# Salience Weighted Utility over Presentations & The Etiology of Risky and Intertemporal Choice

Jonathan W. Leland<sup>1</sup>

Mark Schneider<sup>2</sup>

August 21, 2014

### Abstract

It has long been recognized that decisions are influenced by the framing of alternatives. However, little progress has been made in formally modeling such effects. In this paper, we introduce a general definition of a choice presentation and a model of salience weighted utility over presentations (SWUP) which provides a calculus for analyzing the impact of framing on choice. We show that SWUP predicts deep parallels between decisions under risk and time as arising from the same basic properties of human salience perception. We apply SWUP to explain some of the major framing effects which have been documented in the literature, as well as to explain other commonly observed features of preferences in risky and intertemporal choice.

# JEL Code: D81

# Keywords: Salience; Framing Effects; Allais Paradox; Present Bias

<sup>&</sup>lt;sup>1</sup>Consumer Financial Protection Bureau and The National Science Foundation,<u>jleland@nsf.gov</u>. The views expressed are those of the author and do not necessarily reflect those of the Consumer Financial Protection Bureau, The National Science Foundation or the United States.

<sup>&</sup>lt;sup>2</sup>School of Business, University of Connecticut, 2100 Hillside Road Unit 1041, Storrs, CT 06269, <u>mark.schneider@uconn.edu</u>.

### 1. Introduction

Observed behaviors frequently contradict the predictions of models of rational choice. Confronted with such evidence, the tradition in economics has been to posit a nonlinearity in the mapping of objective quantities into subjective values. Bernoulli (1738) accommodated the fact that people would not pay a large amount to play a game of infinite expected value by positing that their utility functions were concave. Friedman and Savage (1948) explained why some people gamble, some insure, and some do both by positing a utility function over final wealth levels with multiple inflection points. To explain why people would exhibit risk aversion for choices framed as gains but risk seeking when they were restated in terms of losses, Kahneman and Tversky (1979) proposed a value function that is concave for gains and convex for losses. To explain risk seeking (averse) behavior in choices between gains (losses) at low probabilities, they posited that objective likelihoods map into subjective decision weights in a non-linear fashion, over-weighting certainty and low probabilities.

In this paper we propose, instead, that it is non-linearity in the *perception of differences* in objective magnitudes across alternatives that is responsible for behaviors departing from the requirements of rational models. We also show, somewhat surprisingly, that this same approach predicts situations where behaviors will accord with standard axioms of rational choice. We develop our approach in two steps. The predictions of a model of choice that involves comparisons of magnitudes – payoffs, probabilities, dates of receipt or payment – will be sensitive to which attributes are being compared<sup>3</sup>. This, in turn, depends on how the choices are framed. As such, our first task, undertaken in Section 2, is to develop a theory of the framing of alternatives. We considertwo plausible ways choices might be presented or perceived by decision makers. In one, which we refer to as a *parallel* presentation, the options are presented in a way reminiscent of the decision matrices Savage (1954) introduced in discussing state-dependent choices and used by Loomes, Sugden and colleagues to testLoomes and

<sup>&</sup>lt;sup>3</sup> For discussions and demonstrations of this point, see Leland (1994, 1998, 2010).

Sugden's (1982) Regret theory.<sup>4</sup> In another choice presentation, referred to as a *minimalist* presentation, the choices look like the risky and intertemporal prospects as they are usually described in experimental surveys and journal articles.

In Section 2, we also introduce a model of risky and intertemporal choice where, for a given presentation, agents choose based on an evaluation procedure that involves cross-lottery comparisons of payoffs and probabilities or payoffs and time periods, respectively. The key assumption of the model is that larger differences in payoffs, probabilities and times of receipt or payment are perceived as notably dissimilar or salient and are disproportionately overweighted as a consequence<sup>5</sup>. We close Section 2 by proposing a set of properties that characterize human salience perception, building on and extendingBordalo et al. (2012). We refer to the resulting model as Salience Weighted Utility over Presentations (SWUP).

Section 3 applies the formalism developed in Section 2 to explain commonly observed behaviors under risk and over time. We close Section 3 by noting how SWUP explains a variety of framing effects such as the "hidden zero" effect in intertemporal choice, which are not explained by existing models.

Section 4 takes stock of our predictions and observes which behavioral properties of preferences can be derived from basic properties of human salience perception. Section 5 re-examines the data leading to the classic characterization of the fourfold pattern of risk attitudes from the perspective of SWUP. Section 6 discusses how SWUP relates to leading models of risky and intertemporal choice in the literature. Section 7 concludes.

<sup>&</sup>lt;sup>4</sup> Although in our approach, the representation has no implications regarding statistical dependence or independence of outcomes – it is simply a way to depict the choices.

<sup>&</sup>lt;sup>5</sup>This approach to context-dependent evaluation follows on work beginning with Tversky (1969), Gonzalez-Vallejo (2002), and Scholten andRead (2010), among others in psychology, and Loomes and Sugden (1982), Rubinstein (1988, 2003), Leland (1994, 1998 and 2002), Bordalo et al (2012) and Koszegi and Szeidl(2013), among others, in economics. Models of risky choice by Reyna and Brainerd (1991) and research on heuristics by Gigerenzer et al (1999) are also related.

### 2. A Model of Salience-Weighted Utility over Presentations

To develop a model of context-dependent choice, we first develop a characterization of how decisions are framed through the introduction of a *choice presentation*. We then develop a model of comparative evaluation over those presentations to predict choices between two alternatives, where the central feature of the evaluation process is that larger differences in attribute values (payoffs, probabilities, time delays) are perceived as more salient than smaller differences and are systematically over-weighted.

### 2.1 A Model of Choice Presentations

An *alternative*,  $\mathbf{A} \equiv (\mathbf{x}, \mathbf{a})$ , is an n-dimensional vector of possible consequences,  $\mathbf{x}$ , and an n-dimensional vector of attributes  $\mathbf{a}$ , where the i<sup>th</sup> attribute corresponds to the i<sup>th</sup> consequence. Denote the outcome space by  $\mathbf{X}$ . For decisions under risk,  $\mathbf{a} \equiv \mathbf{p}$  is a probability vector, in which case  $(\mathbf{x}, \mathbf{p})$  is a lottery. Denote the set of all lotteries by  $\mathbf{\Delta}(\mathbf{X})$ . For decisions over time,  $\mathbf{a} \equiv \mathbf{r}$  is a vector of time periods, in which case  $(\mathbf{x}, \mathbf{r})$  is a consumption plan. Denote the set of all consumption plans by  $\mathcal{M}$ . In our analysis, we will denote alternatives by f and g when referring to lotteries. Arbitrary consumption plans are denoted *cor d*. We will denote consumption plans involving a 'smaller sooner' payoff and a 'larger later' payoff by SS and LL, respectively.

A choice presentation or frame, **F**, is an  $m \times k$  matrix consisting of alternatives  $\{A_1, A_2, ..., A_m\}$ , where k = 2n. We denote the set of all presentationsby  $\mathcal{F}$ . A generic choice presentation is displayed in Figure 1. We may think of the gridlines as 'framing' the alternatives and attributes in the presentation.

	$x_1$	$u_1$	•••	X <sub>i</sub>	$a_i$	•••	$x_n$	$u_n$
$A_1$	<i>x</i> <sub>11</sub>	<i>a</i> <sub>11</sub>		$x_{1i}$	$a_{1i}$		$x_{1n}$	$a_{1n}$
$A_j$	<i>x</i> <sub>j1</sub>	a <sub>j1</sub>		x <sub>ji</sub>	a <sub>ji</sub>		x <sub>jn</sub>	$a_{jn}$
•••		•••		•••	•••		•••	
$A_m$	<i>x</i> <sub><i>m</i>1</sub>	$a_{m1}$		$x_{mi}$	$a_{mi}$		x <sub>mn</sub>	$a_{mn}$

**Figure 1: A generic Choice Presentation** 

For two alternatives,  $(\mathbf{x}, \mathbf{a})$ , and  $(\mathbf{y}, \mathbf{b})$ , we consider a decision maker who compares the i<sup>th</sup> prize of  $\mathbf{x}$  with the i<sup>th</sup> prize of  $\mathbf{y}$ , and compares the i<sup>th</sup> attribute of  $\mathbf{a}$  with the i<sup>th</sup> attribute of  $\mathbf{b}$ . In conventional treatments of choice, the index i would be interpreted as indexing different states of the world, or different time periods. For reasons to become clear shortly, here the index serves instead as a guide to the way agents are assumed to perceive and evaluate alternatives.

Consider a choice between two lotteries,  $f \equiv (\mathbf{x}, \mathbf{p})$  and  $g \equiv (\mathbf{y}, \mathbf{q})$ , where  $\mathbf{p}$  and  $\mathbf{q}$  are probability vectors with corresponding outcome vectors **x** and **y**. The (binary) presentation for this decision is given in Figure 2a. Analogously, Figure 2b displays the binary presentation for two consumption plans  $\mathbf{c} \equiv (\mathbf{x}, \mathbf{r})$  and  $\mathbf{d} \equiv (\mathbf{y}, \mathbf{t})$  with vectors of time periods **r** and **t**. To make clear what is being compared, we focus on presentations of the form in Figures 2a and 2b in which both row-vectors have the same cardinality.

Figure 2a: A binary choice presentation for lotteries

	$(x_1, y_1)$	$(p_1, q_1)$	$(x_2, y_2)$	$(p_2, q_2)$	 $(x_i, y_i)$	$(p_i, q_i)$	 $(x_n, y_n)$	$(p_n,q_n)$
f	$x_1$	$p_1$	<i>x</i> <sub>2</sub>	$p_2$	 $x_i$	$p_i$	 $x_n$	$p_n$
g	$y_1$	$q_1$	$y_2$	<i>q</i> <sub>2</sub>	 $y_i$	$q_i$	 $\mathcal{Y}_n$	$q_n$

Figure 2b: A binary choice presentation for consumption plans

	$(x_1, y_1)$	$(r_1, t_1)$	$(x_2, y_2)$	$(r_2, t_2)$	 $(x_i, y_i)$	$(r_i, t_i)$	 $(x_n, y_n)$	$(r_n, t_n)$
С	$x_1$	$r_1$	<i>x</i> <sub>2</sub>	$r_2$	 $x_i$	$r_i$	 $x_n$	$r_n$
d	$y_1$	$t_1$	$y_2$	$t_2$	 $y_i$	t <sub>i</sub>	 $\mathcal{Y}_n$	$t_n$

To highlight the comparisons being made in the examples provided in Sections 3 and 4, we will present choices between alternatives in the form of Figure 2. In each case, we will highlight the modal choice typical of experimental subjects in bold.

In our analysis, we focus on two types of presentations. The first type which we refer to as the *parallel presentation* sets  $p_i = q_i$  (for lotteries), or sets  $t_i = r_i$  (for consumption plans) for all i = 1, 2, ... n. The second type, which we refer to as the *minimalist* or *efficient presentation* contains the minimum number of cells in the presentation necessary to represent the choice alternatives. In other words, the minimalist presentation has no redundancy in that the same outcome does not appear in the same row-vector more than once (for a lottery), and the same time period does not appear in the same row-vector more than once (for a consumption plan). We will also consider versions of these presentationswhich are *monotonic in outcomes* in the sense that all outcomes are ordered to be weakly increasing  $(x_1 \le x_2 \le \dots \le x_n \text{ and } y_1 \le y_2 \le \dots \le y_n)$  or weakly decreasing, as well as presentations which are *monotonic in time periods* in the sense that all time periods are ordered to be weakly increasing.<sup>6</sup>

In binary choices involving a degenerate lottery, it is natural to frame the decision in a *monotone parallel* (MP) presentation, since each outcome in the non-degenerate lottery can only be compared with the single outcome in the degenerate lottery when making comparisons between lotteries. In contrast, when decisions involve two non-degenerate lotteries or two intertemporal prospects, they are naturally perceived in a *minimalist* presentation, which is the simplest representation of a decision.

#### 2.2 A Model of Salience-weighted Evaluation

In standard treatments of expected utility theory (EUT) or discounted utility theory (DUT), monotone parallel and monotone minimal descriptions of risky or intertemporal prospects are considered to be irrelevant. Considering first the case of risky choice and denoting a binary relation on  $\Delta(\mathbf{X})$ by  $\gtrsim$ , an expected utility maximizer with von Neumann-Morgenstern utility index,  $U: \mathbb{R} \to \mathbb{R}$  haspreferences between lotteries (**x**,**p**) and (**y**,**q**) characterized by

(1) 
$$(\mathbf{x},\mathbf{p}) \gtrsim (\mathbf{y},\mathbf{q}) \Leftrightarrow \sum_{i}^{n} p_{i} U(x_{i}) \ge \sum_{i}^{n} q_{i} U(y_{i}).$$

where i = 1, 2, ..., n indexes the location of each attribute in the presentation in Figure 2a. Under EUT, the decision maker's preferences over lotteries are represented by an

<sup>&</sup>lt;sup>6</sup> Time has a natural forward direction which makes it implausible that time periods are framed in a decreasing monotonic presentation. One may also consider a presentation which is monotonic in probabilities, but such a presentation is not needed in our analysis.

expected utility functional which computes the value of each lottery separately. Such a decision maker satisfies her preferences by selecting the lottery with the highest expected utility. Leland and Sileo (1998) point out that the choice could also be based on a procedure involving a particular kind of between-attribute comparison. Specifically, (1) can be rewritten as:

(2) 
$$(\mathbf{x},\mathbf{p}) \gtrsim (\mathbf{y},\mathbf{q}) \Leftrightarrow \sum_{i}^{n} [p_{i}U(x_{i}) - q_{i}U(y_{i})] \geq 0,$$

which can, in turn, be written as:

(3) 
$$(\mathbf{x},\mathbf{p}) \gtrsim (\mathbf{y},\mathbf{q}) \Leftrightarrow \sum_{i}^{n} [(p_{i} - q_{i})(U(x_{i}) + U(y_{i}))/2 + (U(x_{i}) - U(y_{i}))(p_{i} + q_{i})/2] \ge 0.$$

The individual choosing as specified in (3) decides between lotteries by considering the probability differences associated with outcomes weighted by the average utility of those outcomes, plus utility differences of outcomes weighted by their average probability of occurrence. Equation (3) enables us to consider a decision maker who compares the i<sup>th</sup> prize of **x** with the i<sup>th</sup> prize of **y**, and compares the i<sup>th</sup> probability of **p** with the i<sup>th</sup> probability of **q**, where i indexes the i<sup>th</sup> column vector in the choice presentation, as in Figure 2a. The behavior of an individual who chooses as in (3) will be identical to that of an expected utility maximizer who chooses as in (1) and evaluates each prospect separately. The behavior of an agent choosing according to either (1) or (3)will also be invariant to whether the options are described in a monotone parallel presentation or a monotone minimalist presentation.

Now suppose that in the process of comparing risky alternatives by (3), an agent notices when the payoff in one alternative is "a lot more money" than the payoff in anotherand when one alternative offers "a much better chance" of receiving an outcome than the other. In these cases, we will assume that large differences in attribute values (payoffs, probabilities) across different alternatives are perceived as particularly salient or attract disproportionate attention and are overweighted in the evaluation process. Under such an evaluation process the way the choice is presented can make a difference since the choice presentation determines which differences are overweighted.More precisely, for decisions under risk, we incorporate salience weights  $\phi(p_i, q_i)$  on probability differences and  $\mu(x_i, y_i)$  on payoff differences. Thus, we modify (3) such that

(4) 
$$(\mathbf{x},\mathbf{p}) \gtrsim (\mathbf{y},\mathbf{q}) \Leftrightarrow \sum_{i}^{n} [\phi(p_{i},q_{i})(p_{i}-q_{i})(U(x_{i})+U(y_{i}))/2 +\mu(x_{i},y_{i})(U(x_{i})-U(y_{i}))(p_{i}+q_{i})/2] \ge 0.$$

If  $\phi(p_i, q_i) = \mu(x_i, y_i)$  for all *i*, the model coincides with EUT. Since, as we will show in Section 3.2,(4) predicts characteristics of observed risk attitudes even with linear utility for gains and losses, and to avoid the calibration problems identified by Rabin  $(2000)^7$ , we let U(x) = x in our analysis. Thus, the salience weights on attribute differences constitute the only channel through which the predictions of our model can diverge from that of a risk-neutral expected utility maximizer.

The model extends analogously to choices over time. Let  $c \equiv (\mathbf{x}, \mathbf{r})$  and  $d \equiv (\mathbf{y}, \mathbf{t})$  be consumption plans. A decision maker who maximizes the discounted utility of future consumption streams with respect to a constant discount factor  $\delta \in [0,1]$ , where index *i* denotes the position of the i<sup>th</sup> attribute in the presentation in Figure 2b, coincides with an agent who weakly prefers *c* over *d* if and only if (5) holds:

(5) 
$$\sum_{i}^{n} \delta^{r_{i}}[U(x_{i})] \geq \sum_{i}^{n} \delta^{t_{i}}[U(y_{i})].$$

Note that (5) can be rewritten as (6):

$$(6)\sum_{i}^{n} \left[ \left( \delta^{r_{i}} - \delta^{t_{i}} \right) \left( U(x_{i}) + U(y_{i}) \right) / 2 + \left( U(x_{i}) - U(y_{i}) \right) \left( \delta^{r_{i}} + \delta^{t_{i}} \right) / 2 \right] \ge 0.$$

Analogous to our assumptions for decisions under risk, we assume agents not only notice when one alternative offers "a lot more money" than the payoff in another, but are sensitive to the fact that the payoff offered in one option arrives "a lot later" than the payoff in another. These assumptions are captured byplacing salience weights $\pi(r_i, t_i)$  on time differences, and $\mu(x_i, y_i)$  on payoff differences. We thus modify (6) such that *c* is chosen over *d* if and only if (7) holds:

<sup>&</sup>lt;sup>7</sup> Rabin (2000) shows that the rejection of favorable mixed prospects involving small stakes over a large range of wealth levels implies absurd risk aversion at large stakes under expected utility theory with a concave utility function.

(7) 
$$\sum_{i}^{n} [\pi(r_{i}, t_{i})(\delta^{r_{i}} - \delta^{t_{i}}) (U(x_{i}) + U(y_{i}))/2 + \mu(x_{i}, y_{i})(U(x_{i}) - U(y_{i})) (\delta^{r_{i}} + \delta^{t_{i}})/2] \ge 0$$

We refer to the model presented in (4) and (7) as salience-weighted utility over presentations (SWUP). We refer to a decision maker who chooses based on (4) and (7) as a *focal thinker*<sup>8</sup>. All of our propositions pertain to the behavior of a focal thinker.

#### **2.3** Properties of Salience Perception

The behavior of agents who choose according to SWUP will depend critically on the perceived salience of attribute differences they encounter. As such, we now propose a set of assumptions regarding the properties of human salience perception. We motivate our assumptions by appeal to some simple examples and note findings in the literature on the perception of differences, where relevant.<sup>9</sup>

A long tradition in psychology has studied the sensitivity of the perceptual system to changes in the magnitude of a stimulus. Since the Weber-Fechner law was introduced in the 19<sup>th</sup> century, it has been widely recognized that a fundamental property of the perceptual system is *diminishing absolute sensitivity* (DAS), and that this property applies to a range of sensory modalities including tone, brightness, and distance. In the context of salience perceptions of numerical differences, DAS implies that the difference between the numbers 100 and 0 will be perceived as more salient than the difference between 100 and 200. Cognitive psychologists have also examined other features that characterize human perceptions of numerical differences. One robust finding, first noted by Moyer and Landauer (1967), is that it takes adults longer to correctly respond to questions regarding which of two numbers is larger, the smaller their arithmetic difference<sup>10</sup>. To the extent that differences that are more readily discerned are more salient, this finding

<sup>&</sup>lt;sup>8</sup>Bordalo et al. (2012) use the term 'local thinker' to describe an agent in their model. Koszegi and Szeidl (2013) consider an agent who maximizes 'focus-weighted utility'. All these terms refer to the fact that agents focus too much attention to certain differences in attributes across alternatives.

<sup>&</sup>lt;sup>9</sup>Schley and Peters (2014) contains a nice discussion of this literature as it pertains to decision making.

<sup>&</sup>lt;sup>10</sup>For example, Moyer and Landauer (1967) found that it takes adults longer to answer the question "Which number is larger, 2 or 3?" than to answer the question "Which number is larger, 2 or 7?" This phenomenon is referred to as the "symbolic distance" effect.

would suggest that salience perceptions are *ordered* such that given four numbers x, y, w, z with x > y > w > z, the comparison of x to z is more salient than the comparison of y to w.

Now suppose values are scaled up proportionately, such that 100 versus 0 and 200 versus 100 are scaled to 1000 versus 0 and 2000 versus 1000. In contrast to the impact of an additive shift in values, proportionate increases in magnitudes make differences more salient (e.g., 2000 versus 1000 seems more salient than 200 versus 100) in which case salience perceptions reflect *increasing proportional sensitivity*(IPS). Ordering, DAS, and IPS, constitute the main substantive assumptions of our model.

If a constant, say 1000, is added to numbers 0, 100, and 200, the decrease in salience from comparing 1100 and 1000 relative to comparing 100 and 0 seems greater than the decrease in salience from comparing 1100 and 1200 relative to comparing 100 and 200. If so, this suggests salience perceptions reflect *relative increasing absolute sensitivity*.

In addition, the increase in salience when scaling values up from 100 - 0 to 1000 - 0 may seem greater than the increase in salience from scaling 200 - 100 to 2000 - 1000. If so, salience perceptions exhibit *relative decreasing proportional sensitivity*.

Finally, it seems plausible that salience perceptions reflect the property of *loss* sensitivity – that the salience associated with comparing 0 and 100 is less than that between 0 and -100. Indeed, the familiar characterization that losses 'loom larger' than gains seems to refer directly to an asymmetry in salience perception.

These intuitions, applied to payoffs, probabilities and time delays, can be formalized in the following properties that we will assume our salience functions  $\mu$ ,  $\pi$ , and  $\phi$ obey:

**Definition 1 (Properties of a Salience Function):** A *salience function* $\sigma(x, y)$  is any (non-negative) symmetric and continuous function that satisfies properties 1 - 3:

- **1.** Ordering: If [x', y'] is a subset of [x, y], then  $\sigma(x', y') < \sigma(x, y)$ .
- 2. Diminishing Absolute Sensitivity (DAS):  $\sigma(.)$  exhibits diminishing absolute sensitivity if, for any x, y > 0 and any  $\epsilon > 0$ ,  $\sigma(x + \epsilon, y + \epsilon) < \sigma(x, y)$ .
- 3. Increasing Proportional Sensitivity (IPS): $\sigma(.)$  exhibits increasing proportional sensitivity if, for any x, y > 0 and any  $\alpha > 1, \sigma(x, y) < \sigma(\alpha x, \alpha y)$ .

At times, we also use the following additional properties a salience function might have:

- **4.** Loss Sensitivity (LS): For any  $x, y > 0, \sigma(x, y) < \sigma(-x, -y)$ .
- 5. **Reflection:** For all x, y, x', y' > 0,  $\sigma(x', y') < \sigma(x, y)$  if and only if

$$\sigma(-x',-y') < \sigma(-x,-y).$$

6. Relative increasing absolute sensitivity (RIAS): For all  $y > x \ge y' > x' \ge 0$ , and  $\epsilon > 0$ ,

$$\frac{\sigma(x'+\epsilon,y'+\epsilon)}{\sigma(x',y')} < \frac{\sigma(x+\epsilon,y+\epsilon)}{\sigma(x,y)}$$

7. Relative decreasing proportional sensitivity (RDPS): For all  $y > x \ge y' > x' \ge 0$ , and $\alpha > 1$ ,

$$\frac{\sigma(\alpha x, \alpha y)}{\sigma(x, y)} < \frac{\sigma(\alpha x', \alpha y')}{\sigma(x', y')}$$

Subsets of these properties, or closely related ones, have been considered elsewhere in the literature on choice. Diminishing sensitivity is the continuous analog to the property of *Increasing Absolute Similarity* in Leland (2002) in the context of similarity judgments, and is closely related to Scholten and Read's (2010) *diminishing absolute sensitivity* assumption for delays and outcomesand to the *decreasing absolute sensitivity* property assumed in Prelec andLowenstein's (1991) extension of Prospect Theory to intertemporal choice. Increasing Proportional Sensitivity analogous to the property of *increasing proportional dissimilarity* used by Leland (2002) and is related to Scholten and Read's (2010) *augmenting proportional sensitivity* and tothe property of *increasing proportional sensitivity* in Prelec and Lowenstein (1991). Bordalo et al. (2012) assume properties 1, 2, and 5apply to perceptions of payoff salience but, in fact, properties 1, 2, 3, 5, 6, and 7 are all satisfied by the salience function (8) they proposed<sup>11</sup> (where  $\theta > 0$ ):

(8) 
$$\mu(x_i, y_i) = \left[\frac{|x_i - y_i|}{|x_i| + |y_i| + \theta}\right]$$

<sup>&</sup>lt;sup>11</sup> While Bordalo et al. (2012) did not propose loss sensitivity as a property of salience perception, they effectively assumed it by positing a piece-wise linear utility function which is steeper for losses than gains.

# 3. The Etiology of Risky and IntertemporalChoice

The behavior of expected utility and discounted utility maximizers with linear utility will be invariant to the host of arithmetic manipulations we might apply to the components of risky or intertemporal prospects. For U(x) = x, evaluations using (3) will be invariant to proportional rescaling of probabilities and payoffs. The choices recommended by (3) or (6) will not change if we increment all payoffs or dates of receipt in the offered prospects by a fixed amount. Choices will be invariant to the inclusion or exclusion of *payoff-probability* or *payoff-time*components that are shared across prospects (i.e. common consequences) since such components simply add a constant to the expected or discounted utilities of both options. Risk and time preferences will also be invariant to changes in the sign of outcomes, to changes in the labeling of attributes, and, as noted earlier, to changes in the alignment of attributes in a choice presentation.

In the following sections we show why none of these invariance properties necessarily hold for a focal thinker. We will consider risky prospects presented as in Figure 2a and intertemporal prospects as presented in Figure 2b. We illustrate each type of anomaly with an intuitive example and provide a formal analytical result for that behavior where possible. In our formal results the definitions are frame-dependent (they hold for the choice presentations given in that definition). Throughout our propositions we assume the agent is a focal thinker. We first demonstrate basic properties of parallel frames in which a focal thinker is predicted to obey some classical axioms of rational choice, consistent with empirical observations for choices presented in this format.

# 3.1 Parallel Presentations - Rational Behavior

For monotonic parallel presentations, a focal thinker always obeys first-order stochastic dominance.Since payoffs across alternatives are ordered monotonically, any differences in the evaluation process in (4) will always favor the stochastically dominant option. Compared probabilities across alternatives are always identical in a monotone parallel presentation, so they do not introduce any salience distortions into the evaluation. More formally:

**Definition 2:**(Stochastic Dominance): Lottery f (first-order) stochastically dominates g if  $F(x) \le G(x)$  for all  $x \in X$ , with at least one strict inequality, where F(x) and G(x) are the cumulative distribution functions corresponding to f and g, respectively. We say that *stochastic dominance holds* if f (first-order) stochastically dominates g implies f > g.

**Proposition 1:** Stochastic dominance holds for any monotone parallel presentation.

**Proof:** If *f* first-order stochastically dominates *g*, then in any MP presentation,  $x_i \ge y_i$  for all *i*, and all probability differences are zero. Thus, the salience weights in (4) favor *f* over *g* in each binary comparison for which the differences are not zero.

For similar reasons, the choices of an agent based on (4) will obey the independence axiom of EUT in an MP presentation. To demonstrate, recall:

**Definition 3 (Independence Axiom):** Given any  $f, g, f'' \in \Delta(X)$ , define lotteries f' and g' as follows:  $f' \equiv \alpha f + (1 - \alpha)f''$  and  $g' \equiv \alpha g + (1 - \alpha)f''$  for any  $\alpha \in (0,1)$ . Then  $f \gtrsim g$  if and only if  $f' \gtrsim g'$ .

**Proposition 2:** The independence axiom holds for any monotone parallel presentation.

**Proof:** Let  $f \coloneqq (\mathbf{x}, \mathbf{p}), g \coloneqq (\mathbf{y}, \mathbf{q})$ , and  $f'' \coloneqq (\mathbf{z}, \mathbf{s})$ . In a parallel presentation, we have the following representation of f, g, f' and g':

Figure 3.	The Inde	pendence	Axiom in	Parallel	Frames

f	$x_1$	$p_1$	•••	$x_n$	$p_n$
g	<i>y</i> <sub>1</sub>	$p_1$		$y_n$	$p_n$

<b>f</b> ′	$x_1$	$\alpha p_1$	•••	$x_n$	$\alpha p_n$	<i>z</i> <sub>1</sub>	$(1-\alpha)s_1$	•••	<i>Z</i> <sub>m</sub>	$(1-\alpha)s_m$
$g^{'}$	<i>y</i> <sub>1</sub>	$\alpha p_1$		$y_n$	$\alpha p_n$	$z_1$	$(1-\alpha)s_1$		Z <sub>m</sub>	$(1-\alpha)s_m$

If the frame is parallel, the outcome differences and probability differences are zero for each column vector following the column vector  $(x_n, y_n)$ . If the frame is also monotone, the only nonzero payoff differences in the choice between f' and g' are precisely the same as in the choice between f and g.

Parallel frames are also predicted to induce behavior consistent with rationality properties in intertemporal choice, such as stationarity and cancellation.

**Definition 4 (Stationarity):** The binary relation  $\gtrsim_t$  is *stationary* if for every  $t, r \ge 0, x, y \in \mathbb{R}$  and  $\Delta \ge 0$ ,

$$(x,r) \gtrsim_t (y,r+\Delta) \Leftrightarrow (x,t) \gtrsim_t (y,t+\Delta).$$

**Proposition 3:** *Stationarity holds for any parallel presentation.* 

**Proof:** Consider choice presentations between consumption plans c and d and between c' and d' in Figure 4:

Figure 4. The Stationarity Axiom in Parallel Frames

С	x	r	0	$r + \Delta$	<i>c</i> ′	x	t	0	$t + \Delta$
d	0	r	у	$r + \Delta$	d'	0	t	у	$t + \Delta$

A preference for *c* over dimplies  $\mu(x, 0)x\delta^r + \mu(0, y)(-y)\delta^{r+\Delta} \ge 0$ . A preference for *c'* over *d'*implies  $\mu(x, 0)x\delta^t + \mu(0, y)(-y)\delta^{t+\Delta} \ge 0$ . Note that we can write  $t \equiv r + a$ , for some constant *a*. Then  $\delta^a$  can be factored out and stationarity holds.

Finally, consider the cancellation axiom satisfied by DUT:

**Definition 5 (Cancellation Axiom):** The binary relation  $\gtrsim_t$  satisfies *cancellation* if for every  $r, t, t' \ge 0$ , and  $x, y, z \in \mathbb{R}$ ,

$$(x,r) \gtrsim_t (y,t) \Leftrightarrow (x,r;z,t') \gtrsim_t (y,t;z,t').$$

Cancellation is satisfied by most models of intertemporal choice. However, there are intuitive cases in which it is systematically violated (See Section 3.4). For now, we note that it is straightforward to verify the following result. Let  $t' \neq t$  and  $t' \neq r$ . Then:

**Proposition 4:** The cancellation axiom holds for any parallel presentation.

# **3.2Parallel Presentations–Properties of Risk Attitudes**

Consider a choice between a bet to gain\$x with probability p and \$0 with probability 1 - p versus receiving the bet's expected value with certainty. The MP presentation for this bet is displayed in the left panel of Figure 5. For p = 0.5, and x = \$100, it follows directly from diminishing absolute sensitivity (DAS) that the expected value is always selected for gains because the downside risk associated with receiving \$0 versus \$50 is more salient than the upside risk of winning \$100 versus \$50 (i.e.,  $\mu(50,0) > \mu(100,50)$ ). This observation holds more generally for p > 0.5. For choices involving losses, the reflection property implies that the risk seeking option appears more attractive because  $\mu(-50,0) > \mu(-100,-50)$ ). Thus, any payoff salience function induces risk aversion for gains and risk seeking for losses (at least for moderate and high probabilities).

**Definition 6** (Risk aversion for gains of moderate and high probabilities): Consider the choice presentation in Figure 5, where x > 0, and  $p \in (0,1)$ . Risk aversion for gains holds if E(f) > f.

Figure 5. Risk Aversion for Gains

f	x	p	0	1 - p
E(f)	хp	p	хp	1 – <i>p</i>

**Proposition 5:** For any salience function  $\mu(\cdot, \cdot)$ , E(f) > f for all  $p \in [0.5, 1]$ .

**Proof:**Note that E(f) > f if and only if  $\mu(0, xp) > \mu(x, xp)$ . By symmetry, ordering, and diminishing absolute sensitivity, we have  $\mu(0, xp) \ge \mu(0, x(1-p)) > \mu(x, xp)$  for all  $p \ge 0.5$ .

Note that if  $\mu(\cdot, \cdot)$  satisfies the reflection property, then changing the sign of the payoffs in Figure 5, (i.e., for all x < 0) we have f > E(f) for all  $p \in [0.5,1]$ . That is, the salient comparison is now between losing 0 or losing xp, in which case the decision maker is risk-seeking. This observation, coupled with Proposition 5 yields risk aversion for gains and risk seeking for losses of moderate and high probabilities and implies that if we observe risk-seeking for gains and risk-aversion for losses, such behavior must occur only at some sufficiently low p < 0.5. Indeed, under SWUP, the function for payoff salience, (8), from Bordalo et al. (2012) directly implies the classical fourfold pattern of risk attitudes identified by Tversky and Kahneman (1992). We defer discussion of this pattern of behavior to Section 5.

Next, consider a choice between accepting or rejecting a 50:50 chance of winning or losing x. The salience property of loss sensitivity implies that a focal thinker will abstain from symmetric bets. If we consider generalizing the choices by spreading the outcomes of one or both lotteries symmetrically, we expect to observe rejection of mean preserving spreads consistent with either risk aversion or loss aversion. More formally:

**Definition 7 (Loss Aversion; Kahneman and Tversky, 1979):**Consider the presentation in Figure 6 (where y > x > 0). *Loss aversion* holds if g > f.

у	0.5	-y	0.5
x	0.5	-x	0.5

#### **Figure 6.Loss Aversion**

Definition 7 is essentially the behavioral definition of loss aversion that Kahneman and Tversky(1979, p. 279) originally proposed: aversion to mean-preserving spreads of 50-50 gain-loss bets. It is straightforward to verify the following proposition:

**Proposition 6:** *Loss aversion holds if and only if*  $\mu(\cdot, \cdot)$  *obeysloss sensitivity.* 

f

g

In addition to studying general attitudes toward risk, economists beginning with Pratt (1964) and Arrow (1965), have enquired as to how those attitudes vary with changes in wealth and, more specifically, how the composition of portfolios varies between risky and safe assets as people become wealthier. One question they addressed concerns how an increase in wealth influences the absolute amounts invested in risky assets. Arrow (1970), among others, has proposed that people allocate a larger absolute amount to risky assets as wealth increases, implying that utility functions *exhibit decreasing absolute risk aversion*(DARA).Figure 7 illustrates this issue framed in terms of a choice between a lottery and a sure thing where the agent is confronted with the presentation on the right after a \$1000 increase in wealth. Given these choices, decreasing absolute risk aversionimplies that an agent indifferent between f and g will prefer g' to f'. SWUP predicts DARA for all salience perceptions which exhibit relative increasing absolute sensitivity (RIAS) as illustrated below.

Figure 7.Illustration of Decreasing Absolute Risk Aversion

	$(x_1, y_1)$	p	$(x_2, y_2)$	1 - p		$(x_1, y_1)$	p	$(x_2, y_2)$	1 - p
f	100	0.50	100	0.50	f'	1100	0.50	1100	0.50
g	0	0.50	500	0.50	$oldsymbol{g}'$	1000	0.50	1500	0.50

**Definition 8(Decreasing Absolute Risk Aversion):** Consider the presentations in Figure 8involving choices between f and g and between f' and g', where y > x > 0, and  $p \in (0,1)$ . *Decreasing absolute risk aversion (DARA) holds* if for any w > 0,  $f \sim g \Rightarrow f' < g'$ .

Figure 8.Decreasing Absolute Risk Aversion

f	x	р	x	<b>1</b> – <i>p</i>	f'	x + w	p	<i>x</i> + <i>w</i>	1 - p
g	у	p	0	1 - p	$oldsymbol{g}'$	y + w	р	w	1 – <i>p</i>

**Proposition 7:** *DARA holdsif* $\mu(\cdot, \cdot)$  *satisfies RIAS.* 

**Proof:** For the choice between *f* and *g*,  $f \sim g$  if and only if

(9) 
$$\mu(x,y)(y-x)p = \mu(x,0)(x)(1-p)$$

For the choice between *f* and *g*,  $f' \prec g'$  if and only if

(10) 
$$\mu(x+w,y+w)(y-x)p > \mu(w,x+w)(x)(1-p)$$

DARA holds if (9) implies (10) for all w > 0 which clearly holds under RIAS.

An additional question regardingrisk attitudes, originally posed by Zeckhauser and Keeler (1970), concerns the impact of simple proportional increases in payoffs on choices when wealth is held constant.Consider the choice presentations in Figure 9.

**Figure 9.Illustration of Size of Risk Aversion** 

	$(x_1, y_1)$	p	$(x_2, y_2)$	1 - p		$(x_1, y_1)$	p	$(x_2, y_2)$	1 - p
f	10	0.50	10	0.50	f'	100	0.50	100	0.50
g	0	0.50	25	0.50	g'	0	0.50	250	0.50

For a focal thinker, the property of Relative Decreasing Proportional Sensitivity (RDPS) implies *size of risk aversion*. As the value of the prizes are scaled-up, the salience of the difference between the safe outcome and zero (favoring the risk-averse choice) grows larger relative to the salience of the best outcome compared to the safe one, which favors the risky option (i.e.,  $(\mu(100,0)/\mu(250,100) > \mu(10,0)/\mu(25,10))$ . As a result, an individual indifferent between *f* and *g* will choose *f* over *g*'. More formally:

**Definition 9 (Size of Risk Aversion):** Consider the presentations in Figure 10 for choices between f and g and between f' and g', where y > x > 0, and  $p \in (0,1)$ . *Size of risk aversion (SORA) holds* if for any  $k > 1, f \sim g \Rightarrow f' > g'$ .

<b>T</b> .	10 01	0 70 1	•
HIGHTO	111 \$170		Avorcion
I'IZUIC	10.0120	UL INDA	AVCISIUII

f	x	p	x	1 - p	<b>f</b> '	kx	p	kx	<b>1</b> – <i>p</i>
g	у	p	0	1 – <i>p</i>	g'	ky	р	0	1 - p

**Proposition 8:***SORA holds if*  $\mu(\cdot, \cdot)$  *satisfies RDPS.* 

**Proof:** For the choice between f and g,  $f \sim g$  if and only if

(11) 
$$\mu(x,y)(y-x)p = \mu(x,0)(x)(1-p)$$

For the choice between *f* and *g*,  $f' \prec g'$  if and only if

(12) 
$$\mu(kx, ky)(y - x)p < \mu(kx, 0)(x)(1 - p)$$

SORA holds if (11) implies (12) for all k > 1, which clearly holds under RDPS.

Using analogous reasoning, RDPS also implies increasing relative risk aversion (IRRA). Both DARA and IRRA were observed experimentally by Holt and Laury (2002).

The preceding results demonstrate that some robust properties of risk attitudes (loss aversion, decreasing absolute risk-aversion, increasing relative risk aversion, size of risk aversion and the reflection of risk attitudes) can be derived directly from the properties of salience perception in Definition 1, even if utility is linear. Note also that all of these properties except loss aversion directly follow from salience function (8), and that (8) can be straightforwardly extended to satisfy the property of loss sensitivity. Thus these robust properties of risk preferences can be captured entirely by properties of the salience function, while avoiding Rabin's calibration paradox.

# 3.3 Parallel Presentations - Preferencefor Concentration

Koszegi and Szeidl (2011, 2013) identify a new and general pattern of behavior which they describe as a bias toward concentration: People prefer alternatives with a small number of large advantages relative to options with a large number of small advantages. This property follows from the most basic feature of the SWUP model – that larger differences in attribute values are overweighted relative to smaller differences. This is also implied by the ordering property of salience perception in Definition 1. In the context of intertemporal choice, such a bias toward concentration implies present-biased behavior when the costs of current consumption are distributed over many future dates (such as in addictive behaviors), but it also implies future-biased behavior when the benefit of many periods of effort is concentrated on a single goal (as in a personal or career achievement). Extending this bias to decisions under risk, a bias toward concentration implies risk aversion when losses are distributed over many small outcomes relative to a single large loss (as in purchasing insurance) and risk-seeking behavior when the benefits of a small reward can be foregone for the chance of a single large reward (as in purchasing a lottery ticket). More formally, recall that  $f \equiv (\mathbf{x}, \mathbf{p})$  and  $g \equiv (\mathbf{y}, \mathbf{q})$ , where  $\mathbf{x}, \mathbf{p}, \mathbf{y}$ , and  $\mathbf{q}$  are each n-dimensional vectors. For the binary choice presentation Fin Figure 2a, define sets  $A(\mathbf{x}, \mathbf{y})$  and  $A(\mathbf{p}, \mathbf{q})$  as follows:

$$A(\mathbf{x}, \mathbf{y}) := \{x_i, y_i \in \mathbf{F} : x_i > y_i\}.$$
  
$$A(\mathbf{p}, \mathbf{q}) \coloneqq \{p_i, q_i \in \mathbf{F} : p_i > q_i | x_i > 0\} \cup \{p_i, q_i \in \mathbf{F} : p_i < q_i | x_i, y_i < 0\}.$$

Set  $A(\mathbf{x}, \mathbf{y})$  is the set of column vectors in **F** for which *f* has a payoff advantage over *g*. We call  $A(\mathbf{x}, \mathbf{y})$  the set of *compared payoff advantages* for *f*.Similarly,  $A(\mathbf{p}, \mathbf{q})$  is the set of column vectors in **F** for which *f* has a probability advantage over *g* (e.g., the column vectors for which *f* yields a higher probability of a gain or a lower probability of loss). We call  $A(\mathbf{p}, \mathbf{q})$  the set of *compared probability advantages* for *f*. Analogously, define  $A(\mathbf{y}, \mathbf{x})$  and  $A(\mathbf{q}, \mathbf{p})$  as sets of compared payoff and probability advantages<sup>12</sup> for *g*.Next we extend the notion of a bias toward concentration (Koszegi and Szeidl, 2011) to decisions under risk:

**Definition 10:** Let  $\Delta(a_i, b_i) = |a_i - b_i|$ . Lottery *f* has more concentrated advantages than disadvantages given **F**, if (i) and (ii) hold.

(i) 
$$\min_{i \in A(\mathbf{x}, \mathbf{y})} \Delta(x_i, y_i) > \max_{i \in A(\mathbf{y}, \mathbf{x})} \Delta(x_i, y_i)$$

(ii) 
$$\min_{i \in A(\mathbf{p},\mathbf{q})} \Delta(p_i, q_i) > \max_{i \in A(\mathbf{q},\mathbf{p})} \Delta(p_i, q_i)$$

That is, each absolute compared advantage for f over g is greater than any single compared disadvantage for both payoffs and probabilities. However, g may have many small advantages over f.

**Definition 11:** Let EU(f) = EU(g) and suppose f has more concentrated advantages than disadvantages given **F**. A *bias toward concentration* (BTC) holds if f > g given **F**.

A type of bias toward concentration follows from the ordering property of salience perception in Definition 1. First, we provide the following definition:

**Definition 12:** Lottery *finterval-dominatesg* given  $\mathbf{F}$  if $[x_j, y_j] \subset [y_i, x_i]$  for all  $i \in A(\mathbf{x}, \mathbf{y}), j \in A(\mathbf{y}, \mathbf{x}), \text{ and } [p_j, q_j] \subset [q_i, p_i]$  for all  $i \in A(\mathbf{p}, \mathbf{q}), j \in A(\mathbf{q}, \mathbf{p})$ .

<sup>&</sup>lt;sup>12</sup>A choice presentation constrains which comparisons will be made by the perceptual system. By 'compared advantage' we refer to the pairs of attributes whose differences are computed in (4) (i.e., those comparisons which are induced by the presentation).

Under Definition 12, *f* interval-dominates *g* if the interval between payoffs (probabilities) for any compared advantage of *g* over *f* is a subset of the interval between payoffs (probabilities) for any compared advantage of *f* over *g*. For example, *f* interval-dominates *g* in the presentation in Figure 11. Note that in the example in the figure E(f) = E(g), although this is not a necessary condition for interval dominance. Also note that, for U(x) = x, BTC holds in Figure 11 if f > g.

Figure 11. Illustration of Interval Dominance and Bias Toward Concentration

	$(x_1, y_1)$	$(p_1, q_1)$	$(x_2, y_2)$	$(p_2, q_2)$	$(x_3, y_3)$	$(p_3, q_3)$	$(x_4, y_4)$	$(p_4, q_4)$
f	100	0.25	40	0.25	50	0.25	50	0.25
g	0	0.25	80	0.25	80	0.25	80	0.25

If EU(f) = EU(g) and f interval-dominates g, then f has a few large (compared) advantages over g, whereas g has a larger number of smalleradvantages over f.

**Proposition 9:** *Suppose f interval-dominates ggiven* **F***. Then BTC holds for any parallel presentation and any salience function* $\mu(\cdot, \cdot)$ *.* 

**Proof:** For EU(f) = EU(g), in the absence of salience weights, we have:

$$\sum_{i=1}^{n} (U(x_i) - U(y_i)) \left[ \frac{p_i + q_i}{2} \right] + (p_i - q_i) \left[ \frac{U(x_i) + U(y_i)}{2} \right] = 0$$

We can rewrite this relationship as follows,

$$\sum_{i \in A(\mathbf{x}, \mathbf{y})} (U(x_i) - U(y_i)) \left[ \frac{p_i + q_i}{2} \right] + \sum_{i \in A(\mathbf{p}, \mathbf{q})} (p_i - q_i) \left[ \frac{U(x_i) + U(y_i)}{2} \right] = \sum_{i \in A(\mathbf{y}, \mathbf{x})} (U(y_i) - U(x_i)) \left[ \frac{p_i + q_i}{2} \right] + \sum_{i \in A(\mathbf{q}, \mathbf{p})} (q_i - p_i) \left[ \frac{U(x_i) + U(y_i)}{2} \right].$$

Note that, for a parallel frame, the sets  $A(\mathbf{p}, \mathbf{q})$  and  $A(\mathbf{q}, \mathbf{p})$  are empty. By the ordering property of a salience function, if *f* interval-dominates*g* given **F**, then  $\min_{i \in A(\mathbf{x}, \mathbf{y})} \mu(x_i, y_i) > \max_{i \in A(\mathbf{y}, \mathbf{x})} \mu(x_i, y_i)$ . This yields:

$$\sum_{i \in A(\mathbf{x}, \mathbf{y})} \mu(x_i, y_i) (U(x_i) - U(y_i)) \left[ \frac{p_i + q_i}{2} \right] > \sum_{i \in A(\mathbf{y}, \mathbf{x})} \mu(x_i, y_i) (U(y_i) - U(x_i)) \left[ \frac{p_i + q_i}{2} \right].$$

Thus,  $f \succ g$ . It is straightforward to verify that if f interval-dominates g given  $\mathbf{F}$  then f has more concentrated advantages than disadvantages given  $\mathbf{F}$ .

Defining concentrated advantages, BTC, and interval dominance analogously to Definitions 10, 11, and 12 for choices over time yields a similar result to Proposition 9 for consumption plans:

**Proposition 10:** Suppose consumption plan c interval-dominates dgiven  $\mathbf{F}$ . Then BTC holds for any parallel presentation and any salience function $\mu(\cdot, \cdot)$ .

Note that BTC may apply in a wide range of consumer contexts such as choosing to purchase a single product (e.g., a lap top, television, or car) with many small monthly payments rather than in a few large installments, choosing to purchase a newspaper or magazine subscription framed in terms of the price per day, rather than an identical offer framed in terms of the price per year, or choosing to pay only the monthly minimum fee on a credit card rather than payingthe balance in full.

#### 3.4 Minimalist Presentations – Anomalies in Risky and Intertemporal Choice

Choices between smaller sooner (SS) and a larger later (LL) rewardmap naturally into a minimalist presentation as shown on the left in Figure 12. Likewise, risky choices involving binary lotteries are naturally framed in the minimalist format on the right.

**Figure 12.** Minimalist Presentations for Simple Consumption Plans and Lotteries

	$(x_1, y_1)$	$(r_1, t_1)$		$(x_1, y_1)$	$(p_1, q_1)$	$(x_2, y_2)$	$(p_2, q_2)$
SS	x	r	f	x	p	0	1 - p
LL	у	t	g	у	q	0	1 - q

In a minimalist presentation, the salience of payoff differences trades off against the salience of either probability or time differences. Theoretically inconsequential arithmetic manipulations of prospect components may alter choices by changing the perceived salience of these prospect attributes. To illustrate, consider first simple intertemporalchoices between two payoffs, separated by a period of one year, as shown in Figure 13. The stationarity axiom of discounted utility theory requires that individuals either choose SS and SS' or LL and LL' –delay influences the decision onlythroughtheabsolutedifferenceinthetimeperiods. Strotz(1955) was the first to conjecture, and subsequent experimental work (as cited in Frederick et al. 2002) has

confirmed, that people will instead choose the impatient option*SS* in the choice on the left but the patient option *LL'* in the choice on the right in Figure 13. Strotz noted that these choices implied dynamically inconsistent preferences – an individual stating a preference for SS and LL' today will always prefer to reverse the latter decision 10 years hence. This pattern of behavior has been labeled the *common difference effect* or, when the payoff in SS is immediate, as*present bias* (Prelecand Loewenstein1991). Standard modelsattribute this choice pattern to hyperbolic or quasi-hyperbolic discounting (e.g., Loewenstein and Prelec 1992, Laibson 1997).

Figure 13. Illustration of Present Bias

	$(x_1, y_1)$	Period		$(x_2, y_2)$	Period
SS	100	0	SS'	100	10
LL	120	1	LL'	120	11

In the SWUP model, SS is chosen over LL because the salience weight associated with the difference in dates of receipt (favoring SS) outweighs the weight associated with the difference in payoffs (favoring LL). However, as a consequence of diminishing absolute sensitivity, the salience weight associated with the time delays is smaller in the choice between SS' and LL' making LL appear more attractive. More generally, we have:

**Definition 13 (Common Difference Effect):**Consider the frames in Figure 14, where y > x > 0 and  $t > r \ge 0, \Delta > 0$ .

# Figure 14.The Common Difference Effect

	$(x_1, y_1)$	Period		$(x_2, y_2)$	Period
SS	x	r	SS'	x	$r + \Delta$
LL	у	t	LL'	у	$t + \Delta$

The common difference effect holds if  $SS \sim LL \Rightarrow SS' \prec LL'$ .

2. 3.

**Proposition 11:** *The common difference effect holds for any salience function* $\pi(r, t)$ *.* 

**Proof:** For the choice between SS and LL, SS ~ LL under SWUP if and only if

$$\mu(x,y)\left(U(y) - U(x)\right)\left[\frac{\delta^r + \delta^t}{2}\right] = \pi(r,t)(\delta^r - \delta^t)\left[\frac{U(y) + U(x)}{2}\right]$$

For the choice between SS' and LL',  $SS' \prec LL'$  if and only if

$$\mu(x,y)\big(U(y) - U(x)\big)\bigg[\frac{\delta^r + \delta^t}{2}\bigg] > \pi(r + \Delta, t + \Delta)(\delta^r - \delta^t)\bigg[\frac{U(y) + U(x)}{2}\bigg]$$

Given  $SS \sim LL$ , the common difference effect holds for all  $\Delta > 0$  if and only if  $\pi(r, t) > \pi(r + \Delta, t + \Delta)$  for all  $\Delta > 0$  (that is, if and only if  $\pi(r, t)$  satisfies DAS).

In addition to requiring that choices between SS and LL be invariant to the addition of a constant delay to both options, the discounted utility model requires they beinvariant to proportional increases in payoffs.Thaler (1981), followed by many others,<sup>13</sup>found instead that proportional increases in payoffs tend to promote increasing patience – a phenomenon called the absolute magnitude effect, or simply the 'magnitude effect.'

# Figure 15.Illustration of the Magnitude Effect

	$(x_1, y_1)$	Period		$(x_1, y_1)$	Period
SS	75	0 years	SS'	750	0 years
LL	100	1 year	LL'	1000	1 year

An illustrative example of the magnitude effect is provided in Figure 15. In the choice presentation on the left, the agent is impatient if the salience of the difference between 75 and 100 is overwhelmed by the difference between the times of receipt. For such an agent, however, when given a choice between SS' and LL', the salience of the payoff difference is enhanced as a consequence of the proportionate scaling and leads to a shift in preference toward more patient behavior. Thus, the magnitude effect holds if salience perceptions satisfy increasing proportional sensitivity from Definition 1. More formally:

**Definition 14 (Magnitude Effect):** Consider the two presentations in Figure 16(with  $y > x > 0, t > r \ge 0, k > 1$ ):

# **Figure 16.The Magnitude Effect**

<sup>&</sup>lt;sup>13</sup> For an extensive discussion of the literature pertaining to the magnitude effect, see Read (2003).

	$(x_1, y_1)$	Period		$(x_1, y_1)$	Period
SS	x	r	SS'	kx	r
LL	у	t	LL'	ky	t

The magnitude effect holds if, for all k > 1,  $SS \sim LL \Rightarrow SS' \prec LL'$ .

**Proposition 12:** *The magnitude effect holds for any salience function* $\mu(\cdot, \cdot)$ . **Proof:** For the choice between *SS* and *LL*, *SS* ~ *LL* under SWUP if and only if

$$\mu(x,y)(U(y) - U(x))\left[\frac{\delta^r + \delta^t}{2}\right] = \pi(r,t)(\delta^r - \delta^t)\left[\frac{U(y) + U(x)}{2}\right]$$

For the choice between SS' and LL',  $SS' \prec LL'$  if and only if

$$\mu(kx,ky)(U(y) - U(x))\left[\frac{\delta^r + \delta^t}{2}\right] > \pi(r,t)(\delta^r - \delta^t)\left[\frac{U(y) + U(x)}{2}\right]$$

Given  $SS \sim LL$ , the magnitude effect holds if and only if  $\mu(kx, ky) > \mu(x, y)$  for all k > 1, which holds if and only if  $\mu(\cdot, \cdot)$  satisfies IPS.

A third robust pattern of intertemporal choice occurs when the sign of payoffs is reversed. In particular, people are often more patient for losses than for gains, an observation known as the sign effect (Loewenstein and Prelec, 1991). The presentation in Figure 17 provides an illustrative example of the sign effect, in which a person who prefers to receive \$25 in one year over \$40 in two years will also prefer to pay \$25 in one year rather than pay \$40 in two years. Essentially, the salience of paying \$40 relative to paying \$25 is greater than the salience of gaining\$40instead of gaining \$25.

# **Figure 17.Illustration of the Sign Effect**

	$(x_1, y_1)$	Period		$(x_1, y_1)$	Period
SS	25	1	SS'	-25	1
LL	40	2	LL'	-40	2

Formally, we have the following definition and result:

**Definition 15(Sign Effect):** Consider the presentations in Figure 18 (with  $y > x > 0, t > r \ge 0$ ): The sign effect holds if  $SS \sim LL \Rightarrow SS' > LL'$ .

# **Figure 18.The Sign Effect**

	$(x_1, y_1)$	Period		$(x_1, y_1)$	Period
SS	x	r	SS'	-x	r
LL	у	t	LL'	-y	t

**Proposition 13:** *The sign effect holds if and only if*  $\mu(\cdot, \cdot)$  *satisfies loss sensitivity.* **Proof:** For the choice between *SS* and *LL*, *SS* ~ *LL* under SWUP if and only if

 $\mu(x,y)(y-x)((\delta^r+\delta^t)/2)=\pi(r,t)(\delta^r-\delta^t)((x+y)/2)$ 

For the choice between SS' and LL', SS' > LL' if and only if

$$\mu(-x, -y)(y - x)((\delta^r + \delta^t)/2) > \pi(r, t)(\delta^r - \delta^t)((x + y)/2)$$

Given that  $SS \sim LL$ , it is clear that the above inequality holds if and only if  $\mu(-x, -y) > \mu(x, y)$ .

A proportional scaling of probabilities should be of no consequence to an EU maximizer, just as aproportional scaling of payoffs should not impact the choices of a DU maximizer, However, as first shown by Allais (1953), people are frequently riskseeking in a choice between lotteries involving small probabilities but risk-averse when the probabilities are scaled up proportionately, a finding known as the common ratio effect. A popular variant due to Kahneman and Tversky(1979) is presented in Figure 19.

f	3000	0.90	0	0.10	f'	3000	0.002	0	0.998
g	6000	0.45	0	0.55	$oldsymbol{g}'$	6000	0.001	0	0.999

Figure 19. Illustration of the Common Ratio Effect

In the common ratio effect, scaling probabilities up, holding payoffs fixed, leads to more risk-averse behavior. More formally, we have:

**Definition 16 (Common Ratio Effect):** Consider Figure 20, with  $y > x > 0, 1 \ge p > q \ge 0$  and  $\alpha \in (0,1)$ . The common ratio effect holds if  $f \sim g \implies g' > f'$ .

f	x	р	0	1 - p	f'	x	αp	0	$1 - \alpha p$
g	у	q	0	1 - q	$oldsymbol{g}'$	у	αq	0	$1-\alpha q$

# Figure 20.The Common Ratio Effect

**Proposition 14:** *The common ratio effect holds for any salience function*  $\phi(p, q)$ *.* 

**Proof:** For the choice between f and g,  $f \sim g$  under SWUP if and only if

$$\mu(x,y)(U(y) - U(x))\left[\frac{p+q}{2}\right] = \phi(p,q)(p-q)\left[\frac{U(y) + U(x)}{2}\right]$$

For the choice between f' and g', g' > f' if and only if

$$\mu(x,y)\big(U(y)-U(x)\big)\Big[\frac{p+q}{2}\Big] > \phi(\alpha p, \alpha q)(p-q)\Big[\frac{U(y)+U(x)}{2}\Big].$$

By increasing proportional sensitivity, scaling  $\alpha p$  and  $\alpha q$  each by  $\frac{1}{\alpha}$  leads to  $\phi(\alpha p, \alpha q) < \phi(p, q)$  for all  $\alpha \in (0, 1)$ .

### **3.5Framing Effects Explained by SWUP**

To this point we have shown that SWUP explains when observed behaviors are in accordance with strong rationality axioms in both risky and intertemporal settings (when choices are framed in parallel presentations), and when the same person may systematically violate these axioms (when choices are framed in minimalist presentations).We now turn to situations where choices differ given different, but logically equivalent, representations of the same pair of options. These framing effects and violations of descriptive invariance are among the most vexing of choice anomalies precisely because they cannot be explained if we treat choices as inevitably preference revealing, as assumed in most axiomatic decision theories.

In the SWUP model, framing effects result from two sources –changes in the arrangement of attributes (*alignment effects*) and changes in the salience of attribute labels (*labeling effects*). We consider each in turn.

# 3.5.1 Alignment Effects

To illustrate alignment effects consider the two depictions of risky prospects f and gshown on the left in Figure 21. The pairs of lotteries are the same, but they are given in

a minimalist presentation (top) and in a parallel presentation (bottom). In the minimalist presentation, the salience of the difference between 2500 and 2400 trumps the 0.01 difference in probabilities, producing a preference for the risky alternative, f. In the parallel presentation, it is the salience of receiving a payoff 0 versus 2400 that drives the choice of g. Support for this shift in observed risk preference can be found in experimental studies due to Leland (2001) and Bordalo et al. (2012). Leland (2001) provided 37 experimental subjects with the choice between f and g in the minimalist presentation in Figure 21, and provided 37 other subjects with the choice between f and g in the parallel presentation in Figure 21. In the minimalist frame, 32% of experimental subjects<sup>14</sup> chose the safe optiong whereas 73% chose the safe optiong when lotteries were given in the parallel presentation.<sup>15</sup>

Figure 21. Framing Effects due to Changes in the Alignment of Attributes

	Hidden	Zero E	ffect (R	isk)	Hie	dden Ze	ro Effe	ect (Tim	e)		
f	$(x_{1},y_{1})$ <b>2500</b>	(p <sub>1</sub> ,q <sub>1</sub> ) <b>0.33</b>	(x <sub>2</sub> ,y <sub>2</sub> )	(p <sub>2</sub> ,q <sub>2</sub> ) <b>0.67</b>			SS	$(x_{1},y_{1})$ <b>100</b>	$(r_{1,t_{1}})$ 0	]	
g	2400	0.34	0	0.66			LL	120	1	]	
	(x <sub>1</sub> ,y <sub>1</sub> )	$(p_{1},q_{1})$	(x <sub>2</sub> ,y <sub>2</sub> )	$(p_{2},q_{2})$	(x <sub>3</sub> ,y <sub>3</sub> )	$(p_{3},q_{3})$		(x <sub>1</sub> ,y <sub>1</sub> )	$(r_{1,}t_{1})$	(x <sub>2</sub> ,y <sub>2</sub> )	$(r_{2}, t_{2})$
f	2500	0.33	0	0.66	0	0.01	SS	100	0	0	1
g	2400	0.33	0	0.66	2400	0.01	LL	0	0	120	1

# **Stochastic Dominance Framing Effect**

	$(x_{1}, y_{1})$	$(p_{1},q_{1})$	$(x_{2}, y_{2})$	$(p_{2}, q_{2})$	$(x_{3}, y_{3})$	$(p_{3}, q_{3})$	$(x_{4}, y_{4})$	$(p_{4}, q_{4})$	$(x_{5}, y_{5})$	$(p_{5}, q_{5})$
f	0	0.9	45	0.06	45	0.01	-10	0.01	-15	0.02
g	0	0.9	45	0.06	30	0.01	-15	0.01	-15	0.02
	$(x_{1}, y_{1})$	$(p_{1}, q_{1})$	$(x_{2}, y_{2})$	$(p_{2}, q_{2})$	$(x_{3}, y_{3})$	$(p_{3}, q_{3})$	$(x_{4}, y_{4})$	$(p_{4}, q_{4})$		
f	0	0.9	45	0.07	-10	0.01	-15	0.02		
g	0	0.9	45	0.06	30	0.01	-15	0.03		
-										

<sup>&</sup>lt;sup>14</sup>Subjects wereCarnegie MellonUniversityundergraduates.

<sup>&</sup>lt;sup>15</sup>Along similar lines, Bordalo et al. (2012) report that given choices presented in the minimalist format 46% chose the risky option. When the choices are depicted as in the parallel presentation, a significant majority, 78%, are risk-averse.

Now consider the two depictions of the intertemporal choices between smaller-sooner and larger-later options shown on the upper right side in Figure 21. In the minimalist presentation of these alternatives, receiving a payoff immediately ratherthan after a oneyear delay is more salient than the difference between a payoff of 100 or 120, producing a preference for the smaller sooner (SS) reward. However, in the parallel presentation, the difference between 120 and 0 is more salient than the difference between 100 and 0, producing a preference for the larger later (LL) reward. Response patterns such as these have been observed experimentally by Magen et al. (2008) and referred to as the "hidden zero" effect. The SWUP model explains this and makes the prediction that there will also be a hidden zero effect under risk analogous to the one observed in intertemporal choice.

An additional alignment effect concerns choices where one lottery stochastically dominates another. Tversky and Kahneman (1986) showed that while people choose the stochastically dominant lottery when the dominance relationship is transparent, they may systematically violatestochastic dominance when the dominance relation is masked. Kahneman and Tversky presented subjects with the parallel presentation between lotteries f and g in the bottom panel of Figure 21. For the parallel presentation, all subjects chose f thereby satisfying dominance. But when choosing between f and g in the minimalist presentation, 58% chose the stochastically dominated lottery,g. The intuition under the SWUP model is that the comparison between payoffs of \$30 and -\$10 in the minimalist presentation is more salient than the comparison of the 1% difference in the probability of receiving \$45. Note that under the SWUP model, the three alignment effects in Figure 21 each emerge from the same mechanism – a switch in the framing of alternatives between parallel and minimalist presentations.

The framing effects analyzed in Figure 21 imply that violations of rational choice theory can be systematically 'turned on' or 'turned off' depending on the structure of the choice presentation. Thus, SWUP has important implications for the robustness of rational choice violations, predicting how they are sensitive to the framing of alternatives. This observation is illustrated in the context of the Allais paradox and present bias in

Figure 22.The alignment effects observed in Figure 21 suggest that the most famous rational choice violations for decisions under risk (the Allais paradox) and over time (present bias) will be sensitive to whether alternatives are framed in parallel or minimalist presentations. For instance, while SWUP predicts the Allais paradox and present biased behavior when choices are presented in the usual format given to experimental subjects (top panel of Figure 22), the paradoxes are predicted to largely disappear when both alternatives are displayed in parallel presentations (bottom of Figure 22). For decisions under risk, Leland (2001) and Bordalo et al. (2012) found experimental support for thesepredictions using the version of the Allais paradox (Allais, 1953) due to Kahneman and Tversky (1979) which is presented in Figure 22.

In the variant of the Allais paradox in Figure 22, EUT predicts a decision maker to choose either lotteries f and f' or g and g'. However, when presented as in the top panel of Figure 22, most people choose f and g' (Kahneman and Tversky, 1979). Kahneman and Tversky (1979) attribute that latter choice to a certainty effect where sure outcomes are overweighted. Along similar lines, Camerer (1992) and more recently Andreoni and Sprenger (2012)have proposed that these choice patterns reflect boundary effects which occur when one of the choice options is certain to obtain.

(x <sub>1</sub> ,y <sub>1</sub> )	$(p_{1},q_{1})$	$(x_2, y_2)$	$(p_{2},q_{2})$	(x <sub>3</sub> ,y <sub>3</sub> )	$(p_{3},q_{3})$		(x <sub>1</sub> ,y <sub>1</sub> )	$(r_{1},t_{1})$
2500	0.33	2400	0.66	0	0.01	SS	100	0
2400	0.33	2400	0.66	2400	0.01	LL	120	1
$(x_{1}, y_{1})$	$(p_{1}, q_{1})$	$(x_{2}, y_{2})$	$(p_{2}, q_{2})$	_			$(x_{1}, y_{1})$	$(r_{1}, t_{1})$
2500	0.33	0	0.67			SS	100	10
2400	0.34	0	0.66			LL	120	11
	$(x_{1},y_{1})$ 2500 2400 $(x_{1},y_{1})$ 2500 2400	$\begin{array}{c} (x_1,y_1) & (p_1,q_1) \\ \hline 2500 & 0.33 \\ \hline 2400 & 0.33 \\ \hline (x_1,y_1) & (p_1,q_1) \\ \hline 2500 & 0.33 \\ \hline 2400 & 0.34 \\ \end{array}$	$\begin{array}{c ccccc} (x_1,y_1) & (p_1,q_1) & (x_2,y_2) \\ \hline 2500 & 0.33 & 2400 \\ \hline 2400 & 0.33 & 2400 \\ \hline (x_1,y_1) & (p_1,q_1) & (x_2,y_2) \\ \hline 2500 & 0.33 & 0 \\ \hline 2400 & 0.34 & 0 \\ \hline \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$				

# Figure 22. Framing Effects and Violations of Rational Choice Theory

#### **Consistent Preferences under Risk**

**The Allais Paradox** 

#### **Consistent Preferences over Time**

**Present-Biased Preferences** 

	$(x_{1}, y_{1})$	$(p_{1}, q_{1})$	$(x_{2}, y_{2})$	$(p_{2},q_{2})$	$(x_{3}, y_{3})$	$(p_{3}, q_{3})$		$(x_{1}, y_{1})$	$(r_{1},t_{1})$	$(x_{2}, y_{2})$	$(r_{2},t_{2})$
f'	2500	0.33	2400	0.66	0	0.01	SS	100	0	0	1
$oldsymbol{g}'$	2400	0.33	2400	0.66	2400	0.01	LL	0	0	120	1

	$(x_{1}, y_{1})$	$(p_{1}, q_{1})$	$(x_{2}, y_{2})$	$(p_{2},q_{2})$	$(x_{3}, y_{3})$	$(p_{3},q_{3})$		$(x_{1}, y_{1})$	$(r_{1},t_{1})$	$(r_{2},t_{2})$	$(r_{2},t_{2})$
f	2500	0.33	0	0.66	0	0.01	SS	100	10	0	11
g	2400	0.33	0	0.66	2400	0.01	LL	0	10	120	11
		-							•		

The Allais paradox choice pattern is naturally explained by SWUP when the lotteries are presented as in the top of Figure 22. For the choice between f' and g', g' > f' under SWUP if and only if

 $\mu(2400,0)(2400-0)(0.01) > \mu(2400,2500)(2500-2400)(0.33)$ 

For the choice between *f* and *g*, f > g if and only if

 $\mu(2400,2500)(2500 - 2400)(0.335) > \phi(0.34, 0.33)(0.34 - 0.33)(2450).$ 

SWUP thus explains the Allais paradox if the comparison between payoffs of 2400 and 0 is sufficiently more salient than the comparison between 2400 and 2500, and if the comparison between 2400 and 2500 is sufficiently more salient than the 0.01 difference in the probability of winning.<sup>16</sup>

Empirical violations analogous to the Allais paradox (also known as the common consequence effect) in risky choice have been observed in choices over time. In one example of cancellation violations, Rao and Li (2011) presented experimental subjects with the two binary choices illustrated in Figure 23 (where payoffs are in apples):

**Figure 23. Violations of Cancellation** 

SS	5	0 days	-6	1 day	0	7 days
LL	0	0 days	-6	1 day	8	7 days

<sup>&</sup>lt;sup>16</sup>The explanation for the choice of g' over f' is the same here as in Bordalo et al. (2012). However, their explanation for the choice of f over g assumes that agents interpret the choice as between statistically independent prospects, and evaluate options f and g as if they recast the choices as involving comparisons of 2500 and 2400, 2500 and 0, 2400 and 0, and 0 and 0. It seems implausible that all possible comparisons are made (and particularly those between 2500 and 0 and between 2400 and 0) given the representation of f and g in Figure 22. Still more problematic are results reported in Leland (1994) demonstrating that choices vary systematically with theoretically inconsequential changes in the way they are described (and, in particular, due to changes in the alignment of payoffs and probabilities across alternatives) but are insensitive to whether the alternatives were statistically dependent or independent.

SS'	5	0 days
LL'	8	7 days

Cancellation implies that the outcome of losing 6 apples should not affect the decision since it is common to both options. In their study, Rao and Li observed that 84% of respondents chose *LL*, while only 34% of respondents chose *LL'*, thereby violating cancellation. This finding is problematic for major models of intertemporal choice (e.g. discounted utility theory and models of quasi-hyperbolic discounting) since these models satisfy cancellation. The choices of *LL* and *SS'* follow plausibly as a consequence of salience. In the first choice, the salience of receiving a payoff of 8 rather than 0 drives the choice of the patient option *LL* (just as the presence of a chance of ending up with \$0 rather than \$2400 drove the risk averse choice in the common consequence example). When the common outcome of -6 in 1 day is eliminated and the frame is changed to a minimalist presentation, it is the salience of the one week difference in time delays that dominates the evaluation.<sup>17</sup>

Finally, in the domain of consumer choice, recall from Section 3.3 that SWUP predicts a bias toward concentration (BTC) – a tendency to choose options with a few large advantages over alternatives with a larger number of smaller advantages (or reject options with a few large disadvantages in favor of alternatives with many small disadvantages). In a wide range of consumer contexts the seller has opportunities to manipulate the profile of advantages and disadvantages the consumer encounters. BTC implies, for example, that a personconsidering a newspaper or magazine subscription will be more likely to subscribe if the opportunity is framed in terms of the price per day, rather than an identical offer framed in terms of the price per year, as illustrated in Figure 24. In the figure, a decision maker chooses between an annual newspaper subscription with a fee of \$219 per year (option c) and an equivalent subscription offer advertised as 60 cents per day (option c'). For either option, suppose the full payment is made at the same time and the only difference between the offers is how they are presented to the

<sup>&</sup>lt;sup>17</sup> Violations of cancellation as in Figure 23 are also plausibly explained even if both pairs of consumption plans are framed as minimalist presentations.

consumer. BTC suggests that the agent will strictly prefer c' over c, even though the options are logically equivalent. Rather than changing the alignment of probabilities or time delays, Figure 24 changes the aggregation of payoffs by concentrating payments in option c and spreading payments in c'.

	$(x_{1}, y_{1})$	$(r_{1},t_{1})$	$(x_{2}, y_{2})$	$(r_{2},t_{2})$	$(x_{3}, y_{3})$	$(r_{3}, t_{3})$		$(x_{365}, y_{365})$	$(r_{365}, t_{365})$
С	-219	1	0	2	0	3		0	365
<i>c</i> ′	-0.60	1	-0.60	2	-0.60	3	•••	-0.60	365
				3.5.2	Labellin	ng Effec	ts		

**Figure 24.Consumer Purchase Framing Effect** 

Perhaps the most famousand troubling example of a framing effectisTversky and Kahneman's (1981) Asian Flu problem in which respondents are told that the U.S. is preparing for the outbreak of an epidemic which is expected to kill 600 people. Some are told that policy makers need to choose between two disease prevention strategies: Program A saves 200 lives. Program B has 1/3 chance of saving 600 people and 2/3 chance of saving no one. The parallel frame of this decision problem is given in the top of Figure 25.

A different group of respondents is toldthat policy makers need to choose between Programs C and D, and that if Program C is taken, then 400 people will die. If Program D is taken, then there is a 1/3 chance that no people will die and a 2/3 chance that all 600 people will die. The parallel frame of this decision is also given in Figure 25.

Programs A and C differ only in how the outcomes are labeled (as lives saved or lives lost) and are thus considered to be logically equivalent. The same observation holds for Programs B and D. These equivalences notwithstanding, most people chose Program A over B and Program D over C(Tversky and Kahneman, 1981) thereby exhibiting a framing effect.

For the SWUP model, the most salient column in the 'lives saved' presentation is the case where zero lives are saved, inducing risk aversion, and the most salient column in the 'lives lost' presentation is where zero lives are lost, inducing risk-seeking behavior.

Both of these observations plausibly follow from diminishing absolute sensitivity and the reflection property of the salience function and can easily be shown to hold under the Bordalo et al. (2012) salience function, (8).

SWUP can also explain other labeling effects such as the date-delay effect observed in Read et al. (2005): People are more patient when time is expressed in calendar dates rather than equivalent delays in weeks (or months). For example, 86% of their subjects preferred the smaller sooner (SS) payoff of \$370 in 17 weeks over the larger later (LL) payoff of \$450 in 56 weeks. However, 60% of subjects preferred the delayed \$450 when the same time periods were presented as calendar dates. These date and delay presentations are shown in the bottom panel of Figure 25. For a decision maker indifferent between SS and LL in the date frame, SWUP explains the shift in time preference in the date-delay effect if the comparison between 17 and 56 weeks is more salient than the comparison between the two calendar dates (which follows, for instance, as a consequence of ordering and DAS if salience weights for calendar dates are based on comparing the years 2003 and 2004).

	Gain-L	oss Framing	gEffect	
$(x_{1},y_{1})$		$(p_{1},q_{1})$	$(x_{2}, y_{2})$	$(p_{2},q_{2})$
Program A	200 lives saved	1/3	200 lives saved	2/3
Program B	600 lives saved	1/3	0 lives saved	2/3
(x <sub>1</sub>	y <sub>1</sub> )	$(p_{1}, q_{1})$	$(x_{2}, y_{2})$	$(p_{2},q_{2})$
Program C	400 people will die	1/3	400 people will die	2/3
Program D	0 people will die	1/3	600 people will die	2/3

Figure 25. Framing Effects due to Changes in the Labeling of Attributes

# **Date-Delay Framing Effect**

	$(x_1, y_1)$	Delay (Weeks)	(x <sub>1</sub>	1, <b>y</b> 1)	Date
SS	370	17	SS	370	September 26, 2003
LL	450	56	LL	450	June 25, 2004
					•

The observations in Figures 21 - 25suggest that SWUP may also provide the foundation for a theory of choice architecture since it predicts how different behaviors may be systematically elicited given changes in the alignment, aggregation, or labeling of attributes.

### 4. Taking Stock

The model of choice presentations introduced in Section 2 provides a formalization of how alternatives and attributes (payoffs, probabilities, time delays) are systematically framed, thereby making the analysis of framing effects tractable. Building on this visual representation of alternatives, we provided a model of context-dependent evaluation over presentations and introduced general properties of human salience perception which influence how attribute differences are evaluated. At the beginning of Section 3, we also identified simplealgebraic manipulations that can be performed on a choice presentation, thereby providing the basis for a calculus of framing effects.

Given a choice presentation between two alternatives, one can perform a small set of basic operations to produce other presentations. For a risk-neutral expected utility maximizer or a discounted utility maximizer with linear utility, risk and time preferences are invariant to all of the following elementary transformations of a presentation: (i) multiplying payoffs or attributes by the same positive scalar, (ii) adding the same constant to all payoffs or attribute values,(iii) changing the sign of payoffs, (iv) adding a common consequence to each alternative, (iv) changing the alignment of attributes in apresentation, and (vi) changing the labeling of payoffs or attribute values. However, systematic shifts in risk and time preferences have been observed under each of these transformations.

The basic arithmetic manipulations of a choice presentation noted in the previous paragraph provide a compact means of organizing the predictions of SWUP, and also provide a natural nomenclature for classifying many of the diverse behaviors observed for decisions under risk and over time. Table 1 provides a summary of the predictions of SWUP for choice alternatives framed as presented in Sections 3.2 - 3.5. The table also notes the basic arithmetic manipulation of a choice presentation to which the observed behavior corresponds.

As can be seen, Table 1 provides a short, but surprisingly comprehensive, summary of observed properties of risk and time preferences<sup>18</sup>.

Elementary Operation	The Etiology of Choice under Risk
Scale Payoffs by Constant	Increasing Relative Risk Aversion
Scale Probabilities by Constant	Common Ratio Effect (Risk)
Add Constant to Payoffs	Decreasing Absolute Risk Aversion
Change Sign of Payoffs	Reflection of Risk Attitudes
Add Common Consequence	Allais Paradox
Change Cell Alignment	Hidden Zero Effect (Risk)
Change Attribute Labels	Gain-Loss Framing Effect
Elementary Operation	The Etiology of Choice over Time
Elementary Operation Scale Payoffs by Constant	The Etiology of Choice over Time Absolute Magnitude Effect
Elementary Operation         Scale Payoffs by Constant         Scale Time Periods by Constant	The Etiology of Choice over TimeAbsolute Magnitude EffectCommon Ratio Effect (Time)
Elementary OperationScale Payoffs by ConstantScale Time Periods by ConstantAdd Constant to Payoffs	The Etiology of Choice over TimeAbsolute Magnitude EffectCommon Ratio Effect (Time)Proportional Magnitude Effect
Elementary OperationScale Payoffs by ConstantScale Time Periods by ConstantAdd Constant to PayoffsAdd Constant to Time Periods	The Etiology of Choice over TimeAbsolute Magnitude EffectCommon Ratio Effect (Time)Proportional Magnitude EffectCommon Difference Effect
Elementary OperationScale Payoffs by ConstantScale Time Periods by ConstantAdd Constant to PayoffsAdd Constant to Time PeriodsChange Sign of Payoffs	The Etiology of Choice over TimeAbsolute Magnitude EffectCommon Ratio Effect (Time)Proportional Magnitude EffectCommon Difference EffectSign Effect
Elementary OperationScale Payoffs by ConstantScale Time Periods by ConstantAdd Constant to PayoffsAdd Constant to Time PeriodsChange Sign of PayoffsAdd Common Consequence	The Etiology of Choice over TimeAbsolute Magnitude EffectCommon Ratio Effect (Time)Proportional Magnitude EffectCommon Difference EffectSign EffectViolations of Cancellation

Table 1. Elementary Operations and Behaviors Predicted by SWUP

<sup>&</sup>lt;sup>18</sup>Two behaviors in Table 1 which were not discussed earlier are the common ratio effect and the proportional magnitude effect for choices over time. The common ratio effect for time occurs when time periods are scaled up by a constant, in which case time preferences become less patient. The proportional magnitude effect occurs when a constant is added to all payoffs, in which case time preferences become less patient. Both are properties of all discounting models (Scholten and Read, 2010) and thus hold under SWUP even without relying on IPS or DAS.

### 4.1 Deriving Preferences from Salience Perception

Rather than derive a representation of preferences, we started with models of expected utility, discounted utility and, in a companion paper, subjective expected utility<sup>19</sup>in which utility is linear, the discount factor is constant, andthe subjective probability distribution is unique. We then showed that deviations from these models determined only by well-defined and plausible properties of salience perception can be used to formally derive behavioral properties that are typically attributed to the structure of preferences. Table 2 displays the resulting correspondence between properties of salience perception and properties of preferences. All results in Table 2 hold generally for any salience functions satisfying the specified properties, as shown in Propositions 5 - 14.

|--|

<b>Properties of Salience Perception</b>	<b>Properties of Preferences</b>
Ordering	Bias Toward Concentration (Risk)
Ordering	Bias Toward Concentration (Time)
Diminishing Absolute Sensitivity	Risk Aversion
Diminishing Absolute Sensitivity	Ambiguity Aversion
Diminishing Absolute Sensitivity	Present Bias
Increasing Proportional Sensitivity	Allais Common Ratio Effect
Increasing Proportional Sensitivity	Absolute Magnitude Effect
	Loss Aversion

<sup>&</sup>lt;sup>19</sup> The generalization of (4) to the Anscombe-Aumann (1963)subjective expected utility framework for choices under uncertainty is considered in a companion paper (Schneider and Leland, 2014) in which SWUP is applied to explain ambiguity aversion and the Ellsberg paradoxes as arising from the property of diminishing absolute sensitivity of the probability salience function.

Loss Sensitivity	
Loss Sensitivity	Sign Effect
Reflection	Reflection of Risk Attitudes
Relative Increasing Absolute Sensitivity	Decreasing Absolute Risk Aversion
Relative Decreasing Proportional Sensitivity	Increasing Relative Risk Aversion
Relative Decreasing Proportional Sensitivity	Size of Risk Aversion

#### 5. The Fourfold Pattern of Risk Attitudes

One robust behavioral pattern we have not yet explored is the fourfold pattern of risk attitudes identified by Tversky and Kahneman (1992) as one of the defining features of cumulative prospect theory (CPT). Tversky and Kahneman describe this pattern as "risk aversion for gains and risk-seeking for losses of high probability; risk-seeking for gains and risk aversion for losses of low probability" (p. 297). Under CPT, the fourfold pattern arises due to non-linear probability weighting, coupled with the prospect theory value function. Under SWUP, the fourfold pattern (FFP) arises due to the agent being risk-averse when the downside of taking a risk is salient and risk-seeking when the upside is salient, a property also shared with the salience model of Bordalo et al. (2012).

In this section we re-examine the data from Tversky and Kahneman (1992),demonstrating the fourfold pattern from the perspective of SWUP, and ask whether it can be plausibly explained by the simpler two-fold pattern of risk-seeking behavior when a lottery's upside is salient and risk-aversion when the downside is salient.

The data from Tversky and Kahneman (1992) isbased on responses of 25 experimental subjects who each madechoices between 56 binary lotteries and sure payoffs. Twenty-eight lotteries involved only non-negative outcomes; the remaining twenty-eight lotteries were obtained by reversing the sign of payoffs of the first twenty-eight lotteries. Tversky and Kahnemanpresentedboth the binary lottery and its expected value to each subject and asked subjects to choose between one of seven sure outcomes that they would prefer over the lottery. The subjects were subsequently asked to choose between a narrower range of certain payoffs and the same lottery to identify an indifference point.

In their data, Tversky and Kahneman (1992) found the median responses to be riskaverse for twenty-six lotteries, risk-seeking for twenty-eight lotteries, and risk-neutral for the remaining two lotteries. Thus, EUT with a concave utility function cannot explain more than half of the modal responses.

Here we apply SWUP to predict whether the modal response of the experimental subjects inTversky and Kahneman (1992) is either risk-averse or risk-seeking for each of

the 56 lotteries. This information can be represented as achoice between each lottery and its expected value. As noted in Sections 2 and 3, such a choice is naturally framed in a monotone parallel presentation since it involves a degenerate lottery. For U(x) = x, the payoffs and probabilities cancel in (4) and the predictions of SWUP are driven entirely by the salience weights for the upside of the lottery as compared to the salience weights for the downside of the lottery. Using the salience function (8) due to Bordalo et al. (2012) as a simple and plausible parametric form which satisfies basic properties of a salience function in Definition 1, we can compute an index reflecting the net upside or downside of a lottery. We do so by computing the salience weight favoring the lottery minus the salience weight favoring the expected value. SWUP predicts that when this net difference is positive, the upside of taking the risk exceeds the downside and the agent will be riskseeking (and will be risk-averse when the net difference is negative). The results of these calculations are depicted in Figure 26 which plots the net difference between the salient upside and salient downside of the lottery (for  $\theta = 1$ ) on the vertical axis and indicates whether the modal response of experimental subjects was risk-seeking or risk-averse. The lotteries for gains were each assigned a number 1 through 28, and the corresponding reflected lottery for losses was assigned the same number. This lottery number is plotted on the horizontal axis. The two green circles in the figure correspond to median responses which were risk-neutral. The lotteries above (below) the horizontal axis are predicted by SWUP to yield risk-seeking (risk-averse) choices.

Using just the parameter  $\theta$  in (8), SWUP predicts 26 of the 28 risk-seeking responses (92.8%) and 24 of the 26 risk-averse responses (92.3%) for  $\theta = 1$ .For  $\theta = 2$ , SWUP predicts 27 of the 28 risk-seeking responses (96.4%) and 25 of the 26 risk-averse responses (96.1%). When the five parameters of cumulative prospect theory were fit to the same data, CPT explains 91% of the median responses (Brandstatter et al., 2006). The data suggests that the fourfold pattern of risk attitudes can be more parsimoniously summarized as a twofold pattern.



Figure 26: A Fourfold Pattern or a Twofold Pattern of Risk Attitudes?

### 6. Related Literature

The leading descriptive model for decisions under risk is widely recognized to be Tversky and Kahneman's (1992) cumulative prospect theory (CPT). While CPT can explain most of the phenomena identified for decisions under risk, it does not explain alignment effects such as the hidden-zero effect. As a consequence, CPT cannot explain both the Allais paradox when presented as in the top panel of Figure 21 and the disappearance of the paradox when framed as a parallel presentation. CPT also does not explain behavior for decisions over time.<sup>20</sup>

The hyperbolic discounting model of Loewenstein and Prelec (1992) and the tradeoff model of Scholten and Read (2010) can account for present bias, the magnitude effect, and the sign effect, but cannot explain violations of cancellation or the hidden zero effect identified by Magen et al. (2008). Models of quasi-hyperbolic discounting (Laibson1997;

<sup>&</sup>lt;sup>20</sup>The inability to explain these types of behaviors will plague any model that assumes an absolute,rather than a comparative, evaluation process.

O'Donoghue and Rabin, 1999) explain present bias, but do not explain the other features of time preferences identified in Section 3.

Tversky (1969), Rubinstein (1988), and Leland (1994, 1998) among others, have proposed models of risky choice that involve ignoring small or similar differences between attributes across alternatives and attending to large, dissimilar ones. Like SWUP, these models imply that choices will be sensitive to the way alternatives are framed to the extent framing determines what is being compared and what attribute differences are perceived as decisive in determining choices. Leland (2002)and Rubinstein (2003) have demonstrated that similarity reasoning extends naturally to intertemporal choice. However, while this class of models provides a plausible explanation for many anomalous behaviors observed in risky and intertemporal choice, the models are too imprecise in that they do not clearly determine when two attributes are similar or dissimilar, and are too non-compensatory to provide an adequate depiction of behavior in general.

Recent models by Bordalo et al. (2012) and Koszegi and Szeidl (2013) are closely related to SWUP in that they assume that the salience of differences across alternatives influences choices through their impact on expected and discounted utility, respectively. The model in Bordalo et al. (2012) predicts many of the behaviors under risk predicted by SWUP while Koszegi and Szeidl's (2013) focus model is similar to SWUP in that it predicts bias toward concentration. However, neither model considers the possibility that agents' decisions under risk or over time might be swayed by the perceived salience of differences in probabilities or dates of receipt, respectively. As a result, they cannot explain behaviors such as the Ellsberg paradox or the common-difference effect that arise from salience perceptions on these dimensions.

#### 7. Discussion

We conclude by summarizing the main contributions of this paper. We began by introducing a new model of choice presentations and a context-dependent evaluation procedure that generalizes EUT and DUT by accounting for endogenously defined salient

payoffs, probabilities, and time delays. We demonstrated that observed behaviors toward risk and time, commonly attributed to the structure of preferences, can be formally 'derived' from basic properties of human salience perception, as shown in Propositions 5 - 14 and summarized in Table 2. In this respect, the principle of diminishing absolute sensitivity emerges as a unifying principle for decision making as it can be used to formally derive commonly observed risk aversion, present bias, and ambiguity aversion if it is satisfied by the salience functions for payoffs, time delays, and probabilities, respectively. SWUP also provides an explanation for why commonly observed choice anomalies, like the common ratio and common difference effect, arise when options are presented in minimalist form. We then turned to the problem of framing effects. Kreps (1988) notes that very little progress has been made in modeling framing effects, despite their potentially large influence on decisions. Our third contribution was to apply the model of choice presentations to help make the analysis of framing effects tractable. We applied SWUP to explain major framing effects observed in the literature such as the gain-loss framing effect typified in Tversky and Kahneman's (1981) Asian Flu problem, the stochastic dominance framing effect from Tversky and Kahneman (1986), Magen et al.'s (2008) hidden zero effect for choices over time, and an analogous hidden zero effect for choices under risk. Moreover, our analysis demonstrates that the latter three effects arise from the same underlying mechanism -a switch between parallel versus minimal presentations. By extension we have shown how violations of EUT and DUT can be systematically 'turned on' or 'turned off' by switching between parallel and minimalist frames, consistent with observations. Finally, we applied SWUP to experimental data from Tversky and Kahneman (1992) and found that it provides a simpler explanation of the fourfold pattern of risk attitudes than the one originally proposed.

In sum, we have presented a model that highlights deep parallels in observed behaviors for risky and intertemporal choice. In particular, we demonstrated that these behaviors arise from a common frame-dependent evaluation procedure influenced by basic properties of human salience perception.

#### References

- Allais, M. (1953). "Le comportement de l'hommerationneldevant le risque: critique des postulatsetaxiomes de l'écoleAméricaine". *Econometrica*, 21 (4): 503–546.
- Andreoni, J. and C. Sprenger (2012). "Risk Preferences Are Not Time Preferences". American Economic Review, 102(7):3357-3376.
- Anscombe, F. J., and R. Aumann (1963). "A Definition of Subjective Probability". *Annals* of Mathematical Statistics, 34(1): 199-205.
- Arrow, K.(1965). "The theory of risk aversion". In Aspects of the Theory of Risk Bearing, by YrjoJahnssoninSaatio, Helsinki.
- Arrow, K. (1970). "Aspects of the Theory of Risk-Bearing." In Essays in the Theory of Risk-Bearing. Amsterdam: North-Holland.
- Bernoulli, D. (1968). "Exposition of a New Theory of the Measurement of Risk". In Alfred Page, ed., Utility Theory: A Book of Readings. (originally publishedin 1738)
- Bordalo, P., N. Gennaioli, and A. Shleifer (2012). "Salience Theory of Choice under Risk". Quarterly Journal of Economics, 127: 1243-1285.
- Brandstätter, E., G. Gigerenzer, and R. Hertwig, (2006). "The priority heuristic: Making choices without trade-offs". Psychological Review. 2006. 113: 409-432.
- Camerer, C. and M. Weber (1992). "Recent developments in modeling preferences: Uncertainty and ambiguity". *Journal of Risk and Uncertainty*, 5 (4): 325-370.
- Frederick, S., G. Loewenstein, and T. O'Donoghue (2002). "Time discounting and time preference: A critical review". Journal of Economic Literature, 40: 351-401.
- Friedman, M. and L. Savage (1948). "The Utility Analysis of Choices Involving Risk". Journal of Political Economy, 56: 279-304.
- Gigerenzer, G., Todd, P. M., and the ABC Group (1999).Simple heuristics that make us smart. New York: Oxford University Press
- González-Vallejo, C. (2002). "Making trade-offs: A probabilistic and context-sensitive model of choice behavior". *Psychological Review*, 109(1):137-155.
- Holt, C. and S. Laury (2002)."Risk Aversion and Incentive Effects". American Economic Review, 92(5): 1644-1655.

- Kahneman, D. and A. Tversky, (1979). "Prospect Theory: An Analysis of Decisions Under Uncertainty". *Econometrica*, 47: 263-291.
- Koszegi, B. and A. Szeidl (2011)."A model of focusing in economic choice".
- Manuscript. UC Berkeley.
- Koszegi, B. and A. Szeidl (2013)."A model of focusing in economic choice". *Quarterly Journal of Economics*. 128, 53-104.
- Kreps. D. M. (1988). "Notes on the Theory of Choice." Boulder. Westview press.
- Laibson, D. (1997). "Golden Eggs and Hyperbolic Discounting". Quarterly Journal of Economics, 112(2): 443-478.
- Leland, J.(1994). "Generalized Similarity Judgments: An Alternative Explanation for
- Choice Anomalies". Journal of Risk and Uncertainty, 9:151-172.
- Leland, J.(1998). "Similarity Judgments in Choice Under Uncertainty: A
- Reinterpretation of Regret Theory". Management Science, 44(5): 1-14.
- Leland, J. (2001). "Similarity, Uncertainty and Time Tversky (1969) Revisited". IBM Research Report.RC22120 (99107).
- Leland, J. (2002). "Similarity Judgments and Anomalies in Intertemporal Choice". *Economic Inquiry*, 40(4), 574-581.
- Leland, J. (2010). "The Hunt for a Descriptive Model of Choice Under Risk a view from the road not taken". Journal of Socio-Economics - Special Issue: Topics on Contemporary Behavioral Economics, 39(5): 568-577.
- Leland, J. and P. Sileo (1998)."The "Nothing to Gain/Nothing to Lose" Effect Gambling and Insurance Purchase."Manuscript.
- Loomes, G. and R. Sugden, (1982). "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty". *The Economic Journal*, 92: 805-824.
- Loewenstein, G. and D. Prelec (1992). "Anomalies in Intertemporal Choice: Evidence and Interpretation". *The Quarterly Journal of Economics*,107: 573-597.
- Magen, E., C. Dweck, and J. Gross (2008). "The hidden-zero effect". *Psychological Science*, 19:648–649.

- Moyer, R. and T. Landauer (1967). "The Time Required For Judgments of Numerical Inequality". *Nature*, 215:1519-1520.
- O'Donoghue, T. and M. Rabin (1999). "Incentives for Procrastinators". *Quarterly Journal* of Economics, 114(3): 769-816
- Pratt, J. (1964). "Risk aversion in the small and in the large". *Econometrica* 32, January–April: 122–136.
- Prelec, D. and G. Loewenstein. (1991). "Decision Making Over Time and Under Uncertainty: A Common Approach". *Management Science*, 37(7): 770-786.
- Rabin, M. (2000). "Risk Aversion and Expected-utility Theory: A Calibration Theorem". *Econometrica*, 68: 1281–1292.
- Rao, L-L.& Li, S. (2011). "New paradoxes in intertemporal choice". Judgment and Decision Making. 6(2): 122-129.
- Read (2003). "Intertemporal Choice." Manuscript. The London School of Economics. Working Paper No.LSEOR 03.58.
- Read, D., S. Frederick, B. Orsel, and J. Rahaman (2005). "Four score and seven years from now: the date/delay effect in temporal discounting". *Management Science*.51(9): 1326-1335.
- Reyna, V. and C. Brainerd (1991). "Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion". *Journal of Behavior and Decision Making*, 4: 249–262.
- Rubinstein, A. (1988)."Similarity and Decision-making Under Risk (Is There a Utility Theory Resolution to the Allais Paradox?)" *Journal of Economic Theory*, 46, 145-153.
- Rubinstein, A. (2003). "Economics and Psychology"? The Case of Hyperbolic Discounting, *International Economic Review* 44, 1207-1216.
- Savage, L. J. (1954). The Foundations of Statistics.New York, Wiley.
- Schley, D. and E. Peters (in press). "Assessing 'economic value": Symbolic-Number Mappings Predict Risky and Riskless Valuations". *Psychological Science*.
- Schneider, M. and J. Leland (2014). "Salience, Ambiguity, and the Ellsberg Paradox,"

Manuscript.

- Scholten, M. and D. Read (2010). "The psychology of intertemporal tradeoffs". *Psychological Review*, 117(3): 925-944.
- Strotz, R. H. (1955). "Myopia and Inconsistency in Dynamic Utility Maximization". *Review of Economic Studies*, 23 (3): 165–180.
- Thaler, R. (1981). "Some Empirical Evidence on Dynamic Inconsistency". *Economic Letters*. 8(3): 201-207.
- Tversky, A. (1969) "The Intransitivity of Preferences". Psychological Review, 76, 31-48.
- Tversky A. and D. Kahneman (1981)." The framing of decisions and the psychology of choice". *Science*, 211 (4481): 453-458.
- Tversky A. and D. Kahneman (1986)."Rational Choice and the Framing of Decisions". *Journal of Business*, 59: S251-S278.
- Tversky, A. and D. Kahneman (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty". *Journal of Risk and Uncertainty*, 5(4): 297–323.
- Zeckhauser R. and E. Keeler (1970)."Another Type of Risk Aversion".*Econometrica*, 38(5): 661-665.