-74.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. American Economic Review, 93(5), 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(6), 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(1), 193-206.

Knetsch, J. (1989). The endowment effect and evidence of nonreversible indifference curves. American Economic Review, 79(5), 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(3), 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(4), 1133-1165.

Köszegi, B., & Rabin, M. (2007). Reference-Dependent Risk Attitudes. American Economic Review, 97(4), 1047-1073.

Köszegi, B., & Rabin, M. (2009). Reference-Dependent Consumption Plans. American Economic Review, 99(3), 909-936.

Köszegi, B., & Szeidl, A. (2013). A Model of Focusing in Economic Choice. Quarterly Journal of Economics. 128(1), 53-104.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual Fixations and the Computation and Comparison of Value in Simple Choice. Nature Neuroscience, 13(10), 1292-1298.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics, 112(2), 443-478.

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(37), 13214-13223.

Loewenstein, G., & Adler, D. (1995). A Bias in the Prediction of Tastes. The Economic Journal, 105(431), 929-937.

Loomes, G., & Sugden, R. (1986). Disappointment and Dynamic Consistency in Choice under Uncertainty. Review of Economic studies, 53(2), 271-82.

Masatlioglu, Y., Nakajima, D., & Ozbay, E. (2012). Revealed Attention. American Economic Review, 102(5), 2183-2205.

Munro, A., & Sugden, R. (2003). On the Theory of Reference-Dependent Preferences. Journal of Economic Behavior & Organization, 50(4), 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(2), 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(9), 1142-1150.

Nayakankuppam, D., & Mishra, H. (2005). The Endowment Effect: Rose-Tinted and Dark-Tinted Glasses. Journal of Consumer Research, 32(December), 390-395.

Novemsky, N., & Kahneman, D. (2005). The Boundaries of Loss Aversion. Journal of Marketing Research, 43(2), 119-128.

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(4), 1108-1116.

Rabin, M. (2000). Risk Aversion and Expected-utility Theory: A Calibration Theorem. Econometrica 68(5), 1281-1292.

Reutskaja, E., Camerer, C. F., & Rangel, A. (2011). Search Dynamics in Consumer Choice Under Time Pressure. American Economic Review, 101(2), 900-926.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy 82, 34-55.

Simon, H. (1955). A Behavioral Model of Rational Choice. Quarterly Journal of Economics, 69(1), 99-118.

Strack, F., & Mussweiler, T. (1997). Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility. Journal of Personality and Social Psychology, 73(3), 437-446.

Thaler, R. (1980). Toward a Positive Theory of Consumer Choice. Journal of Economic Behavior & Organization, 1(1), 39-60.

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. Quarterly Journal of Economics, 106(4), 1039-1061.

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(3), 303-324.