2017

Risk Preference, Time Preference, and Salience Perception

Jonathan Leland

Mark Schneider
Chapman University

Risk Preference, Time Preference, and Salience Perception

Comments
Working Paper 17-16

Risk Preference, Time Preference, and Salience Perception

Jonathan Leland **Mark Schneider**

A model of decision making is introduced that provides a unified approach to choices under risk and over time. This model of salience weighted utility over presentations (SWUP) predicts systematic departures from expected utility and discounted utility using the same mathematical structure and the same psychological intuition. SWUP explains six parallels between risk and time: (i) the common ratio effect and common difference effect; (ii) the common consequence effect and cancellation effect; (iii) the peanuts effect and magnitude effect; (iv) the fourfold pattern of risk attitudes and bias toward concentration; (v) a downside risk framing effect and opportunity cost framing effect; and (vi) the gain-loss labeling effect and the date-delay labeling effect. SWUP assigns weights on utility differences that depend on their importance (their probabilities of occurrence and their proximities of occurrence) and their salience or similarity, and so provides a bridge between rational and heuristic representations of decision making.

1. Introduction

In 1937 Paul Samuelson proposed that individuals might choose between consumption plans such that they maximize their discounted payoff or "utility" (DU) from future consumption. A decade later, von Neumann and Morgenstern (1947) proposed that people choose between risky prospects as if they maximize their expected utility (EU). To this day, these models constitute the standard theories of rational choice over time and under risk in the social and behavioral sciences. Nevertheless, almost from their inception, persistent questions have been raised regarding their descriptive accuracy. Friedman and Savage (1948) objected to the standard assumption in Expected Utility theory that people are uniformly risk averse while Allais (1953) raised questions regarding the descriptive accuracy of the EU independence axiom. Strotz (1955) critiqued Discounted Utility theory, expressing concern that people would be unable to commit themselves to future plans resulting in time-inconsistent choices.

The challenges to the descriptive adequacy of EU and DU raised by Allais, Strotz, and others have defined the research agenda on individual decision making for over a half a century. Alternative models have been proposed for risky choice while a very different set of models has been proposed for intertemporal choice to accommodate the EU and DU anomalies. However, relatively few models have emerged to explain behaviors in both domains, potentially because risk and time preferences are thought to reveal very different aspects of behavior: Time preferences reveal some measure of impatience, while risk preferences reveal an individual characteristic of risk-tolerance.

More recently, Prelec and Loewenstein (1991) have pointed out parallels between anomalies under risk and over time, suggesting that there might be a fundamental link between behaviors in these domains. Subsequent work suggests the link involves the way differences in payoffs and probabilities and payoffs and dates of receipt across alternatives are evaluated and perceived. Rubinstein (1988) noted that an

individual choosing to receive \$3000 over a lottery offering an 80% chance of winning \$4000 might, in violation of the independence axiom, choose a lottery offering a 20% chance of \$4000 over a 25% chance to win \$3000 because the probabilities of 0.25 and 0.20 appear similar in value. Leland (2002) noted that, by the same reasoning, an individual choosing to receive \$20 in 1 month over receiving \$25 in 2 months might, in violation of the stationarity axiom of DU, choose \$25 in 12 months over \$20 in 11 months if a delay of 11 months appears similar to a delay of 12 months. Recent work by Bordalo et al. (2012) on the influence of salient payoff differences in risky choice and Koszegi and Szeidl (2013) on the impact of focusing on large attribute differences in intertemporal choice lends further credence to the idea that differential weighting of attribute differences may connect risky and intertemporal behavior.

Building on this possibility, we propose here a unified approach to modeling risky and intertemporal choice in which behavior in both domains is driven by the interaction between *properties of preferences* (which are assumed to satisfy expected utility for risk and discounted utility for time), *properties of salience perception* (which are assumed to satisfy diminishing sensitivity to absolute changes and increasing sensitivity to proportional changes in attribute values), and the *frame* in which choices are presented (which determines what differences are compared across alternatives). We then show that simple operations on a frame generate a comprehensive set of behaviors observed under risk and over time.

2. Salience Weighted Utility over Presentations

We proceed to model choices in three steps, first providing a formal theory for the framing of risky and intertemporal choices, then specifying the evaluation process or computational decision algorithm agents employ to choose between alternatives, and then characterizing the nature of salience perceptions that drive the evaluation. Taken together, we refer to our model of the choice process, characterization of the perceptual system, and treatment of frames as Salience Weighted Utility over Presentations (SWUP).

2.1 Frames for Lotteries and Consumption Plans

For the purposes of developing a comparative model of risky and intertemporal choice we begin by representing the options in a *presentation* or *frame* as shown in Figure 1.

Figure 1. Presentations or "Frames" for Decisions under Risk and over Time

Choice Frame 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)
p	x_1	p_1	x_2	p_2	…	x_i	p_i	…	x_n	p_n
q	y_1	q_1	y_2	q_2	…	y_i	q_i	…	y_n	q_n

Choice Frame 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)
r	x_1	r_1	x_2	r_2	…	x_i	r_i	…	x_n	r_n
t	y_1	t_1	y_2	t_2	…	y_i	t_i	…	y_n	t_n

In our setup, for decisions under risk, there is a finite set, X, of outcomes. A *lottery* is a mapping $p: X \rightarrow [0,1]$ such that $\sum_{x \in X} p(x) = 1$. Denote the set of all lotteries by $\Delta(X)$. We consider one-dimensional arrays $\mathbf{p}$ and $\mathbf{q}$ which represent lotteries p and q (in a well-defined sense) that offer a finite and equal number of outcomes denoted $\mathbf{x_i}$ and $\mathbf{y_i}$, $i = 1,2, \dots, n$, where each $\mathbf{x_i}$ occurs with probability $\mathbf{p_i}$ and each $\mathbf{y_i}$ occurs with probability $\mathbf{q_i}$. Denote the support of a lottery, p (the set of outcomes for which $p(x) > 0$) by $supp(p)$, and denote the number of outcomes in the support of a lottery p by $|supp(p)|$.

For decisions over discrete time periods t ∈ {0, 1, 2, …, T}, a *consumption plan, r,* is a sequence of dated outcomes in X. Denote the set of consumption plans by C. We study choices between consumption plans *r* and *t* where $r := (x_0, x_1, \dots, x_T)$ and $t := (y_0, y_1, \dots, y_T)$. We consider one-dimensional arrays **r** and **t** which represent consumption plans r and t that offer a finite and equal number of outcomes, where each outcome $\mathbf{x_i}$ occurs in time period, $\mathbf{r_i}$ and each $\mathbf{y_i}$ occurs in period $\mathbf{t_i}$, for all i. We use bold font for attributes in an array and italicized font for attributes in the support of a lottery or consumption plan.

To represent a lottery, p, we employ a one-dimensional array, **p**, consisting of $n(\mathbf{p})$ outcomes and $n(\mathbf{p})$ corresponding probabilities. Denote the ith outcome in **p** and the ith corresponding probability by $\mathbf{x_i}$ and $\mathbf{p_i}$, respectively. Notice that outcome-probability pairs appear in no particular order in the array and some outcomes could be repeated in the array. Hence, many arrays could represent the same lottery.

Definition 1 (Representation of a lottery): An array **p** is a *representation* of lottery p if (i) and (ii) hold:

(i) For i =1,2,…,$n(\mathbf{p})$, $\sum_{\mathbf{i}} \mathbf{p_i} = 1$

(ii) For all i such that $\mathbf{x_i} = x$, $\sum_{\mathbf{i}} \mathbf{p_i} = p(x)$.

Note that a representation **p** of lottery p differs from the lottery itself since it permits the same outcome to appear more than once in the array provided that the corresponding probabilities sum to the overall probability of that outcome. Representations of two lotteries presented jointly, constitute a frame.

Definition 2 (Frame for lotteries): A *presentation* or *frame*, $[\![\mathbf{p}, \mathbf{q}]\!]$ of lotteries, p and q, is a matrix containing a representation **p** of p and a representation **q** of q.

In our analysis, we will consider cases where both representations in a frame have the same dimension, although this dimension can vary across frames.

Similarly, let any $r \in C$ be represented in a one-dimensional array, **r**, consisting of $m(\mathbf{r})$ outcomes and $m(\mathbf{r})$ periods. Denote the ith outcome in **r** and the ith corresponding period by $\mathbf{x_i}$ and $\mathbf{r_i}$, respectively.

Definition 3 (Representation of a consumption plan): Array **r** is a *representation* of $r \in C$ if (i)-(iii) hold:

(i) For any dated outcome, $x_t \neq 0$, the pair $(\mathbf{x}, \mathbf{t})$ is in **r** if and only if x_t is in r.

(ii) For any dated outcome $x_t = 0$, if the pair $(\mathbf{x}, \mathbf{t})$ is in **r** then x_t is in r.

(iii) If there are T dated outcomes in r, then dim(**r**) $\leq$ 2T.

A representation $\mathbf{r}$ of consumption plan r differs from the consumption plan itself since it permits periods of zero consumption to be 'compressed' (an outcome of zero and its corresponding time period might not be in $\mathbf{r}$). In addition, there is no restriction on the order in which the outcomes in r appear in $\mathbf{r}$. Finally, property (iii) implies that $\mathbf{r}$ contains at most all T outcomes and the corresponding T periods. When two representations of different consumption plans are presented together, they constitute a frame.

Definition 4 (Frame for consumption plans): A *presentation* or *frame*, $[\![\mathbf{r},\mathbf{t}]\!]$ of consumption plans, r and t, is a matrix containing a representation $\mathbf{r}$ of r and a representation $\mathbf{t}$ of t.

We next consider the framing of degenerate lotteries (those yielding a single outcome with probability 1). Consider a choice between a lottery p yielding *x* with certainty and a non-degenerate lottery q. It seems almost unavoidable that one compares each outcome in q to the unique outcome in p. We adopt the convention that a choice with a degenerate lottery is framed as in Figure 2.

Figure 2. Choice Frame with a Degenerate Lottery

	$(\mathbf{x_1},\mathbf{y_1})$	$(\mathbf{p_1},\mathbf{q_1})$	$(\mathbf{x_2},\mathbf{y_2})$	$(\mathbf{p_2},\mathbf{q_2})$	…	$(\mathbf{x_i},\mathbf{y_i})$	$(\mathbf{p_i},\mathbf{q_i})$	…	$(\mathbf{x_n},\mathbf{y_n})$	$(\mathbf{p_n},\mathbf{q_n})$
$\mathbf{p}$	$\mathbf{x}$	$\mathbf{p_1}$	$\mathbf{x}$	$\mathbf{p_2}$	…	$\mathbf{x}$	$\mathbf{p_i}$	…	$\mathbf{x}$	$\mathbf{p_n}$
$\mathbf{q}$	$\mathbf{y_1}$	$\mathbf{p_1}$	$\mathbf{y_2}$	$\mathbf{p_2}$	…	$\mathbf{y_i}$	$\mathbf{p_i}$	…	$\mathbf{y_n}$	$\mathbf{p_n}$

We present a model of decision making in which choices are predicted given the frame. In illustrating our model, we clearly specify the frame, employing the simplest presentation of choices for each example where outcomes of a non-degenerate lottery are monotonically ordered for choices under risk and in which time periods are monotonically ordered for choices over time.

2.2 Salience Weighted Utility for Decisions under Risk

Let there be a preference relation, $\succ$, over $\Delta(X)$ reflecting the decision maker's preferences over lotteries. The decision maker is assumed to have standard expected utility preferences: For all lotteries $p, q \in \Delta(X)$,

(1) $$p \succ q \text{ if and only if } \sum_{x\in X} p(x)u(x) > \sum_{y\in X} q(y)u(y).$$

This is the conventional approach. Next, we ask how might an individual evaluate *presentations* of lotteries like the ones in Figure 1? An extensive literature in economics and psychology has demonstrated that even small changes in presentations can have consequential effects on behavior. To allow for this possibility, let i = 1, 2, …,n index the location of the i^{th} attribute (payoff, probability) in a frame.

The decision maker is given two arrays and is asked to choose one. Here we consider a second relation $\gtrsim$ over representations conditional on the frame which may be viewed as a 'perceptual relation' (rather than a preference relation). That is, $\mathbf{p} \gtrsim \mathbf{q}$ ($\mathbf{p} \approx \mathbf{q}$) means that lottery p 'looks strictly better than' ('looks equally as good as') lottery q when a frame $[\![\mathbf{p},\mathbf{q}]\!]$ presents the lottery pair $\{p, q\}$.

For the generic frame in Figure 1, given (1), an unbiased perceptual relation can be expressed as:

(2) $\mathbf{p} \mathrel{\hat{\succ}} \mathbf{q}$ if and only if $\sum_{i=1}^{n} \mathbf{p_i} u(\mathbf{x_i}) > \sum_{i=1}^{n} \mathbf{q_i} u(\mathbf{y_i})$,

for all $\mathbf{p}, \mathbf{q}$, such that $\mathbf{p}$ is a *representation* of p and $\mathbf{q}$ is a *representation* of q and for all $p, q \in \Delta(X)$.

To account for the role of salience perception in decision making, we let the decision maker use a "salient" evaluation defined over frames when deciding between alternatives. In particular, salient comparisons between alternatives may frustrate the expression of the decision maker's true preferences. The decision maker may instead be systematically swayed by changes in the arrangement of attributes in a frame which systematically make some comparisons focal and which make others inconsequential. Rather than assuming that choices always 'reveal preferences,' our approach admits two possibilities whenever the relations $\succ$ and $\hat{\succ}$ are not equivalent: In any given situation (i) choices may reveal preferences (consistent with $\succ$) or (ii) they may reveal systematic deviations from preference maximization due to biases in salience perception (consistent with $\hat{\succ}$).

To further motivate the possibility that frames may frustrate the expression of preference, note that equations (1) and (2) provide an *alternative-based evaluation* - one lottery is strictly preferred to another, if and only if it yields a greater expected payoff to the decision maker. Here we allow for the possibility that agents choose by making across-lottery comparisons of payoffs and their associated probabilities of occurrence. Building on Leland and Sileo (1998), note that the alternative-based evaluation in (2) may be written equivalently as an *attribute-based evaluation* such that $\mathbf{p} \mathrel{\hat{\succ}} \mathbf{q}$ if and only if (3) holds:

(3) $$\sum_{i=1}^{n} [(\mathbf{p_i} - \mathbf{q_i})(u(\mathbf{x_i}) + u(\mathbf{y_i}))/2 + (u(\mathbf{x_i}) - u(\mathbf{y_i}))(\mathbf{p_i} + \mathbf{q_i})/2] > 0.$$

Note that the evaluation procedures in (2) and (3) operate over frames rather than over lotteries directly. Nevertheless, (2) and (3) both characterize frame-invariant preferences (they are not sensitive to changes in frames). The attribute-based evaluation in (3) computes 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. A decision maker who chooses among representations according to the attribute-based evaluation in (3) will be indistinguishable from one who chooses according to the alternative-based evaluation in (2). But now suppose that in the process of comparing risky alternatives an agent notices when the payoff in one alternative is "a lot more money" than the payoff in another and 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 across different alternatives are perceived as particularly salient or attract disproportionate attention and are overweighted in the evaluation process. To capture this intuition that larger differences in attributes are over-weighted or attract disproportionate attention, we place salience weights $\phi(\mathbf{p_i}, \mathbf{q_i})$ on probability differences and $\mu(\mathbf{x_i}, \mathbf{y_i})$ on payoff differences, yielding the following model of *salience-weighted evaluation,* in which $\mathbf{p} \mathrel{\hat{\succ}} \mathbf{q}$ if and only if (4) holds:

(4) $$\sum_{i=1}^{n} [\phi(\mathbf{p_i}, \mathbf{q_i})(\mathbf{p_i} - \mathbf{q_i})(u(\mathbf{x_i}) + u(\mathbf{y_i}))/2 + \mu(\mathbf{x_i}, \mathbf{y_i})(u(\mathbf{x_i}) - u(\mathbf{y_i}))(\mathbf{p_i} + \mathbf{q_i})/2] > 0.$$

We refer to model (4) as Salience Weighted Utility over Presentations (SWUP). Note that SWUP has a dual interpretation as a model of *salience-based choice* that overweights large differences and as a model of *similarity-based choice* that underweights small differences.

2.3 Salience Weighted Utility for Decisions over Time

The model extends analogously to choices over time such as those in the lower panel of Figure 1. Let $i = 1,2,\ldots,n$ denote the position of the i^th^ attribute (payoff, time period) in a frame. For decisions over time periods $t \in \{0, 1, 2, \ldots, T\}$, a *consumption plan, r*, is a sequence of dated outcomes in X. Denote the set of consumption plans by C. Let $r := (x_0, x_1, \ldots, x_T)$ and $t := (y_0, y_1, \ldots, y_T)$. As in the discounted utility model, a decision maker has a preference relation, $\succ_t$, over consumption plans such that for all $r, t \in C$,

(5) $$r \succ_t t \text{ if and only if } \sum_{t=0}^{T} \delta^t u(x_t) > \sum_{t=0}^{T} \delta^t u(y_t),$$

where $\delta \in (0,1]$ is a constant discount factor.

The decision maker is now given two arrays and is asked to choose one. Here we consider the relation $\hat{\succ}_t$ over representations of consumption plans, given the frame. The relation $\hat{\succ}_t$ may be viewed as a 'perceptual relation' (i.e., $\mathbf{r} \hat{\succ}_t \mathbf{t}$ ($\mathbf{r} \hat{\sim}_t \mathbf{t}$) means that r 'looks strictly better than' ('looks equally as good as') t when a frame $[\![\mathbf{r}, \mathbf{t}]\!]$ presents the pair of consumption plans $\{r, t\}$).

For the generic frame in Figure 1, given (5), a perceptual relation can be expressed as follows:

(6) $$\mathbf{r} \hat{\succ}_t \mathbf{t} \text{ if and only if } \sum_{i=1}^{m} \delta^{\mathbf{r_i}}[u(\mathbf{x_i})] > \sum_{i=1}^{m} \delta^{\mathbf{t_i}}[u(\mathbf{y_i})],$$

for all $\mathbf{r}, \mathbf{t}$, such that $\mathbf{r}$ is a *representation* of r and $\mathbf{t}$ is a *representation* of t and for all $r, t \in C$.

The alternative-based evaluation in (6) is equivalent to an attribute-based evaluation in which $\mathbf{r} \hat{\succ}_t \mathbf{t}$ if and only if (7) holds:

(7) $$\sum_{i}^{m}[(\delta^{\mathbf{r_i}} - \delta^{\mathbf{t_i}})(u(\mathbf{x_i}) + u(\mathbf{y_i}))/2 + (u(\mathbf{x_i}) - u(\mathbf{y_i}))(\delta^{\mathbf{r_i}} + \delta^{\mathbf{t_i}})/2] > 0.$$

Placing salience weights $\theta(\mathbf{r_i}, \mathbf{t_i})$ on time differences and $\mu(\mathbf{x_i}, \mathbf{y_i})$ on payoff differences, yields a salience-weighted evaluation in which $\mathbf{r} \hat{\succ}_t \mathbf{t}$ if and only if:

(8) $$\sum_{i}^{m}[\theta(\mathbf{r_i}, \mathbf{t_i})(\delta^{\mathbf{r_i}} - \delta^{\mathbf{t_i}})(u(\mathbf{x_i}) + u(\mathbf{y_i}))/2 + \mu(\mathbf{x_i}, \mathbf{y_i})(u(\mathbf{x_i}) - u(\mathbf{y_i}))(\delta^{\mathbf{r_i}} + \delta^{\mathbf{t_i}})/2] > 0.$$

We refer to an agent who chooses according to salience-based evaluation models (formulas 4 and 8) as a *focal thinker* since such an agent focuses on salient differences in attributes. Such an agent chooses the alternative which 'looks better', with respect to that agent's perceptual system. Note that whether one alternative 'looks better' than another to an agent, according to $\hat{\succ}$, depends on the agent's risk preferences and time preferences (as measured by u and δ, respectively) *and* the salience agents ascribe to large versus small differences in payoffs, time delays, and probabilities (as measured by μ, θ, and ϕ, respectively). In this respect, SWUP provides a bridge between economic and psychological approaches to decision making by modeling choice as dependent on both properties of preferences and properties of the perceptual system.

In SWUP, differences in attributes are weighted by their salience and by their importance. The earlier similarity models (e.g., Rubinstein 1988, 2003; Leland 1994, 2002) evaluated differences only by their similarity or salience and so are very non-compensatory. In contrast, the classical economic models weight attributes by their 'importance', where payoffs are effectively assigned higher importance if they are more likely to occur (represented by their probability weight in expected utility theory) or if they are nearer to the present (represented by their discount factor in discounted utility theory). In SWUP, the probability differences $(\mathbf{p_i} - \mathbf{q_i})$ are assigned an 'importance weight' of $(u(\mathbf{x_i}) + u(\mathbf{y_i}))/2$ (probability differences are more important if they correspond to larger average payoffs), and likewise utility differences are weighted by $(\mathbf{p_i} + \mathbf{q_i})/2$ (assigning greater importance to payoffs with larger average probabilities) for decisions under risk and by $(\delta^{\mathbf{r_i}} + \delta^{\mathbf{t_i}})/2$ (assigning greater importance to payoffs with larger average discount factors) for decisions over time. In this respect, SWUP provides a unified approach to heuristic models of attribute-based choice and to classical models of rational choice since SWUP includes both a 'salience weight' and an 'importance weight' on differences in payoffs, probabilities, and delays.

2.4 The Nature of Salience Perceptions

The salience functions μ, ϕ and θ determine the only ways in which a focal thinker differs from a rational agent who chooses over arrays according to formulas 2 and 6. We assume a salience function exhibits the two properties of the perceptual system in Definition 5, defined by Bordao et al. (2013).

Definition 5 (Salience Function): A *salience function* $\sigma(\mathbf{a}, \mathbf{b})$ is any (non-negative), symmetric[1] and continuous function that satisfies the following two properties:

1. **Ordering:** If $[\mathbf{a}', \mathbf{b}'] \subset [\mathbf{a}, \mathbf{b}]$ then $\sigma(\mathbf{a}', \mathbf{b}') < \sigma(\mathbf{a}, \mathbf{b})$.
2. **Diminishing Absolute Sensitivity (DAS):** For any $\mathbf{a}, \mathbf{b}, \boldsymbol{\epsilon} > 0$, $\sigma(\mathbf{a} + \boldsymbol{\epsilon}, \mathbf{b} + \boldsymbol{\epsilon}) < \sigma(\mathbf{a}, \mathbf{b})$.

While properties 1 and 2 define a salience function, consistent with Bordalo et al. (2013), we will also allow for the possibility that a salience function satisfies a third property, increasing proportional sensitivity:

Increasing Proportional Sensitivity (IPS): For any $\mathbf{ab} > 0$ and any $\boldsymbol{\alpha} > 1$, $\sigma(\boldsymbol{\alpha}\mathbf{a}, \boldsymbol{\alpha}\mathbf{b}) > \sigma(\mathbf{a}, \mathbf{b})$.

There is an intuitive relationship between these properties: DAS implies for a fixed absolute difference, that the perceptual system is more sensitive to larger ratios. IPS implies for a fixed ratio, that perception is more sensitive to larger absolute differences. Ordering implies that perception is more sensitive to larger intervals (which have both a larger absolute difference and a larger ratio than the smaller intervals they contain).

Ordering, DAS, and IPS are also supported by the psychology literature. Ordering is consistent with the "symbolic distance" effect - it takes longer to answer which of two numbers is larger, the smaller

[1] A function $f(\mathbf{x}, \mathbf{y})$ is symmetric if $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$.

their arithmetic difference.[2] Since the Weber-Fechner law was introduced in the 19th century, it has also been recognized that diminishing (absolute) sensitivity is a fundamental property of the perceptual system that applies across a range of sensory modalities including tone, brightness, and distance (Stevens, 1957). Schley and Peters (2014) provide experimental support that diminishing sensitivity also characterizes how the brain maps symbolic numbers onto mental magnitudes. Support for the IPS property can be found in the marketing literature, where IPS is referred to as "the unit effect". Pandelaere et al. (2011), for example, find that the perceived difference between ratings of 7 and 9 on a 0-10 scale appears smaller than the difference between 700 and 900 on a 0-1,000 scale. Similarly, Wertenbroch et al. (2007) report that, even when \$1 (in US currency) equals S\$1.70 in Singapore currency, "a target price of S\$1.70 will appear as less expensive when evaluated against a target budget of S\$17.00 than \$1 against \$10" (p. 3).

3. Parallels between Behavior under Risk and over Time

We now consider the implications of SWUP for risky choice and intertemporal choice. We show that even a *parameter-free version* of the model simultaneously predicts a wide variety of analogous behaviors between risk and time. To begin, we let $u(x) = x$, and we adopt the following parameter-free specifications for the salience functions that satisfy ordering and diminishing sensitivity:

$$\text{(9)} \qquad \mu(\mathbf{x}, \mathbf{y}) = \frac{|\mathbf{x} - \mathbf{y}|}{|\mathbf{x}| + |\mathbf{y}| + \mathbf{1}}, \qquad \phi(\mathbf{p}, \mathbf{q}) = \frac{|\mathbf{p} - \mathbf{q}|}{\mathbf{p} + \mathbf{q}}, \qquad \theta(\mathbf{r}, \mathbf{t}) = \frac{|\mathbf{r} - \mathbf{t}|}{\mathbf{r} + \mathbf{t}}$$

The payoff salience function μ is the only one of the salience functions that may take negative values as its arguments, and there also might be two zero payoff values in the same column vector of a frame. The specification for μ ensures that the salience function[3] is defined even if $\mathbf{x} = \mathbf{y} = 0$. It also satisfies increasing proportional sensitivity. In our analysis, we do not embed a constant in the salience functions for probabilities or delays for simplicity, although doing so would not affect our qualitative results. Each of the specifications in (9) is a special case of the salience function $\sigma(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|/(|\mathbf{x}| + |\mathbf{y}| + \lambda)$ proposed by Bordalo et al. (2012), in which $\lambda \geq 0$. Having fixed the utility and salience functions, the time discount factor δ (which we let be the annual discount factor) is the only parameter that may vary, and so we report the full range of δ values consistent with each behavior for each example. Somewhat surprisingly, our 'parameter-free' specification of the salience functions is sufficient to predict all behaviors studied in this paper, even with linear utility, and even if δ is fixed at a plausible annual discount factor (e.g., $\delta = 0.9$). We provide a more formal treatment of the behaviors discussed in this section in the appendix.

[2] 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?"

[3] The time salience function might also compare two payoffs in period 0 and so we define $\theta(\mathbf{r}, \mathbf{t})$ as in (9) when it is not the case that $\mathbf{r} = \mathbf{t} = 0$ and we set $\theta(\mathbf{0}, \mathbf{0}) = 0$.

3.1 The Common Ratio Effect and the Common Difference Effect

We begin with two of the best-known anomalies for choices under risk and over time: the special case of the common ratio effect known as the *certainty effect* and the special case of the common difference effect known as *present bias*. The certainty effect is a systematic violation of expected utility theory. A classic example due to Kahneman and Tversky (1979) is illustrated in the left panel of Figure 3.

Figure 3. The Common Ratio Effect (Risk) and the Common Difference Effect (Time)

Common Ratio Effect

	$	**prob.**	**$**	**prob.**
p	**3000**	**0.80**	**3000**	**0.20**
q	4000	0.80	0	0.20
p′	3000	0.25	0	0.75
q′	**4000**	**0.20**	**0**	**0.80**

Common Difference Effect

	$	**month**
SS	**510**	**0**
LL	530	2
SS′	510	24
LL′	**530**	**26**

In the figures in Section 3, the typical choice patterns are highlighted in bold font. In Figure 3, a decision maker chooses between lotteries p and q and between p′ and q′. Lottery p offers $3000 with certainty. Lottery q offers an 80% chance of receiving $4000. Expected utility theory predicts a decision maker will choose either p and p′ or q and q′. However, most subjects choose p and q′. Under SWUP, this pattern of choices is predicted since the salient comparison is between $3000 and $0 in the choice between p and q which favors p, and the salient comparison is between 3000 and 4000 in the choice between p′ and q′ which favors q′. The preference for p and q′ is also predicted by our parameter-free specification in (9).

The right panel of Figure 3 displays an example of present bias based on Read and Scholten (2012). A decision maker chooses between a ‘smaller sooner’ (SS) option and a ‘larger later’ (LL) option when the SS payment is immediate and when both the SS and LL payments are delayed by 24 months. Discounted utility theory predicts a decision maker who prefers $510 today (SS) to $530 in two months (LL) will also prefer $510 in 24 months (SS′) to $530 in 26 months (LL′). However, in such choices, most people choose SS and LL′. Under SWUP, the salient comparison is between receiving a payment in period 0 versus in two months for the choice between SS and LL, but the salient comparison is between $510 and $530 for the choice between SS′ and LL′. The choice of SS and LL′ is also predicted by our parameter-free specification in (9) for all $\delta \in [0.90, 0.99]$. We define the common difference effect more generally in the appendix and prove the following result which holds for any utility functions and any salience functions.

Proposition 1: *A focal thinker exhibits the common difference effect if and only if* θ *satisfies DAS.*

Note that under SWUP, the certainty effect and present bias are both driven by a comparison against 0 (comparing $3000 to $0 and comparing a zero time delay to a 2 month delay).

3.2 The Common Consequence Effect and the Cancellation Effect

A second parallel between risk and time preferences concerns the common consequence effect for decisions under risk (Allais, 1953) and the cancellation effect for choices over time (Rao and Li, 2011). Figure 4 displays a prototypical example of each effect. The common consequence effect violates the independence axiom of expected utility, while the cancellation effect violates a similar independence condition (the cancellation axiom) of discounted utility theory.

Figure 4. The Common Consequence Effect (Risk) and the Cancellation Effect (Time)

Common Consequence Effect (Risk)

	$	**prob.**	**$**	**prob.**	**$**	**prob.**
p	2500	0.33	2400	0.66	0	0.01
q	**2400**	**0.33**	**2400**	**0.66**	**2400**	**0.01**

p′	**2500**	**0.33**	**0**	**0.67**
q′	2400	0.34	0	0.66

Cancellation Effect (Time)

	$	**month**	**$**	**month**	**$**	**month**
SS	510	0	-520	1	0	2
LL	**0**	**0**	**-520**	**1**	**530**	**2**

SS′	**510**	**0**
LL′	530	2

In the common consequence example from Kahneman and Tversky (1979), a decision maker chooses between lotteries p and q and between lotteries p′ and q′. Lottery p offers $2400 with certainty, and q offers a 33% chance of $2500, a 66% chance of $2400 and a 1% chance of $0. Expected utility theory predicts a decision maker will choose either p and p′ or q and q′. However, most subjects choose p and q′. Under SWUP, this pattern of choices is predicted since the salient comparison is between $2400 and $0 in the choice between p and q, and the salient comparison is between 2500 and 2400 in the choice between p′ and q′. That is, the comparison between $2400 and $0 that is cued in the first choice is not cued in the second choice. The choice of p and q′ is also predicted by our parameter-free specification in (9).

The right panel of Figure 4 displays a prototypical example of the cancellation effect. In such examples, SWUP predicts a shift in preference from the sooner to the delayed option when the common consequence of -$520 in 1 month is added to both options. Similar preference reversals were observed by Rao and Li (2011). A violation of cancellation holds if SS′ is chosen over LL′, but LL is chosen over SS. The preference for SS′ and LL is also predicted[4] by our specification in (9) for all $\delta \in [0.80, 0.99]$.

[4] Our results are robust to whether the choice between SS and LL is represented as in the more compact frame shown below:

SS	510	0	-520	1
LL	**-520**	**1**	**530**	**2**

3.3 The Peanuts Effect and the Magnitude Effect

To further investigate the relationship between risky and intertemporal choices as formalized by SWUP, we turn to an unresolved paradox in the literature: In their ‘common approach’ to risk and time, Prelec and Loewenstein (1991) could not simultaneously explain the peanuts effect in risky choice (Markowitz, 1952) and the magnitude effect in intertemporal choice. Prototypical examples of these effects are illustrated in Figure 5. In the peanuts effect, behavior shifts from risk-seeking at small stakes to risk aversion at larger stakes. In the magnitude effect, behavior shifts toward greater patience at larger stakes.

Figure 5. Illustrations of the Peanuts Effect and the Magnitude Effect

The Peanuts Effect (Risk)

	$	**prob.**	**$**	**prob.**
R	**1**	**0.10**	**0**	**0.90**
S	0.10	0.10	0.10	0.90
R'	1000	0.10	0	0.90
S'	**100**	**0.10**	**100**	**0.90**

The Magnitude Effect (Time)

	$	**period**
SS	**2**	**0**
LL	3	1
SS'	200	0
LL'	**300**	**1**

The difficulty in reconciling these two effects has been that the peanuts effect reveals decreasing sensitivity to payoffs at larger stakes, whereas the magnitude effect reveals increasing sensitivity to payoffs at larger stakes. Even for decisions under risk, the most widely used specification of prospect theory (Tversky and Kahneman, 1992) does not explain the peanuts effect. The literature following Prelec and Loewenstein has not resolved the dilemma and the difficulty in reconciling the peanuts effect and the magnitude effect has been viewed as strong evidence that the same utility function cannot apply to both risky and intertemporal decisions. SWUP provides a partial resolution – when one of the options in the peanuts effect is a certain option (the type of situation where the peanuts effect is most likely to be observed), SWUP simultaneously explains the peanuts and magnitude effects. The left panel of Figure 5 illustrates the peanuts effect in which a person prefers a 10% chance of winning $1 over a guaranteed ten cents, but prefers a guaranteed $100 over a 10% chance of winning $1000. The peanuts effect in Figure 5 is driven by a conflict between decreasing absolute sensitivity and increasing proportional sensitivity. At low stakes, the influence of increasing proportional sensitivity dominates (the comparison of a dime versus a dollar is more salient than that between a dime versus nothing). However, at larger stakes, the influence of decreasing absolute sensitivity dominates (the comparison of receiving $0 instead of a guaranteed $100 can be more salient than that between $100 and a small chance of $1000). For choices over time, the right panel of Figure 5 illustrates the magnitude effect in which a person prefers $2 now to $3 in one year, but prefers $300 in one year to $200 now. Both effects hold under our parameter-free specification from (9) for all $\delta \in [0.86, 0.93]$.

3.4 Skewness Preference for Risk and Time

In this section, we demonstrate that SWUP predicts a preference for positive skewness and an aversion to negative skewness in the distribution of payoffs for a lottery and in the distributions of payoffs for a consumption plan. This preference for skewness is predicted by SWUP to generate the fourfold pattern for choices under risk (Tversky and Kahneman, 1992) and to generate a bias toward concentration for choices over time (Koszegi and Szeidl, 2013). The preference for skewness under risk explains the simultaneous purchasing of lottery tickets and insurance policies. The preference for skewness over time explains, for instance, the perception, that a newspaper subscription cost of 60 cents per day seems less expensive than a subscription of $219 per year.

Skewness Preference for Time (Bias toward Concentration): We show that the ordering property of salience perception can explain a pattern of behavior that Koszegi and Szeidl (2013), 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. In the context of intertemporal choice, such a bias toward concentration may explain behaviors like addiction in situations where the benefits of consumption are immediate and the costs of current consumption are distributed over many future dates. Conversely, it implies future-biased behavior when the benefit of many periods of effort is concentrated on a single goal (as in a career achievement).

To illustrate such behavior, consider a decision maker given the choice between a large immediate payment of $219 for an annual newspaper subscription, or daily future payments of $0.60 for a year. Considering the undiscounted case (i.e., $\delta = 1$), the two options both have a value of $219, but a bias toward concentration implies that a decision maker will strictly prefer the financing option of $0.60 per day. As a second example, consider a decision maker deciding between two donation options to a charity he supports: One option recommends that he make a donation of $240 today (Plan A). Another option recommends he donates just $20 per month (Plan B). These two options are illustrated in Figure 6.

Figure 6. Charity Donation Framing Effect (Undiscounted case)

	(x_0,y_0)	**(r_0,t_0)**	**(x_1,y_1)**	**(r_1,t_1)**	**(x_2,y_2)**	**(r_2,t_2)**		**(x_{12},y_{12})**	**(r_{12},t_{12})**
A	-240	0	0	1	0	2	…	0	12
B	**0**	**0**	**-20**	**1**	**-20**	**2**	**…**	**-20**	**12**

A focal thinker deciding between Plan A and Plan B in Figure 6 will favor Plan B. Note that this inclination will hold, even if the large immediate payment is lower than $240. For instance, suppose the payment in Plan A is $225. Then under our parametric specification from (9), a bias toward concentration (choice of Plan B) holds for all $\delta \in [0.88, 0.97]$ even though Plan A offers higher discounted utility. To

see why the bias toward concentration arises, note that Plan A has one large disadvantage and many small advantages relative to Plan B. In contrast, Plan B has a single large advantage of avoiding the immediate payment. For the undiscounted case, SWUP predicts a strict preference for B if and only if (10) holds:

$$\mu(-240{,}0)(-240) + \mu(-20{,}0)(20)(12) < 0. \tag{10}$$

Inequality (10) holds if and only if $\mu(-20{,}0) < \mu(-240{,}0)$. Since $[-20{,}0] \subset [-240{,}0]$, the preference for Plan B follows directly from the ordering property in Definition 5. We can formalize a bias toward concentration more generally as follows:

Definition 6 (Bias Toward Concentration): For frame $[\![\mathbf{A}, \mathbf{B}]\!]$ presenting consumption plans $\{A, B\}$ and frame $[\![\mathbf{A}', \mathbf{B}']\!]$ presenting plans $\{A', B'\}$ in Figure 7, a *bias toward concentration* (BTC) holds if there exists $\mathbf{n}^* \geq 1$ such that for all $\mathbf{n} > \mathbf{n}^*$, $\mathbf{A} \succnsim \mathbf{B}$ and $\mathbf{A}' \succnsim \mathbf{B}'$ for all $x > 0$, while $\mathbf{B} \succnsim \mathbf{A}$ and $\mathbf{B}' \succnsim \mathbf{A}'$ for all $x < 0$.

Figure 7. Bias Toward Concentration

	(x_0,y_0)	(r_0,t_0)	(x_1,y_1)	(r_1,t_1)	(x_2,y_2)	(r_2,t_2)		(x_n,y_n)	(r_n,t_n)
A	**nx**	**0**	**0**	**1**	**0**	**2**	**…**	**0**	**n**
B	**0**	**0**	**x**	**1**	**x**	**2**	**…**	**x**	**n**

	(x_1,y_1)	(r_1,t_1)	(x_2,y_2)	(r_2,t_2)		(x_n,y_n)	(r_n,t_n)	(x_{n+1},y_{n+1})	(r_{n+1},t_{n+1})
$\mathbf{A}'$	**0**	**1**	**0**	**2**	**…**	**0**	**n**	**nx**	$\mathbf{n+1}$
$\mathbf{B}'$	**x**	**1**	**x**	**2**	**…**	**x**	**n**	**0**	$\mathbf{n+1}$

A bias toward concentration implies a form of present bias when concentrated gains are immediate, and future bias when concentrated gains are remote. The example of a bias toward concentration in the lower panel of Figure 7 may be characteristic of a form of goal-seeking behavior, such as the case of a junior faculty member who foregoes small payoffs from leisure and social activities over many periods while focusing on the goal of obtaining tenure. For the undiscounted case, such behavior follows from ordering.

Proposition 2: *Let $\delta = 1$ and $u(\mathbf{x}) = \mathbf{x}$. Then a focal thinker exhibits BTC if μ satisfies ordering.*

Skewness Preference for Risk (The Fourfold Pattern of Risk Attitudes): A robust property of observed choices under risk is the fourfold pattern of risk attitudes. A decision maker who exhibits the fourfold pattern displays "risk aversion for gains and risk seeking for losses of high probability and risk seeking for gains and risk aversion for losses of low probability" (Tversky and Kahneman 1992, p. 297). A prototypical example is given in Figure 8. This pattern is predicted by SWUP under the specification in (9). We also define the fourfold pattern more generally in the appendix and prove the following result:

Proposition 3: *Let* $u(\mathbf{x}) = \mathbf{x}$*. Then a focal thinker exhibits the fourfold pattern of risk attitudes if* μ *satisfies ordering, DAS, and IPS.*

Under SWUP, the fourfold pattern can be summarized as a *preference for positively skewed lotteries* (as they have a large upside potential) *and an aversion to negatively skewed lotteries* (since they have a large downside risk). Analogously, the bias toward concentration can be summarized as a *preference for positively skewed consumption sequences and an aversion to negatively skewed consumption sequences*.

Figure 8. The Fourfold Pattern of Risk Preferences

Risk Aversion for high-probability Gains

(i)	**(x_1,y_1)**	**(p_1,q_1)**	**(x_2,y_2)**	**(p_2,q_2)**
p	$\mathbf{99}$	$\mathbf{0.99}$	$\mathbf{99}$	$\mathbf{0.01}$
q	100	0.99	0	0.01

Risk Seeking for high-probability Losses

(ii)	**(x_1,y_1)**	**(p_1,q_1)**	**(x_2,y_2)**	**(p_2,q_2)**
p	-99	0.99	-99	0.01
q	$\mathbf{-100}$	$\mathbf{0.99}$	$\mathbf{0}$	$\mathbf{0.01}$

Risk Seeking for low-probability Gains

(iii)	**(x_1,y_1)**	**(p_1,q_1)**	**(x_2,y_2)**	**(p_2,q_2)**
p	1	0.01	1	0.99
q	$\mathbf{100}$	$\mathbf{0.01}$	$\mathbf{0}$	$\mathbf{0.99}$

Risk Aversion for low-probability Losses

(iv)	**(x_1,y_1)**	**(p_1,q_1)**	**(x_2,y_2)**	**(p_2,q_2)**
p	$\mathbf{-1}$	$\mathbf{0.01}$	$\mathbf{-1}$	$\mathbf{0.99}$
q	-100	0.01	0	0.99

Under the fourfold pattern for risk, a decision maker smooths payoffs (or hedges risk) when a lottery or investment opportunity is negatively skewed (as in frames (i) and (iv) in Figure 8), and spreads payoffs (or embraces risk) when the lottery is positively skewed (as in frames (ii) and (iii) in Figure 8). This can explain both a preference to purchase insurance against low-probability losses and a preference to purchase lottery tickets for low-probability gains. Similarly, for choice over time, this prediction can rationalize both a preference for installment plans (paying $19.99 per month instead of a single payment of $225 per year, even though the latter is less expensive) and a preference for lumps-sum payments over annuities (even if the former are heavily taxed relative to the latter). SWUP thus makes the novel prediction that the degree of consumption smoothing will differ for gains and losses, and that this will be analogous to how the propensity for hedging risk differs for upside and downside risk.

3.5 Framing Opportunity Cost and Downside Risk

Frederick et al (2009) find that people do not spontaneously consider opportunity costs, but will do so when prompted. One recent example in intertemporal choice is the finding by Magen et al. (2008) that people do not naturally consider the opportunity costs of immediate consumption, but they are more likely to do so, when that opportunity cost is salient. Magen et al. refer to this finding as the 'hidden zero effect'. We will also illustrate an analogous hidden zero effect predicted by SWUP for decisions under risk.

Framing the Opportunity Costs of Impatience (The Hidden Zero Effect for Time): Recent work in intertemporal choice (Magen et al., 2008; Read and Scholten, 2012; Read et al., 2017) has observed that making the opportunity costs of choosing the sooner payoff more salient systematically shifts choices toward more patient behavior. This 'hidden zero effect' is difficult to explain in a standard model of context-independent behavior. However, SWUP provides a natural explanation for this framing effect. To illustrate, consider the consumption plan frames in Figure 9, where the top frame conceals the zero payoffs and the bottom frame reveals them. In the simpler (top) frame in Figure 9, the comparisons cued by the frame are between 510 and 530 and between receiving a payoff in period 0 versus in 2 periods from now. Since the salience of time differences (now versus 2 periods) is greater than the salience of payoff differences, a focal thinker will be inclined to select the sooner option. In contrast, the comparisons cued by the 'salient opportunity cost' (bottom) frame are between payoffs of 510 and 0 and between 530 and 0 (highlighting the opportunity cost of choosing the sooner option and receiving 0 instead of 530), while time differences are less salient. Since the most salient difference (530 versus 0) favors the delayed option, a focal thinker will be inclined to select the more patient option. Under our running parametric specification from (9), the hidden zero effect for time holds for all $\delta \in [0.80, 0.99]$. The explanation of the hidden zero effect given by SWUP is also consistent with the finding that shifting attention from time delays to monetary amounts induces more patient behavior (Fisher and Rangel, 2014).

Framing Downside Risk as Probabilities or Outcomes (The Hidden Zero Effect for Risk): An analogous hidden zero effect is predicted by SWUP for decisions under risk. To illustrate, in Figure 9, the *downside risk* of choosing p over q in the simpler (top) frame is represented by the comparison between probabilities (the reduction in probability of winning from 0.34 to 0.33). In the 'salient downside risk' (bottom) frame, the downside risk is represented by the comparison between outcomes (the risk of receiving \$0 instead of \$2400). The origin of the hidden zero effect is implicit in Savage's (1954) classic book on subjective expected utility theory. Savage proposes that Allais paradox violations of expected utility can be reduced if the prospect format used by Allais is changed into a decision matrix which makes the normative appeal of the independence axiom more transparent. Savage's reframing hypothesis implies the hidden zero effect since the choice between p' and q' in the top frame in Figure 9 is also part of a classic example of the common consequence effect used by Kahneman and Tversky. Replacing the top frame in Figure 9 with the bottom frame produces the type of presentation format that Savage suggested will reduce violations of expected utility. Evidence consistent with the hidden zero effect for risk was observed in a number of papers (e.g., Harless (1992), Leland (2010)). In such cases, SWUP predicts that framing downside risk as outcomes will induce more risk-averse behavior than framing downside risk as probabilities. This effect also holds for the specification in (9).

Figure 9. Hidden Zero Effects predicted by SWUP for choices under Risk and over Time

Hidden Zero Effect (Risk)
(Downside Risk Framing)

	(x_1,y_1)	(p_1,q_1)	(x_2,y_2)	(p_2,q_2)
p′	**2500**	**0.33**	**0**	**0.67**
q′	2400	0.34	0	0.66

	(x_1,y_1)	(p_1,q_1)	(x_2,y_2)	(p_2,q_2)	(x_3,y_3)	(p_3,q_3)
p′	2500	0.33	0	0.01	0	0.66
q′	**2400**	**0.33**	**2400**	**0.01**	**0**	**0.66**

Hidden Zero Effect (Time)
(Opportunity Cost Framing)

	(x_1,y_1)	Months
SS	**510**	**0**
LL	530	2

	(x_1,y_1)	Months	(x_2,y_2)	Months
SS	510	0	0	2
LL	**0**	**0**	**530**	**2**

3.6 Labeling Effects for Risk and Time Preferences

Even though a focal thinker has stable preferences, SWUP also predicts systematic label framing effects in which risk and time preferences shift in a predicted direction when changing the labeling of outcomes from gains to losses or when changing the labeling of time periods from time delays to calendar dates.

Framing Effects between Positive and Negative Frames: Consider one of the most famous examples of a framing effect, due to Tversky and Kahneman (1981), in which respondents are told that the U.S. is preparing for the outbreak of an epidemic which is expected to kill 600 people. Policy makers must choose between two prevention strategies: Program A saves 200 lives. Program B has a 1/3 chance of saving 600 people and a 2/3 chance of saving no one. A different group of respondents was given a choice between Programs C and D, and were told 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 frames of both decisions are given in Figure 10.

Programs A and C differ only in how the outcomes are labeled (as lives saved or lives lost) and are thus logically equivalent. This observation also holds for Programs B and D. However, most people chose A over B and chose D over C, thereby exhibiting a framing effect.

The traditional explanation for this framing effect is due to a value function that is concave for gains and convex for losses. SWUP provides a different perspective. Under SWUP, Program A is chosen over Program B if saving 0 lives versus saving 200 lives is more salient than saving 200 lives versus 600 lives. In addition, D is chosen over C if having 0 deaths versus 400 deaths is more salient than having 400 deaths versus 600 deaths. Formally, for $u(x) = x$, a focal thinker chooses Program A over Program B if $\mu(200,0) > \mu(200,600)$. Analogously, the focal thinker chooses Program D over Program C if $\mu(-400,-600) < \mu(-400,0)$. Both of these inequalities hold for the parameter-free specification from (9). SWUP thus yields the choice for A over B in the 'lives saved' frame and the choice for D over C in the 'lives lost' frame, as observed by Tversky and Kahneman.

Figure 10. Framing Effect between Positive and Negative Frames

	(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_1,y_1)	(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**

Framing Effects between Date and Delay Frames: The format in which time periods are presented can also significantly affect the degree of discounting. For instance, Read et al. (2005) identified a robust framing effect in which they observed more patient behavior when time periods were presented as calendar dates than when they were presented as delays (numbers of weeks or months) prior to the receipt of a monetary reward. An example of this 'date-delay effect' from their experiment is presented in Figure 11. Read et al. found that subjects typically opted for \$370 in 17 weeks instead of \$450 in 56 weeks in the delay frame, but preferred the delayed \$450 payment when the same time interval was presented in calendar dates. If the salience of calendar dates can be approximated in this example by the comparison between 2003 and 2004, then under SWUP, the framing effect is explained, if $\theta(17{,}56) > \theta(2003{,}2004)$, which can be shown to be implied by the properties of ordering and diminishing sensitivity of the time salience function θ. Thus, SWUP not only predicts a date-delay framing effect, but it predicts the effect in the observed direction. Moreover, under SWUP, this context-dependent shift in discounting is driven by changes in salience perception and is predicted even if the decision maker has stable time preferences. Under our running parameter-free specification, the date-delay effect is predicted by SWUP for all $\delta \in [0, 0.95]$.

Figure 11. The Date-Delay Framing Effect in Intertemporal Choice

	(x_1,y_1)	Delay (Weeks)
SS	**370**	**17**
LL	450	56

	(x_1,y_1)	Date
SS	370	September 26, 2003
LL	**450**	**June 25, 2004**

4. The Etiology of Choice under Risk and over Time

In addition to showing in Section 2 that SWUP applies the *same* mathematical structure and the *same* psychological intuition to choices under risk and over time, we demonstrated in Section 3 that SWUP predicts a set of analogous behaviors across the two domains. These behaviors are summarized in Table 1. The predicted set of behaviors has strong empirical support and has not previously been explained in a unified model. In addition, the same simple parameter-free specification explains each of these effects.

Although relatively short, the list of analogous behaviors between risk and time summarized in Table 1 is surprisingly comprehensive. Moreover, all of these behaviors can be explained by SWUP as salience effects. Thus all behaviors in Table 1 can be explained through the same underlying mechanism – systematic overweighting of differences in payoffs, probabilities, and time delays, where the weights are endogenous and satisfy basic properties of human salience perception. Interestingly, all of the context-dependent shifts in risk-taking and time discounting summarized in Table 1 are explained even if the decision maker has stable (context-independent) risk and time preferences given by u and δ.

Table 1. Parallels between Risk and Time Predicted by SWUP

Operations on Lottery Frames	**The Etiology of Choice under Risk**	**Representative Reference**
Add Common Consequence	Common Consequence Effect	Allais (1953)
Scale Probabilities by Constant	Common Ratio Effect (Certainty Effect)	Kahneman & Tversky (1979)
Scale Outcomes by Constant	Peanuts Effect	Markowitz (1952)
Change Skewness of Distribution	Fourfold Pattern of Risk Attitudes	Tversky & Kahneman (1992)
Change Salience of Downside Risk	Hidden Zero Framing Effect (Risk)	Savage (1954)
Change Attribute Labels	Gain-Loss Framing Effect	Tversky & Kahneman (1981)
Operations on Temporal Frames	**The Etiology of Choice over Time**	**Representative Reference**
Add Common Consequence	Cancellation Effect	Rao & Li (2011)
Add Constant to Time Periods	Common Difference Effect (Present Bias)	Laibson (1997)
Scale Outcomes by Constant	Magnitude Effect	Prelec & Loewenstein (1991)
Change Skewness of Distribution	Bias Toward Concentration	Koszegi & Szeidl (2013)
Change Salience of Opportunity Cost	Hidden Zero Framing Effect (Time)	Magen et al. (2008)
Change Attribute Labels	Date-Delay Framing Effect	Read et al. (2005)

Table 1 also displays the formal operations one can perform on a lottery or consumption plan to transform it into a different lottery or consumption plan and the corresponding behaviors predicted by SWUP to result from each operation. Regarding the anomalies in Table 1, note that the certainty effect and standard present-biased behavior in smaller-sooner versus larger-later choices are special cases of the common ratio effect and the common difference effect, respectively.

5. Related Literature

Prelec and Loewenstein (1991) were the first to recognize parallels between the major anomalies for risk and time. For instance, they recognized an intuitive relationship between the common ratio effect and the common difference effect, and between the peanuts effect and the magnitude effect. They also

recognized the role of IPS and DAS in explaining some of these anomalies. However, the model they proposed does not explain the peanuts effect or any of the anomalies in sections 3.2, 3.4, 3.5 or 3.6. In the present paper, we have extended the discussion of parallels between risk and time to include the common consequence effect and the cancellation effect, the fourfold pattern of risk attitudes and a bias toward concentration, the hidden zero effects for risk and time, and framing effects due to changing the labeling of attributes. In addition to the striking similarity in these behaviors for risk and time, we have shown that these effects are each predicted by SWUP.

Aside from the early contribution of Prelec and Loewenstein, the standard approach has been to develop domain-specific models. For instance, the major models for choice under risk such as rank-dependent utility (Quiggin, 1982), regret theory (Loomes and Sugden, 1982; Bell, 1982), and prospect theory (Kahneman and Tversky, 1979) were developed only for risk preferences. Similarly, the major models for choice over time such as the hyperbolic discounting model of Loewenstein and Prelec (1992), Laibson's (1997) model of quasi-hyperbolic discounting, and the recent advance by Koszegi and Szeidl (2013) were developed only for time preferences.

The strong parallels that SWUP predicts for risky and intertemporal choice enables SWUP to explain behavior in choices over time that cannot be explained by cumulative prospect theory (CPT) due to Tversky and Kahneman (1992) which does not apply to intertemporal choice. Even for choice under risk, there are notable differences between SWUP and CPT. For instance, CPT does not explain the downside risk framing effect depicted in Figure 9, and the most commonly applied version of CPT with a power value function cannot explain the peanuts effect, even allowing for any probability weighting function.

To provide a unified explanation of choices under risk and over time, models by Leland (2002) and Rubinstein (2003) advocated an approach based on similarity judgments. That approach can explain present bias, and is analogous to the earlier models of similarity judgments that explain the common ratio effect (Rubinstein, 1988; Leland, 1994). Although these similarity models were based on a strong psychological intuition, they were largely qualitative, based on the notion of a similarity relation without formalizing when two payoffs or time delays were similar or different. Bordalo et al. (2012) pioneered an approach to choice under risk which provided a quantitative approach to formalizing aspects of salience perception. A fundamental structural difference between the model of Bordalo et al. and SWUP is that the Bordalo et al. model only accounts for the salience of payoffs, whereas SWUP incorporates both the salience of payoffs and the salience of probabilities. In addition, Koszegi and Szeidl's model accounts only for the salience of payoffs, whereas SWUP incorporates both the salience of payoffs and the salience of time delays. By accounting for probability salience, Leland, Schneider, and Wilcox (2017) extend SWUP to choice under uncertainty and demonstrate that diminishing sensitivity of the probability salience function explains ambiguity aversion in Ellsberg's (1961) paradox which is not accounted for by the model of Bordalo et al.

(2012). Similarly, by accounting for time salience, we establish in Proposition 1 that diminishing sensitivity of the time salience function explains standard present bias behavior as illustrated by the common difference effect that is not accounted for by the model of Koszegi and Szeidl (2013).

Of behaviors in Table 1, quasi-hyperbolic discounting explains only present bias and the model of Koszegi and Szeidl explains only a bias toward concentration. Koszegi and Szeidl write:

> "Yet because our theory does not match evidence in single choices, it is not a complete theory of choice over time, and a complete theory would have to incorporate hyperbolic discounting or a related model as well." (Koszegi and Szeidl, 2013, p. 85).

The SWUP model provides a unified explanation for the six behaviors for intertemporal choice summarized in Table 1 as resulting from a perceptual system that is biased to focus on salient differences in payoffs and time delays and also predicts a directly analogous set of six broad empirical findings for choice under risk. Rather than relying on quite distinct domain specific models (e.g., prospect theory for risk, quasi-hyperbolic discounting for time), SWUP provides a unified approach to choices under risk and over time. However, SWUP is also an incomplete theory as it does not account for interaction effects between risk and time that have been observed for choices involving both risk and delays (Keren and Roelofsma, 1995; Baucells and Heukamp, 2010; Abdellaoui et al., 2011; Hardisty and Pfeffer, 2017).

In the recent literature, attention has turned to integrating risk and time preferences. This literature was pioneered by Halevy (2008) who developed a model that accounts for some systematic interactions between risk and time (based on experimental results that risk affects time preference). Perhaps the most successful model to date in providing a unified treatment of risk and time preference is the probability time tradeoff (PTT) model of Baucells and Heukamp (2012). Notably, the PTT model accounts for three systematic interaction effects between risk and time preferences when choices involve both risk and time delays as well as some other fundamental behaviors such as the common ratio and common difference effects. Although SWUP does not account for interactions between risk and time, it encompasses a broader set of parallels between risk and time than previous approaches and explains all 12 behaviors in Table 1.

6. Conclusion

Choices under risk and over time have traditionally been studied separately, and the notions of impatience and risk tolerance are distinct concepts. The model of salience weighted utility over presentations (SWUP) proposed here generalizes the classic expected and discounted utility models to predict endogenous changes in the salience of payoffs, probabilities, and time delays. Surprisingly, under SWUP, the same mathematical structure and the same psychological intuition predict a wide range of analogous behaviors between risky and intertemporal choice and unify these behaviors as examples of salience effects. Moreover, the 12 context-dependent shifts in risk taking and time discounting we have studied are predicted by SWUP, even if salience weights are parameter-free, and even if the decision maker has stable risk and time preferences.

Appendix: Formal Results for Behaviors under Risk and over Time

In this appendix, we provide more formal discussions of the behaviors illustrated by example in Section 3. For our results, we consider the case where both options in the frame have equal expected utilities (for risk) and equal discounted utilities (for time) to highlight the predictions arising from differences in salience. For a focal thinker, we write the expected utility of lottery p (the agent's assumed preferences in the absence of salience distortions) as $EU(p)$ and the discounted utility of consumption plan r as $DU(r)$. Also, in our analyses to follow, the definitions of behaviors are adapted to our setting (using the perceptual relation $\hat{\succ}$ from Section 2 which depends on both the agent's preferences and the agent's salience perception, instead of the agent's preference relation $\succ$).

A.1 The Common Consequence Effect

We first provide a more formal definition of the common consequence effect

Definition 7: (Common Consequence Effect): Consider the frames in Figure 13, where $\mathbf{x} > \mathbf{y} > \mathbf{0}, \mathbf{q} > \mathbf{p},$ and $\mathbf{q}, \mathbf{p} \in (0,1)$. The *common consequence effect holds* if $\mathbf{q} \hat{\succ} \mathbf{p}$ and $\mathbf{p}' \hat{\succ} \mathbf{q}'$.

Figure 13. The Common Consequence Effect

p	**x**	**p**	**y**	$\mathbf{1-q}$	**0**	$\mathbf{q-p}$
q	**y**	**p**	**y**	$\mathbf{1-q}$	**y**	$\mathbf{q-p}$

$\mathbf{p}'$	**x**	**p**	**0**	$\mathbf{1-p}$
$\mathbf{q}'$	**y**	**q**	**0**	$\mathbf{1-q}$

Proposition 4: *Let $EU(p) = EU(q)$. Then a focal thinker exhibits the common consequence effect if and only if $\mu(\mathbf{0}, \mathbf{y}) > \mu(\mathbf{x}, \mathbf{y}) > \phi(\mathbf{p}, \mathbf{q})$.*

Remark: For both the classic Allais example in which $(x, y, p, q) = (\$5\ Million, \$1\ Million, 0.10, 0.10)$, and for the Kahnemman and Tversky example in Figure 4, the condition in Proposition 4 holds for the salience functions in (9). For the Kahneman and Tversky example, the necessary and sufficient condition in Proposition 4 becomes $\mu(0,2400) > \mu(2500,2400) > \phi(0.33,0.34)$. That is, the common consequence effect holds if the perceived downside of receiving \$0 instead of \$2400 outweighs the upside of receiving \$2500 instead of \$2400, but this \$100 difference in payoffs outweighs the 0.01 difference in the probability of winning. The explanation of SWUP for the choice of q over p is similar to Bordalo et al. (2012), with both explanations based on payoff salience. However, the Bordalo et al. explanation for the choice of p' over q' is based only on payoff salience and assumes that agents interpret the choice as between statistically independent prospects, and evaluate options as if they compare \$2500 and \$2400, \$2500 and \$0, \$2400 and \$0, and \$0 and \$0). It seems more likely to us that the choice between p' and q' is determined by the salience of payoffs (\$2500 versus \$2400) contrasted with the salience of probabilities (0.33 versus 0.34).

A.2 The Common Ratio Effect

We next provide a definition of the common ratio effect:

Definition 8 (Common Ratio Effect): Consider the frames in Figure 14, where $\mathbf{x} > \mathbf{y} > \mathbf{0}$, and $\mathbf{c}, \mathbf{p} \in (0,1)$. The *common ratio effect holds* if $\mathbf{q} \gtrsim \mathbf{p}$ and $\mathbf{p}' \gtrsim \mathbf{q}'$.

Figure 14. The Common Ratio Effect

$\mathbf{p}$	$\mathbf{x}$	$\mathbf{p}$	$\mathbf{0}$	$\mathbf{1-p}$	$\mathbf{p}'$	$\mathbf{x}$	$\mathbf{cp}$	$\mathbf{0}$	$\mathbf{1-cp}$
$\mathbf{q}$	$\mathbf{y}$	$\mathbf{p}$	$\mathbf{y}$	$\mathbf{1-p}$	$\mathbf{q}'$	$\mathbf{y}$	$\mathbf{c}$	$\mathbf{0}$	$\mathbf{1-c}$

Proposition 5: *Let $EU(p) = EU(q)$. Then a focal thinker exhibits the common ratio effect if and only if* $\mu(\mathbf{0}, \mathbf{y}) > \mu(\mathbf{x}, \mathbf{y}) > \phi(\mathbf{cp}, \mathbf{c})$.

Remark: For the example of the common ratio effect in Figure 3, the necessary and sufficient condition in Proposition 5 becomes $\mu(0, 3000) > \mu(4000, 3000) > \phi(0.20, 0.25)$. That is, the common consequence effect holds if the perceived downside of receiving 0 instead of 3000 outweighs the perceived upside of receiving 4000 instead of 3000, but this \$1000 difference in payoffs outweighs the 0.05 difference in the probability of winning. These conditions also hold for the parameter-free specification in (9).

A.3 The Peanuts Effect

Next, we provide a definition of the peanuts effect:

Definition 9 (Peanuts Effect): Consider Figure 15, where $\mathbf{x} > \mathbf{0}$ and $\mathbf{p} \in (0,1)$. The *peanuts effect holds* if $\mathbf{p} \gtrsim \mathbf{q}$ and there is some $\mathbf{k} > \mathbf{1}$ such that $\mathbf{q}' \gtrsim \mathbf{p}'$.

Figure 15. The Peanuts Effect

$\mathbf{p}$	$\mathbf{x}$	$\mathbf{p}$	$\mathbf{0}$	$\mathbf{1-p}$	$\mathbf{p}'$	$\mathbf{kx}$	$\mathbf{p}$	$\mathbf{0}$	$\mathbf{1-p}$
$\mathbf{q}$	$\mathbf{xp}$	$\mathbf{p}$	$\mathbf{xp}$	$\mathbf{1-p}$	$\mathbf{q}'$	$\mathbf{kxp}$	$\mathbf{p}$	$\mathbf{kxp}$	$\mathbf{1-p}$

Proposition 6: *Let $EU(p) = EU(q)$. Then a focal thinker exhibits the peanuts effect if and only if* $\mu(\mathbf{xp}, \mathbf{x}) > \mu(\mathbf{xp}, \mathbf{0})$ *and* $\mu(\mathbf{kxp}, \mathbf{0}) > \mu(\mathbf{kxp}, \mathbf{kx})$.

A.4 The Fourfold Pattern of Risk Attitudes

We next demonstrate that, for linear utility, SWUP implies the fourfold pattern of risk attitudes for any salience function μ that satisfies ordering, diminishing absolute sensitivity, and increasing proportional sensitivity. Our approach is to demonstrate that half of the fourfold pattern (risk aversion for gains and risk seeking for losses of high probability) is implied by ordering and DAS. We then demonstrate that the other half of the fourfold pattern (risk seeking for gains and risk aversion for losses of low probability) is implied by IPS. We first provide a simple behavioral definition of commonly observed risk preferences:

Definition 10 (Risk preferences at moderate to high probabilities): Consider the frames in Figure 16 of a choice between a representation $\mathbf{p}$ of a lottery p and its expected value, $\mathbf{E}(\mathbf{p})$, where $\mathbf{x} > \mathbf{y} \geq 0$**.**

(i) *Risk aversion for gains at moderate to high probabilities* holds if $\mathbf{E}(\mathbf{p}) \gtrsim \mathbf{p}$ for all $\mathbf{p_1} \in$**[0.5,1)**.

(ii) *Risk seeking for losses at moderate to high probabilities* holds if $\mathbf{p}' \gtrsim \mathbf{E}(\mathbf{p}')$ for all $\mathbf{p_1} \in$**[0.5,1)**.

Figure 16. Diminishing Absolute Sensitivity and Attitudes toward Risk

	Risk Aversion for Gains					Risk Seeking for Losses			
(i)	**(x_1,y_1)**	**(p_1,q_1)**	**(x_2,y_2)**	**(p_2,q_2)**	**(ii)**	**(x_1,y_1)**	**(p_1,q_1)**	**(x_2,y_2)**	**(p_2,q_2)**
$\mathbf{E}(\mathbf{p})$	$\mathbf{E}(\mathbf{p})$	$\mathbf{p_1}$	$\mathbf{E}(\mathbf{p})$	$\mathbf{1-p_1}$	$\mathbf{E}(\mathbf{p}')$	$\mathbf{E}(\mathbf{p}')$	$\mathbf{p_1}$	$\mathbf{E}(\mathbf{p}')$	$\mathbf{1-p_1}$
$\mathbf{p}$	$\mathbf{x}$	$\mathbf{p_1}$	$\mathbf{y}$	$\mathbf{1-p_1}$	$\mathbf{p}'$	$-\mathbf{x}$	$\mathbf{p_1}$	$-\mathbf{y}$	$\mathbf{1-p_1}$

To isolate the role of salience perceptions in governing risk attitudes, we state the following proposition for the case where the utility function is linear.

Lemma 1: *Consider the frames in Figure 16 and let* $u(\mathbf{x}) = \mathbf{x}$*. Then for all* $\mathbf{p_1} \in [\mathbf{0.5}, \mathbf{1}]$ *a focal thinker is risk-averse for gains and risk-seeking for losses if* μ *satisfies ordering and DAS.*

Proof: For $u(\mathbf{x}) = \mathbf{x}$, $\mathbf{E}(\mathbf{p}) \gtrsim \mathbf{p}$ if and only if $\mu\big(\mathbf{y}, (\mathbf{x}-\mathbf{y})\mathbf{p_1} + \mathbf{y}\big) > \mu\big(\mathbf{x}, (\mathbf{x}-\mathbf{y})\mathbf{p_1} + \mathbf{y}\big)$.

For $p = 0.5$, DAS implies $\mu(\mathbf{y}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})) > \mu\big(\mathbf{x}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})\big)$. To see that DAS implies that this inequality holds, let $\boldsymbol{\epsilon} = 0.5(x-y)$. Then $\mu(\mathbf{y}+\boldsymbol{\epsilon}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})+\boldsymbol{\epsilon}) = \mu(\mathbf{0.5}(\mathbf{x}+\mathbf{y}), \mathbf{x})$, which by symmetry, equals $\mu\big(\mathbf{x}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})\big)$. For $\mathbf{p_1} \in [0.5,1]$, note that $\big[\mathbf{x}, \mathbf{x}\mathbf{p_1} + \mathbf{y}(\mathbf{1}-\mathbf{p_1})\big] \subseteq [\mathbf{x}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})]$ and thus ordering implies $\mu\big(\mathbf{x}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})\big) \geq \mu\Big(\mathbf{x}, \mathbf{x}\mathbf{p_1} + \mathbf{y}\big(\mathbf{1}-\mathbf{p_1}\big)\Big)$. Thus, $\mu(\mathbf{y}, \mathbf{0.5}(\mathbf{x}+\mathbf{y})) > \mu\Big(\mathbf{x}, \mathbf{x}\mathbf{p} + \mathbf{y}\big(\mathbf{1}-\mathbf{p_1}\big)\Big)$ for all $\mathbf{p_1} \in [0.5,1]$. The result for losses follows analogously. ■

For the case where $\mathbf{p_1} = 0.5$, Lemma 1 requires only DAS. In conjunction with DAS, ordering implies that the result extends to all $\mathbf{p_1} \in [\mathbf{0.5}, \mathbf{1}]$. This stronger result implies that if we observe risk-seeking for gains or risk aversion for losses, such behavior will occur at low probabilities, consistent with the experimental observation of the fourfold pattern of risk attitudes (Tversky and Kahneman, 1992).

Whereas DAS and ordering have implications for risk preferences at moderate and high probabilities, IPS has implications for risk preferences over low probability outcomes. Consider the frames in Figure 17, where we set $\mathbf{x} > 0, \mathbf{k} \geq \mathbf{2}$, and $\mathbf{E}(\mathbf{p}) = \mathbf{E}(\mathbf{q})$**.** Frames (i) and (ii) each involve a choice between a sure option and a mean-preserving spread. Setting $\mathbf{k} \geq 2$ ensures that the payoff with the largest absolute value occurs at low probabilities. For $\mathbf{k} = 2$, we have $\mathbf{p_1} = 1/3$, independent of $\mathbf{x}$, and $\mathbf{p_1}$ decreases toward 0 as $\mathbf{k}$ increases, holding $\mathbf{x}$ fixed. In the figure, $\mathbf{p_1}$ is the unique value determined by the requirement that $\mathbf{E}(\mathbf{p}) = \mathbf{E}(\mathbf{q})$**.** For $\mathbf{x} = \$1$, and $\mathbf{k} = 1{,}000{,}000$, the choice in frame (i) resembles the decision of purchasing a lottery ticket ($\mathbf{p_1}$ ~ 0.000001), and the choice in frame (ii) resembles the decision

to insure against low-likelihood disasters. For such large values of $\mathbf{k}$ relative to $\mathbf{x}$, the payoff $\mathbf{x/k}$ is essentially 0 and its inclusion may be viewed as merely a technical condition. Thus, for large $\mathbf{k}$ and small $\mathbf{x}$, the frames in Figure 16 approximate a choice between a lottery with one non-zero outcome occurring with probability $\mathbf{p_1}$, and its expected value. Formally, we have the following definition:

Definition 11 (Risk Preference at low probabilities): Consider frames (i) and (ii) in Figure 17, where $\mathbf{x} > 0, \mathbf{k} \geq 2$, and $\mathbf{E}(\mathbf{p}) = \mathbf{E}(\mathbf{q})$:

(i) *Risk-seeking for gains at low probabilities* holds if $\mathbf{q} \succsim \mathbf{p}$.

(ii) *Risk aversion for losses at low probabilities* holds if $\mathbf{p'} \succsim \mathbf{q'}$

Figure 17. Increasing Proportional Sensitivity and Attitudes toward Risk

	Lottery Purchase					**Insurance Purchase**			
(i)	$(\mathbf{x_1},\mathbf{y_1})$	$(\mathbf{p_1},\mathbf{q_1})$	$(\mathbf{x_2},\mathbf{y_2})$	$(\mathbf{p_2},\mathbf{q_2})$	**(ii)**	$(\mathbf{x_1},\mathbf{y_1})$	$(\mathbf{p_1},\mathbf{q_1})$	$(\mathbf{x_2},\mathbf{y_2})$	$(\mathbf{p_2},\mathbf{q_2})$
$\mathbf{p}$	$\mathbf{x}$	$\mathbf{p_1}$	$\mathbf{x}$	$\mathbf{1-p_1}$	$\mathbf{p'}$	$-\mathbf{x}$	$\mathbf{p_1}$	$-\mathbf{x}$	$\mathbf{1-p_1}$
$\mathbf{q}$	$\mathbf{kx}$	$\mathbf{p_1}$	$\mathbf{x/k}$	$\mathbf{1-p_1}$	$\mathbf{q'}$	$-\mathbf{kx}$	$\mathbf{p_1}$	$-\mathbf{x/k}$	$\mathbf{1-p_1}$

Lemma 2: *Let* $u(\mathbf{x}) = \mathbf{x}$. *Then for frames* (i) *and* (ii) *in Figure 17, a focal thinker is risk-seeking for gains at low probabilities and risk-averse for losses at low-probabilities if* μ *satisfies IPS.*

Proof: We show that IPS implies risk-seeking for low-probability gains. Risk aversion for low probability losses follows analogously. Given $u(\mathbf{x}) = \mathbf{x}$, a focal thinker always chooses $\mathbf{q}$ over $\mathbf{p}$ if and only if $\mu(\mathbf{x},\mathbf{kx})(\mathbf{x}-\mathbf{xk})\mathbf{p_1} + \mu\left(\mathbf{x},\frac{\mathbf{x}}{\mathbf{k}}\right)\left(\mathbf{x}-\frac{\mathbf{x}}{\mathbf{k}}\right)\left(\mathbf{1}-\mathbf{p_1}\right) < 0$. Since $E(\mathbf{p}) = E(\mathbf{q})$, we have: $\mathbf{p_1} = \frac{\mathbf{1}-(\mathbf{1}/\mathbf{k})}{\mathbf{k}-(\mathbf{1}/\mathbf{k})}$.

Thus, $\mathbf{q}$ is strictly preferred to $\mathbf{p}$ if and only if

$$\mu(\mathbf{x},\mathbf{kx})(\mathbf{x}-\mathbf{xk})\left[\frac{\mathbf{1}-(\mathbf{1}/\mathbf{k})}{\mathbf{k}-(\mathbf{1}/\mathbf{k})}\right] + \mu(\mathbf{x},\mathbf{x}/\mathbf{k})(\mathbf{x}-\mathbf{x}/\mathbf{k})\left[\frac{\mathbf{k}-\mathbf{1}}{\mathbf{k}-(\mathbf{1}/\mathbf{k})}\right] < 0,$$

which can be rewritten as $\mu(\mathbf{x},\mathbf{x}/\mathbf{k})\mathbf{x}\left[\frac{(\mathbf{k}-\mathbf{1})(\mathbf{1}-(\mathbf{1}/\mathbf{k})}{\mathbf{k}-(\mathbf{1}/\mathbf{k})}\right] < \mu(\mathbf{x},\mathbf{kx})\mathbf{x}\left[\frac{(\mathbf{k}-\mathbf{1})(\mathbf{1}-(\mathbf{1}/\mathbf{k}))}{\mathbf{k}-(\mathbf{1}/\mathbf{k})}\right]$, which implies $\mu(\mathbf{x},\mathbf{x}/\mathbf{k}) < \mu(\mathbf{x},\mathbf{kx})$. Then by symmetry and increasing proportional sensitivity, we have $\mu(\mathbf{x},\mathbf{kx}) = \mu(\mathbf{kx},\mathbf{x}) = \mu\left(\mathbf{kx},\mathbf{k}\left(\frac{\mathbf{x}}{\mathbf{k}}\right)\right) > \mu\left(\mathbf{x},\frac{\mathbf{x}}{\mathbf{k}}\right)$. The result for losses follows analogously. ∎

Definition 12 (Fourfold pattern): An agent who is risk-averse for gains and risk-seeking for losses at high probabilities, as defined in Definition 10, and is risk-seeking for gains and risk-averse for losses at low probabilities, as in Definition 11, exhibits the *fourfold pattern of risk preferences*.

Proposition 3 (Fourfold Pattern of Risk Preferences): *Let* $u(\mathbf{x}) = \mathbf{x}$. *Then, by Lemmas 1 and 2, a focal thinker exhibits the fourfold pattern of risk preferences if* μ *satisfies ordering, DAS, and IPS.*

A.5 The Hidden Zero Effect for Choice under Risk

We next provide a formal definition of the hidden zero effect for risk:

Definition 13: (Hidden Zero Effect for Risk): Consider the two choice frames in Figure 18, for which $\mathbf{x} > \mathbf{y} > \mathbf{0}, \mathbf{q} > \mathbf{p}$, and $\mathbf{p}, \mathbf{q} \in (0,1)$. The *Hidden Zero Effect for risk* holds if $\mathbf{p} \succsim \mathbf{q}$ in the simpler (top) frame and $\mathbf{q} \succsim \mathbf{p}$ in the "salient downside risk" (bottom) frame.

Figure 18. The Hidden Zero Effect for Risk

P	$\mathbf{x}$	$\mathbf{p}$	$\mathbf{0}$	$\mathbf{1-p}$
q	$\mathbf{y}$	$\mathbf{q}$	$\mathbf{0}$	$\mathbf{1-q}$

p	$\mathbf{x}$	$\mathbf{p}$	$\mathbf{0}$	$\mathbf{1-q}$	$\mathbf{0}$	$\mathbf{q-p}$
q	$\mathbf{y}$	$\mathbf{p}$	$\mathbf{0}$	$\mathbf{1-q}$	$\mathbf{y}$	$\mathbf{q-p}$

Proposition 7: *Let $EU(p) = EU(q)$. Then a focal thinker exhibits the hidden zero effect for risk if and only if $\mu(\mathbf{0}, \mathbf{y}) > \mu(\mathbf{x}, \mathbf{y}) > \phi(\mathbf{p}, \mathbf{q})$.*

A.6 The Cancellation Effect

The Cancellation Effect in intertemporal choice can be defined as follows:

Definition 14: Consider the frames in Figure 19 where $\mathbf{y} > \mathbf{x} > \mathbf{0}$, and $\mathbf{t} > \mathbf{r} \geq \mathbf{0}$. The *Cancellation effect* holds if $\mathbf{SS} \succsim \mathbf{LL}$ and $\mathbf{LL'} \succsim \mathbf{SS'}$.

Figure 19. The Cancellation Effect

SS	$\mathbf{x}$	$\mathbf{r}$	SS′	$\mathbf{x}$	$\mathbf{r}$	$\mathbf{z}$	$\mathbf{s}$	$\mathbf{0}$	$\mathbf{t}$
LL	$\mathbf{y}$	$\mathbf{t}$	LL′	$\mathbf{0}$	$\mathbf{r}$	$\mathbf{z}$	$\mathbf{s}$	$\mathbf{y}$	$\mathbf{t}$

Proposition 8: *Let $DU(SS) = DU(LL)$. Then a focal thinker exhibits the cancellation effect if and only if $\theta(\mathbf{r}, \mathbf{t}) > \mu(\mathbf{x}, \mathbf{y})$.*

The cancellation effect also requires $\mu(\mathbf{x}, \mathbf{0}) < \mu(\mathbf{0}, \mathbf{y})$ but this inequality is guaranteed by the ordering property. For the example in Figure 4, the condition in Proposition 8 becomes $\theta(0, 2) > \mu(520{,}530)$ which holds for the parameter-free specification in (9) and is consistent with intuition that the difference between a 0 time delay and a 2-month delay is more salient than the difference between \$520 and \$530.

A.7 The Common Difference Effect

We next provide a formal definition of the common difference effect.

Definition 15 (Common Difference Effect): Consider frames (i) and (ii) in Figure 20, for any $\mathbf{y} > \mathbf{x} > 0$, $\mathbf{t} > \mathbf{r} \geq 0$, and any $\boldsymbol{\Delta} > 0$. The *common difference effect* holds if $\mathbf{SS} \precsim \mathbf{LL}$ implies $\mathbf{LL'} \succsim \mathbf{SS'}$.

Figure 20. Properties of Time Preferences Predicted by SWUP

Simple Consumption Plans			The Common Difference Effect			The Magnitude Effect		
(i)	(x_1,y_1)	Period	(ii)	(x_1,y_1)	Period	(iii)	(x_1,y_1)	Period
SS	x	r	SS′	x	r + Δ	SS′	kx	r
LL	y	t	LL′	y	t + Δ	LL′	ky	t

Proposition 1: *A focal thinker exhibits the common difference effect if and only if* θ *satisfies DAS.*

Proof: Note that a focal thinker views **SS** and **LL** to look equally good if and only if (11) holds:

(11) $$\mu(\mathbf{x},\mathbf{y})\big(u(\mathbf{y})-u(\mathbf{x})\big)(\delta^{\mathbf{r}}+\delta^{\mathbf{t}})=\theta(\mathbf{r},\mathbf{t})(\delta^{\mathbf{r}}-\delta^{\mathbf{t}})(u(\mathbf{y})+u(\mathbf{x})).$$

A focal thinker always chooses $\mathbf{LL'}$ over $\mathbf{SS'}$ if and only if inequality (12) holds:

(12) $$\mu(\mathbf{x},\mathbf{y})\big(u(\mathbf{y})-u(\mathbf{x})\big)\big(\delta^{\mathbf{r+\Delta}}+\delta^{\mathbf{t+\Delta}}\big)>\theta(\boldsymbol{r}+\boldsymbol{\Delta},\boldsymbol{t}+\boldsymbol{\Delta})\big(\delta^{\mathbf{r+\Delta}}-\delta^{\mathbf{t+\Delta}}\big)\big(u(\mathbf{y})+u(\mathbf{x})\big).$$

By factoring out δ^{Δ} and by substitution, we see that the common difference effect holds if and only if we have $\theta(\mathbf{r},\mathbf{t})>\theta(\mathbf{r}+\boldsymbol{\Delta},\mathbf{t}+\boldsymbol{\Delta})$, which holds for any $\Delta>0$ if and only if θ satisfies DAS. ■

For the typical example in Figure 3, the common difference effect holds for the specification in (9).

A.8 The Magnitude Effect

The magnitude effect can be defined as follows:

Definition 16 (Magnitude Effect): Consider frames (i) and (iii) in Figure 20. The *magnitude effect* holds if for all $\mathbf{k}>1$, SS $\gtrsim$ LL and LL′ $\gtrsim$ SS′.

Proposition 9: *Let* $DU(SS) = DU(LL)$. *Then a focal thinker exhibits the magnitude effect if and only if*

$$\mu(\mathbf{kx},\mathbf{ky})>\theta(\mathbf{r},\mathbf{t})>\mu(\mathbf{x},\mathbf{y}).$$

Note that if μ *satisfies IPS, we have* $\mu(\mathbf{kx},\mathbf{ky})>\mu(\mathbf{x},\mathbf{y})$ *for all* $\mathbf{k}>\mathbf{1}$.

Next, we state and prove our finding on the bias toward concentration from Section 3.4, and then demonstrate a more formal result for the hidden zero effect over time.

A.9 The Bias Toward Concentration

Proposition 2 (Bias Toward Concentration): *Let* $\delta=1$ *and* $u(\mathbf{x})=\mathbf{x}$. *Then if* μ *satisfies ordering, a focal thinker exhibits a bias toward concentration.*

Proof: In the undiscounted case, for $\mathbf{x}<0$, we have $\mu(\mathbf{nx},\mathbf{0})\mathbf{nx}<\mu(\mathbf{x},\mathbf{0})\mathbf{nx}$ in the choice between A and B and also in the choice between A′ and B′, which hold for all $\mathbf{n}>1$ if μ satisfies ordering. Conversely, for $\mathbf{x}>0$, we have $\mu(\mathbf{nx},\mathbf{0})>\mu(\mathbf{x},\mathbf{0})$ in the choice between A and B, and in the choice between A′ and B′, which hold for all $\mathbf{n}>1$ if μ satisfies ordering. ■

A.10 The Hidden Zero Effect for Choice over Time

Definition 17: Consider Figure 21, where $\boldsymbol{y} > \boldsymbol{x} \geq \mathbf{0}$, and $\mathbf{t} > \mathbf{r} \geq \mathbf{0}$. The *Hidden Zero effect for time* holds if **SS** $\succsim$ **LL** in the left frame and **LL** $\succsim$ **SS** in the 'salient opportunity cost' (right) frame.

Figure 21. The Hidden Zero Effect

SS	**x**	**r**
LL	**y**	**t**

SS	**x**	**r**	**0**	**t**
LL	**0**	**r**	**y**	**t**

Proposition 10: *Let $DU(SS) = DU(LL)$. Then a focal thinker exhibits the hidden zero effect for time if and only if $\theta(\mathbf{r}, \mathbf{t}) > \mu(\mathbf{x}, \mathbf{y})$.*

For the example in Figure 9, the condition in the proposition holds for the specification in (9).

References.

Abdellaoui, M., Diecidue, E., and Onculer, A. (2011). Risk preferences at different time periods: An experimental investigation, *Management Science*, 57, 975-987.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critiques des postulats et axioms de l'ecole Americaine. *Econometrica, 21*, 503-546.

Baucells, M., and Heukamp, F.H. (2010). Common ratio using delay. *Theory and Decision*, 68, 149-158.

Baucells, M., and Heukamp, F.H. (2012). Probability and Time Trade-off. *Management Science*, 58, 831-842.

Bell, D.E. (1982). Regret in decision making under uncertainty. *Operations Research, 30,* 961-981.

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

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and Consumer Choice. *Journal of Political Economy, 121,* 803-843.

Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics, 75,* 643-669.

Fisher, G., and Rangel, A. (2014). Intertemporal Discount Rates Are Mediated By Relative Attention. Manuscript.

Frederick, S., Novemsky, N., Wang, J., Dhar, R., and Nowlis, S. (2009). Opportunity cost neglect. *Journal of Consumer Research*, 36, 553-561.

Friedman, M., and Savage, L.J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, 56, 279-304.

Hardisty, D.J., and Pfeffer, J. (2017). Intertemporal uncertainty avoidance: When the future is uncertain, people prefer the present, and when the present is uncertain, people prefer the future," *Management Science*, 63, 519-527.

Harless, D. W. (1992). Actions versus prospects: The effect of problem representation on regret. *American Economic Review*, 82, 634-649.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk *Econometrica*, 47, 263-291.

Keren, G., and Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice. *Organizational Behavior and Human Decision Processes*, 63, 287-297.

Koopmans, T.C. (1960). Stationarity Ordinal Utility and Impatience. *Econometrica*, 28, 287-309.

Koszegi, B., and Szeidl, A. (2013). A model of focusing in economic choice. *Quarterly Journal of Economics. 128,* 53-104.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112, 443-478.

Leland, J. (1994). Generalized similarity judgments: An alternative explanation for choice anomalies. *Journal of Risk and Uncertainty*. 9 (2) 151-172.

Leland, J. (2002). Similarity Judgments and Anomalies in Intertemporal Choice. *Economic Inquiry, 40,* 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,* 568-577.

Leland, J., Schneider, M., and Wilcox (2017). Minimal and Transparent frames for Risk, Time, and Uncertainty. Manuscript. https://www.chapman.edu/research/institutes-and-centers/economic-science-institute/_files/WorkingPapers/leland-schneider-wilcox-swup-2017.pdf

Leland, J. and Sileo, P. (1998). The "Nothing to Gain/Nothing to Lose" Effect – Gambling and Insurance Purchase. Manuscript.

Loewenstein, G., and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and interpretation. *Quarterly Journal of Economics, 107,* 573-597.

Loomes, G., and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92,* 805-824.

Magen, E., Dweck, C. and Gross, J. (2008). The hidden-zero effect. *Psychological Science*, *19,* 648–649.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, 60, 151-158.

Moyer, R. and Landauer, T. (1967). The Time Required For Judgments of Numerical Inequality. *Nature*, 215, 1519-1520.

Pandelaere, M., Briers, B., and Lembregts, C. (2011). How to make a 29% increase look bigger: The unit effect in option comparisons. *Journal of Consumer Research*, *38,* 308-322.

Prelec, D., and Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science, 37,* 770-786.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3, 323-343.

Rao, L-L. and Li, S. (2011). New paradoxes in intertemporal choice. *Judgment and Decision Making*, *6*, 122-129.

Read, D., Frederick, S., Orsel, B., and Rahman, J. (2005). Four score and seven years from now: The Date/Delay effect in Temporal Discounting. *Management Science,* 51, 1326-1335.

Read, D., and Scholten, M. (2012). Tradeoffs between sequences: Weighing accumulated outcomes against outcome-adjusted delays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1675-1688.

Read, D., Olivola, C.Y., Hardisty, D.J. (2017). The value of nothing: Asymmetric attention to opportunity costs drives intertemporal decision making. Forthcoming, *Management Science.*

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.

Samuelson, P. (1937). A note on measurement of utility. *Rev. Econ. Stud.* 4. 155-161.

Savage, L. J. (1954). *The Foundations of Statistics*. New York, Wiley.

Schley, D. R., and Peters, E. (2014). Assessing “economic value”: Symbolic-number mappings predict risky and riskless valuations. *Psychological science*, *25*, 753-761.

Scholten, M., and Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychological Review*, 117, 925-944.

Stevens, S. S. (1957). On the psychophysical law. *Psychological review*, *64*, 153-181.

Strotz, R.H. (1955). Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23, 165-180.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.

Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297-323.

Wertenbroch, K., Soman, D. and Chattopadhyay (2007). On the perceived value of money: The reference dependence of currency numerosity effects. *Journal of Consumer Research*, *34*, 1 – 10.

Von Neumann, J., and Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton.