An Information-Processing Theory of Loss Aversion*

J. Miguel Villas-Boas
(University of California, Berkeley)

October, 2021

*Comments and suggestions by Mark Dean, David Gal, Emir Kamenica, Leif Nelson, Drazen Prelec, Derek Rucker, Jakub Steiner, and Birger Wernerfelt on a previous version of the paper are gratefully appreciated. E-mail address: villas@haas.berkeley.edu.

An Information-Processing Theory of Loss Aversion

Abstract

This paper considers a model where a risk-neutral decision-maker receives both a signal about whether an outcome is above a certain threshold (a reference point) and a signal on the value of the outcome. Given these two signals, the expected value function of the outcome exhibits diminishing sensitivities both above and below the reference point, and loss aversion if the reference point is not too low. The paper shows how the informativeness of each signal affects the declining sensitivities and loss aversion effects, and reduces to risk-neutral decision-making when the signal on the value of outcome is perfectly informative. The loss aversion effects occur for high reference points because of the signal on the value of the outcome, and can occur for low reference points because the decision-maker may pay greater attention to the less likely low outcomes.

1. Introduction

In the last few decades, since Kahneman and Tversky (1979), there has been a large literature documenting the existence of reference point effects, with losses affecting decision-makers more than gains, loss aversion. This has been documented in the context of lotteries, riskless choice (e.g., Thaler 1980, Tversky and Kahneman 1991, Koszegi and Rabin 2006), with both experiments and field evidence (e.g., Camerer 2000). See O'Donoghue and Sprenger (2018) for a recent survey of the evidence provided and variations on this general effect.

A general set-up of the theory, "prospect theory," involves both diminishing sensitivities and loss aversion, where the utility with respect to some outcome x, $V(x)$ is normalized to zero at the reference point, and satisfies $V''(x) < 0, V(x) < -V(-x)$, and $V'(x) < V'(-x)$ for $x > 0$, and $V''(x) > 0$ for $x < 0$. As noted in O'Donoghue and Sprenger (2018) most of the literature has focused on the loss aversion effects, with a two-part linear formulation, with $V(x) = x$ for $x > 0$, and $V(x) = \lambda x$ for $x < 0$, with $\lambda > 1$.

This paper takes the perspective that information processing on the value of an outcome is imperfect, and that decision-makers receive information (potentially also imperfect) about whether the outcome is above or below a certain threshold, which will be the reference point. This later component captures the idea that decision-makers may find it easier to process information in terms of discrete categories, than capturing information about some continuous variable, and information just above or below a threshold simplifies that processing of information to only two categories. The paper considers that the decision-maker is risk neutral and receives a noisy signal about the value of the outcome, and a noisy signal about whether the outcome is above or below the threshold. Given those signals the decision-maker updates her beliefs about the value of the outcome, and obtains an expected value of the outcome given the signals.

Going one step further, given a value of the outcome and the data generating process for the signals, the decision-maker can compute the expected value over the possible signals that the decision-maker will receive of the expected value of the outcome given the signals. Given the noisy signal about the value of the outcome, this expected value over the possible signals is going to be less steep than the outcome, leading to the possible diminishing sensitivities effect. The signal about whether the outcome is above or below the threshold will lead the expected value over the possible signals to be steeper close to the threshold (which is an implication of prospect theory). Furthermore, under some conditions, we can obtain loss aversion effects, such that the expected value over signals increases less with the outcome for outcomes above the threshold than for outcomes below the threshold. The loss aversion effects can occur for high reference points because of the signal on the value of the outcome, and the set of outcomes above the reference having lower range, and can occur for low reference points because the decision-maker may pay greater attention to the less likely low outcomes (in a rational inattention framework with binary actions).

Some benchmarks are interesting to study in this set-up. First, if the signal on the value of the outcome is fully informative, the model above reduces to expected utility theory. Second, if the signal on the value of the outcome does not have any information, all the action comes from the signal about whether the outcome is above or below the threshold, and in that case we can still obtain diminishing sensitivities and loss aversion effects. Third, another interesting benchmark is the case is which the signal about whether the outcome is above or below the threshold is fully informative. In that case, the expected value over the signals has a discontinuity (it is vertical) at the reference point, being the extreme case of the value function being steeper close to the reference point.

To give further intuition on the effect presented here related to loss aversion, consider an example in which an individual may reject a gamble that pays either \$15 or −\$10 with equal probability. The individual would receive signals on each of these two outcomes and could come to the conclusion that −\$10 is worse than \$15 is good (with a reference point of \$0). This may occur because the individual could confuse −\$10 with the average loss, and \$15 with the average gain, and the average loss may be worse than the average gain is good. Regarding the comparison between the average loss and average gain, consider the potential range of each. Regarding the possible loss, one can pick the least preferred extreme misery from some form of a health problem, to social ostracism, to a car accident, to passing away. In terms of possible gains, an individual has at most 24 hours per day, and therefore what is gained has to be enjoyed during that limited period of time. What is the extra utility gained of having more friends, mansions, aircraft, or yachts? This range of possible outcomes may suggest that the range of losses is bigger than the range of gains, with, possibly, the average loss being worse than the average gain is good.[1] In the model considered here, this effect is obtained when the individual has a higher reference point. Another question is the extent to which the evaluation of the two outcomes is independent. In fact, the individual may have a good sense that the number 15 is 50% greater than the number 10, but when evaluating

[1]Baumeister et al. (2001) discusses how the worse outcomes may be stronger than the better outcomes, and notes that "Survival requires urgent attention to bad outcomes, but it is less urgent with regard to good ones" (p. 325).

the benefit of the gain of $15, uses the individual's perception of what a gain could be, and when evaluating the effects of a loss, the individual understands what a loss may be.

In some sense, the interpretation presented here can be seen as almost semantic, as when evaluating the "loss of $10" the individual gains information from both the word "loss" and the number "10." The final evaluation depends on how much weight to put on each piece of information. Prospect theory gives a specific effect to the word "loss," and this paper can be seen as providing a justification about why the word "loss" has that effect.

We note that there has also been recent research providing evidence that losses may, in some situations, have a lower impact than gains of the same size (e.g., Gal and Rucker 2018). In relation to that literature this paper illustrates that the reference point may affect whether losses have or have not a greater impact than gains (see also Martin, Reimann, and Norton 2016, Higgins and Liberman 2018, Wardley and Alberhasky 2021). Considering the signal to be above or below a threshold as optimal for a decision-maker with limited attention in a binary action problem, Woodford (2012) presents diminishing sensitivity effects in the value function. In relation to that paper, here I consider also the possibility of a signal on the value of the outcome, and investigate loss aversion effects. See also Khaw, Li, and Woodford (2020) on the possibility of errors in cognition leading to small-stakes risk aversion. There is also a literature on obtaining diminishing sensitivities based on ease of comparison with most frequent alternatives, and which leads to loss aversion effects if the distribution of alternatives is positively skewed (e.g., Friedman 1989, Stewart, Chater, and Brown 2006). This paper is also related to work that accounts for errors in perception of problems, that may lead to optimal biases in analysis to better counteract those errors in perception (e.g., Steiner and Stewart 2016, Gossner and Steiner 2018). This paper can also be seen in the spirit of offering additional foundations for the loss aversion affects (e.g., Fudenberg 2006).

The remainder of the paper is organized as follows. The next section presents a simple model that illustrates the main effects without formally modeling the signals observed by the decision-makers. Section 3 presents the full general model, and Section 4 considers the benchmark in which the signal about the value of the outcome provides no information (completely noisy signal). Section 5 considers the benchmark in which the signal about whether the outcome is above or below the threshold is fully informative, and Section 6 considers a full model where both signals are partially informative. Section 7 presents concluding remarks.

2. A Simple Model

This Section presents a simple model that illustrates the forces at work, without formally modeling the probability distribution of the signals that the decision-maker can receive. The following Sections discuss conditions on the probability distribution of the signals such that this model holds.

Let x be an outcome that the decision-maker cares about. For example, this could be the amount received by the decision-maker in a lottery, or the benefit of having a certain object. Let $H(x)$ be the cumulative distribution function of x prior to the consumer receiving any signals. The decision-maker is assumed to be risk-neutral with respect to x in its true utility, $U(x) = x$. This "true utility" is the utility in the case in which the decision-maker has full information about x, by which we also mean that the decision-maker has fully processed what x represents. Risk-neutrality on x allows us to focus on the curvature of the value to the decision-maker of x when she is making decisions, $V(x)$, The curvature on $V(x)$ will come from the imperfect information processing of x, generated by not fully informative signals.

Let r be the reference point of the decision-maker, let S_1 be the signal that the outcome x is above or equal to the reference point, and let S_0 be the signal that the outcome x is below the reference point. For the purposes of this Section let us consider that S_1 and S_0 are fully informative. That is, $\Pr(x \geq r|S_1) = 1$ and $\Pr(x < r|S_0) = 1$.

Consider now that the value of x for the decision-maker, $V(x)$, is a weighted average of x and the expected value of x given that the decision-maker observed either S_1 or S_0. That is,

$$V(x) = \begin{cases} \beta x + (1-\beta)E(x|x \geq r) & \text{if } x > r \\ r & \text{if } x = r \\ \beta x + (1-\beta)E(x|x < r) & \text{if } x < r \end{cases} \tag{1}$$

where β is the weight on x. The idea would be that the weight on x would be greater if the signal received on the outcome x would be more informative, and we investigate this idea in the following Sections. When $\beta \to 1$, the decision-maker processes fully the information on x and would behave like an expected utility maximizer. When $\beta \to 0$, the signal on x would not provide any information, and the decision-maker has just to rely on the information whether x is above or below the reference point.

Note that this formulation has diminishing sensitivities as $\beta < 1$, and that the value function is steeper (in fact, vertical) at the reference point.

Consider now whether this value function exhibits loss aversion. The value function would be seen as exhibiting loss aversion if $V(r^+) - r < r - V(r^-)$ which results in the condition

$$E(x|x \geq r) + E(x|x < r) < 2r. \tag{2}$$

Note that this condition is satisfied if the reference point r is sufficiently high, and is not satisfied if the reference point r is sufficiently low. That is, this points to the existence of loss aversion if the reference point is sufficiently high.

For example, if x is normally distributed, note that we have $\lim_{r \to +\infty} E(x|x \geq r)/r = 1$ and $\lim_{r \to -\infty} E(x|x < r)/r = 1$, which yields, by (2), that there is loss aversion if the reference point r is sufficiently high, and there is no loss aversion if the reference point r is sufficiently low.

For another specific example, consider that x is uniformly distributed on $[0, 1]$. Then $E(x|x \geq r) = (1+r)/2$ and $E(x|x < r) = r/2$ and condition (2) reduces to $r > 1/2$. That is, there is loss aversion if the reference point r is greater than $1/2$. If we expect that reference points are greater than the unconditional expected value of the outcome, we should then observe loss aversion in practice.

Note also that one may expect that the distribution of the outcome x is negatively skewed, as the decision-maker could be concerned about the very bad outcomes. In the uniform distribution considered we could make it negatively skewed by considering that the distribution of x has a mass point at 0 with mass ζ (with $\zeta \in [0, 1)$) and have a uniform distribution on $[0, 1]$ with probability $(1 - \zeta)$. In that case, condition (2) reduces to

$$r > \frac{1 - \zeta}{2(1 + \zeta)}. \tag{3}$$

As $E(x) = (1 - \zeta)/2$ in that case, we would then have that in the case of negatively skewed distribution, a reference point equal to the unconditional expected value of the outcome would lead to loss aversion effects.

Note that this effect of the existence of loss aversion if the reference point is not too low can be seen as resulting from the smaller range of outcomes above the reference point when the reference point is high. This then leads $V(x)$ not to be able to raise too much above the reference point - which leads to gains having smaller effects on $V(x)$ than losses.

Note that the vertical segment of the value function can be smoothed out (made with positive slope) if the signal of whether the outcome is above or below the reference point is not perfectly informative. This possibility is going to be considered in the next sections.

The existence of the reference point allows for information-processing in categories, in the extreme case of just two categories. One could think of existing further signals about finer categories, which may also affect the value function for decision-makers. The point presented here is that with just two categories and a continuous signal on the value of the outcome, one is able to obtain effects that exhibit diminishing sensitivities and loss aversion.

3. General Set-Up with Signals

Consider now a set-up in which we formally consider the signal structure. As mentioned above, there is a signal about whether the outcome is above or below the reference point, S_1 or S_0, respectively, and a signal about the value of the outcome x. These signals can be interpreted as information processing by the decision-maker about the value of x.

For the signal of whether the outcome is above or below the reference point let

$$\Pr(S_1|x) = f(x, r) \tag{4}$$

where $f(x, r) \in [0, 1]$, increases in x, and decreases in r. By complementarity, we then have $\Pr(S_0|x) = 1 - f(x, r)$. Note given monotonicity on x and that $f(x, r)$ is bounded, we have that $\lim_{x\to+\infty} \frac{\partial f(x,r)}{\partial x} = \lim_{x\to-\infty} \frac{\partial f(x,r)}{\partial x} = 0$, and therefore $f(x, r)$ is S-shaped on x (or has more than two inflection points).

For the signal of the outcome x, which we denote by s, we assume that it has some convex support in $\mathbb{R}$, and with conditional density probability function $g(s|x)$ and conditional cumulative probability function $G(s|x)$, decreasing in x. We assume that signals S_i and s are independent given x. Let $h(x)$ be the prior density probability distribution on the outcome x.

3.1. Expected Value of Outcome Given Signals

We would like to compute the expected value of the outcome given the signals, $E(x|S_i, s)$ for $i = 0, 1$. We then need to obtain the conditional distribution of the outcome x given the signals by Bayes' rule,

$$h(x|S_1, s) = \frac{f(x, r)g(s|x)h(x)}{\int f(t, r)g(s|t)h(t)\, dt}, \tag{5}$$

and obtain $h(x|S_0, s)$ in a similar way.

We can then obtain the expected value of the outcome given the signals as

$$E(x|S_i, s) = \int xh(x|S_i, s)\, dx. \tag{6}$$

Let $v(S_i, s) = E(x|S_i, s)$. One may find that $v(S_i, s)$ already has some of the properties of the value function of prospect theory, if we consider that function as a function of the signals received, for above and below the reference point. However, $v(S_i, s)$ depends on the signals observed, which are not considered in prospect theory. But we can go one step further and compute the expected value of $v(S_i, s)$ given the outcome x, which is discussed in the next subsection.

3.2. Expected Value over Signals

Consider now the expected value, given the outcome, of the expected value of the outcome given the signals. This can be understood as decision-makers being asked, for a given outcome x, what they expect the outcome to be given the signals that they may receive. Obtaining these for different draws of the signals, we can take the expected value over the signals. This can also be seen as the average evaluation of the outcome across individuals, given the signals that these individuals may receive.

In fact, in the presentation of prospect theory in Kahneman and Tversky (1979) there is the description of several experiments in which subjects were given the choice between gambles, and reports of the fraction of subjects who prefer one gamble over the other. In the example of the introduction, this could be, for instance, that 70% of the subjects preferred not to take that gamble. The set-up presented here can then justify this different choice by the subjects, by the fraction of individuals who received signals on the possible outcomes such that those signals make the individual choose not to take the gamble. We note that this formulation is not exactly part of prospect theory, as it just presents the value function of the outcome, but it can also be seen as consistent with prospect theory.

Formally, we would then define the value function over the outcome, $V(x)$, as

$$V(x) = f(x, r) \int v(S_1, s)g(s|x)\, ds + [1 - f(x, r)] \int v(S_0, s)g(s|x)\, ds. \tag{7}$$

Intuitively, if the outcome x is large, we would expect the signal s to indicate that x is on the large side, but conservatively because s is not fully informative about x. This could be a region of diminishing sensitivities for x. The same argument can be done for x small. For x close to the reference point r the value function $V(x)$ could be moving relatively fast between $v(S_1, s)$ and $v(S_0, s)$ being relatively steep on x. Note that with $f(x, r)$ being S-shaped in x (or having more than two inflection points) we have diminishing sensitivities if the signal s is not informative on x. The next sections considers several particular cases of this model. One particular straightforward benchmark is that if s is fully informative about x, we then have $V(x) = x$, the expected utility framework.

3.3. Additional Motivation for Signal on Threshold with Binary Actions

One additional motivation for the existence of the signal on the threshold comes from the literature on rational inattention when the decision-maker faces a binary action. Suppose that the decision-maker has to take one of two actions based on the signal the decision-maker receives about x. What would be the information-processing behavior which the decision-maker would like to have to make better decisions while taking into account the costs of information-processing?

In order to consider this question one has to have a particular specification of the costs of information-processing. One particular form of considering those costs is the rational inattention framework in which the costs of processing information are related to the difference between the entropy of the priors and the expected entropy of the posteriors (see, e.g., Sims 1998). Under this framework Matějka and McKay (2015) shows that optimal information-processing behavior is to have only two signals, one in which the decision-maker takes one action, and the other in which the decision-maker takes the other action, and that the probability of choosing each action is consistent with a generalization of the logit model.[2]

In the set-up of this paper suppose that the decision-maker is considering whether to take this alternative that gives uncertain value x or an alternative that gives value r for certain. Then we have (from Matějka and McKay 2015)

$$f(x, r) = \frac{\widehat{e}_1 e^{\alpha x}}{\widehat{e}_1 e^{\alpha x} + (1 - \widehat{e}_1) e^{\alpha r}} \tag{8}$$

for some $\alpha > 0$, and where $\widehat{e}_1$ is the unconditional probability of the decision-maker receiving the signal S_1, $\widehat{e}_1 = \int_x f(x, r) h(x)\, dx$.

[2] See also Woodford 2008, Yang 2011, Gentzkow and Kamenica 2014, Caplin, Dean, and Leahy 2019, 2021.

This formulation assumes that the decision-maker can perfectly design the information structure that the decision-maker desires, and that the costs of that information structure are related to the change in entropy from the priors to the posteriors. In fact, the decision-maker may not be able to perfectly choose the information structure of the signals that the decision-maker receives. For example. it may be possible that the decision-maker also receives for free a signal s on the continuous value of the outcome x, in which case the optimal function $f(x, r)$ is no longer given by (8). Similarly, if the costs of information processing are not related to the change in entropy from the priors to the posteriors, the exact specification of the function $f(x, r)$ may not be this generalized logit model. But this rational inattention framework points to the important incentive of the decision-maker getting a signal about whether the outcome is above or below a certain threshold, the reference point.

4. Signal s Not Informative on Outcome

Consider the case when the signal s provides no information on the outcome. That is, all the inference about the value of the outcome comes from the signal about whether the outcome is above or below the reference point. We consider first the case in which the probability of the decision-maker receiving a signal that the outcome is above a certain threshold is the logit model, and then consider the case in which that probability is the one of the rational inattention framework under a binary action.

4.1. Logit Model

To obtain sharper results consider the logit representation of the function $f(x, r)$

$$f(x, r) = \frac{e^{\alpha x}}{e^{\alpha x} + e^{\alpha r}} \tag{9}$$

where α is a constant which represents the degree of informativeness of the signal S_i. When α is greater the signal becomes more and more informative about whether the outcome is above or below the reference point. Note also that $f(r, r) = 1/2$.

We can then obtain that

$$v(S_1) = \frac{\int x \frac{e^{\alpha x}}{e^{\alpha x}+e^{\alpha r}} h(x)\, dx}{\int \frac{e^{\alpha x}}{e^{\alpha x}+e^{\alpha r}} h(x)\, dx} \tag{10}$$

$$v(S_0) = \frac{\int x \frac{e^{\alpha r}}{e^{\alpha x}+e^{\alpha r}} h(x)\, dx}{\int \frac{e^{\alpha r}}{e^{\alpha x}+e^{\alpha r}} h(x)\, dx}. \tag{11}$$

This yields that $v(S_1) > v(S_0)$ with both $v(S_1)$ and $v(S_0)$ increasing in the reference point r, and $v(S_1) - v(S_0)$ increasing in α. We also have that $\lim_{\alpha\to 0} v(S_1) = \lim_{\alpha\to 0} v(S_0) = E(x)$ and $\lim_{\alpha\to\infty} v(S_1) = E(x|x > r)$ and $\lim_{\alpha\to\infty} v(S_0) = E(x|x < r)$.

Using these expected payoff of the outcome given the signals we can then obtain the value function as

$$V(x) = \frac{e^{\alpha x}}{e^{\alpha x} + e^{\alpha r}}[v(S_1) - v(S_0)] + v(S_0), \tag{12}$$

from which one immediately obtains diminishing sensitivities.

Consider now the question of loss aversion. Note that $V(x)-V(r) = V(r)-V(r-(x-r))$, such that with the current specification for $f(x, r)$ there is no loss (or gain) aversion.

However, note that if we allow the decision-maker to be more accurate at identifying that the outcome x is below the reference point r, than at identifying that it is above the reference point we can obtain that the value function exhibits loss aversion. That is, if

$$f(x) = \begin{cases} \dfrac{e^{\alpha x}}{e^{\alpha x} + e^{\alpha r}} & \text{if } x \geq r \\ \dfrac{e^{\alpha' x}}{e^{\alpha' x} + e^{\alpha' r}} & \text{if } x < r \end{cases} \tag{13}$$

and $\alpha' > \alpha$ we obtain that the value function $V(x)$ exhibits loss aversion. In the next Sections we show that we can obtain loss aversion effects even if the function $f(x, r)$ is symmetric on x around the reference point r.

4.2. Rational Inattention Framework with Binary Actions

Consider now the rational inattention framework with binary actions. Suppose that the decision-maker is considering whether to take this alternative that gives uncertain value x or an alternative that gives value r for certain.

The decision-maker can also decide the information structure, which in this case reduces to choosing function $f(x, r)$, and the costs of processing information are proportional to the difference between the entropy of the priors and the expected entropy of the posteriors. As shown in Matějca and McKay (2015) the optimal function $f(x, r)$ in this case satisfies (8) noted above. In the same way as in the previous subsection we can then obtain $v(S_1), v(S_0)$, and $V(x)$. In order to fully obtain $f(x, r)$ we still need to specify the prior distribution of x. To obtain sharper results let the prior distribution of x be a uniform distribution on the segment $[0, 1]$. From this we can obtain the unconditional probability of the decision-maker receiving signal $S_1, \widehat{e}_1$ implicitly by

$$\widehat{e}_1(e^{\alpha(1-\widehat{e}_1)} - 1) - (1 - \widehat{e}_1)(e^{\alpha r} - e^{\alpha(r-\widehat{e}_1)}) = 0. \tag{14}$$

We can get then that when $r = 1/2$ we have $\widehat{e}_1 = 1/2$. We can also obtain using Proposition 3 in Matějca and McKay (2015) that $\widehat{e}_1$ is decreasing in r, and therefore $\widehat{e}_1 > 1/2$ if and only if $r < 1/2$.

Note first that, as in the previous subsection, $V(x)$ exhibits diminishing sensitivities. Consider now the question of loss aversion. We know that if $V(x)$ is concave at $x = r$, we have that $V(x)$ exhibits loss aversion close to $x = r$. We can check that $V(x)$ is concave at $x = r$ if and only if $\widehat{e}_1 > 1/2$, i.e., if and only if $r < 1/2$. We state this result in the following proposition.

Proposition 1. *Suppose that the signal s on the outcome is completely uninformative and consider the rational inattention framework with binary actions. Then, the value function $V(x)$ exhibits loss aversion if the reference point is sufficiently low.*

Note that this result is the opposite of the result in the simple model of Section 2, in which the value function exhibited loss aversion if the reference point was not too low. We will also replicate that result below in the case in which the decision-maker receives also a continuous signal on the value of the outcome. That result was due to the fact that when the reference point is sufficiently high, there is a lower probability of the outcome being above the reference point, and therefore, the potential for gain above the reference point is smaller than the potential for loss below the reference point.

The opposite effect obtained in the rational inattention framework uncovers a new mechanism that works in the opposite direction of the mechanism of the simple model of Section 2. In the rational inattention framework, the decision-maker puts greater effort to discriminate whether the outcome x is above or below the reference point r, for the values of the outcome close to the reference point. When the reference point is low, the decision-maker is more likely to take the uncertain alternative x, and therefore puts more effort in trying to see if the outcome x is below the reference point. That results in the value function being steeper below than above the reference point, that is, the value function exhibiting loss aversion.

5. Perfectly Informative Signal on the Relation to the Reference Point

We now consider the benchmark in which the signal about whether the outcome is above or below the reference point is perfectly informative, and the signal on the outcome provides some information.

To get sharper results (and to connect with the results in the next Section) we concentrate the presentation on the case in which the prior distribution of x is uniform on $[0,1]$ and the distribution of s given x has a mass point at x of size $p \in [0,1]$, and otherwise is distributed uniformly on $[0,1]$.[3] The parameter p is an index of how informative the signal on the outcome x is. When $p = 0$ that signal is completely uninformative. When $p = 1$ that signal is completely informative.

The signal about whether the outcome is above or below the reference points is fully informative, $\Pr(S_1|x > r) = 1$ and $\Pr(S_1|x < r) = 0$. When $x = r$, we assume that both signals S_1 and S_0 are equally likely, the natural limit when the signal is not fully informative, as discussed in the next Section. That is, $Pr(S_1|x = r) = 1/2$.

We can then obtain using Bayes' rule the expected value of the outcome given the signals,

$$E(x|s, S_1) = v(S_1, s) = \begin{cases} \dfrac{ps + (1-p)(1-r^2)/2}{p + (1-p)(1-r)} & \text{if } s \geq r \\ \dfrac{1+r}{2} & \text{if } s < r. \end{cases} \tag{15}$$

$$E(x|s, S_0) = v(S_0, s) = \begin{cases} \dfrac{r}{2} & \text{if } s \geq r \\ \dfrac{ps + (1-p)r^2/2}{p + (1-p)r} & \text{if } s < r \end{cases} \tag{16}$$

where we use the facts that $E(x|x > r) = (1 - r^2)/2, E(x|x < r) = r^2/2$, and $H(r) = r$.

Taking the expected value over the signals given the outcome x we can then obtain the

[3] A case with a normal distribution of x and of the signal s given x is presented in the Appendix.

value function,

$$V(x) = \begin{cases} p\dfrac{px + (1-p)(1-r^2)/2}{p + (1-p)(1-r)} + (1-p)\dfrac{1+r}{2} & \text{if } x > r \\ \dfrac{p}{2}\left(\dfrac{pr + (1-p)(1-r^2)}{p + 2(1-p)(1-r)} + \dfrac{pr + (1-p)r^2}{p + 2(1-p)r}\right) + (1-p)\left(\dfrac{1}{4} + \dfrac{r}{2}\right) & \text{if } x = r \\ p\dfrac{px + (1-p)r^2/2}{p + (1-p)r} + (1-p)\dfrac{r}{2} & \text{if } x < r. \end{cases} \tag{17}$$

Note the value function exhibits diminishing sensitivities as $V'(x) < 1$ for both $x > r$ and $x < r$.

Consider now the question of loss aversion. To check that we can obtain that we have loss aversion if $V(r) - V(r^-) > V(r^+) - V(r)$, which holds if $r > 1/2$ (with derivation presented in Appendix), which we state in the following proposition.

Proposition 2. *Consider that the signal of whether the outcome is above or below the reference point is fully informative. Given that the signal s on the outcome is partially informative, $p > 0$, the value function exhibits loss aversion if and only if $r > 1/2$. If the signal s on the outcome is completely uninformative, $p \to 0$, there are no loss aversion effects.*

The condition of the reference point not being too low for the value function to exhibit loss aversion can be seen as relatively intuitive. If the reference point is not too low, the possible range of values when the outcome is above the reference point is smaller, and therefore, the potential gain of an average outcome above the reference point is smaller than the potential loss of having an average outcome below the reference point. Note that this result is the opposite of the one in the previous Section in which there was no signal on the value of the outcome, and we allowed the decision-maker to choose the information structure when facing a binary action problem. The result in this Section points to a mechanism which generates loss aversion with high reference points. The result in the previous Section pointed to a different mechanism which generates loss aversion with low references points. Note that these results present empirical implications that are not present on the description of loss aversion effects in the literature, and which would be interesting to further explore.

Note also that this condition of $r > 1/2$ for the existence of loss aversion is the same as the one in the simple model of Section 2. However, note that there are some significant differences of the model considered here and the simple model of Section 2.

First, in Section 2 we assumed that $V(r) = r$, which does not hold here in general. In fact, we can obtain that $V(r) = r$ in this case when $p = 1$, when the signal on the outcome is fully informative.

Second, in Section 2 we assumed that the diminishing sensitivities had the same slope above and below the reference point (a slope of β). On the hand here we obtain that the slope of the diminishing sensitivities is steeper for outcomes above the reference point than for outcomes below the reference point if and only if $r > 1/2$, and the slope is only the same above and below the reference point for the particular case of $r = 1/2$.

Finally, note that if the signal s on the outcome is completely uninformative then there are no loss aversion effects. This further clarifies that the continuous signal on the outcome is important to get loss aversion effects if the reference point is not too low.

6. Both Signals are Partially Informative

We now consider the case in which both the signal on the outcome and the signal about whether the outcome is above or below the reference point is partially informative.

We consider the stochastic structure of the signal on the outcome as in the previous Section.

For the signal about whether the outcome is above or below the reference point we assume the following structure:

$$f(x,r) = \begin{cases} 1 - \dfrac{1}{2}e^{-\alpha(x-r)} & \text{if } x \geq r \\ \dfrac{1}{2}e^{-\alpha(r-x)} & \text{if } x < r, \end{cases} \tag{18}$$

where α is a parameter that captures the degree of informativeness of the signal about whether the outcome is above or below the reference point. If $\alpha \to \infty$ that signal is perfectly informative and we are in the case of the previous Section.[4]

Recall that $\widehat{e}_1$ is the probability of the decision-maker receiving signal S_1, and let $\widetilde{e}_1$ be the expected value of the outcome x given that the decision-maker received signal S_1

[4]This expression for $F(x,r)$ is different than the one in Section 4 to facilitate computations, and has similar properties.

multiplied by $\widehat{e}_1$. That is, for this case we have

$$\widehat{e}_1 = \int_0^r \frac{1}{2}e^{-\alpha(r-x)}\,dx + \int_r^1 [1-\frac{1}{2}e^{-\alpha(r-r)}]\,dx \tag{19}$$

$$\widetilde{e}_1 = \int_0^r \frac{1}{2}xe^{-\alpha(r-x)}\,dx + \int_r^1 x[1-\frac{1}{2}e^{-\alpha(r-r)}]\,dx. \tag{20}$$

Similarly, let $\widehat{e}_0$ be the probability of the decision-maker receiving signal S_0, and $\widetilde{e}_0$ be the expected value of the outcome x given that the decision-maker received signal S_0 multiplied by $\widehat{e}_0$,

$$\widehat{e}_0 = \int_0^r [1-\frac{1}{2}e^{-\alpha(r-x)}]\,dx + \int_r^1 \frac{1}{2}e^{-\alpha(r-r)}\,dx \tag{21}$$

$$\widetilde{e}_0 = \int_0^r x[1-\frac{1}{2}e^{-\alpha(r-x)}]\,dx + \int_r^1 \frac{1}{2}xe^{-\alpha(r-r)}\,dx. \tag{22}$$

We can obtain the values of $\widehat{e}_1, \widehat{e}_0, \widetilde{e}_1$, and $\widetilde{e}_0$ as a function of α and r.[5] When α converges to infinity, we have $\widehat{e}_1, \widehat{e}_0, \widetilde{e}_1$, and $\widetilde{e}_0$, converging to $1-r, r, \frac{1-r^2}{2}$, and $r^2/2$, respectively.

This then yields the expected value of the outcome as a function of the signals as

$$E(x|s,S_1) = v(s,S_1) = \begin{cases} \dfrac{ps[1-\frac{1}{2}e^{-\alpha(s-r)}]+(1-p)\widetilde{e}_1}{p[1-\frac{1}{2}e^{-\alpha(s-r)}]+(1-p)\widehat{e}_1} & \text{if } s \geq r \\ \dfrac{ps\frac{1}{2}e^{-\alpha(r-s)}+(1-p)\widetilde{e}_1}{p\frac{1}{2}e^{-\alpha(r-s)}+(1-p)\widehat{e}_1} & \text{if } s < r \end{cases} \tag{23}$$

$$E(x|s,S_0) = v(s,S_0) = \begin{cases} \dfrac{ps\frac{1}{2}e^{-\alpha(s-r)}+(1-p)\widetilde{e}_0}{p\frac{1}{2}e^{-\alpha(s-r)}+(1-p)\widehat{e}_0} & \text{if } s \geq r \\ \dfrac{ps[1-\frac{1}{2}e^{-\alpha(r-s)}]+(1-p)\widetilde{e}_0}{p[1-\frac{1}{2}e^{-\alpha(r-s)}]+(1-p)\widehat{e}_0} & \text{if } s < r \end{cases} \tag{24}$$

With this function $v(s,S_i)$ we can then construct the value function $V(s)$ as described in Section 3. Let A_1 be the expected value of $v(s,S_1)$ given x and given that $s \neq x$, and let

[5]We obtain $\widehat{e}_1 = 1-r+\frac{1}{2\alpha}[e^{-\alpha(1-r)}-e^{-\alpha r}], \widetilde{e}_1 = \frac{1-r^2}{2} - \frac{2-e^{-\alpha r}}{2\alpha^2} + \frac{1+\alpha}{2\alpha^2}e^{-\alpha(1-r)}, \widehat{e}_0 = 1-\widehat{e}_1$, and $\widetilde{e}_0 = 1/2-\widetilde{e}_1$.

A_0 be the expected value of $v(s, S_0)$ given x and given that $s \neq x$. That is,

$$A_1 = \int_0^r \frac{ps\frac{1}{2}e^{-\alpha(r-s)} + (1-p)\widetilde{e}_1}{p\frac{1}{2}e^{-\alpha(r-s)} + (1-p)\widehat{e}_1} ds + \int_r^1 \frac{ps[1-\frac{1}{2}e^{-\alpha(s-r)}] + (1-p)\widetilde{e}_1}{p[1-\frac{1}{2}e^{-\alpha(s-r)}] + (1-p)\widehat{e}_1} ds \quad (25)$$

$$A_0 = \int_0^r \frac{ps[1-\frac{1}{2}e^{-\alpha(r-s)}] + (1-p)\widetilde{e}_0}{p[1-\frac{1}{2}e^{-\alpha(r-s)}] + (1-p)\widehat{e}_0} ds + \int_r^1 \frac{ps\frac{1}{2}e^{-\alpha(s-r)} + (1-p)\widetilde{e}_0}{p\frac{1}{2}e^{-\alpha(s-r)} + (1-p)\widehat{e}_0} ds. \quad (26)$$

We can get that $\lim_{\alpha\to\infty} A_1 = (1+r)/2$ and $\lim_{\alpha\to\infty} A_0 = r/2$. The expressions of A_1 and A_0 can be solved and presented as function of dilogarithms as detailed in the Appendix.

We can then obtain the value function as

$$V(x) = \begin{cases} (1-\frac{1}{2}e^{-\alpha(x-r)})\Big(p\frac{px(1-\frac{1}{2}e^{-\alpha(x-r)}) + (1-p)\widetilde{e}_1}{p(1-\frac{1}{2}e^{-\alpha(x-r)}) + (1-p)\widehat{e}_1} + (1-p)A_1\Big) + \\ \frac{1}{2}e^{-\alpha(x-r)}\Big(p\frac{px\frac{1}{2}e^{-\alpha(x-r)} + (1-p)\widetilde{e}_0}{p\frac{1}{2}e^{-\alpha(x-r)} + (1-p)\widehat{e}_0} + (1-p)A_0\Big) & \text{if } x \geq r \\ \frac{1}{2}e^{-\alpha(r-x)}\Big(p\frac{px\frac{1}{2}e^{-\alpha(r-x)} + (1-p)\widetilde{e}_1}{p\frac{1}{2}e^{-\alpha(x-r)} + (1-p)\widehat{e}_1} + (1-p)A_1\Big) + \\ (1-\frac{1}{2}e^{-\alpha(r-x)})\Big(p\frac{px(1-\frac{1}{2}e^{-\alpha(r-x)}) + (1-p)\widetilde{e}_0}{p(1-\frac{1}{2}e^{-\alpha(r-x)}) + (1-p)\widehat{e}_0} + (1-p)A_0\Big) & \text{if } x < r. \end{cases} \quad (27)$$

When α converges to infinity, this value function converges to the value function in the previous Section, with now a positive slope close to the reference point. Given the results from the previous Section we also have now that if $r > 1/2$ the value function exhibits loss aversion if α is sufficiently large.

Figure 1 presents the value function for some parameter values, illustrating both loss aversion and diminishing sensitivities. Figure 2 illustrates how the value function varies with the degree of informativeness α of the signal about whether the outcome is above or below the reference point, and how the value function approaches the limiting case of the previous Section with a discontinuity at the reference point r. Figure 3 illustrates how the value function varies with the reference point r. Figure 4 presents examples of the value function for various degrees of informativeness of the signal of the outcome, p, illustrating

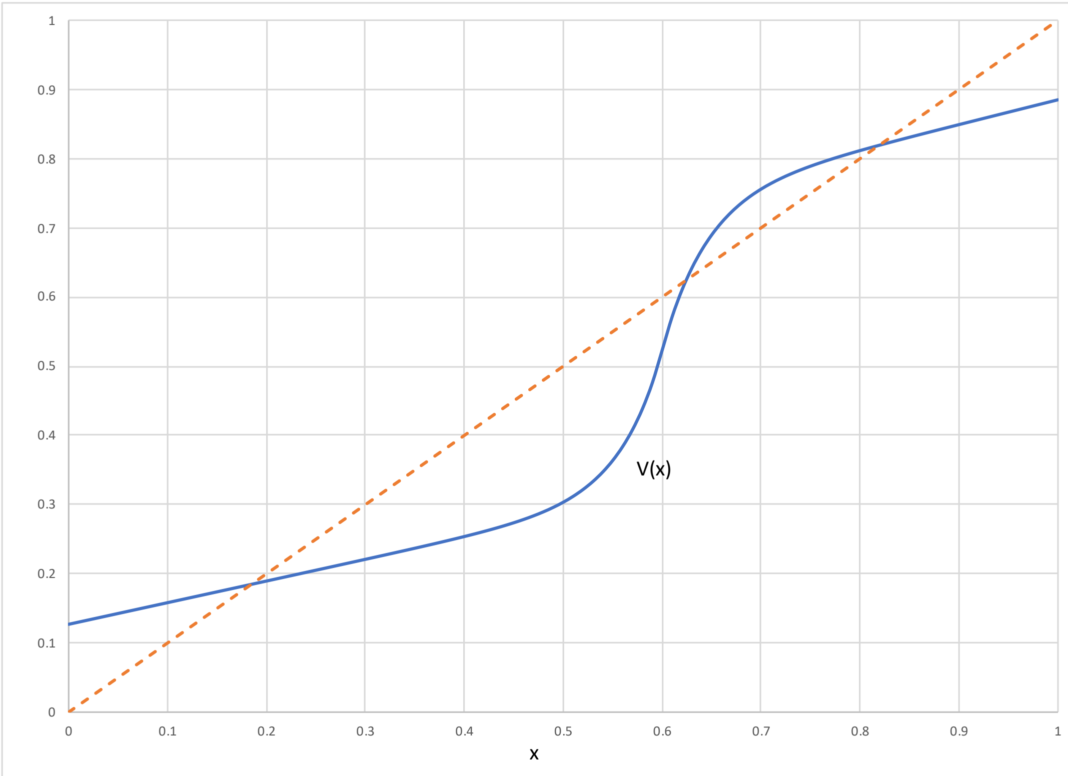

Figure 1: Example of value function for $r = .6, p = .5$, and $\alpha = 25$.

that for large p we get closer to expected utility theory.

7. Conclusion

This paper presents a model of loss aversion and diminishing sensitivities in which the decision-maker receives both a signal on the outcome and a signal about whether the outcome is below or above a reference point. The signal on the outcome being partially informative together with the signal about whether the outcome is above or below the reference point generates diminishing sensitivities above and below the reference point. When the signal on the outcome is informative, if the reference point is not too low then the model generates loss aversion, as there is a larger range for the outcome below the reference point. In a set-up with rational inattention under binary actions, the decision-maker pays more attention to

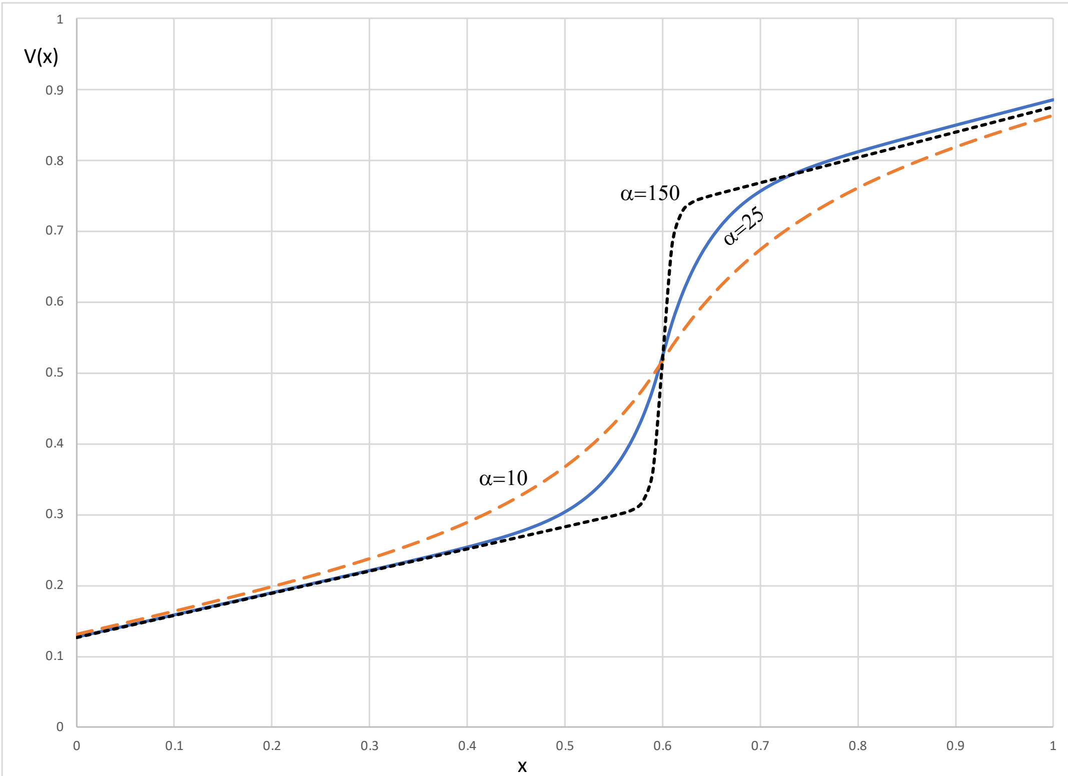

Figure 2: Example of value functions for $r = .6$, and $p = .5$, for various α ($\alpha = 10, 25, 150$).

the outcomes below the reference point when the reference point is sufficiently low, leading to loss aversion when the reference point is low. It would be interesting to investigate which of the two opposing effects is empirically more prevalent.

The existence of a signal about whether the outcome is above or below a reference point captures the idea that it may be sometimes easier to put the outcome in categories than to process information about a continuous variable, and the case of a reference point represents the simplest case of only two categories.

We could imagine that in some situations having more than two categories may be likely. For example, the decision-maker could have four categories, above or below a reference point, extremely high, and extremely low. It would be interesting to explore such situations in experimental settings, and in a variation of the model presented here.

Another potential interesting possibility is that the decision-maker can decide the in-

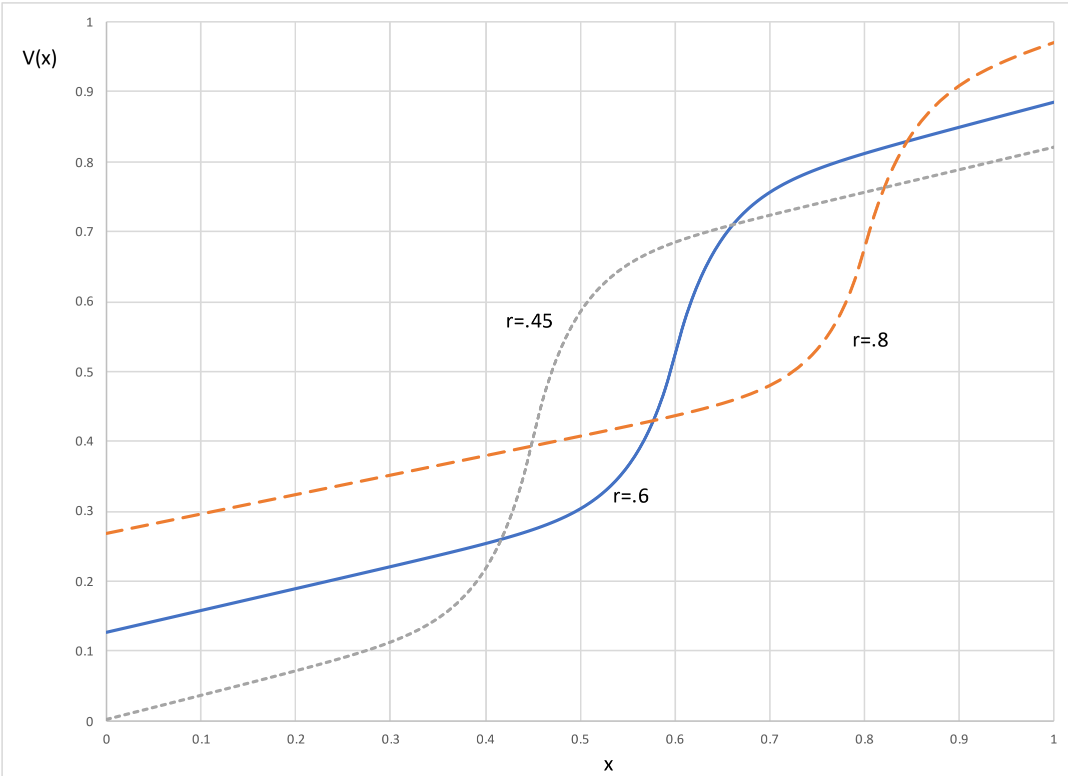

Figure 3: Example of value functions for $p = .5$, and $\alpha = 25$, for various r ($r = .45, .6, .8$).

formativeness, at a cost, of the signals received. It would then be interesting to study the informativeness decisions as a function of the stakes in the problem considered.

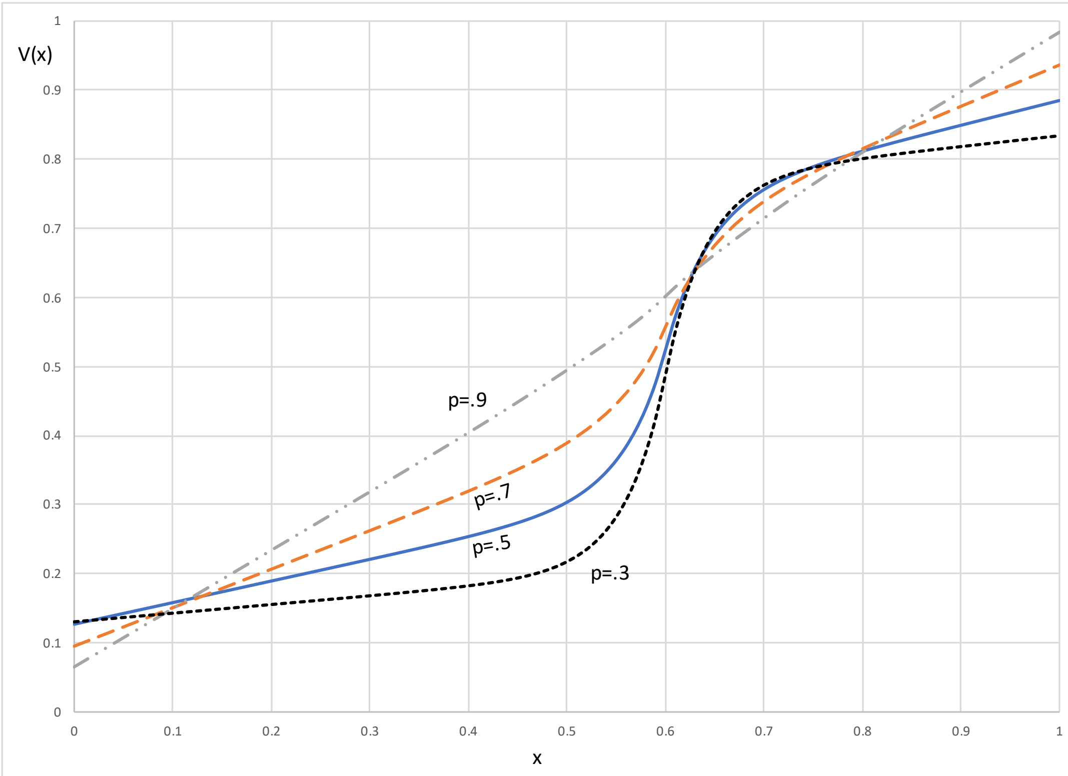

Figure 4: Example of value functions for $r = .6$, and $\alpha = 25$, for various p ($p = .3, .5, .7, .9$).

APPENDIX

Perfectly Informative Signal on Relation to Reference Point with Normal Distribution on x and s :

Consider the case of a perfectly informative signal on the relation to reference point, and that the prior distribution on x is normal, and that the distribution on the signal s given x is also normal.

Let $E(x) = \mu$, and the variance of the prior of x be σ^2. Furthermore, let $E(s|x) = x$, and the variance of s given x be ω^2. Let $\widetilde{\mu}(s)$ and $\widetilde{\sigma}^2$ be the expected value and variance of x given s. From standard calculations on the distribution of the conditional normnal, we know that

$$
\begin{aligned}
E(x|s) &= \widetilde{\mu}(s) = \frac{s\sigma^2 + \mu\omega^2}{\sigma^2 + \omega^2} && \text{(i)} \\
E[(x - \widetilde{\mu}(s))^2|s] &= \widetilde{\sigma}^2 = \frac{\sigma^2\omega^2}{\sigma^2 + \omega^2}. && \text{(ii)}
\end{aligned}
$$

Given the calculations for expected value of truncated normal distributions we can then obtain:

$$
\begin{aligned}
E(x|s, x > r) &= \widetilde{\mu}(s) + \widetilde{\sigma}\phi\left(\frac{r - \widetilde{\mu}(s)}{\widetilde{\sigma}}\right) \Big/ \left[1 - \Phi\left(\frac{r - \widetilde{\mu}(s)}{\widetilde{\sigma}}\right)\right] && \text{(iii)} \\
E(x|s, x < r) &= \widetilde{\mu}(s) - \widetilde{\sigma}\phi\left(\frac{r - \widetilde{\mu}(s)}{\widetilde{\sigma}}\right) \Big/ \Phi\left(\frac{r - \widetilde{\mu}(s)}{\widetilde{\sigma}}\right), && \text{(iv)}
\end{aligned}
$$

where $\phi()$ and $\Phi()$ are the density and cumulative distribution function, respectively, of the standard normal. From this we can then obtain the value function as

$$
V(x) = \begin{cases} E_s[E(x|s, x > r)|x] = \widetilde{\mu}(x) + \widetilde{\sigma}E_s\left[\phi\left(\frac{r-\widetilde{\mu}(s)}{\widetilde{\sigma}}\right) \Big/ \left[1 - \Phi\left(\frac{r-\widetilde{\mu}(s)}{\widetilde{\sigma}}\right)\right] |x\right] & \text{if } x \geq r \\ E_s[E(x|s, x < r)|x] = \widetilde{\mu}(x) - \widetilde{\sigma}E_s\left[\phi\left(\frac{r-\widetilde{\mu}(s)}{\widetilde{\sigma}}\right) \Big/ \Phi\left(\frac{r-\widetilde{\mu}(s)}{\widetilde{\sigma}}\right) |x\right] & \text{if } x < r \end{cases} \quad \text{(v)}
$$

Derivation of Loss Aversion Condition in the Case of Fully Informative Signal on whether Outcome is above Reference Point:

The condition for the existence of loss aversion is that $V(r) - V(r^-) > V(r^+) - V(r)$, which can be reduced to $V(r) > \frac{1}{2}[V(r^+) + V(r^-)]$. Note now that we can write $V(r)$ as

$$
\begin{aligned}
V(r) = {} & \frac{1}{2}[V(r^+) + V(r^-)] + \frac{p}{2}\Big(\frac{pr + (1-p)(1-r^2)}{p + 2(1-p)(1-r)} + \frac{pr + (1-p)r^2}{p + 2(1-p)r} \\
& - \frac{2pr + (1-p)(1-r^2)}{2p + 2(1-p)(1-r)} - \frac{2pr + (1-p)r^2}{2p + 2(1-p)r}\Big) \qquad \text{(vi)}
\end{aligned}
$$

from which we can obtain that $V(r) > \frac{1}{2}[V(r^+) + V(r^-)]$ if and only if

$$
\frac{1 - r^2}{[p + 2(1-p)(1-r)][p + (1-p)(1-r)]} > \frac{r^2}{[p + 2(1-p)r][p + (1-p)r]}. \quad \text{(vii)}
$$

This condition can be reduced to

$$
\left(\frac{1}{1-r} - \frac{1}{r}\right)\left[p^2\left(\frac{1}{1-r} + \frac{1}{r}\right) + 3(1-p)\right] < 0 \quad \text{(viii)}
$$

which holds if and only if $r > 1/2$.

Expressions for A_1 and A_0:

Consider the first term of A_1. Note that it can be seen as being composed of B_1 and B_2 where

$$B_1 = \int_0^r \frac{s}{1+2\frac{1-p}{2}\widehat{e}_1 e^{\alpha(r-s)}}\,ds \tag{ix}$$

$$B_2 = \int_0^r \frac{(1-p)\widetilde{e}_1}{p\frac{1}{2}e^{-\alpha(r-s)}+(1-p)\widehat{e}_1}\,ds. \tag{x}$$

Integrating B_1 by parts one obtains

$$B_1 = \frac{r^2}{2} + \frac{r}{\alpha}\ln\left(1+2\frac{1-p}{p}\widehat{e}_1\right) + \frac{1}{\alpha^2}[Li_2(-2\frac{1-p}{p}\widehat{e}_1 e^{\alpha r}) - Li_2(-2\frac{1-p}{p}\widehat{e}_1)] \tag{xi}$$

where $Li_2(z)$ is the dilogarithm function (also known as Spence's function), $Li_2(z) = -\int_0^z \frac{\ln(1-u)}{u}\,du$. Note also that if $|z| < 1$, we also have $Li_2(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^2}$, and if $z \leq -1$, we have $Li_2(z) = -\frac{\pi^2}{6} - \frac{1}{2}\ln(-z)^2 - \sum_{k=1}^{\infty} \frac{1}{k^2 z^k}$, which allows us to compute approximations of $Li_2(z)$.

We can also obtain

$$B_2 = \frac{\widetilde{e}_1}{\widehat{e}_1}\left[r - \frac{1}{\alpha}\ln\frac{p+2(1-p)\widehat{e}_1}{pe^{-\alpha r}+2(1-p)\widehat{e}_1}\right]. \tag{xii}$$

Using similar techniques on the second term of A_1 and on A_0 we can then obtain

$$
\begin{aligned}
A_1 \;=\;& \frac{1}{2}+\frac{r}{\alpha}\ln\left(1+2\frac{1-p}{p}\widehat{e}_1\right)+\frac{1}{\alpha^2}[Li_2(-2\frac{1-p}{p}\widehat{e}_1e^{\alpha r})-Li_2(-2\frac{1-p}{p}\widehat{e}_1)]+\\
& \frac{\widetilde{e}_1}{\widehat{e}_1}\left[r-\frac{1}{\alpha}\ln\frac{p+2(1-p)\widehat{e}_1}{pe^{-\alpha r}+2(1-p)\widehat{e}_1}\right]+\frac{p}{p+(1-p)\widehat{e}_1}\Big(\frac{1-r^2}{2}+\frac{1}{\alpha}\ln\left(1-\frac{1}{2}\frac{pe^{-\alpha(1-r)}}{p+(1-p)\widehat{e}_1}\right)-\\
& \frac{r}{\alpha}\ln\left(1-\frac{1}{2}\frac{p}{p+(1-p)\widehat{e}_1}\right)+\frac{1}{\alpha^2}\left[Li_2\left(\frac{1}{2}\frac{p}{p+(1-p)\widehat{e}_1}\right)-Li_2\left(\frac{1}{2}\frac{pe^{-\alpha(1-r)}}{p+(1-p)\widehat{e}_1}\right)\right]\Big)+\\
& +\frac{1}{\alpha}\ln\left(2\frac{p+(1-p)\widehat{e}_1}{p}e^{\alpha(1-r}-1\right)-\frac{r}{\alpha}\ln\left(2\frac{p+(1-p)\widehat{e}_1}{p}-1\right)+\\
& \frac{1}{\alpha^2}\Big[\ln\left(2\frac{p+(1-p)\widehat{e}_1}{p}e^{\alpha(1-r)}-1\right)\ln\left(2\frac{p+(1-p)\widehat{e}_1}{p}e^{\alpha(1-r)}\right)-\ln\Big(2\frac{p+(1-p)\widehat{e}_1}{p}\\
& -1\Big)\ln\left(2\frac{p+(1-p)\widehat{e}_1}{p}\right)\Big]+\frac{1}{\alpha^2}\Big[Li_2\left(1-2\frac{p+(1-p)\widehat{e}_1}{p}e^{\alpha(1-r)}\right)-Li_2\left(-1-2\frac{1-p}{p}\widehat{e}_1\right)\Big]\\
& +\frac{(1-p)\widetilde{e}_1}{p+(1-p)\widehat{e}_1}\left[1-r+\frac{1}{\alpha}\ln\frac{2[p+(1-p)\widehat{e}_1]-pe^{-\alpha(1-r)}}{p+2(1-p)\widehat{e}_1}\right] \qquad \text{(xiii)}\\
A_0 \;=\;& \frac{1}{2}+\frac{r}{\alpha}\ln\left(1+2\frac{1-p}{p}\widehat{e}_0\right)+\frac{1}{\alpha^2}[Li_2(-2\frac{1-p}{p}\widehat{e}_0e^{\alpha(1-r)})-Li_2(-2\frac{1-p}{p}\widehat{e}_0)]+\\
& \frac{\widetilde{e}_0}{\widehat{e}_0}\left[1-r-\frac{1}{\alpha}\ln\frac{p+2(1-p)\widehat{e}_0}{pe^{-\alpha(1-r)}+2(1-p)\widehat{e}_0}\right]+\frac{p}{p+(1-p)\widehat{e}_0}\Big(\frac{r^2}{2}-\\
& \frac{r}{\alpha}\ln\left(1-\frac{1}{2}\frac{p}{p+(1-p)\widehat{e}_0}\right)+\frac{1}{\alpha^2}\left[Li_2\left(\frac{1}{2}\frac{pe^{-\alpha r}}{p+(1-p)\widehat{e}_0}\right)-Li_2\left(\frac{1}{2}\frac{p}{p+(1-p)\widehat{e}_0}\right)\right]\Big)+\\
& -\frac{1}{\alpha}\ln\left(\frac{p+2(1-p)\widehat{e}_0e^{\alpha(1-r)}}{p}\right)+\frac{r}{\alpha}\ln\left(\frac{p+2(1-p)\widehat{e}_0}{p}\right)\\
& +\frac{1}{\alpha^2}\Big[\ln\left(2\frac{p+(1-p)\widehat{e}_0}{p}\right)\ln\left(\frac{p+2(1-p)\widehat{e}_0}{p}\right)\\
& -\ln\left(2\frac{p+(1-p)\widehat{e}_0}{p}e^{\alpha r}\right)\ln\left(2\frac{p+(1-p)\widehat{e}_0}{p}e^{\alpha r}-1\right)\Big]+\frac{1}{\alpha^2}\Big[Li_2\left(1-2\frac{p+(1-p)\widehat{e}_0}{p}e^{\alpha r}\right)\\
& -Li_2\left(-1-2\frac{1-p}{p}\widehat{e}_0\right)\Big]+\frac{(1-p)\widetilde{e}_0}{p+(1-p)\widehat{e}_0}\left[r+\frac{1}{\alpha}\ln\frac{2[p+(1-p)\widehat{e}_0]-pe^{-\alpha r}}{p+2(1-p)\widehat{e}_0}\right] \qquad \text{(xiv)}
\end{aligned}
$$

REFERENCES

Baumeister, R.F., E. Bratslavsky, C. Finkenauer, and K.D. Vohs (2001), "Bad Is Stronger Than Good," *Review of General Psychology,* **5(4)**, 323-370.

Camerer, C. (2000), "Prospect Theory in the Wild: Evidence from the Field," in *Kahneman, D., and A. Tversky (eds.), Choices, Values, and Frames,* Cambridge University Press and Russell Sage Foundation.

Caplin, A., M. Dean, and J. Leahy (2019), "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice," *Review of Economic Studies,* **86(3)**, 1061-1094.

Caplin, A., M. Dean, and J. Leahy (2021), "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," *working paper,* .

Gal, D., and D.D. Rucker (2018), "The Loss of Loss Aversion: Will it Loom Larger than its Gain?," *Journal of Consumer Psychology,* **28(3)**, 497-516.

Friedman, D. (1989), "The S-Shaped Function Value Function as a Constrained Optimum," *American Economic Review,* **79(5)**, 1243-1248.

Fudenberg, D. (2006), "Advancing Beyond Advances in Behavioral Economics," *Journal of Economic Literature,* **XLIV**, 694-711.

Gentzkow, M., and E. Kamenica (2014), "Costly Persuasion," *American Economic Review: Papers & Proceedings,* **104(5)**, 457-462.

Gossner, O., and J. Steiner (2018), "On the Cost of Misperception: General Results and Behavioral Application," *Journal of Economic Theory,* **177**, 816-847.

Higgins, E.T., and N. Liberman (2018), "The Loss of Loss Aversion: Paying Attention to Reference Point," *Journal of Consumer Research,* **28(3)**, 523-532.

Kahneman, D., and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica,* **47(2)**, 263–291.

Khaw, M.W., Z. Li, and M. Woodford (2020), "Cognitive Imprecision and Small-Stakes Risk Aversion," *Review of Economic Studies,* forthcoming.

Koszegi, B., and M. Rabin (2006), "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics,* **121(4)**, 1133–1165.

Martin, J.M., M. Reimann, and M.I. Norton (2016), "Experience Theory, or How Desserts Are Like Losses," *Journal of Experimental Psychology: General,* **145(11)**, 1460-1472.

Matějka, F., and A. McKay (2015), "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review,* **105(1)**, 272-298.

O'Donoghue, T., and C, Sprenger (2018), "Reference-Dependent Preferences," in *Bernheim, B.D., S. Dellavigna, and D. Laibson (eds.), Handbook of Behavioral Economics: Applications and Foundations 1,* pp.1-77, Elsevier.

Sims, C.A. (1998), "Stickiness," *Carnegie-Rochester Conference Series on Public Policy,* **49(1)**, 317-356.

Steiner, J., and C. Stewart (2016), "Perceiving Prospects Properly," *American Economic Review,* **106(7)**, 1601-1631.

Stewart, N., N. Chater, and G.D.A. Brown (2006), "Decision by Sampling," *Cognitive Psychology,* **53**, 1-26.

Thaler, R.H. (1980), "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization,* **1**, 39–60.

Tversky, A., and D. Kahneman (1991), "Loss Aversion in Riskless Choice: A Reference Dependent Model," *Quarterly Journal of Economics,* **106**, 1039–1061.

Wardley, M., and M. Alberhasky (2021), "Framing Zero: Why Losing Nothing Is Better Than Gaining Nothing," *Journal of Behavioral and Experimental Economics,* **90**, 1-8.

Woodford, M. (2008), "Inattention as a Source of Randomized Discrete Adjustment," *working paper,* Columbia University.

Woodford, M. (2012), "Inattentive Valuation and Reference-Dependent Choice," *working paper,* Columbia University.

Yang, M. (2011), "Coordination with Rational Inattention," *working paper,* Princeton University.