“If you can, you must.” Information, utility, and loss aversion*

Greg Kubitz[†] Lionel Page[‡]

June 2024

Abstract

We consider utility as a reward system shaped by evolution to motivate us to make the best decisions possible. An optimal system consists of a reference dependent utility function that adapts to expectations. We show that anticipatory utility and loss aversion emerge as features of this optimal system that encourage the agent to set the reference point to the level of his expectations.

1 Introduction

The fact that subjective satisfaction is reference-dependent is one of the most influential ideas from Prospect Theory (Kahneman and Tversky, 1979). Supported by a large body of empirical evidence, satisfaction is not purely driven by the overall level of success in a given dimension but by how this level compares to a subjective reference point. How this reference point is determined was left open by Kahneman and Tversky who mentioned both the *status quo* and alternatives, like *aspirations* or *expectations*. Somewhat surprisingly, more than four decades after Kahneman and Tversky's seminal contribution, what reference points are and how they are determined is still an open question.

A substantial body of empirical research has looked into the determinants of reference points. They have been found to be influenced by different elements such as past achievements, peers' outcomes, and personal goals (Camerer et al., 1997; Barberis, 2013; O'Donoghue and Sprenger, 2018). This evidence indicates that reference points are not strictly determined by the status quo. They can be influenced by factors under the individual's control like goals or peer groups. This raises a puzzling question from a hedonic perspective: if satisfaction is relative to a reference point, why don't people simply aim to have a low reference point in order to generate high satisfaction?

*We thank Zachary Breig, Moshe Hoffman, Nick Netzer, Jonathan Newton, Luis Rayo, Jakub Steiner, Arthur Robson, and seminar audiences at the University of Melbourne, Deakin, Monash, UTS, the University of Ottawa, Laval, the University of Zürich and various conferences for helpful comments and suggestions.

[†]School of Economics and Finance, Queensland University of Technology. Email: kubitz@qut.edu.au

[‡]School of Economics, University of Queensland. Email: lionel.page@uq.edu.au

That people have some agency in setting their reference point is commonly accepted. Indeed, the idea that having low aspirations and expectations is a pathway to happiness is at least as old as the philosophies of Lao Tzu (6th century BC) and Epictetus (2nd century AD). The notion of "management of expectations" nowadays reflects the idea that people may want to avoid setting expectations too high, as they risk leading to crushing disappointment (Sweeny and Shepperd, 2010; Sweeny, 2018). The possible hedonic gains from keeping ones' aspirations and expectations low seem however at odds with the fact that, in many situations, people seem to set high aspirations that are challenging to achieve (Allen et al., 2017; Anderson and Green, 2018).

In the present paper, we investigate how people set their reference points when they have some agency to do so. We resolve the hedonic puzzle of the reference point's determination by taking an evolutionary perspective of the role of subjective satisfaction. We show that a hedonic utility system that encourages a choice of reference point that is well-calibrated to the current context helps people make fitness maximizing decisions.

To characterize such a utility system we adopt the *principal-agent metaphor* as a methodological approach (Binmore, 1994; Robson, 2001*a*,*b*; Rayo and Becker, 2007). Nature (she) is thought of as a principal aiming to design a system of hedonic utility for an agent (he) to make decisions. Given information about his expected potential–the highest level of success he can achieve–the agent can set his reference point, conceived as an aspiration level. A utility system that leads the agent to have miscalibrated aspirations would negatively impact his ability to reach his potential. For instance, an agent opting for aspirations much lower than his potential would be content with outcomes far below his potential. As a consequence, he would fail to identify choices that allow him to reach his potential. Such an agent would underachieve relative to agents with aspirations near their potential. Therefore, for Nature to incentivize the agent to make fitness-maximizing decisions, she must also incentivize him to choose a reference point that is well-calibrated to his information.

In our framework, the hedonic utility system is a mechanism designed by Nature with the aim of maximizing the agent's fitness. The agent's information about his potential and choice of reference point correspond to his type and reported type respectively. An incentive compatible utility system induces the agent to choose a reference point that reflects his private information about his potential. Given the choice of reference, the agent is assigned a hedonic utility function over fitness outcomes and additionally may experience immediate subjective satisfaction that depends on the reference choice. The agent then faces a decision problem where each choice is associated with a level of fitness and fitness is mapped to utility using the assigned utility function. The agent is forward-looking when choosing his reference point, maximizing his expected utility given the information about his potential fitness levels.

This framework highlights the role of two well known behavioral patterns: the experience of *anticipatory utility*, whereby the agent experiences positive satisfaction associated with having high aspirations, and the existence of *loss aversion* relative to the reference point. *Anticipatory utility* can eliminate the agent's incentive to choose a low reference point by directly rewarding the agent for having higher aspirations. But the existence of anticipatory utility may create an incentive to have overly optimistic aspirations. We show that this tension requires utility functions to be *asymmetric*. The agent should be less rewarded for overachieving relative to his anticipated potential than penalized for failing to reach this potential. This asymmetry in the changes of utility below and above this anticipated potential is what characterizes *loss aversion.* Loss aversion improves the agent's incentives to set aspirations correctly by both decreasing the benefits of over-pessimism and increasing the cost of over-optimism.

Our findings contribute to several important strands of literature. First, our paper extends the literature on the evolutionary foundations of preferences (Robson and Samuelson, 2011). This approach uses "reverse engineering" to investigate whether plausible models of the environment exist where our present preferences would constitute good solutions for making effective decisions, starting from the features of those preferences (Samuelson, 2004). This perspective contributes to the growing interest in the adaptiveness of human cognition (Bénabou and Tirole, 2016). In that perspective, we build on the literature explaining the existence of a S-shape hedonic utility function. Previous contributions have considered two reasonable biological constraints Nature faces when designing a hedonic utility function: the utility function must be bounded as hedonic utility cannot be infinite, and the agent cannot discriminate perfectly between two choices that lead to similar utility levels as hedonic utility is noisy. Under these constraints, the curvature of the fitness maximizing hedonic utility function should follow the cumulative distribution of the agent's possible levels of success Robson (2001*a*); Rayo and Becker (2007); Netzer (2009). Such a pattern minimizes the cost of errors by allowing the agent to better discriminate between those levels of success most likely to be within reach. These results align with the idea of *efficient coding* that emerged independently in neuroscience, positing how a neuron should optimally react to the distribution of a stimulus (Laughlin, 1981).

We add to this literature by examining how reference-dependent utility adapts when the agent faces a novel context. We model this situation as the agent having some private information about the distribution of his potential relative to Nature. This private information reflects the fact that Nature is not able to ex ante map all the causal structure of the world in a reward system that automatically identifies the agent's possible potentials in any given context (Samuelson and Swinkels, 2006).[1] Previously, Rayo and Becker (2007) considered this situation and showed that Nature should smooth the hedonic utility function to reflect her uncertainty about the agent's potential. We show that Nature can instead create a mechanism that encourages the agent to use this information to calibrate their utility function to be more discerning between levels of success near their potential.

Second, our research significantly contributes to the field of behavioral economics by demonstrating that three extensively examined behavioral patterns—anticipatory utility, reference point determination, and loss aversion—are adaptive and interrelated in addressing a common problem: helping us make good decisions, given the inherent imperfection in our subjective perception of the value of the options we face. In addition, our results cast a new light on two other behavioural questions that are closely connected the the formation of reference points: how we set our goals and how we choose our peer groups.

Anticipatory utility. We provide a justification for the existence of *anticipatory utility*, the utility that comes from expectations of future achievements (Loewenstein and Elster,

[1]Since evolution is a learning process, the vast array of possible contexts in the world can be viewed as creating a "curse of dimensionality." It is unfeasible for evolution to converge on an accurate estimate of the distribution of outcomes in every possible context.

1992; Caplin and Leahy, 2001; Kőszegi, 2010). In our model, anticipatory utility is necessary to motivate the agent to set his reference point according to his expected potential. It rewards the agent for being ambitious and setting his aspirations high. The existence of anticipatory utility however may incentivise the agent to form biased beliefs such as overoptimism (Brunnermeier and Parker, 2005), engage in wishful thinking (Caplin and Leahy, 2019) or distort his past memories (Chew, Huang and Zhao, 2020). To avoid such a discrepancy between ex-ante beliefs and ex-post outcomes, Kőszegi (2010) proposes that a rational agent's beliefs and outcomes should respect a *personal equilibrium*: the ex-ante beliefs that generate anticipatory utility should shape behavior in a way that creates outcomes consistent with those beliefs. Our approach provides a justification for the existence of a personal equilibrium between the agent's reference point (set at his expectations) and his likely achievements.

Reference point determination. This in-built consistency also helps improve our understanding of how reference points can be determined given the information held by the agent. The indeterminate level of the reference point in Kahneman and Tversky's Prospect Theory has led to criticisms that the theory has too many degrees of freedom to be falsifiable (Farber, 2005; Pesendorfer, 2006) In two of the most influential papers on reference-dependent preferences, Kőszegi and Rabin (2006, 2007) resolve this problem by proposing to pinpoint the reference point as the agent's expectation. They also added a constraint on the consistency of this expectation by requiring that the agent's behavior—given this expectation and the corresponding distribution of outcomes—should align with that expectation. This concept is again grounded in the notion of personal equilibrium. In our setting, this consistency is a property of an incentive compatible utility system which calibrates the agent's reference point to his information about his potential.

Loss aversion. Another key contribution of the paper is providing an explanation for *loss aversion*, one of the most important deviations from the classical economic model of rational behavior introduced by Kahneman and Tversky (Barberis, 2013). Its existence has been "demonstrated by hundreds of experiments" (Camerer and Loewenstein, 2003) and it has been used to explain a wide range of economic phenomena from the equity-premium puzzle on financial markets (Benartzi and Thaler, 1995) to the behavior of New-York cab drivers (Camerer et al., 1997). We propose here that loss aversion is not so much a bias leading us astray as a feature of our subjective utility system designed to help us to make good decisions. Relative to a symmetric S-shape, a utility function aimed at motivating the agent to set an appropriate reference point is marked by a heightened utility for outcomes close to the reference point. Consequently, relative to an S-shaped function, the agent is rewarded more for outcomes near the reference point, as these outcomes indicate that the chosen reference point was more likely to be correct. This leads to smaller utility gains for outcomes exceeding the reference point and larger utility losses for those falling short of it. This pattern, combined with anticipatory utility during the aspiration stage, helps ensure the agent is motivated to set his reference point at the level of his expected potential. Notably, the pattern of loss aversion that emerges from this approach does not feature a kink and is therefore compatible with recent empirical results that only find loss aversion for large stakes but not for small ones.

Goal setting. Our result also contributes to the literature on *goal setting*. In our framework, the agent's choice of a reference point can be seen as making the best of the available information to set a goal about what he can achieve. Our approach gives a foundation to the literature in psychology and behavioral economics modeling goals as reference points (Heath, Larrick and Wu, 1999) and an answer to the question of how goals are set. In our framework, agents are incentivized to set the highest goals they believe they can achieve. This explains a peculiar aspect of goals, the fact that people often choose hard-to-reach goals that motivate them to exert a lot of effort to be successful (Markle et al., 2018). In the process, our approach reshapes our understanding of reference points as expectations *about the best possible achievements we can reach.*

Choice of peer group. Finally, our results contribute to our understanding of how people choose peer groups. Peer groups are often considered to act as reference points (Clark and Oswald, 1996; Clark, Frijters and Shields, 2008; Card et al., 2012). Our approach explains why people do not simply choose the most modest peer groups possible in order to feel good by comparison (Harris, Anseel and Lievens, 2008). Instead, people align themselves with peer groups whose successes align with their beliefs about their potential.

2 Intuition of Main Results

In this section, we present the intuition of our findings regarding anticipatory utility and loss aversion within a fitness maximizing utility system. These elements naturally arise as solutions to a fundamental challenge Nature faces: ensuring the system's incentive compatibility by encouraging the agent to align his reference point with his expectations about his potential, a key requirement for the agent to make good decisions.

Consider an agent that can find himself in two different states of the world. The first one is inauspicious, his expectation about his fitness potential is low. The second state is propitious, his expectations are high. Denote these two states as ω_1 and ω_2 respectively. The agent is able to choose his reference point in the form of an aspiration level $\tilde{\omega}$ that can be equal to either ω_1 or ω_2. This choice determines his utility function over fitness outcomes $y \in \mathbb{R}$, denoted $v_{\tilde{\omega}}(y)$. We assume the utility functions are such that the agent's expected utility is always higher when setting his reference point low: $\mathbb{E}[v_{\omega_1}(y)|\omega_i] > \mathbb{E}[v_{\omega_2}(y)|\omega_i]$ for $i = 1, 2$. Intuitively, choosing a low reference point ensures high hedonic utility because every outcome will look relatively better. Therefore, without any other motivations, the agent always prefers to set $\tilde{\omega} = \omega_1$.

To counterbalance the agent's incentive to have low aspirations, we assume the agent experiences *anticipatory utility* that increases with the level of agent's aspirations, $u(\tilde{\omega})$. Given the state, the agent is then assumed to choose the aspiration level that maximizes $u(\tilde{\omega}) + \mathbb{E}[v_{\tilde{\omega}}(y)|\omega_i]$. To prevent the agent from choosing a low reference point when the state is propitious, the gain in anticipatory utility from setting a high reference point must be at least as large as the hedonic benefit of having a low reference point:

$$\underbrace{u(\omega_2) - u(\omega_1)}_{\text{Anticipatory utility gain}} \geq \underbrace{\mathbb{E}[v_{\omega_1}(y)|\omega_2] - \mathbb{E}[v_{\omega_2}(y)|\omega_2]}_{\text{Hedonic benefit of over-pessimism}} > 0. \tag{1}$$

Remark 1 (Anticipatory utility)**.** *Experiencing anticipatory utility, with higher aspirations being associated with higher anticipatory utility, helps the agent not to set his aspirations too low.*

However, the existence of anticipatory utility generates a new incentive problem: the agent may prefer overly optimistic aspirations. To ensure he correctly sets low aspirations when the state is inauspicious, the difference in anticipatory utility cannot be too large. It must not exceed the expected loss in hedonic rewards from having a high reference point when outcomes are expected to be low:

$$\underbrace{\mathbb{E}[v_{\omega_1}(y)|\omega_1] - \mathbb{E}[v_{\omega_2}(y)|\omega_1]}_{\text{Hedonic cost of over-optimism}} \geq \underbrace{u(\omega_2) - u(\omega_1)}_{\text{Anticipatory utility gain}} > 0. \tag{2}$$

Inequalities (1) and (2) put upper and lower bounds on the difference in anticipatory utility associated with a high versus a low reference point. This reflects that the agent must be rewarded for high aspirations, but not to the extent that he sets excessively high aspirations. Combining these inequalities imposes a constraint on the utility functions themselves:

$$\underbrace{\mathbb{E}[v_{\omega_1}(y)|\omega_1] - \mathbb{E}[v_{\omega_2}(y)|\omega_1]}_{\text{Hedonic cost of optimism}} \geq \underbrace{\mathbb{E}[v_{\omega_1}(y)|\omega_2] - \mathbb{E}[v_{\omega_2}(y)|\omega_2]}_{\text{Hedonic benefit of pessimism}}. \tag{3}$$

This inequality means that the losses from being overly optimistic must be at least as large as the gains from being overly pessimistic. Rearranging Equation (3) shows that these constraints introduce an asymmetry in the variation of hedonic utility associated with deviations from the reference point.

$$\underbrace{\mathbb{E}[v_{\omega_2}(y)|\omega_2] - \mathbb{E}[v_{\omega_2}(y)|\omega_1]}_{\text{Losses from failing to reach aspirations}} \geq \underbrace{\mathbb{E}[v_{\omega_1}(y)|\omega_2] - \mathbb{E}[v_{\omega_1}(y)|\omega_1]}_{\text{Gains from exceeding aspirations}} \tag{4}$$

Remark 2 (Asymmetry of reward function constraints)**.** *To select the correct level of aspirations, the losses incurred by the agent for falling short of a high reference point must be at least as large as the gains from surpassing a low reference point.*

We show in Section 4.2 that the constraints induced by the incentive compatibility problem are incompatible with an oddly symmetric S-shaped utility function when there are more than two states associated with different levels of expected potential. Differences in anticipatory utility large enough to prevent overly pessimistic aspirations with a symmetric S-shaped utility function will inevitably generate incentives for overly optimistic ones.

For the utility system to be incentive-compatible, the utility functions associated with each state must be higher around the expected fitness potential in that state. This distortion of the utility function reduces the benefits of low reference points and increases the costs of high reference points, as illustrated in Figure 1. A result of this distortion relative to the S-shape utility is loss aversion. This result is formally given in Proposition 9 in Section 4.2.

Remark 3 (Loss aversion)**.** *Loss aversion around the reference point of the utility function improves the agent's incentives to set aspirations correctly. It does so by both decreasing the benefits of over-pessimism and increasing the cost of over-optimism.*

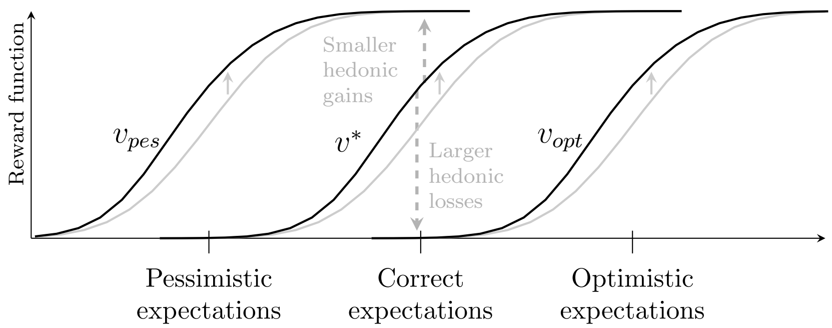


Figure 1: The optimal reward function is elevated around the level of expectations. This elevation reduces the hedonic gains from pessimism and increases the hedonic losses from optimism, incentivizing the agent to choose the correct level of expectations.

3 Model

3.1 The situation faced by the agent

An agent faces a decision problem where he chooses an $x \in \mathbb{R}$. Each choice is mapped to an output level $y \in \mathbb{R}$, that corresponds to the agent's level of success. The mapping, $\phi : \mathbb{R} \times S \to \mathbb{R}$, depends on the state, s, that is an element of a measureable space, S. The state can be thought of as the specific environment of the agent and is observed by the agent when he makes his decision. The state determines not only what the agent can achieve, but also how choices map to outputs. We assume that for each s there is a unique $x \in \mathbb{R}$ that maximizes output, denoted $x^*(s)$, and the output for any $x \in \mathbb{R}$ is given by $\phi(x, s) = \phi(x^*(s), s) - \alpha|x - x^*(s)|$ where $\alpha > 0$ is a constant. Then the output function maps to the half line, $(-\infty, \bar{y}(s)]$, where $\overline{y}(s) = \phi(x^*(s), s)$ is denoted as the fitness *potential* of the agent given state s.

Prior to the decision stage, the agent observes some information in the form of a signal about his potential $\bar{y}$, which we denote by $\omega \in \Omega$. We assume that there are a finite number of signals, $|\Omega| = N < \infty$, which are ordered $\omega_1, \ldots, \omega_N$ so that $\overline{y}^r_{\omega_1} < \cdots < \overline{y}^r_{\omega_N}$, where $\overline{y}^r_{\omega_n} = \mathbb{E}[\overline{y}|\omega_n]$ is the expected potential of the agent given signal ω_n. Signal ω is observed by the agent with probability p_n. We denote the conditional distribution of potential given information ω_n as $F_{\omega_n}(\cdot)$ with corresponding density f_{ω_n} which is taken to be positive and continuous everywhere on the real line. It is assumed that for any $n > m$, $F_{\omega_n} >_{FOSD} F_{\omega_m}$. The ex-ante distribution of potential is denoted F with density f. The information about the potential can be thought of as a coarse signal that the agent observes about his likely fitness potential in the upcoming decision problem.[2]

[2]While we make distribution assumptions on potential, $\bar{y}$, we allow multiple states to map to a single potential. To see this let $S_{\bar{y}} \subset S$ the set of states for which $\bar{y}(s) = \bar{y}$. We take $\bar{Y}$ to be a random variable, $\bar{Y} : S \to \mathbb{R}$ that maps from $S_{\bar{y}} \mapsto \bar{y}$ and $f_{\bar{y}}$ to be the density of the distribution that is induced by $\bar{Y}$.

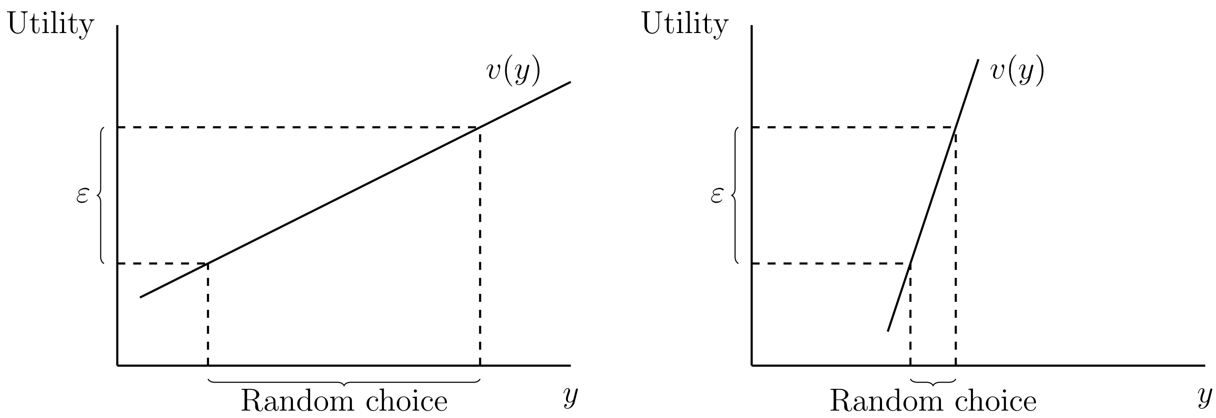


Figure 2: The steepness of the reward function $v(y)$ reduces the propensity to make random mistakes due to perceptual errors ε. A steeper slope reduces the range where two options would not be discriminated.

3.2 Nature's problem

Nature's goal is to maximize the expected fitness of the agent. This is equivalent to minimizing the expected fitness loss relative to the agent's potential: $\mathbb{E}_s[\bar{y}(s) - \phi(s,x)]$. To achieve this goal, Nature can endow the agent with a utility function over outcomes but faces biological constraints. We adopt the two assumptions in Rayo and Becker (2007).

Assumption 1 (Bounded utility)**.** $v(y) \in [0,1]$ *for all* $y \in \mathbb{R}$.

Assumption 2 (Imprecise perception)**.** *For each state* s*, the agent will choose randomly from* $X_{\varepsilon,v}(s) = \{x \in \mathbb{R} | v(\bar{y}(s)) - v(\phi(x,s)) < \varepsilon\}$*, where* $\varepsilon > 0$ *is a constant.*

These constraints make Nature's task non-trivial. The imprecision of perception (Assumption 2) means that the agent is unable to perfectly discriminate between the utility of close outputs. The agent randomly selects between choices that map to outcomes whose utility is within ε of the maximal attainable utility. A hedonic utility function can counteract this imprecision with a steeper utility function (Figure 2) that narrows the range of outputs for which an agent makes arbitrary choices. However, because hedonic utility is bounded (Assumption 1) it cannot be steep everywhere. Nature's challenge is to optimally allocate the slope within its bounds to elicit the most effective decision-making from the agent.

In addition, we assume that the realization of ω is private to the agent and therefore Nature cannot directly use this information when endowing the agent with a utility function.

Assumption 3 (Information asymmetry)**.** *Nature is unable to condition the utility function* v *directly on the signal* ω *observed by the agent.*

Given this informational constraint, if the agent is not able to choose a utility function, Nature is limited to directly endowing the agent with a utility function only using ex-ante distribution of the agent's potential.[3] This optimization problem is given by

$$\begin{aligned} &\min_{v(y)} \quad \mathbb{E}_s[\bar{y}(s) - \phi(s,x) | x \in X_{\varepsilon,v}(s)] \\ \text{such that} \quad & v(y) \in [0,1]. \end{aligned} \tag{5}$$

In the face of this information asymmetry, it is possible for Nature to do more. She can design a utility system that incentivizes the agent to use this information to determine his utility function. We model this hedonic utility system as a mechanism where the agent's reported type, in the form of an aspiration level $\tilde{\omega}$, determines both his allocation in the form of a utility function $v_{\tilde{\omega}}(y)$ and his transfer in the form of anticipatory utility $u(\tilde{\omega})$. The information, action and utility of the agent in the aspiration and decision stages are represented in Figure 3.

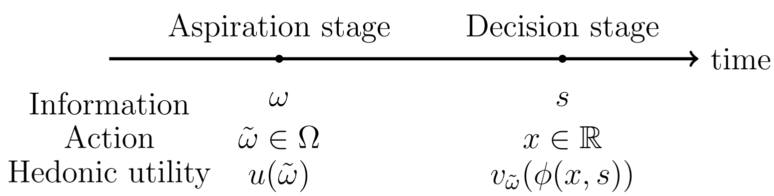


Figure 3: The aspiration and decision problems faced by the agent

We assume the agent is forward looking and chooses $\tilde{\omega} \in \Omega$ that maximizes the sum of the utility in the aspiration and decision stages. By the revelation principle, we restrict attention to implementable utility systems where the agent sends a truthful message to Nature $\tilde{\omega} = \omega$. Denoting $V_\Omega(y) = \{v_\omega(y)\}_{\omega \in \Omega}$ and $U_\Omega = \{u(\omega)\}_{\omega \in \Omega}$, Nature's problem is:

$$\begin{aligned} &\min_{V_\Omega(y), U_\Omega} \quad \mathbb{E}_{s,\omega}[\bar{y}(s) - \phi(s,x) | x \in X_{\varepsilon, v_{\tilde{\omega}}}(s)] \\ \text{such that} \quad & \omega \in \arg\max_{\tilde{\omega} \in \Omega} u(\tilde{\omega}) + \mathbb{E}_s[v_{\tilde{\omega}}(\phi(x,s)) | \omega, x \in X_{\varepsilon, v_{\tilde{\omega}}}(s)] \quad \forall \omega \in \Omega. \\ & v_\omega(y) \in [0,1], \quad \forall \omega \in \Omega. \end{aligned} \tag{6}$$

4 Characterization of Optimal Utility

In this section, we consider the two problems of Nature given by (5) and (6). We approach these problems by proceeding in the following steps that allow for tractability.

We begin by considering an approximation of the problem on a regular partition of $[-a,a] \subset \mathbb{R}$ consisting of J sub-intervals, which we denote $\mathcal{I}_{a,J}$. Sub-intervals are denoted I^i for $i = 1, \ldots, J$, and the upper endpoint of interval I^i is denoted by y^i, so that $I^i = [y^{i-1}, y^i]$. The length of each sub-interval is $c = \frac{2a}{J}$. We approximate the density of potential f over this partition with a step function $\hat{f}_\omega(y) = \sum_{i=1}^{J} \mathbf{1}_{I^i} \hat{f}^i_\omega$ where $\hat{f}^i_\omega = \frac{1}{c} \int_{y^{i-1}}^{y^i} f_\omega(y) dy$. Given this partition, we restrict the choice of utility functions to continuous functions that are linear

[3] Assumption 3 and a problem corresponding to (5) was considered in (Rayo and Becker (2007), Appendix A).

on each sub-interval. We then characterize the limit of a utility system on this partition as $\varepsilon \to 0$, where each utility function in the system satisfies the first order conditions of optimally on this restricted set, and the system satisfies incentive compatability.

Given the limiting utility system for each partition, we use the definition of the Reimann integral to characterize the limit of these utility systems as the partition becomes arbitrarily fine. In this limit, the approximation of the density of potential approaches that of the true density on $[-a, a]$. Moreover, the restriction to continuous utility functions that are piece-wise linear on each sub-interval only places a restriction of continuity on the utility function in this limit. Finally, we take $a \to \infty$ so that the approximate density approaches the true density of potential on the real line.

4.1 Optimal utility using ex-ante distribution of potential

When Nature only uses her information, her problem, formalized in (5), is to choose utility function $v(y)$ which minimizes loss given the ex-ante distribution of the agent's potential, subject to the constraints of bounded utility and imprecise perception given in Assumptions 1 and 2. From Assumption 2, given state s the agent with utility function $v(y)$ randomly chooses $x \in \mathbb{R}$ for which $v(\bar{y}(s)) - v(\phi(x, s)) \leq \varepsilon$. For a given ε, an increasing utility function uniquely determines the largest error $\underline{y}(\bar{y}(s))$ the agent can commit considering potential $\bar{y}(s)$. This output satisfies $v(\bar{y}(s)) - v(\underline{y}(\bar{y}(s))) = \varepsilon$. The expected loss associated with potential $\bar{y}(s)$ is then $(\bar{y}(s) - \underline{y}(\bar{y}(s)))/2$. As the loss only depends on s through the realization of $\bar{y}(s)$, in the following, we refer to potential as simply $\bar{y}$.

Given the approximation of the density of potential on partition $\mathcal{I}$, each potential in interval I^i is equally likely. The expected loss is therefore the sum of the expected loss within each interval weighted by the density approximation associated with that interval. This leads to the following result.

Lemma 1. *Given partition $\mathcal{I}$, $\varepsilon > 0$, approximate density of potential $\hat{f}$ and increasing utility function $v(y)$, the expected loss is given by*

$$L_{\varepsilon,\mathcal{I}}(v(y)) = \sum_{i=1}^{J} \hat{f}^i \int_{y^{i-1}}^{y^i} \frac{\bar{y} - \underline{y}(\bar{y})}{2} d\bar{y}. \tag{7}$$

A continuous utility function that is linear on each sub-interval, I^i, is determined by the choice of a vector $\mathbf{v} \in [0, 1]^{J+1}$. The $[i + 1]$-th component of the vector, denoted by v^i, is the value of the utility function at y^i. Then for $y \in I^i$ the function is determined by

$$v_{\mathcal{I}}(y; \mathbf{v}) = v^{i-1} + (y - y^{i-1}) \frac{v^i - v^{i-1}}{c}. \tag{8}$$

When convenient, we refer to the slope on interval I^i as $v'(I^i) = \frac{v^i - v^{i-1}}{c}$. With some abuse of notation we denote the loss on partition $\mathcal{I}$ given vector $\mathbf{v} \in [0, 1]^{J+1}$ as $L_{\varepsilon,\mathcal{I}}(\mathbf{v}) = L_{\varepsilon,\mathcal{I}}(v_{\mathcal{I}}(y; \mathbf{v}))$. Then the approximate problem on partition $\mathcal{I}$ is:

$$\max_{\mathbf{v} \in [0,1]^{J+1}} -L_{\varepsilon,\mathcal{I}}(\mathbf{v}) \tag{9}$$

On a fixed partition, if any $v^i \leq v^{i-1}$, the expected loss associated with potential on interval I^i will exceed $\hat{f}^i \frac{\varepsilon}{2}$. While for any $\mathbf{v}$ where $v^i - v^{i-1} > 0$ for $i = 1, \ldots, J$, this total expected loss approaches zero as $\varepsilon \to 0$. This follows directly from Lemma 2, and implies that for sufficiently small $\varepsilon > 0$, any loss minimizing utility function will be strictly increasing on $[-a, a]$. The lemma gives the expression for the expected loss for sufficiently small ε and increasing utility.[4]

Lemma 2. *For given partition $\mathcal{I}$ and $\mathbf{v} \in [0,1]^{J+1}$ where $0 < \min\{v^i - v^{i-1} : i = 1, \ldots, J\}$, when $0 < \varepsilon < \min\{v^i - v^{i-1} : i = 1, \ldots, J\}$ the loss is*

$$L_{\varepsilon,\mathcal{I}}(\mathbf{v}) = \sum_{i=1}^{J} \hat{f}^i \frac{c^2 \varepsilon}{2(v^i - v^{i-1})} + o(\varepsilon). \tag{10}$$

To understand the conditions of optimality, consider the impact on the loss $L_{\varepsilon,\mathcal{I}}(\mathbf{v})$ from changing v^i keeping the values at all other endpoints the same. A higher value of v^i increases the slope of $v_{\mathcal{I}}(y; \mathbf{v})$ on $[y^{i-1}, y^i]$, reducing the loss associated with interval I^i. At the same time, it reduces the slope on $[y^i, y^{i+1}]$, increasing the loss stemming from choices in interval I^{i+1}. For each i, the necessary condition of optimally requires that given $\mathbf{v}^{-i}$, which denotes the value of v^j for all $j \neq i$, this change in expected loss is zero. This condition for sufficiently small $\varepsilon > 0$ is given in Lemma 3.

Lemma 3. *For $i = 1, \ldots, J-1$, given an increasing $\mathbf{v}^{-i}$, and sufficiently small ε, $\hat{v}^i(v^{-i})$, the value of v^i that minimizes $L_{\varepsilon,\mathcal{I}}(\mathbf{v})$ given $\mathbf{v}^{-i}$ satisfies*

$$\left.\frac{\partial L_{\varepsilon,\mathcal{I}}(\mathbf{v})}{\partial v^i}\right|_{v^i = \hat{v}^i} = \frac{\hat{f}^{i+1} c^2 \varepsilon}{2(v^{i+1} - \hat{v}^i(\mathbf{v}^{-i}))^2} - \frac{\hat{f}^i c^2 \varepsilon}{2(\hat{v}^i(\mathbf{v}^{-i}) - v^{i-1})^2} + o(\varepsilon) = 0. \tag{11}$$

From Lemma 3 as $\varepsilon \to 0$, the conditions of optimally for $i = 1, \ldots, J$, imply

$$\frac{\hat{f}^{i+1} c^2}{(v^{i+1} - \hat{v}^i(\mathbf{v}^{-i}))^2} = \frac{\hat{f}^i c^2}{(\hat{v}^i(\mathbf{v}^{-i}) - v^{i-1})^2} \Rightarrow \frac{v_{\mathcal{I}}^i - v_{\mathcal{I}}^{i-1}}{c} = \frac{(\hat{f}^i)^{1/2}}{D}, \tag{12}$$

where $\mathbf{v}_{\mathcal{I}}$ is uniquely determined by (12), $v_{\mathcal{I}}^0 = 0$ and $v_{\mathcal{I}}^J = 1$, and D is a positive constant that does not depend on i. Equation (12) requires the slope of the resulting piece-wise linear utility function, $v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ proportional to the square root of the approximated density function on each interval. It therefore has a steeper slope on intervals where $\hat{f}$ is larger, see Figure 4. Proposition 4 provides the limit of $v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ as the approximate density of potential approaches the true density, f.

Proposition 4. *For a given density of potential f, the limit of $v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ as the partition $\mathcal{I}_{a,J}$ gets progressively more refined ($J \to \infty$) and subsequently extends to $\mathbb{R}$ ($a \to \infty$) is*

$$v^*(y) = K \int_{-\infty}^{y} f(z)^{1/2} dz,$$

where $K = (\int_{-\infty}^{\infty} f(z)^{1/2} dz)^{-1}$ is a normalizing constant.

[4] Here $o(\varepsilon)$ is used to describe a function $g(\varepsilon)$ such that $\frac{g(\varepsilon)}{\varepsilon} \to 0$ as $\varepsilon \to 0$.

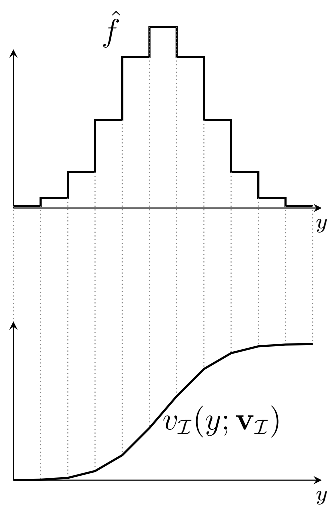


Figure 4: The piece-wise linear utility function $v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ uniquely determined by the limit of the first order conditions of (9) as $\varepsilon \to 0$ has greater slope on the intervals where $\hat{f}$ is higher.

If the distribution of output potential is single-peaked and symmetric around a mean, $\bar{y}^r$, then $v^*(y)$ exhibits odd symmetry around the point $(\bar{y}^r, 1/2)$. In that case, the fitness maximizing utility function is symmetrically S-shaped with an inflection point at the expected output potential, $\bar{y}^r$, which we refer to as the *reference point*.

The result is in line with past literature. Optimal utility functions have previously been identified as integrals of power transformations of the density function of potential output. From Netzer (2009), two optimal utility functions are given by

$$v^*_{Rob}(y) = \int_{-\infty}^{y} f(z)dz = F_\omega(y)$$
$$v^*_{Net}(y) = K \int_{-\infty}^{y} f(z)^{2/3}dz$$

where $K = \left(\int_{-\infty}^{\infty} f(z)^{2/3}dz\right)^{-1}$. The first, $v^*_{Rob}(y)$, was derived in Robson (2001*a*) as the function that minimizes the number of mistakes the agent makes when choosing between two actions with output levels which were drawn independently from f. In the same setting, Netzer (2009) showed that $v^*_{Net}(y)$ maximizes the expected fitness choice.

4.2 Utility system allowing agent to choose a reference point

We now consider the case where Nature designs a utility system that allocates a utility function $v_{\tilde{\omega}}(y)$ and anticipatory utility $u(\tilde{\omega})$ to an agent that reports information $\tilde{\omega}$ and achieves outcome y. To incentivize the agent to reveal his information, the utility system must satisfy the truth-telling condition.

A natural question is whether anticipatory utility can be used to implement the first best utility functions $v^*(y)$ in Proposition 4. The answer is, in general, no. Consider a set of states for which the conditional densities of potential are single-peaked, symmetric, and shifted by their respective means from zero. When there are more than two signals, it is not possible for Nature to incentivize the agent to correctly set his aspiration $\tilde{\omega} = \omega$ with utility functions of the form of $v^*(y)$. Proposition 5 states this result formally.

Proposition 5. *Let $|\Omega| \geq 3$ and f_{ω_n} be single peaked and symmetric around $\bar{y}^r_{\omega_n}$ with $f_{\omega_n}(y - \bar{y}^r_{\omega_n}) = f_{\omega_m}(y - \bar{y}^r_{\omega_m})$. Let $v^*_{\omega_n}$ denote the utility function described in Proposition 4 given density f_{ω_n}. Then there are no anticipatory utilities, $u(\omega_n)$, such that $u(\omega_n) + \mathbb{E}[v^*_{\omega_n}(y)|\omega_n] \geq u(\omega_m) + \mathbb{E}[v^*_{\omega_m}(y)|\omega_n]$ for all m, n.*

To characterize the utility system that satisfies the incentive compatibility constraints given in (6), we again approach the problem by approximating each of the conditional densities, f_{ω_n}, over a partition $\mathcal{I}$. For this partition, in the approximate problem, nature chooses $\mathbf{V} = [\mathbf{v_1}, \ldots, \mathbf{v_N}]$ and $[u_1, \ldots, u_N] \in [0,1]^N$ to minimize loss $L_{\varepsilon,\mathcal{I}}(\mathbf{V}) = \sum_{n=1}^N p_{\omega_n} L_{\varepsilon,\mathcal{I}}(\mathbf{v}_n; \omega_n)$. The problem is given by:

$$\begin{aligned} \max_{\mathbf{v}_n, \mathbf{u}} \quad & -\sum_{n=1}^N p_{\omega_n} L_{\varepsilon,\mathcal{I}}(\mathbf{v}_n; \omega_n) \qquad (13) \\ \text{such that} \quad & v_n^i \in [0,1] \quad \forall n = 1, \ldots, N, \ i = 0, \ldots, J \\ & \mathbb{E}[v_n(y; \mathbf{v}_n)|\omega_n] + u_n \geq \mathbb{E}[v_m(y; \mathbf{v}_m)|\omega_n] + u_m \quad \forall n, m = 1, \ldots, N. \end{aligned}$$

There are two sets of inequality constraints that the vectors $\mathbf{v}_n$ and $\mathbf{u}$ must satisfy. The first is the bounded constraints $0 \leq v_n^i \leq 1$, which have multipliers $\gamma^i_{n,0}(\varepsilon)$ and $\gamma^i_{n,1}(\varepsilon)$ for the lower and upper bounds respectively. The second set is the truth-telling constraints which have multipliers $\lambda_{n,m}(\varepsilon)$ for all $n, m = 1, \ldots, N$.[5] Denoting $h_{n,m}(\mathbf{v}_n, \mathbf{v}_m) = \mathbb{E}[v_n(y; \mathbf{v}_n) - v_m(y; \mathbf{v}_m)|\omega_n]$, the necessary conditions for $\mathbf{v}^*_n, \mathbf{u}^*$ to be a local optimizer are

1. For every $n = 1, \ldots, N$ and $i = 0, \ldots, J$

$$\frac{\partial L_{\varepsilon,\mathcal{I}}(\mathbf{V}^*)}{\partial v_n^i} = \sum_{m=1}^N \lambda_{n,m}(\varepsilon) \frac{\partial h_{n,m}(\mathbf{v}^*_n, \mathbf{v}^*_m)}{\partial v_n^i} + \sum_{m=1}^N \lambda_{m,n}(\varepsilon) \frac{\partial h_{m,n}(\mathbf{v}^*_n, \mathbf{v}^*_m)}{\partial v_n^i} + \gamma^i_{n,0}(\varepsilon) - \gamma^i_{n,1}(\varepsilon), \quad (14)$$

2. for every $n = 1, \ldots, N$

$$\sum_{m=1}^N \lambda_{n,m}(\varepsilon) = \sum_{m=1}^N \lambda_{m,n}(\varepsilon), \qquad (15)$$

3. and $\lambda_{n,m}(\varepsilon)(h_{n,m}(\mathbf{v}^*_n, \mathbf{v}^*_m) + u^*_n - u^*_m) = 0$ for every $n, m = 1, \ldots, N$; $\gamma^i_{n,0}(\varepsilon) v_n^{i*} = 0$ and $\gamma^i_{n,1}(\varepsilon)(1 - v_n^{i*}) = 0$ for all $i = 0, \ldots, J$ and $n = 1, \ldots, N$.

We denote the multipliers as functions of ε to highlight that the expected loss decreases toward zero as $\varepsilon \to 0$, and therefore the impact of relaxing the constraints on this loss will also vanish. As we are interested in the limit of the conditions of optimality as $\varepsilon \to 0$, we will rewrite the conditions in terms of normalized multipliers. We define these as $\lambda_{n,m} \equiv \lim_{\varepsilon \to 0} \lambda_{n,m}(\varepsilon)/\varepsilon$, $\gamma^i_{n,0} \equiv \lim_{\varepsilon \to 0} \gamma^i_{n,0}(\varepsilon)/\varepsilon$, and $\gamma^i_{n,1} \equiv \lim_{\varepsilon \to 0} \gamma^i_{n,1}(\varepsilon)/\varepsilon$.

For given $\mathcal{I}$, we proceed by fixing the limit of the normalized multipliers that satisfy (15) and characterizing the set of vectors, $\mathbf{v}_n$, that satisfy (14). We use λ and γ, as shorthand for $\lambda_{n,m}$ for $n, m = 1, \ldots, N, n \neq m$ and $\gamma^i_{n,k}$ for $n = 1, \ldots, N$, $k = 0, 1$ and $i = 0, \ldots, J$.

[5]We also constrain $u_n \in [0,1]$. These constraints never bind given the truth-telling and limit constraints on $\mathbf{V}$.

Lemma 6. *Denote $\hat{v}_n^i(\mathbf{v}_n^{-i}, \lambda, \gamma)$, as the loss minimizing value of v_n^i that satisfies (14) given λ, γ, and $\mathbf{v}_n^{-i}$.*

1. *For all n and $i = 1, \ldots, J-1$, increasing $\mathbf{v}_n^{-i}$, and sufficiently small ε, $\gamma_{n,0}^i(\varepsilon) = \gamma_{n,1}^i(\varepsilon) = 0$ and $\hat{v}_n^i(\mathbf{v}_n^{-i}, \lambda, \gamma)$ satisfies*

$$\frac{p_{\omega_n} \hat{f}_{\omega_n}^{i+1} c^2 \varepsilon}{\left(v_n^{i+1} - \hat{v}_n^i(\mathbf{v}_n^{-i}, \lambda, \gamma)\right)^2} - \frac{p_{\omega_n} \hat{f}_{\omega_n}^{i} c^2 \varepsilon}{\left(\hat{v}_n^i(\mathbf{v}_n^{-i}, \lambda, \gamma) - v_n^{i-1}\right)^2} - \sum_{m=1}^{N} \lambda_{n,m} \frac{(\hat{f}_{\omega_n}^i + \hat{f}_{\omega_n}^{i+1}) c^2 \varepsilon}{2}$$
$$+ \sum_{m=1}^{N} \lambda_{m,n} \frac{(\hat{f}_{\omega_m}^i + \hat{f}_{\omega_m}^{i+1}) c^2 \varepsilon}{2} + o(\varepsilon) = 0. \qquad (16)$$

2. *For all n and sufficiently small ε, $\gamma_{n,1}^0(\varepsilon)$ must be zero, and $\hat{v}_n^0(v^{-0}, \lambda, \gamma)$ satisfies*

$$\frac{p_{\omega_n} \hat{f}_{\omega_n}^{1} c^2 \varepsilon}{\left(v_n^{1} - \hat{v}_n^0(v^{-0}, \lambda, \gamma)\right)^2} - \sum_{m=1}^{N} \lambda_{n,m} \frac{\hat{f}_{\omega_n}^1 c^2 \varepsilon}{2} + \sum_{m=1}^{N} \lambda_{m,n} \frac{\hat{f}_{\omega_m}^1 c^2 \varepsilon}{2} - \varepsilon \gamma_{n,0}^0 + o(\varepsilon) = 0. \qquad (17)$$

3. *For all n and sufficiently small ε, $\gamma_{n,0}^J(\varepsilon)$ must be zero, and $\hat{v}_n^J(v^{-J}, \lambda, \gamma)$ satisfies*

$$\frac{p_{\omega_n} \hat{f}_{\omega_n}^{J} c^2 \varepsilon}{\left(\hat{v}_n^J(v^{-J}, \lambda, \gamma) - v_n^{J-1}\right)^2} - \sum_{m=1}^{N} \lambda_{n,m} \frac{\hat{f}_{\omega_n}^J c^2 \varepsilon}{2} + \sum_{m=1}^{N} \lambda_{m,n} \frac{\hat{f}_{\omega_m}^J c^2 \varepsilon}{2} - \varepsilon \gamma_{n,1}^J + o(\varepsilon) = 0. \qquad (18)$$

Relative to (11) the additional terms in (16) reflect the impact of increasing v_n^i on the expected utility associated with utility function $v_{\mathcal{I}}(y; \mathbf{v_n})$. It increases the utility function on intervals I^i and I^{i+1} only. For small ε, this impacts the expected utility only when the potential falls in these intervals, which, given ω, occurs with probability $\hat{f}_\omega^i$ and $\hat{f}_\omega^{i+1}$ respectively. In (17) and (18), increasing v_n^0 and v_n^J only impact the utility function on a single interval, I^1 and I^J respectively. These conditions also have additional potentially positive multipliers, which reflect the potential for the lower and upper constraints to bind so that $v_n^0 = 0$ and $v_n^J = 1$ respectively.

Define $\mathbf{v}_{n,\mathcal{I}}$ as the vector that satisfies the conditions of Lemma 6 as $\varepsilon \to 0$. The following lemma characterizes the utility functions derived from these vectors on this partition.

Lemma 7. *Fix $\mathcal{I}$ and let λ and γ be given that satisfy (15) and the complementary slackness conditions. The value of $\mathbf{v}_{n,\mathcal{I}}$ is uniquely determined up to the values of $v_{n,\mathcal{I}}^0$ and $v_{n,\mathcal{I}}^J$, and the resulting piece-wise linear utility functions are*

$$v_{\mathcal{I}}(y; \mathbf{v}_{n,\mathcal{I}}) = v_{n,\mathcal{I}}^0 + (v_{n,\mathcal{I}}^J - v_{n,\mathcal{I}}^0) \frac{c \sum_{j=1}^{i-1} \left(\hat{f}_{\omega_n}^j / D_{\omega_n}^j\right)^{1/2} + (y - y^{i-1}) \left(\hat{f}_{\omega_n}^i / D_{\omega_n}^i\right)^{1/2}}{c \sum_{j=1}^{I} \left(\hat{f}_{\omega_n}^j / D_{\omega_n}^j\right)^{1/2}} \quad \textit{for } y \in I^i$$

$$\textit{where} \quad D_{\omega_n}^i = \sum_{m=1}^{M} \frac{c \lambda_{m,n}}{p_{\omega_n}} \left[\frac{1}{2} \left(\hat{f}_{\omega_n}^i - \hat{f}_{\omega_m}^i \right) + \sum_{j=1}^{i-1} \hat{f}_{\omega_n}^j - \hat{f}_{\omega_m}^j \right] + \frac{\gamma_{n,0}^0}{p_{\omega_n}}.$$

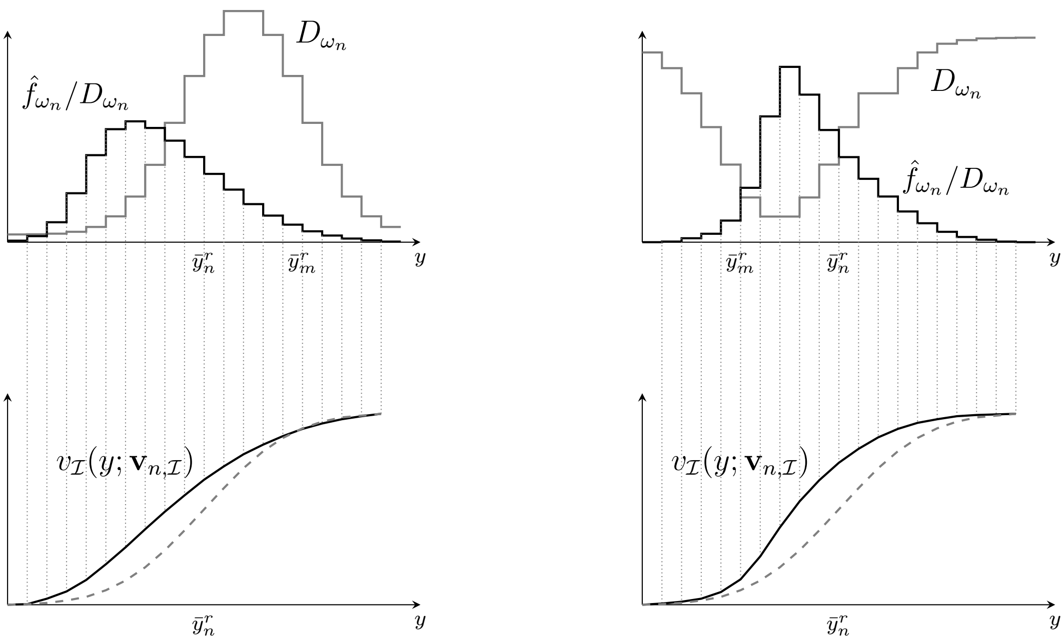


Figure 5: The piece-wise linear utility functions $v_{\mathcal{I}}(y;\mathbf{v}_{n,\mathcal{I}})$ which satisfy (14) as $\varepsilon \to 0$ when $\lambda_{m,n} > 0$ for some $m > n$ (left panel) and when $\lambda_{m,n} > 0$ for some $m < n$ (right panel).

Relative to the condition given in (12), the relationship between the slope of the utility function and the density of the potential on each interval is no longer constant. Instead the ratio of the slope and the density on I^i is given by $D^i_{\omega_n}$ as described in Lemma 7, and leads to the following condition

$$\frac{v_n^i - v_n^{i-1}}{c} = \frac{(\hat{f}^i_{\omega_n})^{1/2}}{D^i_{\omega_n}} \quad \text{for } i = 1, \dots, J \text{ and } n = 1, \dots, N. \tag{19}$$

Given partition $\mathcal{I}$, we define $D_{\omega_n,\mathcal{I}}(y) = \sum_{i=1}^{J} \mathbf{1}_{I^i} D^i_{\omega_n}$. Examples of this function are given in Figure 5. When $\lambda_{m,n} > 0$ for $n < m$ and $\lambda_{n,m} = 0$ for all other m, then $D_{\omega_n,\mathcal{I}}$ has attains its maximum between $\bar{y}^r_{\omega_n}$ and $\bar{y}^r_{\omega_m}$. When $\lambda_{m,n} > 0$ for $n > m$ and $\lambda_{n,m} = 0$ for all other m, then $D_{\omega_n,\mathcal{I}}$ attains its minimum between $\bar{y}^r_{\omega_m}$ and $\bar{y}^r_{\omega_n}$. Then for any set of $0 \le \lambda_{n,m}$, $D_{\omega_n,\mathcal{I}}$ will be a linear combination of the $D_{\omega_n,\mathcal{I}}$ in Figure 5. The resulting $v_{\mathcal{I}}(y;\mathbf{v}_{n,\mathcal{I}})$ will be steeper in intervals below $\bar{y}^r_n$ than those equally far above $\bar{y}^r_n$, generating a pattern of loss aversion when the underlying densities of potential are symmetric about the expected potential.

Proposition 8 characterizes $v_{\mathcal{I}}(y;\mathbf{v}_{n,\mathcal{I}})$ as the approximate density of potential approaches the true density, f_{ω_n}. The property of loss aversion for this limiting function when densities of potential are symmetric is formalized in Proposition 9.

Proposition 8. *Given a set of conditional densities f_{ω_n} for $\omega_n \in \Omega$ that are observed by the agent with probability p_{ω_n} and set of multipliers λ and γ that satisfy (15) and the complementary slackness condition, the limit of $v_{\mathcal{I}}(y,\mathbf{v}_{n,\mathcal{I}})$ as the partition $\mathcal{I}_{a,J}$ gets progressive more refined ($J \to \infty$) and subsequently extends to $\mathbb{R}$ ($a \to \infty$) is*

$$v^*_{\omega_n}(y) = v_n^{-\infty} + \frac{v_n^{\infty} - v_n^{-\infty}}{K_n} \int_{-a}^{y} \frac{(f_{\omega_n}(z))^{1/2}}{D_{\omega_n}(z)} dz.$$

where $D_{\omega_n}(y) = \left(\sum_{m=1}^{M} \frac{\lambda_{m,n}}{p_{\omega_n}} (F_{\omega_n}(y) - F_{\omega_m}(y)) + \frac{\gamma_{n,0}^{-\infty}}{p_{\omega_n}}\right)^{1/2}$, $K_n = \int_{-\infty}^{\infty} \frac{(f_{\omega_n}(z))^{1/2}}{D_{\omega_n}(z)} dz$, $v_n^{-\infty} = \lim_{a\to\infty} \lim_{J\to\infty} v_{n,\mathcal{I}}^0$, $v_n^{\infty} = \lim_{a\to\infty} \lim_{J\to\infty} v_{n,\mathcal{I}}^J$, *and* $\gamma_{n,0}^{\infty}$ *is the multiplier associated with the constraint* $v_n^{-\infty} \geq 0$*.*

Proposition 8 provides the general form of the utility functions given multipliers λ and γ. It remains to characterize the anticipatory utility $\mathbf{u}$ and these multipliers that satisfy (15) and the complementary slackness conditions from problem (13). The set of incentive constraints that bind, if any, will depend on the distributions of potential.

When an incentive constraint between ω_n and ω_m binds and $\lambda_{n,m} > 0$, then $\mathbf{u}$ is such that $u_n - u_m = \mathbb{E}[v^*_{\omega_m}(y) - v^*_{\omega_n}(y)|\omega_n]$. If $n > m$, then positive $\lambda_{n,m}$ impacts the shape of $v^*_{\omega_m}(y)$ by reducing the relative steepness of the utility function where $F_{\omega_m}(y) - F_{\omega_n}(y)$ is the largest. This will decrease the benefit of pessimism for the agent with information ω_n choosing the reference associated with ω_m. That in turn reduces the difference in anticipatory utility that is required to satisfy the incentive constraint. Similarly, if $n < m$, positive $\lambda_{n,m}$ increases the relative steepness of $v^*_{\omega_m}(y)$ where $F_{\omega_n}(y) - F_{\omega_m}(y)$ is the largest. This increases the cost of optimism for the agent with information ω_n choosing the reference associated with ω_m. Again, this reduces the difference in anticipatory utility that is required to satisfy the incentive constraint.

For the remainder of this section we assume that for each $\omega_n, \omega_m \in \Omega$, $f_{\omega_n}(y - \bar{y}^r_{\omega_n}) = f_{\omega_m}(y - \bar{y}^r_{\omega_m})$ and $f_{\omega_n}, f_{\omega_m}$ are single peaked and symmetric around $\bar{y}^r_{\omega_n}$. When the distributions of potential have this shape, then utility functions $v^*(y)$ as characterized in Proposition 4 will be S-shaped and oddly symmetric around the expected potential. In the following Proposition, we show that whenever any incentive compatibility constraint binds and $\lambda_{n,m}$ is positive, then the change in shape of $v^*_m(y)$ introduces loss aversion to the utility function. Moreover, from (15), there necessarily is a positive $\lambda_{k,n}$ for some $k \neq n$ and $v^*_n(y)$ also exhibits loss aversion.

Proposition 9 (Loss aversion)**.** *Let* $\bar{y}^r_n = \mathbb{E}[\bar{y}|\omega_n]$ *denote the reference point of the utility function* $v^*_{\omega_n}(y)$*. Let* f_{ω_n} *be single peaked and symmetric around* $\bar{y}^r_{\omega_n}$ *with* $f_{\omega_n}(y - \bar{y}^r_{\omega_n}) = f_{\omega_m}(y - \bar{y}^r_{\omega_m})$ *and let* $\lambda_{m,n} > 0$ *for some* $m \neq n$*. Then the following inequality holds for all* $y > 0$*:*

$$v^*_{\omega_n}(\bar{y}^r_{\omega_n}) - v^*_{\omega_n}(\bar{y}^r_{\omega_n} - y) > v^*_{\omega_n}(\bar{y}^r_{\omega_n} + y) - v^*_{\omega_n}(\bar{y}^r_{\omega_n}).$$

From Proposition 5 when there a more than two states, then incentive compatibility cannot be satisfied when $\lambda_{n,m} = 0$ for all n, m and the utility functions use the entire range ($v_n^{-\infty} = 0, v_n^{\infty} = 1$). Moreover, if the utility functions do not use the entire range in order to satisfy an incentive compatibility constraint, then there would be a decrease in loss from relaxing this constraint. This implies that $\lambda_{n,m}$ must be greater than 0 for some $n \neq m$, and the utility system will have functions with loss aversion.

4.3 Example: Three state case with normal distributions

Let there be three states $\Omega = \{\omega_1, \omega_2, \omega_3\}$ where the distribution of fitness potential conditional on signal ω_i is normally distributed with mean ω_i and variance one. Given the properties of normal densities we fist show that $v_n^{-\infty} = 0$ for all n, and $v_N^{\infty} = 1$.

Lemma 10. *If for all $m > n$ $\lim_{y \to -\infty} \frac{f_{\omega_m}(y)}{f_{\omega_n}(y)} \to 0$, then $v_n^{-\infty} = 0$, for $n = 1, \ldots, N$ and $v_N^{\infty} = 1$.*

From the discussion above, there is a positive $\lambda_{n,m}$. In the three state case, we can show that local downward IC constraints must bind and that the global upward constraint must also bind in this three state case. There are no other binding constraints. Moreover, from (15), the multipliers for these three constraints must be equal.

Lemma 11. *In the optimal utility system, $\lambda_{2,1} = \lambda_{3,2} = \lambda_{1,3} > 0$, and $\lambda_{1,2} = \lambda_{2,3} = \lambda_{3,1} = 0$.*

Then from Proposition 8, the utility functions that satisfy (14) take the form

$$v_{\omega_1}^*(y) = \frac{v_1^{\infty}}{K_1} \int_{-\infty}^{y} \frac{\left(p_{\omega_n} \varphi(z - \omega_1)\right)^{1/2}}{\left(\lambda_{2,1}(\Phi(z - \omega_1) - \Phi(z - \omega_2)) + \gamma_{1,0}^{-\infty}\right)^{1/2}} dz$$

$$v_{\omega_2}^*(y) = \frac{v_2^{\infty}}{K_2} \int_{-\infty}^{y} \frac{\left(p_{\omega_2} \varphi(z - \omega_2)\right)^{1/2}}{\left(\lambda_{3,2}(\Phi(z - \omega_2) - \Phi(z - \omega_3)) + \gamma_{2,0}^{-\infty}\right)^{1/2}} dz$$

$$v_{\omega_3}^*(y) = \frac{1}{K_3} \int_{-\infty}^{y} \frac{\left(p_{\omega_3} \varphi(z - \omega_3)\right)^{1/2}}{\left(\lambda_{1,3}(\Phi(z - \omega_3) - \Phi(z - \omega_1)) + \gamma_{3,0}^{-\infty}\right)^{1/2}} dz,$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution and K_n are normalising constants given in Proposition 8. A representation of the utility functions is given in Figure 1.

5 Discussion: A unifying explanation of several behavioural patterns

5.1 Anticipatory utility

The idea that higher expectations generate positive satisfaction can at least be traced back to Bentham (1789) and Jevons (1905). The notion of anticipatory utility has now been formally incorporated in economics (Loewenstein, 1987; Kőszegi, 2010; Brunnermeier and Parker, 2005). The fact that subjective satisfaction occurs with the reception of positive news that raises expectations, aligns with findings in neuroscience (Schultz, Dayan and Montague, 1997). Dopaminergic neurons, which release the neurotransmitter dopamine, are believed to be at the core of the brain's subjective reward system. Their firing rates escalate when an unforeseen reward materializes, or upon receiving unanticipated news of a forthcoming reward. These characteristics have led to the "reward prediction error" hypothesis (Schultz, 2016), suggesting that our subjective reward system solely gratifies positive surprises in relation to expectations. This hypothesis seamlessly corresponds with the attributes of the optimal hedonic system delineated here. Favorable news evokes both a sense of anticipatory satisfaction and an adjustment in expectations, resulting in the subsequent (anticipated) reward not being linked to heightened hedonic utility. Only unanticipated positive outcomes at the decision-making juncture yield positive rewards.

We demonstrate that these attributes can form an integral part of an adaptive hedonic reward system, which facilitates the agent's choice of well-calibrated aspirations generating decisions.

5.2 Loss aversion

Loss aversion, one of the most influential concepts in behavioral economics, was introduced by Kahneman and Tversky in their seminal 1979 paper. Despite numerous studies investigating loss aversion and its causes, Novemsky and Kahneman (2005) noted that "a realistic theory of loss aversion is unlikely to be simple." Our model explains the existence of loss aversion as a complement to anticipatory utility providing incentives to the agent when setting his aspirations. The fact that anticipatory utility and loss aversion can generate a trade-off in the choice of an aspiration level has previously been considered by Sarver (2012). In our framework, loss aversion emerges as a feature of an optimal utility system that induces the agent to set his aspirations at the level of his expectations.

In addition to providing a justification for the existence of loss aversion, our model also makes predictions about its features. One feature strongly associated with loss aversion is the idea that the value function exhibits a kink at the reference point, with different slopes to the left and right. Kahneman (2003) explicitly stated that "the core idea of prospect theory [is]—that the value function is kinked at the reference point and loss averse." However, the existence of a kink at zero has not been empirically established. Instead, recent research has stressed that loss aversion seems absent for small payoffs of around \$5-\$12 (Ert and Erev, 2013; Gal and Rucker, 2018; Zeif and Yechiam, 2022; Chapman et al., 2022). And a non-parametric estimation of prospect theory utility functions has found loss aversion overall but no kink at zero (Abdellaoui, Bleichrodt and Kammoun, 2013). A reexamination of past evidence also found that only studies with large payoffs actually showed loss aversion at the time when Kahneman and Tversky published their 1979 paper (Yechiam, 2019).

The fact that loss aversion can be detected for large payoffs but not for small payoffs is compatible with an absence of kink and a smooth value function around the reference point. Differences in slopes above and below the reference point tend to disappear for small payoff variations around the reference point. It is perfectly possible to have both $v'(-x)/v'(x) > 1$ and $\lim_{x\to 0} v'(-x)/v'(x) = 1$, a feature exhibited by our optimal value function.

Our approach also suggests a connection between an agent's familiarity with a setting and the degree of loss aversion exhibited. When an agent encounters a specific type of situation repeatedly, each outcome serves as a sample of observations that can be automatically incorporated into an adaptive utility function (Robson, Whitehead and Robalino, 2021). In that perspective, only novel, context-specific information requires the features of anticipatory utility and loss aversion to induce the agent to adjust his aspirations.

In a comprehensive study involving over 17,000 participants, Mrkva et al. (2020) found that experts are less likely to exhibit loss aversion. Our framework can account for this observation: the asymmetry in the hedonic utility function arises from the agent's *marginal information* that has not yet been incorporated into his reward system. An expert who is highly familiar with a type of situation is less likely to encounter novel, context-specific information. Intuitively, failing to reach ones' reference point may have different hedonic costs for an expert a a novice. For an expert, it may simply reflect an unlucky draw from a known distribution. For a novice, it may come with an additional hedonic cost from the realisation that his reference point was too optimistic in the first place.

5.3 Selection of peer group

It is now well understood that peer group comparisons significantly impact our sense of satisfaction and well-being (Clark and Oswald, 1996; Clark, Frijters and Shields, 2008; Card et al., 2012). These peer groups not only shape our perceptions but also influence our behavior. Physicians have been found to modify their behavior to increase their income (Rizzo and Zeckhauser, 2003), and people witnessing their neighbors win luxury cars in a national lottery increased their spending on cars afterwards (Kuhn et al., 2011).

The effect of peer groups on our well-being raises the possibility for us to choose our peer group in a way so as to improve our well-being (Frank, 1985). Some evidence points to such behavior. For example, in a large resident matching program for medical students, students, on average, preferred to live in places where their relative income would be higher (Bottan and Perez-Truglia, 2022). However, the "keeping up with the Joneses" phenomenon suggests a preference to maintain pace with a high reference group rather than changing it (Harris, Anseel and Lievens, 2008).

Considering the negative impact on our well-being of peer groups with relatively high social status and wealth, and given our degree of freedom in our choice of peer group, a critical question arises: why do we not opt for the most modest peer groups possible? Our result suggests that anticipatory utility can be the force preventing us from choosing convenient reference groups relative to which our achievements look flattering. By incentivizing us to be ambitious, anticipatory utility may drive us to vie for belonging to reference groups with high-achieving members.

5.4 Goal setting

The psychological literature on goal setting (Locke and Latham, 1990) has investigated the role of goals on performance. Setting goals has been found to have a significant motivational impact and increase performance, at least when goals are challenging and are neither too easy nor too hard to reach. Heath, Larrick and Wu (1999) proposed that goals' effect on motivation derives from the fact that they act as a reference point. The steepness of the value function close to a goal, including due to loss aversion, motivates the decision-maker to work hard to achieve the set goal.

How goals are set is however still an open question with both the individual and the surrounding context thought to play a role. In regard to individual choices, goal setting may seem paradoxical. If goals act as reference points the best path to happiness should be to set easily reachable goals. In contradiction with this suggestion, people seem to set themselves ambitious goals. In a study of marathon runners, Allen et al. (2017) found that symbolic thresholds (e.g., finishing in under three hours) act as reference points that increase the performance of runners. Similarly, chess players seem to exert extra effort to improve their best rating, acting as if their personal best is a reference point (Anderson and Green, 2018).

In the process of setting these often challenging goals, it seems that feedback, peer comparison, and past experience contribute to influence how goals are set (Locke and Latham, 2006). But the psychological literature on goal setting has not reached a definite understanding of how these factors contribute to shaping the level of the goals people set for themselves. Our approach provides a unifying explanation for the role of these different factors in the formation of goals. Our optimal reward system is designed to induce us to best use our information to identify our potential and then act to reach it. Personal goals tend to be ambitious because we are rewarded for identifying the highest success level we can reach. In that perspective, external factors such as feedback, peer comparison, or past experience all play a role in shaping our expectations and therefore our reference point.

5.5 Habituation in subjective well-being

The idea that we habituate to changes in circumstances has been proposed under the names adaptation (Helson, 1948), hedonic treadmill (Brickman, 1971), set point theory (Lykken and Tellegen, 1996), and hedonic adaptation (Frederick and Loewenstein, 1999). The idea of habituation predominates in the field of research on subjective well-being (Luhmann and Intelisano, 2018). Our model makes three contributions to this field of research.

First, following existing research on the evolutionary foundation of reference-dependence, it provides an explanation for the mechanism underlying habituation that is still debated among subjective well-being researchers (Luhmann and Intelisano, 2018). Habituation emerges as a solution to help the agent make good decisions by calibrating their perception of subjective value to the range experienced in the context they face. This explanation departs from typical explanations of habituation that have often relied on physiological constraints. Cummins (2013) has, for instance, proposed that subjective well-being follows a homeostatic process like our body temperature because extreme levels of subjective well-being have negative effects on us. Instead, following Robson (2001*a*), we explain habituation as a process required for the agent to have a precise perception of the differences in the values of the options they face when moving across different contexts. The explanation for habituation is enhanced decision-making.

Second, many explanations of how habituation works have been proposed. (Helson, 1948) initially proposed that happiness is a function of the difference between current outcomes and an average of past outcomes. Brickman (1971) extended this idea by arguing that we also compare our current outcomes to those of our peers. Rayo and Becker (2007) integrated these two insights, showing how both past outcomes and peer outcomes can be used to calibrate our subjective value function in the range we are most likely to experience. Our results generalize this perspective by allowing the reference point to be influenced by any relevant piece of information. An important consequence is that habituation can, in part, be forward-looking, an idea suggested by Frederick and Loewenstein (1999). Having information about future outcomes could already influence our reference point before the fact. Frederick and Loewenstein mention, as supporting evidence, that ”the final days of a prison sentence are often regarded as the most frustrating” and that revolutions often happen when conditions begin to improve, but not as fast as expectations.

Finally, our results provide an explanation for what has been a puzzling fact in the subjective well-being literature: the fact that the baseline level of happiness to which we tend to revert is not neutral, but positive (Diener, Lucas and Scollon, 2009). In general, people tend to be moderately happy. Our model provides an explanation for this fact. Unlike in the original prospect theory model, our subjective value function is not assumed to be at a neutral point around our aspirations. Instead, it is pushed upward to reward the agent for setting their aspirations at the level of their expectations. Since this hedonic system is incentive-compatible, it should induce agents to have their aspirations around their expectations most of the time and therefore to experience a level of happiness higher than the midpoint of the range they can experience.

6 Conclusion

In the present study, we proposed an explanation of key features of our hedonic reward system as being adaptive. We consider our reward system as designed by evolution to induce us to make good decisions by ascribing higher subjective values to better options. Taking the point of view that evolution generates pressure toward the optimization of this reward system, we take a mechanism design approach. We ask what would be the features of an optimal reward function designed by Nature (she) to incentivize an agent (he) to make the best decisions possible and for the agent to fully reach his *potential* (his highest possible level of success in terms of fitness). Following previous contributions on this question (Robson, 2001*a*; Rayo and Becker, 2007; Netzer, 2009), we assume two constraints about subjective satisfaction: the agents can make mistakes when faced with options having close subjective values, and the subjective utility function is bounded upward and downward.

Under these two constraints, we retrieve the common result in the literature: the reward function is S-shaped and follows the distribution of the agent's potential as "known" by Nature (encoded in the reward function). The S-shape is optimal because, in order to minimize the agent's errors, the value function is steeper where it matters the most: where the agent's potential is most likely to be. We add to these two constraints a third one about Nature's limited ability to perfectly hard-wire the agent with mechanisms that calibrate his subjective utility function to the distribution of potentials he faces in any given situation. Nature can however incentivize the agent to use his private information about his likely potential in the context he faces to calibrate his utility function. By allowing the agent to choose his reference point given his information about his expected potential Nature can endow the agent with a more discerning utility function based on a more informative distribution about his potential. The utility function will be steeper near the mostly likely values of the agent's potential, on average reducing his loss.

To provide the agent with the incentives to accurately use his information when choosing his reference point, Nature needs to both reward him for having high expectations and provide utility functions that are asymmetric around the agent's expected potential. This asymmetry may reduces the function's discriminating ability at the decision stage, which in these cases reflects a necessary *agency cost* Nature faces as a principal (Laffont and Martimort, 2009).

We find that the utility functions that are part of the optimal utility system are S-shaped, but asymmetric with deviations below the agent's reference point being penalized more than deviations above it are rewarded. This pattern is what characterizes "loss aversion" in

Prospect Theory. The agent's reference point acts as an *aspirations level* or *goal*. Prior to making decisions, the agent experiences subjective satisfaction for setting a higher reference point. This *anticipatory utility* is necessary to compensate for the possible hedonistic benefits from having low aspirations.

Our results, therefore, provide a unifying explanation for several key behavioral patterns identified in behavioral economics: the fact that reference points are set at the level of the agent's expectations, the fact that the value function exhibits loss aversion, and the fact that people experience anticipatory utility prior to decisions. All these patterns appear not so much as irrational biases but as features of an optimal hedonic reward system designed to incentivize the agent to identify his potential and strive to reach it.

References

Abdellaoui, Mohammed, Han Bleichrodt, and Hilda Kammoun. 2013. "Do financial professionals behave according to prospect theory? An experimental study." *Theory and Decision*, 74: 411–429.

Allen, Eric J, Patricia M Dechow, Devin G Pope, and George Wu. 2017. "Reference-dependent preferences: Evidence from marathon runners." *Management Science*, 63(6): 1657–1672.

Anderson, Ashton, and Etan A Green. 2018. "Personal bests as reference points." *Proceedings of the National Academy of Sciences*, 115(8): 1772–1776.

Barberis, Nicholas C. 2013. "Thirty years of prospect theory in economics: A review and assessment." *Journal of Economic Perspectives*, 27(1): 173–96.

Bénabou, Roland, and Jean Tirole. 2016. "Mindful economics: The production, consumption, and value of beliefs." *Journal of Economic Perspectives*, 30(3): 141–64.

Benartzi, Shlomo, and Richard H Thaler. 1995. "Myopic loss aversion and the equity premium puzzle." *The Quarterly Journal of Economics*, 110(1): 73–92.

Bentham, Jeremy. 1789. "An introduction to the principles of morals." *London: Athlone.*

Binmore, Kenneth. 1994. *Game theory and the social contract: just playing.* Vol. 2, MIT press.

Bottan, Nicolas L, and Ricardo Perez-Truglia. 2022. "Choosing your pond: location choices and relative income." *Review of Economics and Statistics*, 104(5): 1010–1027.

Brickman, Philip. 1971. "Hedonic relativism and planning the good society." *Adaptation level theory*, 287–301.

Brunnermeier, Markus K, and Jonathan A Parker. 2005. "Optimal expectations." *American Economic Review*, 95(4): 1092–1118.

Camerer, Colin F, and George Loewenstein. 2003. "Behavioral economics: Past, present, future."

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor supply of New York City cabdrivers: One day at a time." *The Quarterly Journal of Economics*, 112(2): 407–441.

Caplin, Andrew, and John Leahy. 2001. "Psychological expected utility theory and anticipatory feelings." *The Quarterly Journal of Economics*, 116(1): 55–79.

Caplin, Andrew, and John V Leahy. 2019. "Wishful thinking." National Bureau of Economic Research.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. "Inequality at work: The effect of peer salaries on job satisfaction." *American Economic Review*, 102(6): 2981–3003.

Chapman, Jonathan, Erik Snowberg, Stephanie W Wang, and Colin Camerer. 2022. "Looming large or seeming small? Attitudes towards losses in a representative sample." National Bureau of Economic Research.

Chew, Soo Hong, Wei Huang, and Xiaojian Zhao. 2020. "Motivated false memory." *Journal of Political Economy*, 128(10): 3913–3939.

Clark, Andrew E, and Andrew J Oswald. 1996. "Satisfaction and comparison income." *Journal of public economics*, 61(3): 359–381.

Clark, Andrew E, Paul Frijters, and Michael A Shields. 2008. "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles." *Journal of Economic literature*, 46(1): 95–144.

Cummins, Robert A. 2013. "Subjective well-being, homeostatically protected mood and depression: A synthesis." *The exploration of happiness: present and future perspectives*, 77–95.

Diener, Ed, Richard E Lucas, and Christie Napa Scollon. 2009. "Beyond the hedonic treadmill: Revising the adaptation theory of well-being." *The science of well-being: The collected works of Ed Diener*, 103–118.

Ert, Eyal, and Ido Erev. 2013. "On the descriptive value of loss aversion in decisions under risk: Six clarifications." *Judgment and Decision Making*, 8(3): 214–235.

Farber, Henry S. 2005. "Is tomorrow another day? The labor supply of New York City cabdrivers." *Journal of political Economy*, 113(1): 46–82.

Frank, Robert H. 1985. *Choosing the right pond: Human behavior and the quest for status.* Oxford University Press.

Frederick, Shane, and George Loewenstein. 1999. "16 Hedonic adaptation." *D. Kahneman ED, N. Schwarz., Editors. Well-Being the Foundations of Hedonic Psychology*, 302–329.

Gal, David, and Derek D Rucker. 2018. "The loss of loss aversion: Will it loom larger than its gain?" *Journal of Consumer Psychology*, 28(3): 497–516.

Harris, Michael M, Frederik Anseel, and Filip Lievens. 2008. "Keeping up with the Joneses: A field study of the relationships among upward, lateral, and downward comparisons and pay level satisfaction." *Journal of Applied Psychology*, 93(3): 665.

Heath, Chip, Richard P Larrick, and George Wu. 1999. "Goals as reference points." *Cognitive psychology*, 38(1): 79–109.

Helson, Harry. 1948. "Adaptation-level as a basis for a quantitative theory of frames of reference." *Psychological review*, 55(6): 297.

Jevons, Herbert Stanley. 1905. *Essays on economics.* Macmillan and Company, limited.

Kahneman, Daniel. 2003. "Maps of bounded rationality: Psychology for behavioral economics." *American economic review*, 93(5): 1449–1475.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–292.

Kőszegi, Botond. 2010. "Utility from anticipation and personal equilibrium." *Economic Theory*, 44(3): 415–444.

Kőszegi, Botond, and Matthew Rabin. 2006. "A model of reference-dependent preferences." *The Quarterly Journal of Economics*, 121(4): 1133–1165.

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-dependent risk attitudes." *American Economic Review*, 97(4): 1047–1073.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn. 2011. "The effects of lottery prizes on winners and their neighbors: Evidence from the Dutch postcode lottery." *American Economic Review*, 101(5): 2226–2247.

Laffont, Jean-Jacques, and David Martimort. 2009. *The theory of incentives.* Princeton university press.

Laughlin, Simon. 1981. "A simple coding procedure enhances a neuron's information capacity." *Zeitschrift für Naturforschung c*, 36(9-10): 910–912.

Locke, Edwin A, and Gary P Latham. 1990. *A theory of goal setting & task performance.* Prentice-Hall, Inc.

Locke, Edwin A, and Gary P Latham. 2006. "New directions in goal-setting theory." *Current directions in psychological science*, 15(5): 265–268.

Loewenstein, George. 1987. “Anticipation and the valuation of delayed consumption.” *The Economic Journal*, 97(387): 666–684.

Loewenstein, George, and Jon Elster. 1992. “Utility from memory and anticipation.” *Choice over time*, 213–234.

Luhmann, Maike, and Sabrina Intelisano. 2018. “Hedonic adaptation and the set point for subjective well-being.” *Handbook of well-being*, 1–26.

Lykken, David, and Auke Tellegen. 1996. “Happiness is a stochastic phenomenon.” *Psychological science*, 7(3): 186–189.

Markle, Alex, George Wu, Rebecca White, and Aaron Sackett. 2018. “Goals as reference points in marathon running: A novel test of reference dependence.” *Journal of Risk and Uncertainty*, 56(1): 19–50.

Mrkva, Kellen, Eric J Johnson, Simon Gächter, and Andreas Herrmann. 2020. “Moderating loss aversion: Loss aversion has moderators, but reports of its death are greatly exaggerated.” *Journal of Consumer Psychology*, 30(3): 407–428.

Netzer, Nick. 2009. “Evolution of time preferences and attitudes toward risk.” *American Economic Review*, 99(3): 937–55.

Novemsky, Nathan, and Daniel Kahneman. 2005. “The boundaries of loss aversion.” *Journal of Marketing research*, 42(2): 119–128.

O’Donoghue, Ted, and Charles Sprenger. 2018. “Reference-dependent preferences.” In *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 1, 1–77. Elsevier.

Pesendorfer, Wolfgang. 2006. “Behavioral economics comes of age: A review essay on advances in behavioral economics.” *Journal of Economic Literature*, 44(3): 712–721.

Rayo, Luis, and Gary S Becker. 2007. “Evolutionary efficiency and happiness.” *Journal of Political Economy*, 115(2): 302–337.

Rizzo, John A, and Richard J Zeckhauser. 2003. “Reference incomes, loss aversion, and physician behavior.” *Review of Economics and Statistics*, 85(4): 909–922.

Robson, Arthur J. 2001*a*. “The biological basis of economic behavior.” *Journal of economic literature*, 39(1): 11–33.

Robson, Arthur J. 2001*b*. “Why would nature give individuals utility functions?” *Journal of Political Economy*, 109(4): 900–914.

Robson, Arthur J, and Larry Samuelson. 2011. “The evolutionary foundations of preferences.” *Handbook of social economics*, 1: 221–310.

Robson, Arthur J, Lorne A Whitehead, and Nikolaus Robalino. 2021. “Adaptive Cardinal Utility.” *Available at SSRN 4036252*.

Samuelson, Larry. 2004. "Information-based relative consumption effects." *Econometrica*, 72(1): 93–118.

Samuelson, Larry, and Jeroen M Swinkels. 2006. "Information, evolution and utility." *Theoretical Economics*, 1(1): 119–142.

Sarver, Todd. 2012. "Optimal reference points and anticipation." Discussion Paper.

Schultz, Wolfram. 2016. "Dopamine reward prediction-error signalling: a two-component response." *Nature reviews neuroscience*, 17(3): 183–195.

Schultz, Wolfram, Peter Dayan, and P Read Montague. 1997. "A neural substrate of prediction and reward." *Science*, 275(5306): 1593–1599.

Sweeny, Kate. 2018. "On the experience of awaiting uncertain news." *Current Directions in Psychological Science*, 27(4): 281–285.

Sweeny, Kate, and James A Shepperd. 2010. "The costs of optimism and the benefits of pessimism." *Emotion*, 10(5): 750.

Yechiam, Eldad. 2019. "Acceptable losses: The debatable origins of loss aversion." *Psychological research*, 83(7): 1327–1339.

Zeif, Dana, and Eldad Yechiam. 2022. "Loss aversion (simply) does not materialize for smaller losses." *Judgment & Decision Making*, 17(5).

Proof of Lemma 1

From Assumption 2, the agent randomly chooses between choices $x \in \mathbb{R}$ for which $v_{\mathcal{I}}(\bar{y}(s)) - v_{\mathcal{I}}(\phi(x,s)) \leq \varepsilon$. Define $\underline{y}(\bar{y})$ as the output which satisfies $v_{\mathcal{I}}(\bar{y}(s)) - v_{\mathcal{I}}(\underline{y}(\bar{y})) = \varepsilon$. Given potential $\bar{y}(s)$, the expected loss is $\mathbb{E}[\bar{y}(s) - \phi(s,x) | x \in X_{\varepsilon,v_{\mathcal{I}}}(s)]$ where $X_{\varepsilon,v_{\mathcal{I}}}(s)$ is the set of choices for which $\phi(x,s) = \bar{y}(s) - \alpha|x^*(s) - x| \geq \underline{y}(\bar{y}(s))$. Then this set is

$$X_{\varepsilon,v_{\mathcal{I}}}(s) = \left[x^*(s) - \frac{\bar{y} - \underline{y}(\bar{y})}{\alpha}, x^*(s) + \frac{\bar{y} - \underline{y}(\bar{y})}{\alpha}\right],$$

and $\mathbb{E}[\bar{y}(s) - \phi(s,x) | x \in X_{\varepsilon,v_{\mathcal{I}}}(s)] = \frac{\bar{y} - \underline{y}(\bar{y})}{2}$. It follows that the expected loss given the approximate density $\hat{f}$ is

$$L_{\varepsilon,\mathcal{I}}(v(y)) = \sum_{i=1}^{J} \mathbb{E}\left[\frac{\bar{y} - \underline{y}(\bar{y})}{2} \middle| \bar{y} \in I^i\right] \hat{f}^i_\omega = \sum_{i=1}^{J} \hat{f}^i_\omega \int_{y^{i-1}}^{y^i} \frac{\bar{y} - \underline{y}(\bar{y})}{2} d\bar{y}.$$

Proof of Lemma 2

Denote the expected loss associated with potential that lies on interval I^i as $\ell^i_{\varepsilon,\mathcal{I}}(\mathbf{v}) = \hat{f}^i \int_{y^{i-1}}^{y^i} \frac{\bar{y} - \underline{y}(\bar{y})}{2} d\bar{y}$. Then $L_{\varepsilon,\mathcal{I}}(\mathbf{v}) = \sum_{i=1}^{J} \ell^i_{\varepsilon,\mathcal{I}}(\mathbf{v})$. Let $\mathbf{v} \in [0,1]^{J+1}$ where $0 < \min\{v^i - v^{i-1} : i = 1, \ldots, J\}$ and $0 < \varepsilon < \min\{v^i - v^{i-1} : i = 1, \ldots, J\}$ be given. Then for each i, $\varepsilon \leq v^i - v^{i-1}$, and the loss associated with potential on interval I^i is

$$\begin{aligned}
\ell^i_{\varepsilon,\mathcal{I}}(\mathbf{v}) &= \hat{f}^i \left[\int_{y^{i-1} + \frac{\varepsilon}{v'(I^i)}}^{y^i} \frac{\bar{y} - \underline{y}(\bar{y})}{2} d\bar{y} + \int_{y^{i-1}}^{y^{i-1} + \frac{\varepsilon}{v'(I^i)}} \frac{\bar{y} - \underline{y}(\bar{y})}{2} d\bar{y}\right] \\
&= \frac{\hat{f}^i}{2} \left[\left(c - \frac{\varepsilon}{v'(I^i)}\right) \frac{\varepsilon}{v'(I^i)} + \int_{y^{i-1}}^{y^{i-1} + \frac{\varepsilon}{v'(I^i)}} \frac{\varepsilon}{v'(I^{i-1})} + (\bar{y} - y^{i-1}) \frac{v'(I^{i-1}) - v'(I^i)}{v'(I^{i-1})} d\bar{y}\right] \\
&= \frac{\hat{f}^i}{2} \left[\left(c - \frac{\varepsilon}{v'(I^i)}\right) \frac{\varepsilon}{v'(I^i)} + \frac{\varepsilon^2}{v'(I^i) v'(I^{i-1})} + \frac{\varepsilon^2}{2 v'(I^i)^2} \frac{v'(I^{i-1}) - v'(I^i)}{v'(I^{i-1})}\right]. \qquad (20)
\end{aligned}$$

The second line requires calculating $\underline{y}(\bar{y})$ in each of the sub-intervals of Figure 6. For $\bar{y} \in [y^{i-1} + \frac{\varepsilon}{v'(I^i)}, y^i]$, $\bar{y} - \underline{y}(\bar{y}) = \frac{\varepsilon}{v'(I^i)}$; for $\bar{y} \in [y^{i-1}, y^{i-1} + \frac{\varepsilon}{v'(I^i)}]$, $\bar{y} - \underline{y}(\bar{y}) = \frac{\varepsilon}{v'(I^{i-1})} + (\bar{y} - y^{i-1}) \frac{v'(I^{i-1}) - v'(I^i)}{v'(I^{i-1})}$. The sum of (20) for each interval I^i can be represented as (10).

Proof of Lemma 3

Let $\mathbf{v}^{-i}$ be given such that $0 < \min\{(v^{i+1} - v^{i-1})/2, v^j - v^{j-1} \text{ for } j \neq i, i+1\}$, and consider values of ε that are less than this minimum. Then the value of v^i can only impact the loss associated with potential on intervals I^i, I^{i+1} and I^{i+2}. The impact on the loss comes from the slopes of the utility function induced on the intervals $I^i = [y^{i-1}, y^i]$ and $I^{i+1} = [y^i, y^{i+1}]$, as illustrated by panel (a) on Figure 7. For $v^i \in (v^{i-1}, v^{i+1})$, there are three different cases to calculate the loss associated with these three intervals. These cases are represented on panels (b), (c) and (d) on Figure 7. Case 1 arises when $\varepsilon \leq \min\{v^i - v^{i-1}, v^{i+1} - v^i\}$, and

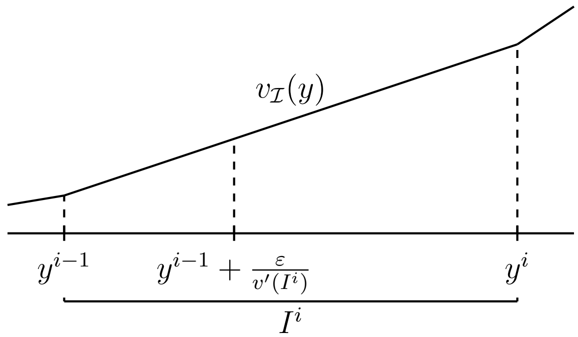


Figure 6: Sub-intervals of I^i.

cases 2 and 3 arise when $v_{\mathcal{I}}(y^i) - v_{\mathcal{I}}(y^{i-1}) < \varepsilon$ or $v_{\mathcal{I}}(y^{i+1}) - v_{\mathcal{I}}(y^i) < \varepsilon$ respectively. In what follows, we show that for sufficiently small ε, for any v^i where Cases 2 and 3 hold will always generate more loss than some value of v^i where Case 1 holds. As a result, the loss minimizing value of v^i given $\mathbf{v}^{-i}$, denoted by $\hat{v}^i(\mathbf{v}^{-i})$, must satisfy (11), the first order condition derived from Case 1.

For v^i where $\varepsilon > v^i - v^{i-1}$, then $I^i \subset [y^{i-1}, y^{i-1} + \frac{\varepsilon}{v'(I^i)}]$ and $\underline{y}(\bar{y}) \in I^{i-1}$ for all $\bar{y} \in I^i$. Then $\bar{y} - \underline{y}(\bar{y}) = \frac{\varepsilon}{v'(I^{i-1})} + (\bar{y} - y^{i-1})\frac{v'(I^{i-1}) - v'(I^i)}{v'(I^{i-1})}$. In this case the loss on I^i is

$$\ell^i_{\varepsilon,\mathcal{I}}(\mathbf{v}) = \hat{f}^i \left[\int_{y^{i-1}}^{y^i} \frac{\bar{y} - \underline{y}(\bar{y})}{2} d\bar{y}\right] = \frac{\hat{f}^i}{2}\left[\frac{c\varepsilon}{v'(I^{i-1})} + \frac{c^2}{2}\frac{v'(I^{i-1}) - v'(I^i)}{v'(I^{i-1})}\right] \geq \frac{\hat{f}^i c^2}{4}. \tag{21}$$

Similarly when $v^i > v^{i+1} - \varepsilon$, the loss on I^{i+1} is bounded below by $\frac{\hat{f}^{i+1}c^2}{2}$. Given $v^i = \frac{v^{i+1}+v^{i-1}}{2}$, $\varepsilon < \min\{v^{i+1} - v^i, v^i - v^{i-1}\}$ the losses on each of the three intervals impacted by v^i are described by (20), and the sum of these losses is

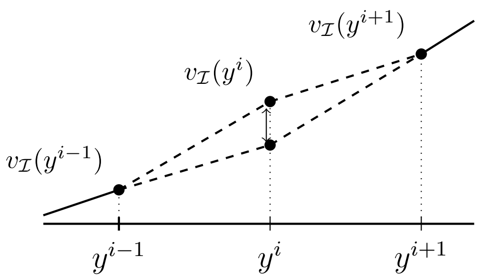
$$\sum_{j=i}^{i+2} \frac{\hat{f}^j}{2}\left[\frac{c^2\varepsilon}{v^j - v^{j-1}} + \frac{c^2\varepsilon^2(v^j - v^{j-2})}{2(v^j - v^{j-1})^2(v^{j-1} - v^{j-2})}\right]. \tag{22}$$

This sum is bounded above by $\frac{c^2\varepsilon}{\hat{\varepsilon}^i}\sum_{j=i}^{i+2}\hat{f}^j$, where $\hat{\varepsilon}^i = \min\{(v^{i+1} - v^{i-1})/2, v^j - v^{j-1}$ for $j \neq i, i+1\}$. Therefore, choosing $\varepsilon^i = \frac{\hat{\varepsilon}^i \min\{\hat{f}^i, \hat{f}^{i+1}\}}{4(\hat{f}^i + \hat{f}^{i+1} + \hat{f}^{i+2})}$ guarantees that for any $\varepsilon < \varepsilon^i$, this upper bound is less than the loss when v^i is such that $\varepsilon^i > \min\{v^i - v^{i-1}, v^{i+1} - v^i\}$. It follows that any loss minimizing choice, $\hat{v}^i(\mathbf{v}^{-i})$, is such that $\varepsilon^i \leq \min\{v^i - v^{i-1}, v^{i+1} - v^i\}$. The sum of the loss on intervals I^i, I^{i+1} and I^{i+2} is given by (22) and the first order condition for $\hat{v}^i(\mathbf{v}^{-i})$ is given by (11).

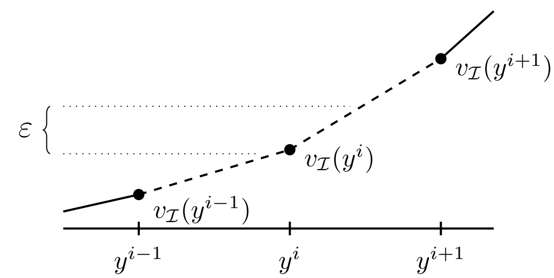
Proof of Proposition 4

For a given $\mathbf{v}^{-i}$, define $\hat{v}^i_{\mathcal{I}}(\mathbf{v}^{-i})$ for $i = 1, \ldots, J-1$ as the value of v^i which satisfies (12) and let $\hat{v}^0_{\mathcal{I}}(\mathbf{v}^{-0}) = 0$ and $\hat{v}^0_{\mathcal{I}}(\mathbf{v}^{-0}) = 1$. Denote the fixed point of the mapping $(v^0, \ldots, v^J) \to (\hat{v}^0_{\mathcal{I}}(\mathbf{v}^{-0}), \ldots, \hat{v}^J_{\mathcal{I}}(\mathbf{v}^{-J}))$ as $\mathbf{v}_{\mathcal{I}}$. From equation (12), for any $i = 1, \ldots, J$, $\frac{\left(v^{i+1}_{\mathcal{I}} - v^i_{\mathcal{I}}\right)^2}{\left(v^i_{\mathcal{I}} - v^{i-1}_{\mathcal{I}}\right)^2} = \frac{\hat{f}^{i+1}_\omega}{\hat{f}^i_\omega}$. Then for any $i, j = 1, \ldots J$,

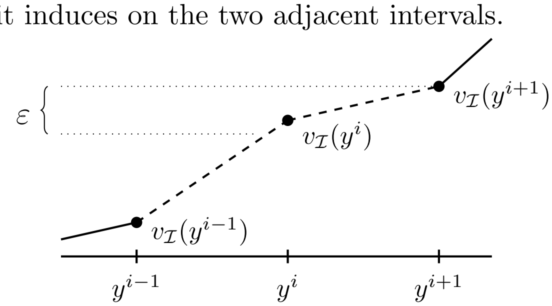
$$v^j_{\mathcal{I}} - v^{j-1}_{\mathcal{I}} = \left(\hat{f}^j_\omega \hat{f}^i_\omega\right)^{1/2}\left(v^i_{\mathcal{I}} - v^{i-1}_{\mathcal{I}}\right). \tag{23}$$



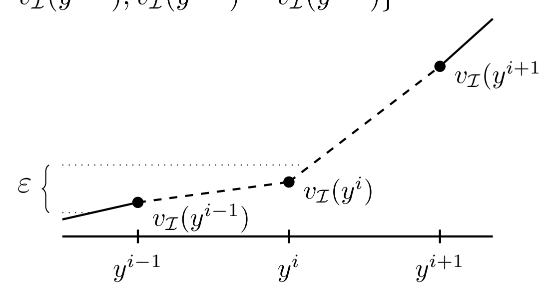
(a) The effect of changing the value of $v_{\mathcal{I}}(y^i)$ on the loss function stems from the different slopes it induces on the two adjacent intervals.



(b) Case 1: the randomization set can overlap with two intervals. $\varepsilon \leq \min\{v_{\mathcal{I}}(y^i) - v_{\mathcal{I}}(y^{i-1}), v_{\mathcal{I}}(y^{i-1}) - v_{\mathcal{I}}(y^{i-2})\}$



(c) Case 2: the randomization set always overlaps with two intervals. $v_{\mathcal{I}}(y^i) - v_{\mathcal{I}}(y^{i+1}) < \varepsilon < v_{\mathcal{I}}(y^i) - v_{\mathcal{I}}(y^{i-1})$



(d) Case 3: the randomization set can overlap with three intervals. $v_{\mathcal{I}}(y^{i-1}) - v_{\mathcal{I}}(y^i) < \varepsilon < v_{\mathcal{I}}(y^i) - v_{\mathcal{I}}(y^{i+1})$

Figure 7: The effect of the level of $v_{\mathcal{I}}(y^i)$ on the loss on the interval I^{i+1}. The figure illustrates that, for a given ε, the loss may depend on the slope on 1, 2, or 3 intervals.

Adding (23) for all $j = 1, \ldots, J$ we have $\sum_{j=1}^{J} v_{\mathcal{I}}^{j} - v_{\mathcal{I}}^{j-1} = \sum_{j=1}^{J} \left(\hat{f}_{\omega}^{j} \hat{f}_{\omega}^{i}\right)^{1/2} \left(v_{\mathcal{I}}^{i} - v_{\mathcal{I}}^{i-1}\right)$. Given $v_{\mathcal{I}}^{0} = 0$ and $v_{\mathcal{I}}^{J} = 1$, $\sum_{j=1}^{J} v_{\mathcal{I}}^{j} - v_{\mathcal{I}}^{j-1} = 1$ and then for $i = 1, \ldots, J$, $v_{\mathcal{I}}^{i} - v_{\mathcal{I}}^{i-1} = \frac{(\hat{f}_{\omega}^{i})^{1/2}}{\sum_{j=1}^{J} (\hat{f}_{\omega}^{j})^{1/2}}$ and $v_{\mathcal{I}}^{i} = \frac{\sum_{j=1}^{i} (\hat{f}_{\omega}^{j})^{1/2}}{\sum_{j=1}^{J} (\hat{f}_{\omega}^{j})^{1/2}}$. Using Equation (8), $v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ is characterized by

$$v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}}) = \left(c \sum_{j=1}^{i-1} \left(\hat{f}_{\omega}^{j}\right)^{1/2} + (y - y^{i-1}) \left(\hat{f}_{\omega}^{i}\right)^{1/2}\right) \Big/ \left(c \sum_{j=1}^{J} \left(\hat{f}_{\omega}^{j}\right)^{1/2}\right). \tag{24}$$

For given $\mathcal{I}$, recall $a = \frac{Jc}{2} > 0$, $y^0 = -a$ and $y^I = a$. Because $\hat{f}^i$ is a value taken by $f(y)$ on I^i, $v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ is a ratio of Riemann sums of $f(y)^{1/2}$ on $[-a, y]$ and $[-a, a]$ over the partition $\mathcal{I}$. Because $f(y)^{1/2}$ is Riemann-integrable over any $[-a, a]$, then by the definition of the Riemann integral

$$\lim_{J \to \infty} v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}}) = K_a \int_{-a}^{y} f(z)^{1/2} dz, \quad \text{where } K_a = \left(\int_{-a}^{a} f(z)^{1/2} dz\right)^{-1}.$$

Using the definition of the improper integral and defining $v^*(y) \equiv \lim_{a \to \infty} \lim_{J \to \infty} v_{\mathcal{I}}(y; \mathbf{v}_{\mathcal{I}})$ gives the desired result.

Proof of Proposition 5

Consider ω_n and ω_m with $n < m$. We first show that the symmetry of the first best solution for the value function, $v^*(y)$, implies that the absolute value of the expected hedonic loss from optimism, $\mathbb{E}[v_{\omega_n}^*(y) - v_{\omega_m}^*(y)|\omega_n]$, is equal to the absolute value of the expected gain from pessimism $\mathbb{E}[v_{\omega_n}^*(y) - v_{\omega_m}^*(y)|\omega_m]$.

From the assumptions on f_{ω_n}, $f_{\omega_n}(y) = f_{\omega_n}(\bar{y}_{\omega_n}^r - (y - \bar{y}_{\omega_n}^r))$ and $f_{\omega_n}(y) = f_{\omega_m}(y + (\bar{y}_{\omega_m}^r - \bar{y}_{\omega_n}^r))$. From Proposition 4, the optimal hedonic utility function $v_{\omega_n}^*$ is S-shaped and oddly-symmetric about a reference point $\bar{y}_{\omega_n}^r$. Therefore, $v_{\omega_n}^*(y) = 1 - v_{\omega_n}^*(\bar{y}_{\omega_n}^r - (y - \bar{y}_{\omega_n}^r))$ and $v_{\omega_n}^*(y) = v_{\omega_m}^*(y + (\bar{y}_{\omega_m}^r - \bar{y}_{\omega_n}^r))$. Define $\Delta v_{nm}^*(y) \equiv v_{\omega_n}^*(y) - v_{\omega_m}^*(y)$.

Using the properties above, it follows that for any $n < m$

$$\begin{aligned}
\mathbb{E}[\Delta v_{nm}^*(y)|\omega_n] &= \int_{-\infty}^{\infty} [\Delta v_{nm}^*(y)] f_{\omega_n}(y) dy \\
&= \int_{-\infty}^{\infty} [v_{\omega_m}(\bar{y}_{\omega_m}^r - (y - \bar{y}_{\omega_m}^r)) - v_{\omega_n}(\bar{y}_{\omega_n}^r - (y - \bar{y}_{\omega_n}^r))] f_{\omega_m}(\bar{y}_{\omega_n}^r + \bar{y}_{\omega_m}^r - y) dy \\
&= \int_{-\infty}^{\infty} [v_{\omega_n}(\bar{y}_{\omega_n}^r - (y - \bar{y}_{\omega_m}^r)) - v_{\omega_m}(\bar{y}_{\omega_m}^r - (y - \bar{y}_{\omega_n}^r))] f_{\omega_m}(\bar{y}_{\omega_n}^r + \bar{y}_{\omega_m}^r - y) dy \\
&= \int_{-\infty}^{\infty} \Delta v_{nm}^*(\bar{y}_{\omega_m}^r + \bar{y}_{\omega_m}^r - y) f_{\omega_m}(\bar{y}_{\omega_n}^r + \bar{y}_{\omega_m}^r - y) dy = \mathbb{E}[\Delta v_{nm}^*(y)|\omega_m]
\end{aligned}$$

Next we show that for any $\ell < n < m$, $\mathbb{E}[\Delta v_{nm}^*(y)|\omega_n] - \mathbb{E}[\Delta v_{nm}^*(y)|\omega_\ell] > 0$. From the properties on the densities, for $\tilde{y} = \frac{1}{2}(\bar{y}_{\omega_n}^r + \bar{y}_{\omega_\ell}^r)$,

$$\begin{aligned}
f_{\omega_n}(\tilde{y} + y) &= f_{\omega_\ell}(\tilde{y} + y - (\bar{y}_{\omega_n}^r - \bar{y}_{\omega_\ell}^r)) = f_{\omega_\ell}(\bar{y}_{\omega_\ell}^r - (\tilde{y} + y - (\bar{y}_{\omega_n}^r - \bar{y}_{\omega_\ell}^r) - \bar{y}_{\omega_\ell}^r)) \\
&= f_{\omega_\ell}\left(\frac{\bar{y}_{\omega_\ell}^r}{2} + \frac{\bar{y}_{\omega_n}^r}{2} - y\right) = f_{\omega_\ell}(\tilde{y} - y).
\end{aligned}$$

Then $f_{\omega_n}(\tilde{y}+y)-f_{\omega_\ell}(\tilde{y}+y) = -[f_{\omega_n}(\tilde{y}-y)-f_{\omega_\ell}(\tilde{y}-y)]$. From Proposition 4, the difference in value functions is $\Delta v^*_{nm}(y) = K\int_{-\infty}^{y} f_{\omega_n}(z)^{1/2} - f_{\omega_m}(z)^{1/2}dz$ where $f_\omega(y)^{1/2}$ is single peaked and symmetric about $\bar{y}^r_\omega$ for $\omega=\omega_n,\omega_m$. Given $\bar{y}^r_{\omega_n} < \bar{y}^r_{\omega_m}$, then $\Delta v^*_{nm}(y) \equiv v^*_{\omega_n}(y) - v^*_{\omega_m}(y)$ is single peaked and symmetric about $\hat{y} = \frac{1}{2}(\bar{y}^r_{\omega_m} + \bar{y}^r_{\omega_n})$. For any $|y_1| > |y_2|$, $\Delta v^*_{nm}(\hat{y}+y_1) = \Delta v^*_{nm}(\hat{y}-y_1) < \Delta v^*_{nm}(\hat{y}+y_2) = \Delta v^*_{nm}(\hat{y}-y_2)$.

Given that $\hat{y} > \tilde{y} > 0$, for any $y > 0$

$$\Delta v^*_{nm}(\tilde{y}+y) = \Delta v^*_{nm}(\hat{y}+y-(\hat{y}-\tilde{y})) > \Delta v^*_{nm}(\hat{y}+y+(\hat{y}-\tilde{y})) = \Delta v^*_{nm}(\tilde{y}-y)$$

where the inequality follows from $|y+\hat{y}-\tilde{y}| > |y-(\hat{y}-\tilde{y})|$ and the last equality comes from the symmetry of $\Delta v^*_{nm}(y)$ which implies that $\Delta v^*_{nm}(\hat{y}+y+(\hat{y}-\tilde{y})) = \Delta v^*_{nm}(\hat{y}-y-(\hat{y}-\tilde{y}))$. The difference in expected differences is

$$\begin{aligned}\mathbb{E}[\Delta v^*_{nm}(y)|\omega_n] - \mathbb{E}[\Delta v^*_{nm}(y)|\omega_\ell] &= \int_{-\infty}^{\infty} \Delta v^*_{nm}(y)(f_{\omega_n}(y) - f_{\omega_\ell}(y))dy \\ &= \int_{-\infty}^{\tilde{y}} \Delta v^*_{nm}(y)(f_{\omega_n}(y) - f_{\omega_\ell}(y))dy + \int_{\tilde{y}}^{\infty} \Delta v^*_{nm}(y)(f_{\omega_n}(y) - f_{\omega_\ell}(y))dy \\ &= \int_{0}^{\infty} (\Delta v^*_{nm}(\tilde{y}+y) - \Delta v^*_{nm}(\tilde{y}-y))(f_{\omega_n}(\tilde{y}+y) - f_{\omega_\ell}(\tilde{y}+y))dy > 0\end{aligned}$$

A similar argument shows that $n < m < \ell$, $\mathbb{E}[\Delta v^*_{nm}(y)|\omega_m] - \mathbb{E}[\Delta v^*_{nm}(y)|\omega_\ell] > 0$.

Using these conditions, we can complete the proof Proposition 5. For any $n < m$, $u(\omega_n) + \mathbb{E}[v^*_{\omega_n}(y)|\omega_n] \geq u(\omega_m) + \mathbb{E}[v^*_{\omega_m}(y)|\omega_n]$ implies $\mathbb{E}[\Delta v^*_{nm}(y)|\omega_m] \leq u(\omega_m) - u(\omega_n) \leq \mathbb{E}[\Delta v^*_{nm}(y)|\omega_n]$. Because the two bounds must be equal, then the difference must equal each bound: $u(\omega_m) - u(\omega_n) = \mathbb{E}[\Delta v^*_{nm}(y)|\omega_m] = \mathbb{E}[\Delta v^*_{nm}(y)|\omega_n]$.

Now take three states $\ell < n < m$ and assume that $u(\omega_\ell), u(\omega_n), u(\omega_m)$ satisfy the incentive constraints. The above conditions imply that $u(\omega_m) - u(\omega_\ell) = \mathbb{E}[\Delta v^*_{\ell m}(y)|\omega_\ell] = \mathbb{E}[\Delta v^*_{\ell n}(y)|\omega_\ell] + \mathbb{E}[\Delta v^*_{nm}(y)|\omega_\ell]$ and $u(\omega_m) - u(\omega_\ell) = \mathbb{E}[\Delta v^*_{\ell n}(y)|\omega_\ell] + \mathbb{E}[\Delta v^*_{nm}(y)|\omega_n]$. However, $\mathbb{E}[\Delta v^*_{nm}(y)|\omega_n] > \mathbb{E}[\Delta v^*_{nm}(y)|\omega_\ell]$ gives a contradiction.

Proof of Lemma 6

Let ε_k be a sequence of such that $\varepsilon_k \to 0$ as $k \to \infty$. Let $\lambda_{n,m}(\varepsilon_k)$ be a sequence where $\lim_{\varepsilon\to 0} \frac{\lambda_{n,m}(\varepsilon_k)}{\varepsilon_k} = \lambda_{n,m} \geq 0$. Similarly, let $\gamma^i_{n,k}$ be a sequence where $\lim_{\varepsilon\to 0} \frac{\gamma^i_{n,k}(\varepsilon_k)}{\varepsilon_k} = \gamma^i_{n,k} \geq 0$. Then for each k, (14) is given by

$$\frac{\partial L_{\varepsilon_k,\mathcal{I}}(\mathbf{V})}{\partial v^i_n} = \sum_{m=1}^{N} \lambda_{n,m}(\varepsilon_k)\frac{\partial h_{n,m}(\mathbf{v}_n,\mathbf{v}_m)}{\partial v^i_n} + \sum_{m=1}^{N} \lambda_{m,n}(\varepsilon_k)\frac{\partial h_{m,n}(\mathbf{v}_n,\mathbf{v}_m)}{\partial v^i_n} + \gamma^i_{n,0}(\varepsilon_k) - \gamma^i_{n,1}(\varepsilon_k).$$

Relative to (11) the first two additional terms can be calculated as

$$\begin{aligned}\frac{\partial h_{n,m}(\mathbf{v})}{\partial v^i_n} &= \frac{\partial}{\partial v^i_n}\mathbb{E}[v_n(y) - v_m(y)|\omega_n] = \frac{\partial}{\partial v^i_n}\int_{y^{i-1}+\frac{c\varepsilon}{v^i_n - v^{i-1}_n}}^{y^i} \mathbb{E}[v_n(y)|\bar{y}]\hat{f}^i_{\omega_n}d\bar{y} \\ &+\frac{\partial}{\partial v^i_n}\int_{y^i+\frac{c\varepsilon}{v^{i+1}_n - v^i_n}}^{y^{i+1}} \mathbb{E}[v_n(y)|\bar{y}]\hat{f}^{i+1}_{\omega_n}d\bar{y} + \frac{\partial}{\partial v^i_n}\sum_{j=i}^{i+2}\int_{y^{j-1}}^{y^{j-1}+\frac{c\varepsilon}{v^j_n - v^{j-1}_n}} \mathbb{E}[v_n(y)|\bar{y}]\hat{f}^j_{\omega_n}d\bar{y}\end{aligned}$$

When $\bar{y}$ and $\underline{y}$ are in the same interval, $\mathbb{E}[v_n(y)|\bar{y}] = \frac{v_n(\bar{y})+v_n(\underline{y})}{2} = v_{\omega_n}(\bar{y}) - \frac{\varepsilon}{2} = v_n^{i-1} + (\bar{y} - y^{i-1})\frac{v_n^i - v_n^{i-1}}{c} - \frac{\varepsilon}{2}$ for $\bar{y} \in I^i$. Then, using Leibniz rule, the first two terms are

$$\frac{\partial}{\partial v_n^i}\int_{y^{i-1}+\frac{c\varepsilon}{v_n^i - v_n^{i-1}}}^{y^i}\left(v_n^{i-1} + (\bar{y} - y^{i-1})\frac{v_n^i - v_n^{i-1}}{c} - \frac{\varepsilon}{2}\right)\hat{f}_{\omega_n}^i d\bar{y} = \frac{\hat{f}_{\omega_n}^i c}{2} + \varepsilon\frac{\hat{f}_{\omega_n}^i c v_n^{i-1}}{(v_n^i - v_n^{i-1})^2}$$

$$\frac{\partial}{\partial v_n^i}\int_{y^i+\frac{c\varepsilon}{v_n^{i+1} - v_n^i}}^{y^{i+1}}\left(v_n^i + (\bar{y} - y^{i-1})\frac{v_n^{i+1} - v_n^i}{c} - \frac{\varepsilon}{2}\right)\hat{f}_{\omega_n}^i d\bar{y} = \frac{\hat{f}_{\omega_n}^{i+1} c}{2} - \varepsilon\frac{\hat{f}_{\omega_n}^{i+1} c v_n^{i+1}}{(v_n^{i+1} - v_n^i)^2}$$

These two terms approach $\frac{c}{2}(\hat{f}_{\omega_n}^i + \hat{f}_{\omega_n}^{i+1})$ as $\varepsilon \to 0$.

For the three terms in the summation, $\underline{y}$ is in the interval to the left of $\bar{y}$. For $\bar{y}$ in this sub-interval of I^i, $\underline{y} \in I^{i-1}$ and $\mathbb{E}[v_n(y)|\bar{y}] = \frac{\bar{y}-y^{i-1}}{\bar{y}-\underline{y}}\frac{v_n(\bar{y})+v_n(y^{i-1})}{2} + \frac{y^{i-1}-\underline{y}}{\bar{y}-\underline{y}}\frac{v_n(y^{i-1})+v_n(\underline{y})}{2}$. Given $v_n(\bar{y}) - v_n(\underline{y}) = \varepsilon$, then $\mathbb{E}[v_n(y)|\bar{y}] = \frac{v_n(\bar{y})+v_n(y^{i-1})}{2} - \frac{\varepsilon}{2}\frac{y^{i-1}-\underline{y}}{\bar{y}-\underline{y}}$. From the calculation in Lemma 3 $\bar{y} - \underline{y} = \frac{\varepsilon}{v'(I^{i-1})} + (\bar{y} - y^{i-1})\frac{v'(I^{i-1})-v'(I^i)}{v'(I^{i-1})}$. Each term in the summation vanishes as $\varepsilon \to 0$.

The second additional term $\frac{\partial h_{m,n}(\mathbf{v})}{\partial v_n^i} = \frac{\partial}{\partial v_n^i}\mathbb{E}[v_m(y) - v_n(y)|\omega_m]$ differs from the first in two ways. First, terms on interval i are multiplied by $\hat{f}_m^i$ rather than f_n^i as the expectation is given ω_m rather than ω_n. And second everything is multiplied by negative 1. The first two terms are $-\frac{\hat{f}_{\omega_m}^i c}{2} - \varepsilon\frac{\hat{f}_{\omega_m}^i c v_n^{i-1}}{(v_n^i - v_n^{i-1})^2}$ and $-\frac{\hat{f}_{\omega_m}^{i+1} c}{2} + \varepsilon\frac{\hat{f}_{\omega_m}^{i+1} c v_n^{i+1}}{(v_n^{i+1} - v_n^i)^2}$ and they approach $\frac{-c}{2}(\hat{f}_{\omega_m}^i + \hat{f}_{\omega_m}^{i+1})$ as $\varepsilon \to 0$. All remaining terms again approach 0 as $\varepsilon \to 0$.

Therefore, for $i = 1, \ldots, J-1$, these additional terms can be written as

$$\sum_{m=1}^{N}\lambda_{n,m}(\varepsilon_k)\frac{\partial h_{n,m}(\mathbf{v}_n, \mathbf{v}_m)}{\partial v_n^i} = \sum_{m=1}^{N}\lambda_{n,m}\left(\frac{(\hat{f}_{\omega_n}^i + \hat{f}_{\omega_n}^{i+1})c^2\varepsilon_k}{2} + \mathcal{O}(\varepsilon_k)\right).$$

For $i = 1, \ldots, J-1$, let and increasing $\mathbf{v}_n^{-i}$ be given. Following the arguments in the proof of Lemma 3, for sufficiently small ε_k, any v_n^i such that $\varepsilon_k > \min\{v_n^{i+1} - v_n^i, v_n^i - v_n^{i-1}\}$ will lead to more loss than for some $v_n^i \in [v_n^{i-1} + \varepsilon_k, v_n^{i+1} - \varepsilon_k]$. Therefore it is sufficient to show that there is a $v_n^i \in [v_n^{i-1} + \varepsilon_k, v_n^{i+1} - \varepsilon_k]$ which satisfies 14. For these values of v_n^i, $\gamma_{n,0}^i = \gamma_{n,1}^i = 0$, and this condition is given by

$$\frac{p_{\omega_n}\hat{f}_{\omega_n}^{i+1}c^2\varepsilon_k}{\left(v_n^{i+1} - \hat{v}_n^i(\mathbf{v}_n^{-i}, \lambda, \gamma)\right)^2} - \frac{p_{\omega_n}\hat{f}_{\omega_n}^i c^2\varepsilon_k}{\left(\hat{v}_n^i(\mathbf{v}_n^{-i}, \lambda, \gamma) - v_n^{i-1}\right)^2} + \mathcal{O}(\varepsilon_k) - \sum_{m=1}^{N}\lambda_{n,m}\frac{(\hat{f}_{\omega_n}^i + \hat{f}_{\omega_n}^{i+1})c^2\varepsilon_k}{2}$$
$$+ \sum_{m=1}^{N}\lambda_{m,n}\frac{(\hat{f}_{\omega_m}^i + \hat{f}_{\omega_m}^{i+1})c^2\varepsilon_k}{2} = 0.$$

Dividing through by ε_k, it is clear that the final three terms are bounded for all ε_k, the first term gets arbitrarily positive for $v_n^i = v_n^{i+1} - \varepsilon_k$ as $\varepsilon_k \to 0$, and the second term gets arbitrarily negative for $v_n^i = v_n^{i+1} - \varepsilon_k$ as $\varepsilon_k \to 0$. All terms are continuous in v_n^i. Then for sufficiently small ε_k there is a $v_n^i \in [v_n^{i-1} + \varepsilon_k, v_n^{i+1} - \varepsilon_k]$ which satisfies this condition and therefore satisfies (14).

For $i = 0$, the additional terms in (14) can be written as

$$\sum_{m=1}^{N} \lambda_{n,m}(\varepsilon_k) \frac{\partial h_{n,m}(\mathbf{v}_n, \mathbf{v}_m)}{\partial v_n^i} = \sum_{m=1}^{N} \lambda_{n,m} \left(\frac{\hat{f}^1_{\omega_n} c^2 \varepsilon_k}{2} + \mathcal{o}(\varepsilon_k) \right).$$

Letting an increasing $\mathbf{v}_n^{-0}$ be given, then again for sufficiently small ε_k, $v_n^0 > v_n^1 - \varepsilon_k$ will lead to more loss than for $v_n^0 \in [0, v_n^1 - \varepsilon]$. For these values of v_n^1, $\gamma^0_{n,1} = 0$ and (14) is

$$\frac{p_{\omega_n} \hat{f}^1_{\omega_n} c^2 \varepsilon_k}{(v_n^1 - \hat{v}_n^0(\mathbf{v}_n^{-0}, \lambda, \gamma))^2} + \mathcal{o}(\varepsilon_k) - \sum_{m=1}^{N} \lambda_{n,m} \frac{\hat{f}^1_{\omega_n} c^2 \varepsilon_k}{2} + \sum_{m=1}^{N} \lambda_{m,n} \frac{\hat{f}^1_{\omega_m} c^2 \varepsilon_k}{2} - \varepsilon_k \gamma^0_{n,0} = 0.$$

Dividing through by ε_k, for sufficiently large $\gamma^0_{n,0}$, there is a v_n^0 that satisfies this condition.

For $i = J$, the additional terms in (14) can be written as

$$\sum_{m=1}^{N} \lambda_{n,m}(\varepsilon_k) \frac{\partial h_{n,m}(\mathbf{v}_n, \mathbf{v}_m)}{\partial v_n^i} = \sum_{m=1}^{N} \lambda_{n,m} \left(\frac{\hat{f}^J_{\omega_n} c^2 \varepsilon_k}{2} + \mathcal{o}(\varepsilon_k) \right).$$

Letting an increasing $\mathbf{v}_n^J$ be given, for sufficiently small ε_k, $v_n^J < v_n^{J-1} + \varepsilon_k$ will lead to more loss that for $v_n^J \in [v_n^{J-1} + \varepsilon_k, 1]$. For these values of v_n^J, $\gamma^J_{n,0} = 0$ and (14) is

$$\frac{p_{\omega_n} \hat{f}^J_{\omega_n} c^2 \varepsilon_k}{(\hat{v}_n^J(v^{-J}, \lambda, \gamma) - v_n^{J-1})^2} + \mathcal{o}(\varepsilon_k) - \sum_{m=1}^{N} \lambda_{n,m} \frac{\hat{f}^J_{\omega_n} c^2 \varepsilon}{2} + \sum_{m=1}^{N} \lambda_{m,n} \frac{\hat{f}^J_{\omega_m} c^2 \varepsilon_k}{2} - \varepsilon_k \gamma^J_{n,1} = 0.$$

Dividing through by ε_k, for sufficiently large $\gamma^J_{n,1}$, there is a v_n^J that satisfies this condition.

Proof of Lemma 7

For a given $\mathbf{v}_n^{-i}$ define $\hat{v}^i_{n,\mathcal{I}}(\mathbf{v}_n^{-i}, \lambda, \gamma)$ as the value which satisfies (16) as $\varepsilon \to 0$. This is

$$\frac{p_{\omega_n} c^2 \hat{f}^i_{\omega_n}}{\left(\hat{v}^i_{n,\mathcal{I}}(\mathbf{v}_n^{-i}, \lambda, \gamma) - v_n^{i-1}\right)^2} - \frac{p_{\omega_n} c^2 \hat{f}^{i+1}_{\omega_n}}{\left(v_n^{i+1} - \hat{v}^i_{n,\mathcal{I}}(\mathbf{v}_n^{-i}, \lambda, \gamma)\right)^2} + \sum_{m=1}^{N} \left(\hat{f}^i_{\omega_n} + \hat{f}^{i+1}_{\omega_n}\right) \frac{c\lambda_{n,m}}{2}$$
$$- \sum_{m=1}^{N} \left(\hat{f}^i_{\omega_m} + \hat{f}^{i+1}_{\omega_m}\right) \frac{c\lambda_{m,n}}{2} = 0. \tag{25}$$

Similarly define $\hat{v}^0_{n,\mathcal{I}}(\mathbf{v}_n^{-0}, \lambda, \gamma)$ and $\hat{v}^J_{n,\mathcal{I}}(\mathbf{v}_n^{-J}, \lambda, \gamma)$ as the values which respectively satisfy

$$-\frac{p_{\omega_n} c^2 \hat{f}^1_{\omega_n}}{\left(v_n^1 - \hat{v}^0_{n,\mathcal{I}}(\mathbf{v}_n^{-0}, \lambda, \gamma)\right)^2} + \sum_{m=1}^{N} \frac{c\lambda_{n,m}}{2} \hat{f}^1_{\omega_n} - \sum_{m=1}^{N} \frac{c\lambda_{m,n}}{2} \hat{f}^1_{\omega_m} + \gamma^0_{n,0} = 0, \text{ and} \tag{26}$$

$$\frac{p_{\omega_n} c^2 \hat{f}^J_{\omega_n}}{\left(\hat{v}^J_{n,\mathcal{I}}(\mathbf{v}_n^{-J}, \lambda, \gamma) - v_n^{J-1})\right)^2} + \sum_{m=1}^{N} \frac{c\lambda_{n,m}}{2} \hat{f}^J_{\omega_n} - \sum_{m=1}^{N} \frac{c\lambda_{m,n}}{2} \hat{f}^J_{\omega_m} - \gamma^J_{n,1} = 0. \tag{27}$$

Denote the ratio of the slope of $v_{\mathcal{I}}(y; v_n)$ and the square root of the density of potential on I^i as:

$$D^i_{\omega_n} = c \left(\hat{f}^i_{\omega_n}\right)^{1/2} / \left(v_n^i - v_n^{i-1}\right).$$

Then from (25), the difference in the square of the ratios between interval $i+1$ and i must satisfy

$$\left(D^{i+1}_{\omega_n}\right)^2 - \left(D^{i}_{\omega_n}\right)^2 = \sum_{m=1}^{N} \frac{c\lambda_{n,m}}{2p_{\omega_n}} \left(\hat{f}^{i}_{\omega_n} + \hat{f}^{i+1}_{\omega_n}\right) - \sum_{m=1}^{N} \frac{c\lambda_{m,n}}{2p_{\omega_n}} \left(\hat{f}^{i}_{\omega_m} + \hat{f}^{i+1}_{\omega_m}\right)$$
$$= \sum_{m=1}^{N} \frac{c\lambda_{m,n}}{p_{\omega_n}} \left[\frac{1}{2}\left(\hat{f}^{i}_{\omega_n} + \hat{f}^{i+1}_{\omega_n}\right) - \frac{1}{2}\left(\hat{f}^{i}_{\omega_m} + \hat{f}^{i+1}_{\omega_m}\right)\right],$$

where the last equality follows from (15). From (26) the squared ratio on the first interval satisfies

$$\left(D^{1}_{\omega_n}\right)^2 = \sum_{m=1}^{N} \frac{c\lambda_{m,n}}{2p_{\omega_n}} \left(\hat{f}^{1}_{\omega_n} - \hat{f}^{1}_{\omega_m}\right) + \frac{\gamma^0_{n,0}}{p_{\omega_n}}. \tag{28}$$

Then from equations (28) and (25) we can find the value of the ratio on each interval.

$$\left(D^{i}_{\omega_n}\right)^2 = \left(D^{1}_{\omega_n}\right)^2 + \sum_{j=1}^{i-1} \left(D^{j+1}_{\omega_n}\right)^2 - \left(D^{j}_{\omega_n}\right)^2$$
$$= \sum_{m=1}^{N} \frac{c\lambda_{m,n}}{p_{\omega_n}} \left[\frac{1}{2}\left(\hat{f}^{i}_{\omega_n} - \hat{f}^{i}_{\omega_m}\right) + \sum_{j=1}^{i-1}\left(\hat{f}^{j}_{\omega_n} - \hat{f}^{j}_{\omega_m}\right)\right] + \frac{\gamma^0_{n,0}}{p_{\omega_n}}.$$

We now characterize the utility function derived from the vector that satisfies the conditions of Lemma 6 as $\varepsilon \to 0$.. Define $R^i_{\omega_n} \equiv \frac{v^i_n - v^{i-1}_n}{c} = \frac{\left(\hat{f}^i_{\omega_n}\right)^{1/2}}{D^i_{\omega_n}}$. Then the sum of these differences can be expressed as:

$$\sum_{j=1}^{J} v^j_{n,\mathcal{I}} - v^{j-1}_{n,\mathcal{I}} = \left(v^i_n - v^{i-1}_{n,\mathcal{I}}\right) \sum_{j=1}^{J} \frac{R^j_{\omega_n}}{R^i_{\omega_n}} = v^J_{n,\mathcal{I}} - v^0_{n,\mathcal{I}}.$$

Therefore $\frac{v^i_{n,\mathcal{I}} - v^{i-1}_{n,\mathcal{I}}}{v^J_{n,\mathcal{I}} - v^0_{n,\mathcal{I}}} = \frac{R^i_{\omega_n}}{\sum_{j=1}^{J} R^j_{\omega_n}}$ and $v^i_{n,\mathcal{I}} = v^0_{n,\mathcal{I}} + (v^J_{n,\mathcal{I}} - v^0_{n,\mathcal{I}})\frac{\sum_{j=1}^{i} R^j_{\omega_n}}{\sum_{j=1}^{J} R^j_{\omega_n}}$. Using (8), for $y \in I^i$ the utility function is $v_{\mathcal{I}}(y; \mathbf{v}_{n,\mathcal{I}}) = v^0_{n,\mathcal{I}} + (v^J_{n,\mathcal{I}} - v^0_{n,\mathcal{I}})\frac{c\sum_{j=1}^{i-1} R^j_{\omega_n} + (y - y^{i-1})R^i_{\omega_n}}{c\sum_{j=1}^{J} R^j_{\omega_n}}$.

Proof of Proposition 8

Using the characterization from Lemma 7, we now find the limit of $v_{\mathcal{I}}(y; \mathbf{v}_{n,\mathcal{I}})$ as the number of intervals increases and the width of the intervals shrinks to zero. Define $D_{\omega_n,\mathcal{I}}(y) = \sum_{i=1}^{J} \mathbf{1}_{\mathcal{I}^i} D^i_{\omega_n}$. For a fixed a, $D_{\omega_n,\mathcal{I}}(y)$ is contained in the $\frac{J(y-a)}{2a} - th$ interval and

$$\lim_{J\to\infty} \left(D_{\omega_n,\mathcal{I}}(y)\right)^2 = \lim_{J\to\infty} \sum_{m=1}^{M} \frac{\lambda_{m,n}}{p_{\omega_n}} \left[\sum_{j=1}^{J(y-a)/2a} \frac{c}{2}\left(\hat{f}^{j}_{\omega_n} + \hat{f}^{j-1}_{\omega_n}\right) - \frac{c}{2}\left(\hat{f}^{j}_{\omega_m} + \hat{f}^{j-1}_{\omega_m}\right)\right] + \frac{\gamma^{-a}_{n,0}}{p_{\omega_n}}$$
$$= \sum_{m=1}^{M} \frac{\lambda_{m,n}}{p_{\omega_n}} \left[\int_{-a}^{y} f_{\omega_n}(z)dz - \int_{-a}^{y} f_{\omega_m}(z)dz\right] + \frac{\gamma^{-a}_{n,0}}{p_{\omega_n}},$$

where $\gamma_{n,0}^{-a}$ is the multiplier associated with the constraint $v_n^{-a} \geq 0$ where $v_n^{-a} = \lim_{J\to\infty} v_{n,\mathcal{I}}^0$. Taking the limit as $a \to \infty$ this is

$$(D_{\omega_n}(y))^2 = \sum_{m=1}^{M} \frac{\lambda_{m,n}}{p_{\omega_n}} \left[F_{\omega_n}(y) - F_{\omega_m}(y)\right] + \frac{\gamma_{n,0}^{-\infty}}{p_{\omega_n}}.$$

For a given $\mathcal{I}$, $\hat{f}_{\omega_n}^i$ is a value taken by f_{ω_n} and $D_{\omega_n,\mathcal{I}}^i$ is a value taken by $D_{\omega_n}(y)$ on I^i. Then $R_{\omega_n}^i$ is a value taken by $f_{\omega_n}(y)^{1/2}/D_{\omega_n}(y)$ on I^i and $v_{\mathcal{I}}(y;\mathbf{v}_{n,\mathcal{I}})$ contains a ratio of Riemann sums of $f_{\omega_n}(y)^{1/2}/D_{\omega_n}(y)$ on $[-a,y]$ and $[-a,a]$. Whenever $f_{\omega_n}(y)^{1/2}/D_{\omega_n}(y)$ is Riemann integrable, then

$$\lim_{J\to\infty} v_{\mathcal{I}}(y;\mathbf{v}_{n,\mathcal{I}}) = v_n^{-a} + \frac{v_n^a - v_n^{-a}}{K_{n,a}} \int_{-a}^{y} \frac{(f_{\omega_n}(z))^{1/2}}{D_{\omega_n}(z)} dz,$$

where $K_{n,a} = \int_{-a}^{a} \frac{(f_{\omega_n}(z))^{1/2}}{D_{\omega_n}(z)} dz$, $v_n^{-a} = \lim_{J\to\infty} v_{n,\mathcal{I}}^0$, and $v_n^a = \lim_{J\to\infty} v_{n,\mathcal{I}}^J$. Using the definition of the improper integral, $v_{\omega_n}^*(y) \equiv \lim_{a\to\infty} \lim_{J\to\infty} v_{\mathcal{I}}(y;\mathbf{v}_{n,\mathcal{I}})$ is given by

$$v_{\omega_n}^*(y) = v_n^{-\infty} + \frac{v_n^{\infty} - v_n^{-\infty}}{K_n} \int_{-a}^{y} \frac{(f_{\omega_n}(z))^{1/2}}{D_{\omega_n}(z)} dz.$$

where $K_n = \int_{-\infty}^{\infty} \frac{(f_{\omega_n}(z))^{1/2}}{D_{\omega_n}(z)} dz$, $v_n^{-\infty} = \lim_{a\to\infty} \lim_{J\to\infty} v_{n,\mathcal{I}}^0$, and $v_n^{\infty} = \lim_{a\to\infty} \lim_{J\to\infty} v_{n,\mathcal{I}}^J$.

Proof of Proposition 9

From Proposition 8

$$v_{\omega_n}^*(\bar{y}_n^r) - v_{\omega_n}^*(\bar{y}_n^r - y) = \frac{v_n^{\infty} - v_n^{-\infty}}{K_n} \int_{\bar{y}_n^r - y}^{\bar{y}_n^r} \left(\frac{p_{\omega_n} f_{\omega_n}(z)}{\left(\sum_{m=1}^{M} \lambda_{m,n} \left(F_{\omega_n}(z) - F_{\omega_m}(z)\right) + \gamma_{n,0}^0\right)} \right)^{1/2} dz$$

and

$$v_{\omega_n}^*(\bar{y}_n^r + y) - v_{\omega_n}^*(\bar{y}_n^r) = \frac{v_n^{\infty} - v_n^{-\infty}}{K_n} \int_{\bar{y}_n^r}^{\bar{y}_n^r + y} \left(\frac{p_{\omega_n} f_{\omega_n}(z)}{\left(\sum_{m=1}^{M} \lambda_{m,n} \left(F_{\omega_n}(z) - F_{\omega_m}(z)\right) + \gamma_{n,0}^0\right)} \right)^{1/2} dz.$$

When $n > m$, the difference $F_{\omega_n}(z) - F_{\omega_m}(z)$ has a unique minimum in $(\bar{y}_{\omega_m}^r, \bar{y}_{\omega_n}^r)$ and is increasing for all y after this minimum. Then for any $\Delta y > 0$, $F_{\omega_n}(\bar{y}_n^r + \Delta y) - F_{\omega_n}(\bar{y}_n^r + \Delta y) > F_{\omega_n}(\bar{y}_n^r - \Delta y) - F_{\omega_n}(\bar{y}_n^r - \Delta y)$.

When $n < m$, the difference $F_{\omega_n}(z) - F_{\omega_n}(z)$ has a unique maximum in $(\bar{y}_{\omega_n}^r, \bar{y}_{\omega_m}^r)$ and is decreasing for all y after this maximum. Then again for any $\Delta y > 0$, $F_{\omega_n}(\bar{y}_n^r + \Delta y) - F_{\omega_n}(\bar{y}_n^r + \Delta y) > F_{\omega_n}(\bar{y}_n^r - \Delta y) - F_{\omega_n}(\bar{y}_n^r - \Delta y)$.

Because f_{ω_n} is symmetric around ω_n, then the integrand is smaller at $z = \bar{y}_n^r + \Delta y$ than $z' = \bar{y}_n^r - \Delta y$ for all $\Delta y > 0$ as long as at least one $\lambda_{m,n}$ is positive.

Proof of Lemma 10

To show $v_n^{-\infty} = 0$, we proceed by showing that for any $c > 0$, there exists $a(c, \lambda)$, such that $\gamma_{n,0}^0 > 0$ for all $a > a(c, \lambda)$ and all $0 < c' < c$. This implies that $\gamma_{n,0}^{-a} > 0$ and by the complementary slackness condition $v_n^{-a} = 0$ for all sufficiently large a. Then $v_n^{-\infty} = 0$.

Given that for all $m > n$ $\lim_{y \to -\infty} \frac{f_{\omega_m}(y)}{f_{\omega_n}(y)} \to 0$, then for all $1 > \delta > 0$, there exists y' such that for all $m > n$ and $y < y'$, $\frac{f_{\omega_m}(y)}{f_{\omega_n}(y)} < \delta$. Then for any interval I^i of partition $\mathcal{I}$ where $y^i < y'$, $\frac{\hat{f}^i_{\omega_m}}{\hat{f}^i_{\omega_n}} < \delta$ for all $m > n$.

We proceed by contradiction by letting $\gamma_{n,0}^0 = 0$. From the definition of $D^i_{\omega_n}$ in Lemma 7, $v^i_{n,\mathcal{I}} - v^{i-1}_{n,\mathcal{I}} = c\frac{\left(\hat{f}^i_{\omega_n}\right)^{1/2}}{D^i_{\omega_n}}$ and therefore

$$v_n^i - v_n^0 = \sum_{j=1}^{i} \frac{c\left(\hat{f}^j_{\omega_n}\right)^{1/2}}{\left(\sum_{m=1}^{N} \frac{c\lambda_{m,n}}{p_{\omega_n}} \left[\sum_{k=1}^{j-1}\left(\hat{f}^k_{\omega_n} - \hat{f}^k_{\omega_m}\right) + \frac{1}{2}\left(f^j_{\omega_n} - \hat{f}^j_{\omega_m}\right)\right]\right)^{1/2}}. \tag{29}$$

For all $m < n$ and any interval I^i where $y^i < y'$, $\hat{f}^i_{\omega_m} > f^i_{\omega_n}$ and therefore for all $j \leq i$,

$$\sum_{m<n} \frac{c\lambda_{m,n}}{p_{\omega_n}} \left[\sum_{k=1}^{j-1}\left(\hat{f}^k_{\omega_n} - \hat{f}^k_{\omega_m}\right) + \frac{1}{2}\left(f^j_{\omega_n} - \hat{f}^j_{\omega_m}\right)\right] \leq 0.$$

Then using (29), the difference $v_n^i - v_n^0$ can be bounded below by

$$v_n^i - v_n^0 \geq \sum_{j=1}^{i} \left(\sum_{m>n} \frac{\lambda_{m,n}}{cp_{\omega_n}} \left[\sum_{k=1}^{j-1} \frac{\hat{f}^k_{\omega_n}}{\hat{f}^j_{\omega_n}}\left(1 - \frac{\hat{f}^k_{\omega_m}}{\hat{f}^k_{\omega_n}}\right) + \frac{1}{2}\left(1 - \frac{\hat{f}^j_{\omega_m}}{\hat{f}^j_{\omega_n}}\right)\right]\right)^{-1/2}.$$

Noting that $\hat{f}^k_{\omega_n} < \hat{f}^j_{\omega_n}$ for all $k < j < i$ due to the single peaked distribution of f_{ω_n}, and letting $\bar{\lambda} = \max\{\lambda_{n,m}\}$, this bound can be written as

$$v_n^i - v_n^0 \geq \sum_{j=1}^{i} \left(\frac{N\bar{\lambda}}{cp_{\omega_n}} \left[j\left(1 - \delta\right)\right]\right)^{-1/2} \geq \left(\frac{icp_{\omega_n}}{N\bar{\lambda}(1-\delta)}\right)^{1/2}.$$

Choose $a(c, \lambda)$ so that $y' - y^0 > \frac{N\bar{\lambda}(1-\delta)}{p_{\omega_n}}$, then there is a $y^i < y'$ where $ci > \frac{N\bar{\lambda}(1-\delta)}{p_{\omega_n}}$. This implies that $v_n^i - v_n^0 > 1$, a contradiction. For any $0 < c' < c$, this choice of a will lead to the same contradiction. Therefore for all $a > a(c, \lambda)$ and $0 < c' < c$, it must be that $\gamma_{n,0}^0 > 0$.

To show that $\gamma_{N,1}^{\infty} = 1$, note that for any m $f^J_{\omega_N} > f^J_{\omega_m}$ for sufficiently large a due to the single peaked distribution of f_{ω_n}. Then to satisfy (27), it must be that $\gamma_{N,1}^J > 0$.

Proof of Lemma 11

From (15), if $\lambda_{m,n} > 0$, then there is a $\lambda_{n,k}$ and a $\lambda_{j,m}$ that also must be positive. If $k = m$, then both $\lambda_{m,n}$ and $\lambda_{n,m}$ would be positive which would imply that $\mathbb{E}[v^*_{\omega_n}(y) - v^*_{\omega_m}|\omega_n] = \mathbb{E}[v^*_{\omega_n}(y) - v^*_{\omega_m}|\omega_m]$. However, this is not possible when both $\lambda_{n,m}$ and $\lambda_{m,n}$ are positive, as the cost of optimism must exceed the benefit of pessimism between these two states. For the same reasons, it cannot be that $j = n$. Therefore $j = k$, and it follows that exactly three constraints bind, as none of the other three opposing constraints can bind. Also, from (15), all three positive multipliers must be equal.

If $\lambda_{1,2} = \lambda_{2,3} = \lambda_{3,1} > 0$, then $\mathbb{E}[v^*_{\omega_1}(y) - v^*_{\omega_2}(y)|\omega_1] = u_2 - u_1$, $\mathbb{E}[v^*_{\omega_2}(y) - v^*_{\omega_3}(y)|\omega_2] = u_3 - u_2$, and $\mathbb{E}[v^*_{\omega_1}(y) - v^*_{\omega_3}(y)|\omega_3] = u_3 - u_1$, which implies that

$$\mathbb{E}[v^*_{\omega_1}(y) - v^*_{\omega_2}(y)|\omega_1] + \mathbb{E}[v^*_{\omega_2}(y) - v^*_{\omega_3}(y)|\omega_2] = \mathbb{E}[v^*_{\omega_1}(y) - v^*_{\omega_3}(y)|\omega_3]$$

However, this is not possible as from the proof of Proposition 5

$$\mathbb{E}[v^*_{\omega_1}(y) - v^*_{\omega_2}(y)|\omega_1] + \mathbb{E}[v^*_{\omega_2}(y) - v^*_{\omega_3}(y)|\omega_2] > \mathbb{E}[v^*_{\omega_1}(y) - v^*_{\omega_3}(y)|\omega_3]$$

when all $\lambda_{n,m} = 0$. When some multipliers are positive, then the cost of optimism is increased, so the two terms on the left hand side are more positive. Also the benefit of pessimism is reduced, reducing the right hand side. Therefore the equality cannot hold, and it follows that $\lambda_{2,1} = \lambda_{3,2} = \lambda_{1,3} > 0$ and $\lambda_{1,2} = \lambda_{2,3} = \lambda_{3,1} = 0$.