Salience Theory of Choice Under Risk

Pedro Bordalo
Nicola Gennaioli
Andrei Shleifer

This version: December 2011
(March 2010)

Barcelona GSE Working Paper Series

Working Paper nº 501

Salience Theory of Choice Under Risk

Pedro Bordalo

Nicola Gennaioli

Andrei Shleifer *

First Draft, March 2010. Revised, December 2011

Abstract

We present a theory of choice among lotteries in which the decision maker's attention is drawn to (precisely defined) salient payoffs. This leads the decision maker to a context-dependent representation of lotteries in which true probabilities are replaced by decision weights distorted in favor of salient payoffs. By specifying decision weights as a function of payoffs, our model provides a novel and unified account of many empirical phenomena, including frequent risk-seeking behavior, invariance failures such as the Allais paradox, and preference reversals. It also yields new predictions, including some that distinguish it from Prospect Theory, which we test.

JEL classification: D03, D81

*Harvard University, CREI and Universitat Pompeu Fabra, Harvard University. We are grateful to Nicholas Barberis, Gary Becker, Colin Camerer, John Campbell, Tom Cunningham, Xavier Gabaix, Morgan Grossman-McKee, Ming Huang, Jonathan Ingersoll, Emir Kamenica, Daniel Kahneman, Botond Koszegi, David Laibson, Pepe Montiel Olea, Drazen Prelec, Matthew Rabin, Josh Schwartzstein, Jesse Shapiro, Jeremy Stein, Tomasz Strzalecki, Dmitry Taubinsky, Richard Thaler, Georg Weiszacker, George Wu and three referees of this journal for extremely helpful comments, and to Allen Yang for excellent research assistance. Gennaioli thanks the Spanish Ministerio de Ciencia y Tecnologia (ECO 2008-01666 and Ramon y Cajal grants), the Barcelona GSE Research Network, and the Generalitat de Catalunya for financial support. Shleifer thanks the Kauffman Foundation for research support.
Corresponding author: Department of Economics, Harvard University, Littauer Center M-9, 1805 Cambridge Street, Cambridge 02138, MA. Email: ashleifer@harvard.edu, Ph: 617-496-2606, Fax: 617-496-1708

1 Introduction

Over the last several decades, social scientists have identified a range of important violations of Expected Utility Theory, the standard theory of choice under risk. Perhaps at the most basic level, in both experimental situations and everyday life, people frequently exhibit both risk loving and risk averse behavior, depending on the situation. As first stressed by Friedman and Savage (1948), people participate in unfair gambles, pick highly risky occupations (including entrepreneurship) over safer ones, and invest without diversification in individual risky stocks, while simultaneously buying insurance. Attitudes towards risk are unstable in this very basic sense.

This systematic instability underlies several paradoxes of choice under risk. As shown by Allais (1953), people switch from risk loving to risk averse choices among two lotteries after a common consequence is added to both, in contradiction to the independence axiom of Expected Utility Theory. Another form of instability is preference reversals (Lichtenstein and Slovic, 1971): in comparing two lotteries with a similar expected value, experimental subjects *choose* the safer lottery but are willing to *pay more* for the riskier one. Camerer (1995) reviews numerous attempts to amend the axioms of Expected Utility Theory to deal with these findings, but these attempts have not been conclusive.

We propose a new psychologically founded model of choice under risk, which naturally exhibits the systematic instability of risk preferences and accounts for the puzzles. In this model, risk attitudes are driven by the salience of different lottery payoffs. Psychologists view salience detection as a key attentional mechanism enabling humans to focus their limited cognitive resources on a relevant subset of the available sensory data. As Taylor and Thompson (1982) put it: "Salience refers to the phenomenon that when one's attention is differentially directed to one portion of the environment rather than to others, the information contained in that portion will receive disproportionate weighting in subsequent judgments." According to Kahneman (2011, p. 324), "our mind has a useful capability to focus on whatever is odd, different or unusual." We call the payoffs that draw the decision maker's attention "salient". The decision maker is then risk seeking when a lottery's upside is salient and risk averse when its downside is salient. More generally, salience allows for a theory of context dependent choice consistent with a broad range of evidence.

We build a model of decision making in which salient lottery payoffs are overweighted. Our main results rely on three assumptions. Two of them, which we label *ordering* and *diminishing sensitivity*, formalize the salience of payoffs. Roughly speaking, a lottery payoff is salient if it is very different in percentage terms from the payoffs of other available lotteries (in the same state of the world). This specification of salience captures the ideas that: i) we attend to differences rather than absolute values (Kahneman, 2003), and ii) we perceive changes on a log scale (Weber's law). Our third assumption states that the extent to which decision weights are distorted depends on the salience of the associated payoffs, and not on the underlying probabilities. This assumption implies (see Proposition 1) that low probabilities are relatively more distorted than high ones, in accordance with Kahneman and Tversky's (1979) observation that people have "limited ability to comprehend and evaluate extreme probabilities". We describe how, under these assumptions, the decision maker develops a context-dependent representation of each lottery. Aside from replacing objective probabilities with decision weights, the decision maker's valuation of payoffs is standard.

At a broad level, our approach is similar to that pursued by Gennaioli and Shleifer (2010) in their study of the representativeness heuristic in probability judgments. The idea of both studies is that decision makers do not take into account fully all the information available to them, but rather over-emphasize the information their minds focus on.[1] Gennaioli and Shleifer (2010) call such decision makers local thinkers, because they neglect potentially important but unrepresentative data. Here, analogously, in evaluating lotteries, decision makers overweight states that draw their attention and neglect states that do not. We continue to refer to such decision makers as local thinkers. In both models, the limiting case in which all information is processed correctly is the standard economic decision maker.

Our model describes factors that encourage and discourage risk seeking, but also leads to an explanation of the Allais paradoxes. The strongest departures from Expected Utility Theory in our model occur in the presence of extreme payoffs, particularly when these occur with a low probability. Due to this property, our model predicts that subjects in the Allais experiments are risk loving when the common consequence is small and attention is drawn

[1] Other models in the same spirit are Mullainathan (2002), Schwartzstein (2009) and Gabaix (2011).

to the highest lottery payoffs, and risk averse when the common consequence is large and attention is drawn to the lowest payoffs. We explore the model's predictions by describing, and then experimentally testing, how Allais paradoxes can be turned on and off. We also show that preference reversals can be seen as a consequence of lottery evaluation in different contexts that affect salience, rather than the result of a fundamental difference between pricing and choosing. The model thus provides a unified explanation of risk preferences and invariance violations based on a psychologically motivated mechanism of salience.

It is useful to compare our model to the gold standard of behavioral theories of choice under risk, Kahneman and Tversky's (KT, 1979) Prospect Theory. Like Prospect Theory, our model incorporates the assumption that decision makers focus on payoffs, rather than on absolute wealth levels, when evaluating risky alternatives (although in our model this happens through payoff salience and not through the value function). Prospect Theory also incorporates the assumption that the probability weights people use to make choices are different from objective probabilities. But the idea that these weights depend on the actual payoffs and their salience is new here. In some situations, our decision weights look very similar to KT's, but in other situations – for instance when small probabilities are *not* attached to salient payoffs or when lotteries are correlated – they are very different. We conduct multiple experiments, both of simple risk attitudes and of Allais paradoxes with correlated states, which distinguish our predictions from KT's, and uniformly find strong support for our model of probability weighting.

The paper proceeds as follows. In Section 2, we provide the basic intuition for how the salience of lottery payoffs shapes risk attitudes in the context of Allais' "common consequence" paradox. In Section 3, we present a salience-based model of choice among two lotteries. In Section 4, we use this model to study risk attitudes, derive from first principles Prospect Theory's weighting function for a class of choice problems where it should apply, and provide experimental evidence for our predictions. In Section 5 we show that our model accounts for the Allais paradoxes, as well as for preference reversals, a phenomenon that Prospect Theory cannot accommodate. We obtain further predictions for context effects (which Prospect Theory also cannot accomodate), such as turning the Allais paradoxes or preference reversals on and off depending on the description of payoff states, and find experimental support for these predictions. In Section 6, we address framing effects, failures of transitivity and extend the model to choice among many lotteries. Section 7 concludes.

2 Salience and the Allais Paradox

The Allais paradoxes (1953) are the best known and most discussed instances of failure of the independence axiom of Expected Utility Theory. Kahneman and Tversky's (1979) version of the "common consequence" paradox asks experimental subjects to choose among two lotteries $L_1(z)$ and $L_2(z)$:

$$L_1(z) = \begin{cases} \$2500 & \text{with prob.} \quad 0.33 \\ \$0 & \qquad\qquad\quad\; 0.01 \\ \$z & \qquad\qquad\quad\; 0.66 \end{cases}, \quad L_2(z) = \begin{cases} \$2400 & \text{with prob.} \quad 0.34 \\ \$z & \qquad\qquad\quad\; 0.66 \end{cases}, \tag{1}$$

for different values of the payoff z. By the independence axiom, an expected utility maximizer should not change his choice as the common consequence z is varied, since z cancels out in the comparison between $L_1(z)$ and $L_2(z)$.

In experiments, for $z = 2400$, most subjects are risk averse, preferring $L_2(2400)$ to $L_1(2400)$:

$$L_1(2400) = \begin{cases} \$2500 & \text{with prob.} \quad 0.33 \\ \$0 & \qquad\qquad\quad\; 0.01 \\ \$2400 & \qquad\qquad\quad\; 0.66 \end{cases} \prec \; L_2(2400) = \left\{ \$2400 \quad \text{with prob.} \quad 1 \right. . \tag{2}$$

When however $z = 0$, most subjects are risk seeking, preferring $L_1(0)$ to $L_2(0)$:

$$L_1(0) = \begin{cases} \$2500 & \text{with prob.} \quad 0.33 \\ \$0 & \qquad\qquad\quad\; 0.34 \end{cases} \succ \; L_2(0) = \begin{cases} \$2400 & \text{with prob.} \quad 0.34 \\ \$0 & \qquad\qquad\quad\; 0.66 \end{cases}. \tag{3}$$

In violation of the independence axiom, z affects the experimental subjects' choices, causing switches between risk averse and risk seeking behavior. Prospect Theory (KT 1979 and TK 1992) explains these switches as follows. When $z = 2400$, the low 0.01 probability of getting zero in $L_1(2400)$ is overweighted, generating risk aversion. When $z = 0$, the extra 0.01 probability of getting zero in $L_1(0)$ is not overweighted, generating risk seeking. This effect is directly built into the probability weighting function $\pi(p)$ by the assumption of subcertainty, e.g. $\pi(0.34) - \pi(0) < 1 - \pi(0.66)$.[2]

Our explanation of the Allais paradox does not rely on a fixed weighting function $\pi(p)$. Rather, it relies on how decision weights change as the payoff z alters the salience of different lottery outcomes. Roughly speaking, in the choice between $L_1(2400)$ and $L_2(2400)$, the downside of \$0 feels a lot lower than the sure payoff of \$2400. The upside of \$2500, however, feels only slightly higher than the sure payoff. Because the lottery's downside is more salient than its upside, the subjects focus on the downside when making their decisions. This focus triggers the risk averse choice.

In contrast, in the choice between $L_1(0)$ and $L_2(0)$, both lotteries have the same downside risk of zero. Now the upside of winning \$2500 in the riskier lottery $L_1(0)$ is more salient and subjects focus on it when making their decisions. This focus triggers the risk seeking choice. The analogy here is to sensory perception: a lottery's salient payoffs are those which differ most from the payoffs of alternative lotteries. The decision maker's mind then focuses on salient payoffs, inflating their weights when making a choice. Section 5 provides a fuller account of the Allais experiment, which also highlights the role played by the level of objective probabilities.

3 The Model

A choice problem is described by: i) a set of states of the world S, where each state $s \in S$ occurs with objective and known probability π_s such that $\sum_{s\in S} \pi_s = 1$, and ii) a choice set $\{L_1, L_2\}$, where the L_i are risky prospects that yield monetary payoffs x_s^i in each state s. For convenience, we refer to L_i as lotteries.[3] Here we focus on choice between two lotteries,

[2] In Cumulative Prospect Theory (Tversky and Kahneman, 1992) the mathematical condition on probability weights is slightly different but carries the same intuition: the common consequence is more valuable when associated with a sure rather than a risky prospect.

[3] Formally, L_i are acts, or random variables, defined over the choice problem's probability space (S, F_S, π), where S is assumed to be finite and F_S is its canonical σ-algebra. However, as we will see in Equation (11),

leaving the general case of choice among $N > 2$ lotteries to Section 6.

The decision maker uses a value function v to evaluate lottery payoffs relative to the reference point of zero.[4] Through most of the paper, we illustrate the mechanism generating risk preferences in our model by assuming a *linear* value function v. In section 6.4, when we focus on mixed lotteries, we consider a piece-wise linear value function featuring loss aversion, as in Prospect Theory. Absent distortions in decision weights, the local thinker evaluates L_i as:

$$V(L_i) = \sum_{s \in S} \pi_s v(x_s^i). \tag{4}$$

The local thinker (LT) departs from Equation (4) by overweighting the lottery's most salient states in S. Salience distortions work in two steps. First, a salience ranking among the states in S is established for each lottery L_i. Second, based on this salience ranking the probability π_s in (4) is replaced by a transformed, lottery specific decision weight π_s^i. To formally define salience, let $\mathbf{x}_s = (x_s^i)_{i=1,2}$ be the vector listing the lotteries' payoffs in state s and denote by x_s^{-i} the payoff in s of lottery $L_j, j \neq i$. Let $x_s^{\min}, x_s^{\max}$ respectively denote the largest and smallest payoffs in $\mathbf{x}_s$.

Definition 1 *The salience of state s for lottery L_i, $i = 1, 2$, is a continuous and bounded function $\sigma(x_s^i, x_s^{-i})$ that satisfies three conditions:*

1) *Ordering. If for states $s, \widetilde{s} \in S$ we have that $[x_s^{\min}, x_s^{\max}]$ is a subset of $[x_{\widetilde{s}}^{\min}, x_{\widetilde{s}}^{\max}]$, then*

$$\sigma\left(x_s^i, x_s^{-i}\right) < \sigma\left(x_{\widetilde{s}}^i, x_{\widetilde{s}}^{-i}\right)$$

2) *Diminishing sensitivity. If $x_s^j > 0$ for $j = 1, 2$, then for any $\epsilon > 0$,*

$$\sigma(x_s^i + \epsilon, x_s^{-i} + \epsilon) < \sigma(x_s^i, x_s^{-i})$$

the decision maker's choice depends only on the L_i's joint distribution over payoffs and not on the exact structure of the state space. Thus we use the term lotteries, in a slight abuse of nomenclature relative to the usual definition of lotteries as probability distributions over payoffs.

[4]This is a form of narrow framing, also used in Prospect Theory. Koszegi and Rabin (2006, 2007) build a model of reference point formation and use it to study shifts in risk attitudes. Their model cannot account for situations where expectations and thus reference points are held fixed (such as lab experiments we consider here). Our approaches are complementary, as one could combine our model of decision weights with Koszegi and Rabin's two part value function.

3) *Reflection. For any two states* $s, \widetilde{s} \in S$ *such that* $x^j_s, x^j_{\widetilde{s}} > 0$ *for* $j = 1, 2$, *we have*

$$\sigma(x^i_s, x^{-i}_s) < \sigma(x^i_{\widetilde{s}}, x^{-i}_{\widetilde{s}}) \quad \textit{if and only if} \quad \sigma(-x^i_s, -x^{-i}_s) < \sigma(-x^i_{\widetilde{s}}, -x^{-i}_{\widetilde{s}})$$

Section 3.1 discusses the connection between these properties and the cognitive notion of salience. The key properties driving our explanations of anomalies are ordering and diminishing sensitivity. The reflection property only plays a role in Section 6.4 when we consider lotteries which yield negative payoffs. To illustrate Definition 1, consider the salience function:

$$\sigma(x^i_s, x^{-i}_s) = \frac{|x^i_s - x^{-i}_s|}{|x^i_s| + |x^{-i}_s| + \theta}, \tag{5}$$

where $\theta > 0$. According to the ordering property, the salience of a state for L_i increases in the distance between its payoff x^i_s and the payoff x^{-i}_s of the alternative lottery. In (5), this is captured by the numerator $|x^i_s - x^{-i}_s|$. Diminishing sensitivity implies that salience decreases as a state's average (absolute) payoff gets farther from zero, as captured by the denominator term $|x^1_s| + |x^2_s|$ in (5). Finally, according to reflection, salience is shaped by the magnitude rather than the sign of payoffs: a state is salient not only when the lotteries bring sharply different gains, but also when they bring sharply different losses. In (5), reflection takes the strong form $\sigma(x^i_s, x^{-i}_s) = \sigma(-x^i_s, -x^{-i}_s)$. These properties are illustrated in Figure 1 below.

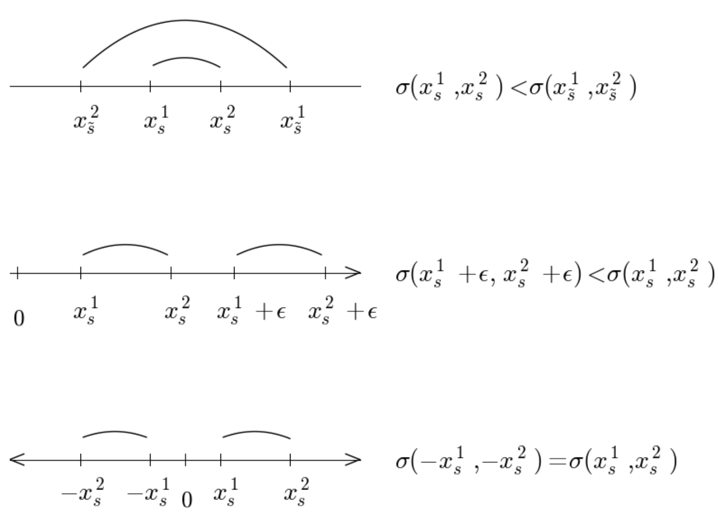

Figure 1: Properties of a salience function, Eq. (5)

The salience function in specification (5) satisfies additional properties besides those of Definition 1. For instance, it is symmetric, namely $\sigma(x_s^1, x_s^2) = \sigma(x_s^2, x_s^1)$, which is a natural property in the case of two lotteries but which is dropped with $N > 2$ lotteries. Although our main results rely only on ordering and diminishing sensitivity, we sometimes use the tractable functional form (5) to illustrate our model.

Consider the choice between $L_1(z)$ and $L_2(z)$ introduced in Section 2. When the common consequence is $z = 2400$, the possible payoff states are $S = \{(2500, 2400),\ (0, 2400), (2400, 2400)\}$. We then have:

$$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400). \tag{6}$$

The inequalities follow from diminishing sensitivity and ordering, respectively, and can be easily verified for Equation (5). The state in which the riskier lottery $L_1(2400)$ loses is the most salient one (which causes risk aversion).[5] A similar calculation shows that, when the common consequence is $z = 0$, the state $(2500, 0)$ in which the risky lottery $L_1(0)$ wins is the most salient one, which points to risk seeking. In short, changing the common consequence affects the salience of lottery payoffs, as described in Section 2. Section 5.1 provides a full analysis of the Allais paradoxes.

3.1 Salience, Decision Weights and Risk Attitudes

Given a salience function σ, for each lottery L_i the local thinker ranks the states and distorts their decision weights as follows:

Definition 2 *Given states $s, \widetilde{s} \in S$, we say that for lottery L_i state s is more salient than $\tilde{s}$ if $\sigma(x_s^i, x_s^{-i}) > \sigma(x_{\widetilde{s}}^i, x_{\widetilde{s}}^{-i})$. Let $k_s^i \in \{1, ..., |S|\}$ be the salience ranking of state s for L_i, with lower k_s^i indicating higher salience. All states with the same salience obtain the same ranking (and the ranking has no jumps). Then, if s is more salient than $\widetilde{s}$, namely if $k_s^i < k_{\widetilde{s}}^i$, the local thinker transforms the odds $\pi_{\widetilde{s}}/\pi_s$ of $\widetilde{s}$ relative to s into the odds $\pi_{\widetilde{s}}^i/\pi_s^i$, given by:*

$$\frac{\pi_{\widetilde{s}}^i}{\pi_s^i} = \delta^{k_{\widetilde{s}}^i - k_s^i} \cdot \frac{\pi_{\widetilde{s}}}{\pi_s} \tag{7}$$

[5] In this example, constructing the state space from the alternatives of choice is straightforward. Section 3.2 describes how the state space S is constructed in more complex cases.

where $\delta \in (0,1]$. *By normalizing* $\sum_s \pi_s^i = 1$ *and defining* $\omega_s^i = \delta^{k_s^i} / \left(\sum_r \delta^{k_r^i} \cdot \pi_r\right)$, *the decision weight attached by the local thinker to a generic state* s *in the evaluation of* L_i *is*:

$$\pi_s^i = \pi_s \cdot \omega_s^i. \tag{8}$$

The local thinker evaluates a lottery by inflating the relative weights attached to the lottery's most salient states. Parameter δ measures the extent to which salience distorts decision weights, capturing the degree of local thinking. When $\delta = 1$, the decision maker is a standard economic decision maker: his decision weights coincide with objective probabilities (i.e., $\omega_s^i = 1$). When $\delta < 1$, the decision maker is a local thinker, namely he overweights the most salient states and underweights the least salient ones. Specifically, s is overweighted if and only if it is more salient than average ($\omega_s^i > 1$, or $\delta^{k_s^i} > \sum_r \delta^{k_r^i} \cdot \pi_r$). The case where $\delta \to 0$ describes the local thinker who focuses only on a lottery's most salient payoffs.

The critical property of Definition 2 is that the parameter δ does not depend on the objective state probabilities. We discuss the cognitive motivations for this assumption in Section 3.1. This specification implies:

Proposition 1 *If the probability of state* s *is increased by* $\mathrm{d}\pi_s = h \cdot \pi_s$, *where* h *is a positive constant, and the probabilities of other states are reduced while keeping their odds constant, i.e.* $\mathrm{d}\pi_{\widetilde{s}} = -\frac{\pi_s}{1-\pi_s} h \cdot \pi_{\widetilde{s}}$ *for all* $\widetilde{s} \neq s$, *then:*

$$\frac{\mathrm{d}\omega_s^i}{h} = -\frac{\pi_s}{1-\pi_s} \cdot \omega_s^i \cdot \left(\omega_s^i - 1\right). \tag{9}$$

Proposition 1 (see the Appendix for proofs) states that an increase in a state's probability π_s reduces the distortion of the decision weight in that state by driving ω_s^i closer to 1. That is, low probability states are subject to the strongest distortions: they are over-weighted if salient and under-weighted otherwise. In contrast to KT's (1979,1992) assumption, low probability (high rank) payoffs are not always overweighted in our model; they are only overweighted if they are salient, regardless of probability (and rank). In accordance with KT, however, the largest distortions of choice occur precisely when salient payoffs are relatively unlikely. This property plays a key role for explaining some important findings such as the common ratio Allais Paradox in Section 5.1.[6]

Given Definitions 1 and 2, the local thinker computes the value of lottery L_i as:

$$V^{LT}(L_i) = \sum_{s \in S} \pi_s^i v(x_s^i) = \sum_{s \in S} \pi_s \omega_s^i v(x_s^i). \tag{10}$$

Thus, L_i's evaluation always lies between the value of its highest and lowest payoffs.

Since salience is defined on the state space S, one may wonder whether splitting states, or generally considering a different state space compatible with the lotteries' payoff distributions, may affect the local thinker's evaluation (10). We denote by X the set of distinct payoff combinations of L_1, L_2 occurring in S with positive probability, and by $S_{\mathbf{x}}$ the set of states in S where the lotteries yield the same payoff combination $\mathbf{x} \in X$, formally $S_{\mathbf{x}} \equiv \{s \in S \,|\mathbf{x}_s = \mathbf{x}\}$. Clearly, $S = \cup_{\mathbf{x} \in X} S_{\mathbf{x}}$. By Definition 1, all states s in $S_{\mathbf{x}}$ are equally salient for either lottery, and thus have the same value of ω_s^i, which for simplicity we denote $\omega_{\mathbf{x}}^i$. Using (8) we can rewrite $V^{LT}(L_i)$ in (10) as:

$$V^{LT}(L_i) = \sum_{\mathbf{x} \in X} \left(\sum_{s \in S_{\mathbf{x}}} \pi_s \right) \omega_{\mathbf{x}}^i v(x_{\mathbf{x}}^i), \tag{11}$$

where $x_{\mathbf{x}}^i$ denotes L_i's payoff in $\mathbf{x}$. Equation (11) says that the state space only influences evaluation through the total probability of each distinct payoff combination $\mathbf{x}$, namely $\pi_{\mathbf{x}} = \sum_{s \in S_{\mathbf{x}}} \pi_s$. This is because salience $\sigma(.,)$ depends on payoffs, and not on the probabilities of different states. Hence, splitting a given probability $\pi_{\mathbf{x}}$ across different sets of states does not affect evaluation (or choice) in our model. There is therefore no loss in generality from viewing S as the "minimal" state space X identified by the set of distinct payoff combinations that occur with positive probability. In the remainder of the paper, we keep the notation of Equation (10), with the understanding that S is this "minimal" state space (and omit the reference to the underlying lotteries).

[6]Proposition 1 can also be stated in terms of payoffs: if lottery L_i yields payoff x_k with probability p_k, then increasing p_k while reducing the probabilities $p_{k'}$ of other payoffs $x_{k'}$ (keeping their odds constant) decreases the distortion of p_k if and only if x_k is more salient than average. That is, in a given choice context, the probabilities of unlikely payoffs are relatively more distorted (see the Appendix for details).

In a choice between two lotteries, Equation (10) implies that - due to the symmetry of the salience function (i.e. $k_s^1 = k_s^2$ for all s) - the local thinker prefers L_1 to L_2 if and only if:

$$\sum_{s \in S} \delta^{k_s} \pi_s \left[v(x_s^1) - v(x_s^2) \right] > 0. \tag{12}$$

For $\delta = 1$, the local thinker's decision weights coincide with the corresponding objective probabilities. For $\delta < 1$, local thinking favors L_1 when it pays more than L_2 in the more salient (and thus less discounted) states.

3.2 Discussion of Assumptions and Setup

Salience and Decision Weights

In our model the choice context shapes decision makers' perception of lotteries through the mechanism of payoff salience. The properties of the salience function seek to formalize features of human perception, which we believe – in line with Kahneman, Tversky, and others – to be relevant for choice under risk. The intensity with which we perceive a signal, such as a light source, increases in the signal's magnitude but also depends on context (Kandel et al, 1991). Analogously, in choice under risk the signals are the differences in lottery payoffs across states. Via the ordering property, the salience function $\sigma(.,.)$ captures the signal's magnitude in a given state. The role of context is captured by diminishing sensitivity (and reflection): the intensity with which payoffs in a state are perceived increases as the state's payoffs approach the status quo of zero, which is our measure of context.[7]

Consistent with psychology of attention, we assume that the decision maker evaluates lotteries by focusing on, and weighting more, their most salient states. The "local thinking" parameter $1/\delta$ captures the strength of the decision maker's focus on salient states, proxying

[7] As in Weber's law of diminishing sensitivity, in which a change in luminosity is perceived less intensely if it occurs at a higher luminosity level, the local thinker perceives less intensely payoff differences occurring at high (absolute) payoff levels. Interestingly, visual perception and risk taking seem to be connected at a more fundamental neurological level. McCoy and Platt (2005) show in a visual gambling task that when monkeys made risky choices neuronal activity increased in an area of the brain (CGp, the posterior cingulate cortex) linked to visual orienting and reward processing. Crucially, the activation of CGp was better predicted by the subjective salience of a risky option than by its actual value, leading the authors to hypothesize that "enhanced neuronal activity associated with risky rewards biases attention spatially, marking large payoffs as salient for guiding behavior (p. 1226)."

for his ability to pay attention to multiple aspects, cognitive load, or simply intelligence. Our assumption of rank-based discounting buys us analytical tractability, but our main results also hold if the distortion of the odds in (7) is a smooth increasing function of salience differences, for instance $\delta^{\left[\sigma(x_s^i, x_s^{-i}) - \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})\right]}$.[8] One benefit of this alternative specification is that it would avoid discontinuities in valuation. However, discontinuities play no role in our analysis, so for simplicity we stick to ranking-based discounting. The main substantive restriction embodied in our model is that the discounting function does not depend on a state's probability, which implies that unlikely states are subject to the greatest distortions. This notion is also encoded in Prospect Theory's weigthing function, in which "highly unlikely events are either ignored or overweighted." (KT 1979). Together with subadditivity, this feature, also present in early work on probability weigthing (Edwards 1962, Fellner 1961), allows KT to account for risk loving behavior and the Allais paradoxes. Quiggin's (1982) rank-dependent expected utility and Tversky and Kahneman's (1992) Cumulative Prospect Theory (CPT) develop weigthing functions in which the rank order of a lottery's payoffs affects probability weighting.[9]

Our theory exhibits two sharp differences from these works. First, in our model the magnitude of payoffs, not only their rank, determines salience and probability weights: unlikely events are overweighted when they are associated with salient payoffs, but underweighted otherwise. As a consequence, the lottery upside may still be underweighted if the payoff associated with it is not sufficiently high. As we show in Section 4, this feature is crucial to explaining shifts in risk attitudes. Second, and more important, in our model decision weights depend on the choice context, namely on the available alternatives as they are presented to the decision maker. In Section 5 we exploit this feature to shed light on the psychological forces behind the Allais paradoxes and preference reversals.

Our main results rely on ordering and diminishing sensitivity of $\sigma(\cdot,\cdot)$, as well as on the comparatively larger distortion of low probabilities. We however sometimes illustrate

[8]A smooth specification would also address a concern with the current model that states with similar salience may obtain very different weights. This implies that i) splitting states and slightly altering payoffs could have a large impact on choice, and ii) in choice problems with many states the (slightly) less salient states are effectively ignored. However, since none of our results is due to these effects, we stick to rank-based discounting for simplicity.

[9]Prelec (1998) axiomatizes a set of theories of choice based on probability weighting, which include CPT. For a recent attempt to estimate the probability weighting function, see Wu and Gonzalez (1996).

the model by using the more restrictive salience function in Equation (5), which offers a tractable case characterized by only two parameters (θ, δ). This allows us to look for ranges of θ and δ that are consistent with the observed choice patterns.

The State Space

Salience is a property of states of nature that depends on the lottery payoffs that occur in each state, as they are presented to the decision maker. The assumption that payoffs (rather than final wealth states) shape the perception of states is a form of narrow framing, consistent with the fact that payoffs are perceived as gains and losses relative to the status quo, as in Prospect Theory.

In our approach, the state space S and the states' objective probabilities are a given of the choice problem.[10] In the lab, specifying a state space for a choice problem is straightforward when the feasible payoff combinations – and their probabilities – are available, for instance when lotteries are explicitly described as contingencies based on a randomizing device. For example, $L_1 \equiv (10, 0.5;\ 5, 0.5)$ and $L_2 \equiv (7, 0.5;\ 9, 0.5)$ give rise to four payoff combinations $\{(10, 7), (10, 9), (5, 7), (5, 9)\}$ if they are played by flipping two separate coins, but only to two payoff combinations $\{(10, 7), (5, 9)\}$ if they are contingent on the same coin flip. In our experiments, we nearly always describe the lotteries' correlation structure by specifying the state space. However, classic experiments such as the Allais paradoxes provide less information: they involve a choice between (standard) lotteries, and the state space is not explicitly described. In this case, we assume that our decision maker treats the lotteries as independent, which implies that the state space is the product space induced by the lotteries' marginal distributions over payoffs.[11] Intuitively, salience detects the starkest payoff differences among lotteries unless some of these differences are explicitly ruled out.

Although for all choice phenomena we study the choice set and thus the state space is unambiguous, in real world applications it may be necessary to make assumptions as to what consideration set the decision maker is actively entertaining. For example, the decision maker may discard universally dominated lotteries from his choice set before evaluating other, more

[10] In particular, we do not address choice problems where outcome probabilities are ambiguous, such as the Ellsberg paradox. This is an important direction for future work. Similarly, the salience-based decision weights are not to be understood as subjective probabilities.

[11] In the Online Appendix (Supplementary Material) we provide experimental evidence consistent with this assumption, as well as details on the information given in the experimental surveys.

attractive, lotteries (see Section 6). As another example, suppose that the payoffs of two lotteries are determined by the roll of the same dice. One lottery pays 1,2,3,4,5,6, according to the dice's face; the other lottery pays 2,3,4,5,6,1. The state in which the first lottery pays 6 and the second pays 1 may appear most salient to the decision maker, leading him to prefer the first lottery. Of course, a moment's thought would lead him to realize that the lotteries are just rearrangements of each other, and recognize them as identical. In the following, we assume that, before evaluating lotteries, the decision maker edits the choice set by discarding all but one of the lottery permutations (at random, thus preserving indifference between the permutations). Both forms of editing are plausibly related to salience itself: in these cases, before comparing payoffs, what is salient to the decision maker are the properties of permutation or dominance of certain lotteries. To focus our study on the salience of lottery payoffs, we do not formally model this editing process, as it plays no role in any of the choice phenomena we study here. However, endogenizing the choice set is an important direction for future work. There is a large literature on consideration set determination in marketing and a growing one in decision theory (e.g., Manzini and Mariotti 2007, Masatlioglu, Nakajima and Ozbay 2010), but a consensus model has not yet emerged. In a similar spirit, the model could be generalized to take into account determinants of salience other than payoff values, such as prior experiences and details of presentation, or even color of font. These may matter in some situations but are not considered here.

Salience and Context Dependent Choice

We are not the first to propose a model of context dependent choice among lotteries. Rubinstein (1988), followed by Aizpurua et al (1990) and Leland (1994), builds a model of similarity-based preferences, in which decision makers simplify the choice among two lotteries by pruning the dimension (probability or payoff, if any), along which lotteries are similar. The working and predictions of our model are different from Rubinstein's, even though we share the idea that the common ratio Allais paradox (see Section 5.1.2) is due to subjects' focus on lottery payoffs. In Regret Theory (Loomes and Sugden 1982, Bell 1982, Fishburn 1982), the choice set directly affects the decision maker's utility via a regret/rejoice term added to a standard utility function. In our model, instead, context affects decisions by shaping the salience of payoffs and decision weights. Regret Theory can account for a certain type of context dependence, such as a role for correlations among lotteries; however, by adopting a traditional utility theory perspective, it cannot capture framing effects or violations of procedural invariance (Tversky, Slovic and Kahneman 1990). Moreover, since Regret Theory does not feature diminishing sensitivity (as it excludes the notion of a reference point), it has a hard time accounting for standard patterns of risk preferences, including risk averse preferences for fair 50-50 gambles over gains and their reflection over losses.

Formal models of context dependent choice (e.g. Fishburn 1982) may be criticized as not being falsifiable because different choice patterns can be justified. We stress that our psychologically based assumptions of ordering and diminishing sensitivity place tight restrictions on the predictions of our model under any value (and salience) function. To give one example, both the ordering and the diminising sensitivity property make strong predictions regarding the conditions for, and the directionality of, the Allais paradox. In particular, they imply that the independence axiom of Expected Utility Theory should hold when the mixture lotteries are correlated (see Section 5.1). To give another example, the distortion of decision weights in Definition 2 implies that pairwise choice among two or three outcome independent lotteries having the same support is transitive (we address intransitivities in Section 6.3), and that choice is consistent with first order stochastic dominance when lotteries are independent (see Online Appendix). In future work, it may be useful to uncover the precise axioms consistent with Definitions 1 and 2.

4 Salience and Attitudes Towards Risk

We first describe how salience affects the risk preferences of a local thinker with linear utility. To do so, consider the choice between a sure prospect $L_0 = (x, 1)$ and a mean preserving spread $L_1 = (x + g, \pi_g; x - l, 1 - \pi_g)$, with $g\pi_g = (1 - \pi_g)l$. All payoffs are positive (we study mixed lotteries in Section 6.4). In this choice, there are two states: $s_g = (x + g, x)$, in which the lottery gains relative to the sure prospect, and $s_l = (x - l, x)$, in which the lottery loses.

Since L_1 is a mean preserving spread of L_0, Equation (12) implies that for any $\delta < 1$, a local thinker with linear utility chooses the lottery if and only if the gain state s_g is more salient than the loss state s_l, i.e. when $\sigma(x + g, x) > \sigma(x - l, x)$. In this case, using the notation of Definition 2, the weight π_g^1 attached to the event of winning under the lottery is higher than the event's probability π_g. As a result, the local thinker perceives the expected value of L_1 to be above that of L_0, and exhibits risk seeking behavior, choosing L_1 over L_0.

Using the fact that $g\pi_g = (1 - \pi_g)l$, the condition for s_g to be more salient than s_l can be written as:

$$\sigma\left(x + \frac{1-\pi_g}{\pi_g} \cdot l, x\right) > \sigma\left(x - l, x\right). \tag{13}$$

The ordering property of salience has two implications. First, when the state s_g is very unlikely, it is also salient: at $\pi_g \simeq 0$ the lottery's upside is very large, its salience is high, and (13) always holds. Second, the salience of s_g decreases in π_g: as the lottery wins with higher probability, its payoff gain g is lower and thus less salient. Thus, Equation (13) is less likely to hold as π_g rises. The diminishing sensitivity property in turn implies that when the lottery gain is equal to the loss (i.e., $g = l$), the loss is salient. As a consequence, when $\pi_g = 1/2$ the state s_g is less salient than s_l, so (13) is violated.

As a result, condition (13) identifies a probability threshold $\pi_g^* < 1/2$ such that: for $\pi_g < \pi_g^*$ the lottery upside is salient, the local thinker overweights it and behaves in a risk seeking way; for $\pi_g > \pi_g^*$ the lottery downside is salient, the local thinker overweights it and behaves in a risk averse way; for $\pi_g = \pi_g^*$ states s_g and s_l are equally salient and the local thinker is risk neutral. Remarkably, these properties of decision weights recover key features of Prospect Theory's inverse S-shaped probability weighting function (KT 1979): over-weighting of low probabilities, and under-weighting of high probabilities. Indeed, Figure 2 shows the decision weight π_g^1 as a function of probability π_g. Low probabilities are over-weighted because they are associated with salient upsides of longshot lotteries. High probabilities are under-weighted as they occur in lotteries with a small, non salient, upside.

Note however that in our model the weighting function is context dependent. In contrast to Prospect Theory, overweighting depends not only on the probability of a state but also on the salience of its payoff in (13). Overweighting is also shaped by the average level of payoffs x. To see this, denote by $r = v^{LT}(L_0) - v^{LT}(L_1)$ the "premium" required by the local thinker to be indifferent between the risky option L_1 and the sure prospect L_0 (r is positive

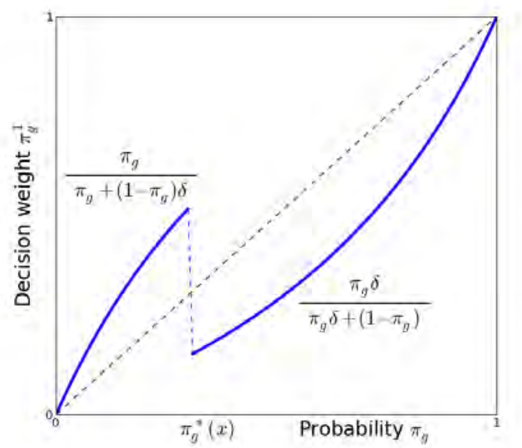

Figure 2: Context dependent probability weighting function

when the local thinker is risk averse). For a rational decision maker with linear utility, $r = 0$ regardless of the payoff level x. To see how the local thinker's risk attitudes depend on x, consider the following definition:

Definition 3 *A salience function is convex if, for any state with positive payoffs* (y, z) *and any* $x, \epsilon > 0$*, the difference* $\sigma(y + x, z + x) - \sigma(y + x + \epsilon, z + x + \epsilon)$ *is a decreasing function of the payoff level* x*. A salience function is concave if this difference increases in* x*.*

A salience function is convex if diminishing sensitivity becomes weaker as the payoff level x rises. The Appendix then proves:

Lemma 1 *If the salience function is convex, then* $r = v^{LT}(L_0) - v^{LT}(L_1)$ *weakly decreases with* x*. Conversely, if the salience function is concave then* r *weakly increases with* x*.*

If convexity holds and diminishing sensitivity becomes weaker with x, then a higher payoff level weakly reduces r, increasing the valuation of the risky lottery L_1 relative to that of the safe lottery L_0. In Equation (13), this increases the threshold π_g^*, boosting risk seeking. If instead diminishing sensitivity becomes stronger with x, a higher payoff level leads to an increase in r, weakly decreasing L_1's valuation relative to that of L_0. In equation (13) this reduces the threshold π_g^*, hindering risk seeking.

The salience function of Equation (5) satisfies convexity. Using this function, the condition (13) for s_g to be more salient than s_l becomes:

$$\left(x + \frac{\theta}{2}\right)(1 - 2\pi_g) > l \cdot (1 - \pi_g), \tag{14}$$

which is indeed more likely to hold for higher x (so long as $\pi_g < 1/2$).

Equation (14) implies that, holding the lottery loss l constant, risk attitudes follow Figure 3 below (where for convenience we set $\theta \cdot l \simeq 0$). As x rises, the threshold π_g^* below which

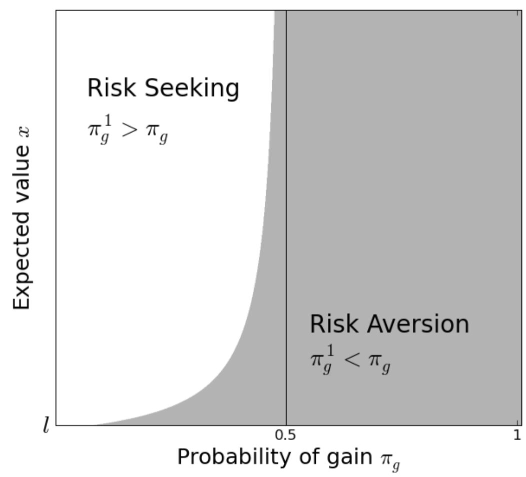

Figure 3: Shifts in risk attitudes

the decision maker is risk seeking increases, so that risk seeking behavior can occur even at relatively high probabilities π_g (but never for $\pi_g > 1/2$, though).

We tested the predictions illustrated in Figure 3 by giving experimental subjects a series of binary choices between a mean preserving spread $L_1 = (x + g, \pi_g;\ x - l, 1 - \pi_g)$ and a sure prospect $L_2 = (x, 1)$. We set the downside of L_1 at $l = \$20$, yielding an upside g of $\$20 \cdot (1 - \pi_g)/\pi_g$. We varied x in $\{\$20, \$100, \$400, \$2100, \$10500\}$ and π_g in $\{.01, .05, .2, .33, .4, .5, .67\}$. For each of these 35 choice problems, we collected at least 70 responses. On average, each subject made 5 choices, several of which held either π_g or x constant. The observed proportion of subjects choosing the lottery for every combination (x, π_g) is reported in Table 1; for comparison with the predictions of Figure 3, the results are shown in Figure 4.

Table 1: Proportion of Risk-Seeking Subjects

Expected value x	0.01	0.05	0.2	0.33	0.4	0.5	0.67
\$10500	0.83	0.65	0.50	0.48	0.46	0.33	0.23
\$2100	0.83	0.65	0.48	0.43	0.48	0.38	0.21
\$400	0.60	0.58	0.44	0.47	0.33	0.30	0.23
\$100	0.58	0.54	0.40	0.32	0.22	0.30	0.13
\$20	0.15	0.2	0.12	0.08	0.10	0.25	0.15

Probability of gain π_g

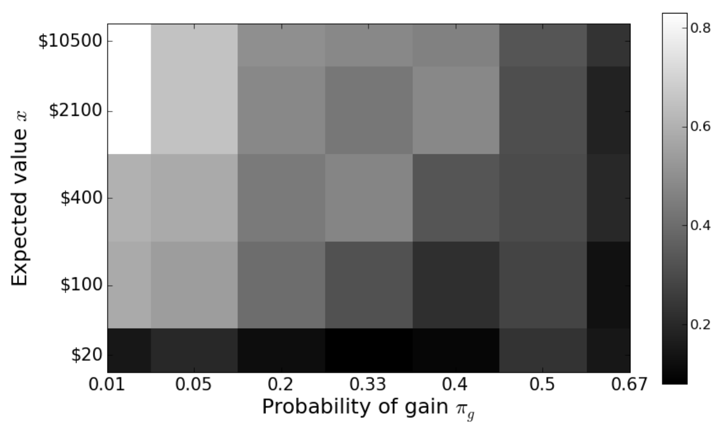

Figure 4: Proportion of Risk-Seeking Subjects

The patterns are qualitatively consistent with the predictions of Figure 3. First, and crucially, for any given expected value x, the proportion of risk takers falls as π_g increases and there is a large drop in risk taking as π_g crosses 0.5. This prediction is consistent with the probability weighting function depicted in Figure 2. Second, for a given $\pi_g < 0.5$, the proportion of risk takers increases with the expected value x. The effect is statistically significant: at $\pi_g = 0.05$ a large majority of subjects (80%) are risk averse when $x = \$20$, but as x increases to \$2100 a large majority (65%) becomes risk seeking. This finding is consistent with the finer hypothesis, encoded in equation (5), that diminishing sensitivity may become weaker at higher payoff levels. The increase in x raises the proportion of risk takers from around 10% to 50% even for moderate probabilities in the range $(0.2, 0.4)$.

Although not a formal test of our theory, these patterns are broadly consistent with the predictions of our model.[12] The Online Appendix describes additional experiments on longshot lotteries whose results are also consistent with out model but inconsistent with Prospect Theory under standard calibrations of the value function. In the Online Appendix we show that using the salience function in (5) the parameter values $\delta \sim 0.7$ and $\theta \sim 0.1$ are consistent with the above evidence on risk preferences, as well as with risk preferences concerning longshot lotteries. These values are not a formal calibration, but we employ them as a useful reference for discussing Allais paradoxes in the next section.

5 Local Thinking and Context Dependence

5.1 The Allais Paradoxes

5.1.1 The "common consequence " Allais Paradox

Let us go back to the Allais paradox described in Section 2. We now describe the precise conditions under which our model can explain it. Recall that subjects are asked to choose between the lotteries:

$$L_1(z) = (2500,\ 0.33; \quad 0,\ 0.01; \quad z,\ 0.66), \quad L_2(z) = (2400,\ 0.34; \quad z,\ 0.66) \tag{15}$$

for different values of z. For $z = 2400$, most subjects are risk averse, preferring $L_2(2400)$ to $L_1(2400)$, while for $z = 0$, most subjects are risk seeking, preferring $L_1(0)$ to $L_2(0)$.

When $z = 2400$, the minimal state space is $S = \{(2500, 2400), (0, 2400), (2400, 2400)\}$. The most salient state is one where the risky lottery L_1^{2400} pays zero because, by ordering

[12]The weighing function of Prospect Theory and CPT can explain why risk seeking prevails at low π_g, but not the shift from risk aversion to risk seeking as x rises. To explain this finding, both theories need a concave value function characterized by strongly diminishing returns. In the Online Appendix we provide further support for these claims by showing that standard calibrations of Prospect Theory cannot explain our experimental findings. For example, the calibration in Tversky and Kahneman (1992) features the value function $v(x) = x^{0.88}$, which is insufficiently concave. Importantly, calibrations of the value function are notoriously unstable: using two other sets of choice data, Wu and Gonzalez (1996) estimate $v(x) = x^{0.5}$ and $v(x) = x^{0.37}$, respectively. The fact that calibration is so dependent on the choice context suggests that choice itself is context dependent.

and diminishing sensitivity we have:

$$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400). \tag{16}$$

By Equation (12), a local thinker then prefers the riskless lottery $L_2(2400)$ provided:

$$-(0.01) \cdot 2400 + \delta \cdot (0.33) \cdot 100 < 0, \tag{17}$$

which holds for $\delta < 0.73$. Although the risky lottery $L_1(2400)$ has a higher expected value, it is not chosen when the degree of local thinking is severe, because its downside of 0 is very salient.

Consider the choice between $L_1(0)$ and $L_2(0)$. Now both options are risky and, as discussed in Section 3, the local thinker is assumed to see the lotteries as independent. The minimal state space now has four states of the world, i.e. $S = \{(2500, 2400), (2500, 0), (0, 2400), (0, 0)\}$, whose salience ranking is:

$$\sigma(2500, 0) > \sigma(0, 2400) > \sigma(2500, 2400) > \sigma(0, 0). \tag{18}$$

The first inequality follows from ordering, and the second from diminishing sensitivity. By Equation (12), a local thinker prefers the risky lottery L_1^0 provided:

$$(0.33) \cdot (0.66) \cdot 2500 - \delta \cdot (0.67) \cdot (0.34) \cdot 2400 + \delta^2 \cdot (0.33) \cdot (0.34) \cdot 100 > 0 \tag{19}$$

which holds for $\delta \geq 0$. Any local thinker with linear utility chooses the risky lottery $L_1(0)$ because its upside is very salient.

In sum, when $\delta < 0.73$ a local thinker exhibits the Allais paradox. This is true for any salience function satisfying ordering and diminishing sensitivity, and thus also for the parameterization $\delta = 0.7$, $\theta = 0.1$ obtained when using (5). It is worth spelling out the exact intuition for this result. When $z = 2400$, the lottery L_2^{2400} is safe, whereas the lottery L_1^{2400} has a salient downside of zero. The local thinker focuses on this downside, leading to risk aversion. When instead $z = 0$, the downside payoff of the safer lottery L_2^0 is also 0. As a result, the lotteries' upsides are now crucial to determining salience. This induces the local thinker to overweight the larger upside of L_1^0, triggering risk seeking. The salience of payoffs thus implies that when the same downside risk is added to the lotteries L_1^{2400} and L_2^{2400}, the sure prospect L_2^{2400} is particularly hurt because the common downside payoff induces the decision maker to focus on the larger upside of the risky lottery, leading to risk seeking behavior. This yields the "certainty effect" of Prospect Theory and CPT (KT 1979 and TK 1992) as a form of context dependence due to payoff salience.

This role of context dependence invites the following test. Suppose that subjects are presented the following correlated version of the lotteries $L_1(z)$ and $L_2(z)$ in Equation (15):

Probability	0.01	0.33	0.66
payoff of $L_1(z)$	0	2500	z
payoff of $L_2(z)$	2400	2400	z

(20)

where the table specifies the possible joint payoff outcomes of the two lotteries and their respective probabilities. Correlation changes the state space *but not* a lottery's distribution over final outcomes, so it does not affect choice under either Expected Utility Theory or Prospect Theory. Critically, this is not true for a local thinker: the context of this correlated version makes clear that the state in which both lotteries pay z is the least salient one, and also that it drops from evaluation in Equation (12), so that the value of z should not affect the choice at all. This is due to the ordering property: states where the two lotteries yield the same payoff are the least salient ones and in fact cancel out in the local thinker's valuation (ordering leads to them being "edited out" by the local thinker). That is, in our model – but not in Prospect Theory – the Allais paradox should not occur when $L_1(z)$ and $L_2(z)$ are presented in the correlated form as in (20).

We tested this prediction by presenting experimental subjects correlated formats of lotteries $L_1(z)$ and $L_2(z)$ for $z = 0$ and $z = 2400$. The observed choice pattern is the following:

	$L_1(2400)$	$L_2(2400)$
$L_1(0)$	7%	9%
$L_2(0)$	11%	73%

The vast majority of subjects do not reverse their preferences (80% of choices lie on the NW-SE diagonal), and most of them are risk averse, which in our model is also consistent with the fact that $(0, 2400)$ is the most salient state in the correlated choice problem (20). Among the few subjects reversing their preference, no clear pattern is detectable. This contrasts with the fact that our experimental subjects exhibit the Allais paradox when lotteries are presented in an uncorrelated form (see Online Appendix, Supplementary Material). Thus, when the lotteries pay the common consequence in the same state, choice is invariant to z and the Allais paradox disappears. Our model accounts for this fact because, as the common consequence z is made evident by correlation, it becomes non-salient. As a result, subjects prune it and choose based on the remaining payoffs.[13]

This result captures Savage's (1972, pg. 102) argument in defense of the normative character of the "sure thing principle", and validates his thought experiment. Other experiments in the literature are consistent with our results. Conlisk (1989) examines a related variation of the Allais choice problem, in which each alternative is given in compound form involving two simple lotteries, with one of the simple lotteries yielding the common consequence z. Birnbaum and Schmidt (2010) present the Allais problem in split form, singling out the common consequence z in each lottery. In both cases, the Allais reversals subside. Our model also rationalizes the disappearance of the Allais paradox in Colinsk's (1989) second example, which uses non-boundary lotteries. See also Harrison (1994) for related work on the common consequence paradox.

[13]We tested the robustness of the correlation result by changing the choice problem in several ways: 1) we framed the correlations verbally (e.g. described how the throw of a common die determined both lotteries' payoffs), 2) we repeated the experiment with uncertain real world events, instead of lotteries, and 3) we varied the ordering of questions, the number of filler questions, and payoffs. As the Online Appendix shows, our results are robust to all these variations. We also ran an experiment where subjects were explicitly presented the lotteries of Equation (15) with $z = 2400$ as uncorrelated, with a state space consisting of the four possible states. The choice pattern exhibited by subjects is: i) very similar to the one exhibited when the state space is not explicitly presented, validating our basic assumption that a decision maker assumes the lotteries to be uncorrelated when this is not specified otherwise, and ii) very different from the choice pattern exhibited under correlation (with 35% of subjects changing their choice as predicted by our model, see the Online Appendix).

5.1.2 The "common ratio" Allais Paradox

We now turn to the "common ratio" paradox, which occurs in the choice between lotteries:

$$L_1(\pi') = (6000,\ \pi';\quad 0,\ 1-\pi'),\quad L_2(\pi) = (\alpha\cdot 6000,\ \pi;\quad 0,\ 1-\pi), \tag{21}$$

where $L_1(\pi')$ is riskier than $L_2(\pi)$ in the sense that it pays a larger positive amount ($\alpha < 1$) with a smaller probability ($\pi' < \pi$). By the independence axiom, an expected utility maximizer with utility function $v(\cdot)$ chooses the safer lottery $L_2(\pi)$ over $L_1(\pi')$ when:

$$v(\alpha\cdot 6000) \geq \frac{\pi'}{\pi}\cdot v(6000) + v(0)\left(1-\frac{\pi'}{\pi}\right). \tag{22}$$

The choice should not vary as long as π'/π is kept constant. A stark case arises when $\pi'/\pi = \alpha$; now the two lotteries have the same expected value and a risk averse expected utility maximizer always prefers the safer lottery $L_2(\pi)$ to $L_1(\pi')$ for any π. Parameter α identifies the "common ratio" between π' and π at different levels of π.

It is well known (KT 1979) that, contrary to the Expected Utility Theory, the choices of experimental subjects depend on the value of π: for fixed $\pi'/\pi = \alpha = 0.5$, when $\pi = 0.9$ subjects prefer the safer lottery $L_2(0.9) = (3000,\ 0.9;\ \ 0,\ 0.1)$ to $L_1(0.45) = (6000,\ 0.45;\ \ 0,\ 0.55)$. When instead $\pi = 0.002$, subjects prefer the riskier lottery $L_1(0.001) = (6000,\ 0.001;\ 0,\ 0.999)$ to $L_2(0.002) = (3000,\ 0.002;\ \ 0,\ 0.998)$. This shift towards risk seeking as the probability of winning falls has provided one of the main justifications for the introduction of the probability weigthing function. In fact, KT (1979) account for this evidence by assuming that this function grows slower than linearly for small π; hence, $\alpha\pi$ is overweighted relatively to π at low values of π, inducing the choice of $L_1(\pi')$ when $\pi = 0.002$.

Consider the choice between $L_1(\pi')$ and $L_2(\pi)$ in our model. For $\alpha = 1/2$ there are four states of the world, $S = \{(6000, 3000), (0, 3000), (6000, 0), (0, 0)\}$. Once more, ordering and diminishing sensitivity suffice to imply that the salience ranking among states is

$$\sigma(6000, 0) > \sigma(0, 3000) > \sigma(6000, 3000) > \sigma(0, 0). \tag{23}$$

It is convenient to express the local thinker's decision as a function of the transformed probabilities of the lottery outcomes (as opposed to those of states of the world).[14] Denoting these transformed probabilities by $\widehat{\pi}'$ and $\widehat{\pi}$, we find that the local thinker evaluates the odds with which the riskier lottery $L_1(\pi')$ pays out relative to the safer one $L_2(\pi)$ as:

$$\frac{\hat{\pi}'}{\hat{\pi}} = \frac{\pi'}{\pi} \cdot \frac{(1-p)+p\delta^2}{(1-\pi')\delta+\pi'^2}. \tag{24}$$

With a linear utility, the local thinker selects the safer lottery $L_2(\pi)$ if and only if $\hat{\pi}'/\hat{\pi} \leq 1/2$. This implies that the local thinker chooses the safer lottery when:

$$\pi \geq \frac{2(1-\delta)}{2-\delta-\delta^2}. \tag{25}$$

As in the common ratio effect, the local thinker is risk averse when π is sufficiently high and risk seeking otherwise. In particular, for $\delta \in (0.22, 1)$, the local thinker switches from $L_2(0.9)$ to $L_1(0.001)$ just as experimental subjects do. This is true for any salience function satisfying ordering and diminishing sensitivity and thus also for the parameterization $\delta = 0.7$, $\theta = 0.1$.

The intuition for this result (see Proposition 1) is that salience exerts a particularly strong effect in low probability states. The upside of the riskier lottery $L_1(\pi')$ is salient at every π, creating a force toward risk seeking. Crucially, however, this force is strong precisely when π is low. In this case, the greater salience of the risky lottery's upside blurs the small probability difference $\pi - \pi' = (1-\alpha)\pi$ between the two lotteries. When instead π is large, the decision maker realizes that the risky lottery is much more likely to pay nothing, inducing him to attach a large weight on the second most salient state $(0, 3000)$. This is what drives the choice of the safe lottery $L_2(\pi)$.

Experimental evidence shows that this common ratio effect is also not robust to the introduction of correlation. KT (1979) asked subjects to choose between two lotteries of the type (23) in a two-stage game where in the first stage there is a 75% probability of the game ending without any winnings and a 25% change of going to stage two. In stage two, the lottery chosen at the outset is played out. The presence of the first stage is equivalent to

[14] From any vector of state-specific decision weights $(\pi_s^i)_{s\in S}$, the decision weight $\pi^i(x)$ attached to lottery i's payoff x is equal to the sum of the decision weights of all states where lottery i pays payoff x. Formally, $\pi^i(x) = \sum_{s\in S_{x^i}} \pi_s^i$ where S_{x^i} is the set of states where i pays x.

reducing by 75% the winning probability for both lotteries, so in terms of final outcomes this setting is equivalent to the setting that leads to the common ratio effect above. Crucially, KT document that in this formulation there is no violation of the independence axiom.

In explaining this behavior, KT informally argue that individuals "edit out" the correlated first stage state where both lotteries pay zero. Our model yields this editing as a consequence of the low salience and cancellation of such state. Adding a correlated state where both lotteries pay 0 neither affects the salience ranking in Equation (23) nor – more importantly – the odds ratios between states. As a result, the local thinker chooses as if he disregards the correlated state and its probability. This is what experimental subjects do.

In sum, our model explains the Allais paradoxes as the by-product of a specific form of context dependence working though the salience of lottery payoffs. Adding a common payoff to all lotteries or rescaling their probabilities changes risk preferences by changing the salience and the weighting of the lotteries' upsides or downsides. These effects depend on how the lotteries are presented. Adding a common payoff or rescaling probabilities by introducing into the lotteries a non-salient correlated state does not affect choice: it is too enticing for subjects to disregard this state and to abide by the independence axiom.

5.2 Preference Reversals

Context dependence in our model can also explain the phenomenon of preference reversal described by Lichtenstein and Slovic (1971) and confirmed by Grether and Plott (1979) and Tversky, Slovic and Kahneman (1990). Subjects are asked to choose between a safer lottery L_π, which has a high probability of a low payoff, and a riskier lottery $L_\$$, which has a low probability of a high payoff (we use conventional notation for the lotteries). Subjects may systematically choose the safer lottery L_π and yet state a higher minimum selling price for the riskier lottery $L_\$$. Preferences as revealed by choice are thus the opposite of preferences as revealed by pricing, leading to claims that choosing and pricing follow two fundamentally different principles. Neither Prospect Theory nor Expected Utility Theory can rationalize preference reversals.

To study preference reversals in our model, consider how a local thinker prices a lottery. Given that in our model valuation is context dependent, the concept of a minimum selling price can be interpreted in two distinct ways.[15] Under the *valuation* approach, in a choice between two lotteries $\{L_1, L_2\}$, the minimum selling price for either of them is the lottery's monetary valuation obtained by using the decision weights determined in $\{L_1, L_2\}$ according to Definitions 1 and 2. Formally, a local thinker with a value function $v(\cdot)$ prices L_1 at:

$$P_{\min}(L_1 \,|L_2\,) = v^{-1}\left[\sum_{s\in S} \pi_s^1 v(x_s^1)\right], \tag{26}$$

where π_s^1 is the decision weights of state s for lottery L_1 in the context of its choice from the set $\{L_1, L_2\}$. With a linear value function, the price $P_{\min}(L_1|L_2)$ is the expected value of L_1 as perceived by the local thinker. If the local thinker is asked to price a lottery in isolation, this approach suggests that he evaluates it in the context of a choice between the lottery itself and the status quo of not having the lottery, namely having zero for sure, $L_0 \equiv (0, 1)$. We see this as a natural way to model the elicitation of minimum selling prices in cases where – as in most preference reversal experiments – subjects must state this price (potentially under an incentive scheme).

Alternatively, under the *revealed preference* approach, the minimum selling price is found by revealed preference: the price of lottery L_1 is then the minimum amount of money c_1 such that adding the sure prospect $(c_1, 1)$ to the choice set makes the local thinker weakly prefer the sure prospect. In this section we adopt the valuation approach to study reversals. In the Appendix, we show that our model can also yield reversals using the revealed preference approach, but under more restricted circumstances than under the valuation approach. Consistent with the difference between the two approaches, experiments that *explicitly* implemented the revealed preference approach found significantly lower levels of reversals (Bostic, Herrnstein and Luce 1989, see also Tversky, Slovic and Kahneman 1990).

In the preference reversal experiments, subjects are first asked to price in isolation, and

[15] Several authors have studied preference reversals by focusing on the details of the experimental procedure, in particular the incentive structures. Thus, Karni and Safra (1987) suggest that Grether and Plott's BDM elicitation mechanism should be interpreted as a choice between two stage-lotteries, and that preference reversals follow from violations of the independence axiom required to interpret elicited prices as certainty equivalents (see also Holt (1986)). Segal (1988) suggests instead it results from violations of the reduction axiom. Evidently, the interpretation of experimental procedures is an important factor even in models which preserve transitivity and are not context dependent.

then to choose among, the following two independent lotteries:

$$L_\$ = \begin{cases} x, & \text{with prob.} \quad \pi' \\ 0, & \quad\quad\quad\quad 1-\pi' \end{cases}, \qquad L_\pi = \begin{cases} \alpha x, & \text{with prob.} \quad \pi \\ 0, & \quad\quad\quad\quad 1-\pi \end{cases}, \tag{27}$$

where typically $\pi'/\pi = \alpha = 1/2$, as in the common ratio experiments. We know from (25) that, with linear utility, the local thinker selects the safer lottery L_π when $\pi > 2(1-\delta)/(2-\delta-\delta^2)$. In the literature, we typically have $\pi > 3/4$, so this constraint holds for any $\delta \geq 2/3$. Thus, when asked to choose, a local thinker having linear utility and $\delta = 0.7$ is risk averse and prefers L_π to $L_\$$, just as most experimental subjects do.

In contrast, when the local thinker is asked to price the lotteries in isolation, he evaluates each lottery relative to $L_0 = (0, 1)$. In this comparison, each lottery's upside is salient. As a consequence, since $\alpha = 1/2$ the local thinker prices the lotteries as:

$$P(L_\pi \,|L_0\,) = \tfrac{x}{2} \cdot \tfrac{\pi}{\pi+(1-\pi)\delta} \ , \quad P(L_\$ \,|L_0\,) = x \cdot \tfrac{\pi/2}{\pi/2+(1-\pi/2)\delta}. \tag{28}$$

For any $\delta < 1$, the local thinker prices $L_\$$ higher than L_π in isolation, i.e.

$$P(L_\$ \,|L_0\,) > P(L_\pi \,|L_0\,).$$

Both lotteries are priced above their expected value, but $L_\$$ is more overpriced than L_π because it pays a higher gain with a smaller probability, and from Proposition 1 we know that lower probabilities are relatively more distorted.[16]

Thus, while in a choice context the local thinker prefers the safer lottery L_π, in isolation he prices the risky lottery $L_\$$ higher, exhibiting a preference reversal. Crucially, this behavior is not due to the fact that choosing and pricing are different operations. In fact, in our model choosing and pricing are the same operation, as in standard economic theory. Preference reversals occur because, unlike in standard theory, evaluation in our model is context dependent. Pricing and choosing occur in different contexts because the alternatives of choice

[16]These predictions are borne out by the literature as well as by our own experimental data. Tversky, Slovic and Kahneman (1990) show that preference reversals follow from overpricing of $L_\$$ in isolation, and that L_π is not underpriced. Our model predicts that decision makers price L_π close to its expected value because it offers an extremely high probability of winning, which is hardly distorted.

are different in the two cases. One noteworthy feature of our model is that it generates preference reversals through violations of "procedural invariance", defined by Tversky, Slovic and Kahneman (1990) as situations in which a subject prices a lottery above its expected value, $P(L_1|L_0) > \mathbb{E}(x_s^1)$, and yet prefers the expected value to the lottery, $L_1 \prec (\mathbb{E}(x_s^1), 1)$. Tversky, Slovic and Kahneman (1990) show that the vast majority of observed reversals follow from the violations of procedural invariance, as predicted by our model. Regret Theory can also generate preference reversals, using the revealed preference approach to determine certainty equivalents (Loomes and Sugden 1983). As a result, these reversals are not due to violations of procedural invariance (in contrast to the evidence), but due to intransitivity in choice instead.

One distinctive implication of our context-based explanation is that reversals between choice and pricing should only occur when pricing takes place in isolation but not if decision makers price lotteries in the choice context itself. We tested this hypothesis by giving subjects a choice between lotteries $L_\$ = (16, 0.31;\ 0, 0.69)$ and $L_\pi = (4, 0.97;\ 0, 0.03)$, which Tversky, Slovic and Kahneman (1990) found to lead to a high rate of preference reversals. Subjects stated their certainty equivalents for the two lotteries, in isolation and in the context of choice.[17] Our model then predicts that preference reversal should occur between choice and pricing in isolation, but not between choice and pricing in the choice context.[18]

Despite considerable variation in subjects' evaluations (which is a general feature of such elicitations, see Grether and Plott (1979), Bostic, Herrnstein and Luce (1990), Tversky, Slovic and Kahneman (1990)), the results are consistent with our predictions. First, among the subjects who chose L_π over $L_\$$, the average (avg) price of L_π in isolation was lower than the average price of $L_\$$ in isolation:

$$avg[P(L_\pi\,|L_0)] = 4.6 < avg[P(L_\$\,|L_0)] = 5.2\ .$$

[17] In our experimental design, each subject priced each lottery only once, and different lotteries were priced in different contexts. This design ensures that subjects do not deform their prices to be consistent with their choices; however, it also implies that preference reversals are not observed within-subject but only at the level of price distributions across subject groups (see the Online Appendix for more details).

[18] This prediction does not change if we allow for the option $L_0 \equiv (0, 1)$ to be included in the choice set. See section 6.2 for details on choice among more than two lotteries.

Thus, our subject pool exhibits the standard preference reversal between choice and average pricing in isolation.[19]

Second, preference reversals subside when we compare choice and pricing in the choice context. In fact, in this context the same subjects priced their chosen lottery L_π higher, on average, than the alternative risky lottery $L_\$$:

$$avg[P(L_\pi|L_\$)] = 4.3 > avg[P(L_\$|L_\pi)] = 4.1$$

As predicted by our model, in the choice context the average price ranking is consistent with choice.[20] One may object that this agreement is caused by the subjects' wish to be coherent when they price just after a choice. However, each subject priced only one of the lotteries in the choice context.[21] It appears to be the act of comparing the lotteries that drives their evaluation during choice, and not (only) an adjustment of value subsequent to choice.

Another potential objection is that our experiments do not elicit true selling prices. It is well known that it is difficult to design price elicitation mechanisms for subjects who violate the independence axiom of Expected Utility Theory. To avoid these problems, Cox and Epstein (1989) study preference reversals by only eliciting the *ranking* of selling prices across lotteries. In their experiments, Cox and Epstein directly compared lotteries to each other, so their procedure can be viewed as eliciting evaluations in the context of choice. They find some evidence of preference reversals, but crucially they show that these reversals are equally likely in both directions (from risk averse choice to risk seeking pricing, and from risk seeking choice to risk averse pricing). Symmetric reversal patterns are typically attributed

[19]This reversal holds not only with respect to average prices but also for the distribution of prices we observe. Assuming that subjects draw evaluations randomly from the price distributions, we estimate that around 54% of the subjects who choose L_π would exhibit the standard preference reversals (see the Online Appendix). The average prices above imply that some subjects priced the safer lottery L_π above its highest payoff. Such overpricing can occur even in a laboratory setting and with incentives schemes (Grether and Plott 1974, Bostic et al 1990), perhaps due to misunderstanding of the pricing task. In the Online Appendix we consider truncations of the data that filters out such overpricing.

[20]In our data, the distribution for $P(L_\pi|L_\$)$ does not dominate that for $P(L_\$|L_\pi)$. This is due to the fact that: i) on average subjects attribute similar values to both lotteries in the choice context, and ii) there is substantial variability in choice (and thus in pricing), as about half the subjects chose each lottery. In the Online Appendix we look in a more detailed way at the manifestation and significance of fact ii) in light on Tversky, Slovic and Kahneman's (1990) analysis of preference reversals.

[21]We ran another version of the survey where we asked the subjects to price the lotteries under comparison but without having to choose between them. These subjects exhibited similar behavior on average, namely pricing $L_\$$ higher than L_π in isolation, but similarly to L_π under comparison.

to arbitrary fluctuations in evaluation, see Bostic et al (1990) (although Cox and Epstein interpreted them as akin to a violation of procedural invariance). Thus we interpret Cox and Epstein's results as consistent with our prediction that systematic preference reversals subside when prices are elicited in a choice context.

These results suggest that choice and pricing may follow the same fundamental principle of context-dependent evaluation. Preferences based on choice could differ from those inferred from pricing *in isolation* because they represent evaluations made in different contexts.

5.3 Taking Stock

We now take stock by summarizing the role of different assumptions in generating our results and by comparing our predictions to those of Prospect Theory. Denote by "Ordering" the ordering property, by "DS" the diminishing sensitivity property, and by "Odds" the property of Definition 2 that distortions do not depend on probability odds. Table 2 summarizes how our model and Prospect Theory account for Allais Paradoxes and preference reversals.[22]

	Salience Theory	Prospect Theory
Allais Common Consequence	Ordering and DS	Sub-certainty of $\pi(p)$
Allais Common Consequence (correlated)	Ordering	Editing (but not explicit)
Allais Common ratio	Ordering, DS and Odds	Sub-additivity of $\pi(p)$
Allais Common ratio (correlated)	Ordering	Isolation effect (Editing)
Preference Reversals	Ordering, DS and Odds	No
Preference Reversals (choice)	Ordering	No

Table 2: Taking stock of Anomalies

Anomalies in our model can be driven by a change in the salience ranking of payoffs (such as the Allais common consequence paradox) or by the differential distortion of small probabilities (such as the Allais common ratio paradox, where the salience ranking does not

[22]Tables 2 and 3 list the main properties that drive each effect in either theory, but are not exhaustive. We use the 1979 version of Prospect Theory, but the cumulative version (TK, 1992) is very similar for the purpose of this comparison.

change). The ordering property plays a crucial role throughout, determining the direction of the choice anomalies.

By providing insight into what drives the anomalies, the model also identifies circumstances where the anomalies disappear. Crucially, these follow from the same properties that cause the anomalies to begin with. Anomalies disappear when the choice problems are set up so that the representation of each lottery is stable for the decision maker across different treatments – e.g. when the common consequence is made evident in the Allais paradox.

The situation with Prospect Theory is very different. Each Allais paradox is explained through a different assumption about the probability weighting function, or on the editing process which is not formally modeled. Finally, Prospect Theory cannot account for preference reversals, since choice follows from context-independent evaluation.

Consider now another important, well-documented choice pattern: the four-fold pattern of risk preferences (TK, 1992): risk aversion (RA) for gains of high probability, risk seeking (RS) for gains of low probability, and the reverse for losses. As shown in Section 4 for gains, and as we will further show in Section 6.1 in the case of negative payoffs, our model reproduces this pattern solely based on the properties of salience (including that diminishing sensitivity depends on the magnitude of payoff level, and not their sign, as encoded in the reflection property of Definition 1). We further predict that risk attitudes should depend on the payoff level x. In light of the experimental results of Section 4, we adopt the convexity property of Definition 3, whereby diminishing sensitivity gets weaker as payoff levels increase. These predictions are summarized in Table 3 below for choices between a sure payoff and a mean-preserving spread.

In Prospect Theory, the main driver of risk attitudes is the curvature of the value function. As discussed in Section 4, different patterns of risk attitudes put different constraints on this function, which can be hard to reconcile. Our context dependent account of risk preferences does not require any assumptions on the curvature of the value function.

In sum, we think that our model provides a parsimonious account of context dependence shifts in risk preference based on psychologically founded assumptions on the nature and the impact of the perceptual salience of lottery payoffs.

for gains	Salience Theory	Prospect Theory
RA for high p	DS	Concave $v(\cdot)$ and sub-certainty of $\pi(p)$
RS for low p, high x	Ordering	$v(\cdot)$ low curvature, $\pi(p) > p$ for small p
x-dependent switch to RS	Ordering, DS	Non-linear $v(\cdot)$
RA for low p, low x	DS, convexity	$v(\cdot)$ very concave for low x

for losses	Salience Theory	Prospect Theory
RS for high p	DS	Convex $v(\cdot)$ and sub-certainty of $\pi(p)$
RA for low p, high $\lvert x \rvert$	Ordering	$v(\cdot)$ low curvature, $\pi(p) > p$ for small p
x-dependent switch to RA	Ordering, DS	Non-linear $v(\cdot)$
RS for low p, low $\lvert x \rvert$	DS, convexity	$v(\cdot)$ very convex for low $\lvert x \rvert$

Table 3: Taking stock of Risk Attitudes (RA: risk aversion, RS: risk seeking)

6 Extensions

6.1 Reflection and Framing Effects

KT (1979) show that experimental subjects tend to shift from risk aversion to risk seeking as gains are reflected into losses. Our model yields these shifts in risk attitudes solely based on the salience of payoffs, without relying on the S-shaped value function of Prospect Theory. To see this, consider the choice between lottery $L_1 = (x_s^1, \pi_s)_{s \in S}$ and sure prospect $L_2 = (x, 1)$, both of which are defined over gains (i.e. $x_s^1, x > 0$) and have the same expected value $\mathbb{E}(x_s^1) = x$. For a local thinker with linear value function:

$$V^{LT}(L_1) = \sum_{s \in S} \pi_s \omega_s^1 x_s^1 = \mathbb{E}(x_s^1) + \text{cov}[\omega_s^1, x_s^1] \tag{29}$$

where $\text{cov}[\omega_s^1, x_s^1] = \sum_{s \in S} \pi_s \left[\omega_s^1 - 1\right] \left[x_s^1 - x\right]$ (recall that $\mathbb{E}(\omega_s^1) = 1$). Thus, the local thinker is risk averse, choosing L_2 over L_1, when $\text{cov}[\omega_s^1, x_s^1] < 0$. If then L_1 and L_2 are reflected into lotteries $L_1' = (-x_s^1, \pi_s)_{s \in S}$ and $L_2' = (-x, 1)$, property 3) in Definition 1 implies that the salience ranking among states does not change. As a result, the same decision maker is risk seeking, choosing L_1' over L_2' when:

$$\operatorname{cov}\left[\omega_s^1, -x_s^1\right] = -\operatorname{cov}\left[\omega_s^1, x_s^1\right] > 0, \tag{30}$$

which is fulfilled if and only if the decision maker was originally risk averse. Intuitively, a salient downside inducing risk aversion in the gain domain becomes a salient upside inducing risk seeking in the loss domain. Our model thus yields the fourfold pattern of risk preferences[23] without assuming a value function that is concave for gains and convex for losses. We next show that with the same logic our model can account for the Tversky and Kahneman (1981) famous framing experiments and the Public Health Dilemma, even with a linear value function.

Note, however, that in our model reflection of risk attitudes is a knife-edge property: it holds only if the decision maker's value function is linear. A concave value function $v(\cdot)$ in the loss domain would play against the reflection of salient payoffs, creating an intrinsic preference for a moderate and certain loss. The distinction between the salience of payoffs and the curvature of the value function can provide insight into findings that reflection of risk attitudes is only partial and decreases with payoff magnitude (Laury and Holt, 2005).

6.2 Choice Among Many Lotteries

We now extend our model to a general choice among $N \geq 2$ of lotteries, which is particularly useful for economic applications. Before doing so, note that preferences over $N \geq 2$ lotteries cannot be inferred from pairwise comparisons because salience changes across comparisons and intransitivities can arise (Section 6.3).

To model choice from an arbitrary set of alternatives $\aleph = \{L_1, \ldots, L_N\}$ defined over a state space S (as in Section 3), we first generalize the notion of payoff salience. Let $x_s = (x_s^1, \ldots, x_s^N)$ be the vector of payoffs delivered in a generic state s, and denote by $x_s^{-i} = \{x_s^j\}_{i \neq j}$ the vector of payoffs excluding x_s^i. The salience of state s for lottery L_i is then captured by a function $\widehat{\sigma}(x_s^i, x_s^{-i})$ which contrasts L_i's payoff x_s^i in s with all other payoffs

[23]The four-fold pattern of risk preferences refers to risk seeking behavior for gambles with small probabilities of gains and gambles with moderate or large probabilities of losses, and risk averse behavior when the signs of payoffs are reversed, see Tversky, Slovic and Kahneman (1990).

$x_{\mathbf{s}}^{-i}$ in the same state. Let $x_s^{-i}+\epsilon$ denote the vector with elements $\{x_s^j+\epsilon\}_{j\neq i}$. In line with Definition 1, we impose the following properties:

Definition 4 *Given a state space S and a choice set $\aleph$, the salience of state s for lottery L_i is given by a continuous and bounded function $\widehat{\sigma}(x_s^i, \mathbf{x}_s^{-i})$ that satisfies three conditions:*
1) *Ordering: if $x_s^i = \max \mathbf{x}_s$, then for any $\epsilon, \epsilon' \geq 0$ (with at least one strict inequality):*

$$\widehat{\sigma}\left(x_s^i+\epsilon, \mathbf{x}_s^{-i}-\epsilon'\right) > \widehat{\sigma}\left(x_s^i, \mathbf{x}_s^{-i}\right).$$

If $x_s^i = \min \mathbf{x}_s$, then for any $\epsilon, \epsilon' \geq 0$ (with at least one strict inequality):

$$\widehat{\sigma}\left(x_s^i-\epsilon, \mathbf{x}_s^{-i}+\epsilon'\right) > \widehat{\sigma}\left(x_s^i, \mathbf{x}_s^{-i}\right).$$

2) *Diminishing sensitivity: if $x_s^j > 0$ for all j, then for any $\epsilon > 0$,*

$$\widehat{\sigma}(x_s^i+\epsilon, \mathbf{x}_s^{-i}+\epsilon) < \widehat{\sigma}(x_s^i, \mathbf{x}_s^{-i})$$

3) *Reflection: for any two states s, $\widetilde{s} \in S$ such that $x_s^j, x_{\widetilde{s}}^j > 0$ for all j, we have*

$$\widehat{\sigma}(x_s^i, \mathbf{x}_s^{-i}) < \widehat{\sigma}(x_{\widetilde{s}}^i, \mathbf{x}_{\widetilde{s}}^{-i}) \quad \textit{if and only if} \quad \widehat{\sigma}(-x_s^i, -\mathbf{x}_s^{-i}) < \widehat{\sigma}(-x_{\widetilde{s}}^i, -\mathbf{x}_{\widetilde{s}}^{-i})$$

When $N>2$, one can construct a salience function satisfying the above requirements by setting:

$$\widehat{\sigma}(x_s^i, \mathbf{x}_{\mathbf{s}}^{-\mathbf{i}}) \equiv \sigma(x_s^i, \mathbf{f}(\mathbf{x}_{\mathbf{s}}^{-\mathbf{i}})), \tag{31}$$

where $\sigma(.,.)$ is the salience function employed in the two lottery case of Section 3, and $f(x_s^{-i}) : R^{N-1} \to R$ is a function of the residual vector x_s^{-i}. Definitons 1 and 3 jointly impose some restrictions on the properties of $f(x_s^{-i})$. One intuitive specification which, together with Definition 1, satisfies Definition 3 is:

$$\mathbf{f}(\mathbf{x}_s^{-i}) = \frac{1}{N-1}\sum_{j\neq i} x_s^j. \tag{32}$$

That is, the salience of a lottery state s depends on the contrast between the lottery's payoff and the average payoff of the other lotteries in s. For $N = 2$, Equation (31) reduces to the salience function of Section 3.

Even though $\sigma(.,.)$ is symmetric, when there are more than two lotteries salience is in general not symmetric because the same state does not necessarily have the same salience for different lotteries. Consider for instance a state where lottery L_i's payoff x_s^i is very different from the payoffs of all the other lotteries in x_s^{-i}, but in turn the payoffs in x_s^{-i} are similar to each other. According to the salience function implied by (31) and (32), then, state s is very salient for L_i but not salient for the other lotteries. In contrast, a state s may be very salient for all lotteries if in that state half the lotteries have a very low payoff and the other half have a very high payoff.

Given a lottery specific salience ranking k_s^i based on the salience function $\widehat{\sigma}$, each state is assigned a decision weight π_s^i according to Equation (8), and a value $V^{LT}(L_i)$ is computed for each lottery L_i according to Equation (10).

One important new effect arising in the choice among $N > 2$ lotteries is that the preference ranking among any two lotteries depends on the remaining alternatives, potentially leading to violations of independence of irrelevant alternatives (IIA). By shaping payoff salience, the choice set is a source of context effects. A detailed analysis of these possibilities can be found in Bordalo (2011) and Bordalo, Gennaioli and Shleifer (2011).

6.3 Intransitivity of pairwise preferences

Systematic intransitivities in choice under risk have been documented by several authors (Tversky 1969, Starmer and Sugden 1996, Luce 2000). Some models accounting for this phenomenon are Tversky's additive difference model and Regret Theory. Our model also yields intransitivities, and the structure imposed by salience allows to place restrictions on the circumstances in which intransitivities can or cannot occur.

As we prove in the Online Appendix our model predicts that intransitivities never occur in choices among independent lotteries sharing the same support with two or three outcomes. For such choices, the state space – and the salience ranking of states – does not change across choices, only the probabilities of states do. The fact that a local thinker is transitive in such choices is consistent with the intuition that intransitivities require shifts in attention and salience from one choice to the next (Tversky, 1969).

To illustrate how intransitive preferences may arise in our model, consider the following three lotteries:

$$L_\pi = \begin{cases} \alpha x, & \pi \\ 0, & 1-\pi \end{cases}, \quad L_\$ = \begin{cases} x, & \alpha\pi \\ 0, & 1-\alpha\pi \end{cases}, \quad L_s = (y, 1), \tag{33}$$

where $x, y > 0$ and $\alpha < 1$. Lotteries $L_\$$ and L_π are of the kind giving rise to the preference reversals of Section 5. In this case, a local thinker prefers the safer lottery L_π to $L_\$$ as long as π is large and δ is not too small. Suppose now that the sure prospect y is such that in the pairwise comparison with $L_\$$ the latter's gain is salient while in that with L_π the latter's loss is salient, i.e. $\sigma(x, y) > \sigma(0, y) > \sigma(\alpha x, y)$. It is then possible to find values (y, δ) such that choices are intransitive:[24]

$$L_\pi \succ L_\$, \quad L_\$ \succ L_s, \quad L_s \succ L_\pi.$$

Intransitivity arises because risk aversion in the direct comparison of L_π with $L_\$$ is reversed to risk seeking when the two lotteries are *indirectly* compared via their pairwise choice against the sure thing L_s. The intuition is as follows. In the direct comparison, $L_\pi \succ L_\$$ because lottery L_π pays off with much higher probability than $L_\$$. In the *indirect* comparison, $L_\pi \prec L_\$$ because the sure thing stresses the upside of the risky lottery and the downside of the safe lottery. This is "as if" in the direct comparison the decision maker chooses based on probabilities, while in the indirect comparison he chooses based on payoffs. This intuition is closely related to Tversky's (1969) account of intransitivities.

Although in our model intransitivities can arise in pairwise choices, choice is well defined within any given choice set. The choice set shapes the salience of each option and thus its valuation according to Definition 3. In particular, the local thinker's choice from $\{L_\pi, L_\$, L_s\}$ is well defined and can be found by applying Definition 3.[25]

[24] One numerical example is $x = 100, \alpha = 1/10, \pi = 3/4, y = 4$ and $\delta = 0.75$.

[25] In particular, any pairwise intransitive choice pattern gives rise to a violation of independence of irrelevant alternatives when compared to choice from the full choice set.

6.4 Mixed Lotteries

We finally apply our model to mixed lotteries, those involving both positive and negative payoffs. To this end, we come back to the KT (1979) piecewise linear value function exhibiting loss aversion, for loss aversion provides an intuitive explanation for risk aversion with respect to small mixed bets. Using the salience function of Equation (5), for which $\sigma(x, y) = \sigma(-x, -y)$ for all x, y, all risk aversion for lotteries symmetric around zero is due to loss aversion. For non-symmetric lotteries, salience and loss aversion interact to determine risk preferences. To see this, consider Samuelson's wager, namely the choice between the lotteries:

$$L_S = \begin{cases} \$200, & 0.5 \\ -\$100, & 0.5 \end{cases}, \qquad L_0 = (\$0,\ 1)\ .$$

In this choice, many subjects decline L_S even though it has a positive and substantial expected value. With a symmetric salience function, we have that $\sigma(200, 0) > \sigma(100, 0) = \sigma(-100, 0)$, implying that in this choice the local thinker focuses on the lottery gain.

Consider now what happens under the following piecewise linear value function:

$$v(x) = \begin{cases} x, & \text{if } x > 0 \\ \lambda x, & \text{if } x < 0 \end{cases},$$

where $\lambda > 1$ captures loss aversion. Now the local thinker rejects L_S provided:

$$200 \cdot \frac{1}{1+\delta} - 100\lambda \cdot \frac{\delta}{1+\delta} < 0.$$

The decision maker rejects L_S when his dislike for losses more than compensates for his focus on the lottery gain, i.e. $\lambda > 2/\delta$.[26] In lotteries whose negative downside is larger in magnitude than the positive upside, salience and loss aversion go in the same direction in triggering risk aversion.

[26]The role of loss aversion can also be gauged by considering the choice between two symmetric lotteries with zero expected value, $L_1 = (-x, 0.5;\ x, 0.5)$ and $L_2 = (-y, 0.5;\ y, 0.5)$, with $x > y$. Since (5) is symmetric, the states $(-x, y)$ and $(x, -y)$ have salience rank 1, whereas states $(-x, -y)$ and (x, y) have salience rank 2, so that L_1 is evaluated at $x(1-\lambda)/2$, and L_1 is evaluated at $y(1-\lambda)/2$. This implies that for any degree of loss aversion $\lambda > 1$, the Local Thinker prefers the safer lottery L_2.

Although our approach can be easily integrated with standard loss aversion, we wish to stress that salience may itself provide one interpretation of the idea that "losses loom larger than gains" (KT 1979) where, independently of loss aversion in the value function, states with negative payoffs are ceteris paribus more salient than states with positive payoffs. The ranking of positive and negative states is in fact left unspecified by Definition 1. One could therefore add an additional property:

4) *Loss salience: for every state s with payoffs $\mathbf{x}_s = (x_s^i)_{i=1,2}$ such that $x_s^1 + x_s^2 > 0$ we* have that

$$\sigma(-x_s^1, -x_s^2) > \sigma(x_s^1, x_s^2).$$

This condition relaxes the symmetry around zero of the salience function of Equation (5) represented in Figure 1, postulating that departures from zero are more salient in the negative than in the positive direction. In this specification, local thinking can itself be a force towards risk aversion for mixed lotteries, complementing loss aversion. In particular, if losses are sufficiently more salient than gains, one can account for Samuelson's wager based on salience alone (and linear utility): if $\sigma(-100, 0) > \sigma(200, 0)$, a local thinker with linear utility rejects Samuelson's bet as long as $200 \cdot \frac{\delta}{1+\delta} - 100 \cdot \frac{1}{1+\delta} < 0$, or $\delta < 1/2$. A specification where risk aversion for mixed lotteries arises via the salience of lottery payoffs may give distinctive implications from standard loss aversion, but we do not investigate this possibility here.

7 Conclusion

Our paper explores how cognitive limitations cause people to focus their attention on some but not all aspects of the world, the phenomenon we call local thinking. We argue that salience, a concept well-known to cognitive psychology, shapes this focus. In the case of choice under risk, this perspective can be implemented in a straightforward and parsimonious way by specifying that contrast between payoffs shapes their salience, and that people inflate the decision weights associated with salient payoffs. Basically, decision makers overweight the upside of a risky choice when it is salient and thus behave in a risk-seeking way, and overweigh the downside when it is salient, and behave in a risk averse way. This approach provides an intuitive and unified explanation of the instability of risk preferences, including the dramatic switches from risk seeking to risk averse behavior resulting from seemingly innocuous changes in the problem, as well as of some fundamental puzzles in choice under risk such as the Allais paradox and preference reversals. It makes predictions for when these paradoxes will and will not occur, which we test and confirm experimentally.

Other aspects of salience have been used by economists to examine the consequences of people reacting to some pieces of data (salient ones) more strongly than to others. For example, Chetty et al. (2009) show that shoppers are more responsive to sales taxes already included in posted prices than to sales taxes added at the register. Barber and Odean (2008) find that stock traders respond to "attention grabbing" news. Perhaps most profoundly, Schelling (1960) has shown that people can solve coordination problems by focusing on salient equilibria based on their general knowledge, without any possibility for communication. Memory becomes a potential source of salient data. Our formal approach is consistent with this work, and stresses that in the specific context of choice under risk the relative magnitude of payoffs is itself a critical determinant of salience.

Our specification of contrast as a driver of salience could be useful for thinking about a variety of economic situations. For example, salience may affect consumer behavior: when considering which of different brands to buy, a consumer might focus on the attributes where the potential brands are most different, neglecting the others (see Tversky and Simonson (1993), Bordalo (2011), Koszegi and Szeidl (2011)). Bordalo, Gennaioli and Shleifer (2011) use a version of this paper's model of salience to investigate consumer choices more broadly. In many applications, the key idea of our approach is that mental frames, rather than being fixed in the mind of the consumer, investor, or voter, are endogenous to the contrasting features of the alternatives of choice. This notion could perhaps provide a way to study how context shapes preferences in many social domains.

Appendices

A Proofs

Proposition 1 If the probability of state s is increased by $d\pi_s = h\pi_s$ and the probabilities of other states are reduced while keeping their odds constant, i.e. $d\pi_{\widetilde{s}} = -\frac{\pi_s}{1-\pi_s}h\pi_{\widetilde{s}}$ for all $\widetilde{s} \neq s$, then for every lottery L_i:

$$\frac{d\omega_s^i}{h} = -\frac{\pi_s}{1-\pi_s} \cdot \omega_s^i \cdot (\omega_s^i - 1)$$

Proof. By definition,

$$\omega_s^i = \frac{\delta^{k_s^i - 1}}{\sum_r \delta^{k_r^i - 1} \cdot \pi_r}$$

Therefore,

$$d\omega_s^i = -\frac{\omega_s^i}{\sum_r \delta^{k_r^i - 1} \cdot \pi_r} \sum_r \delta^{k_r^i - 1} \cdot d\pi_r$$

Replacing $d\pi_s = h\pi_s$ and $d\pi_r = -\frac{\pi_s}{1-\pi_s}h\pi_r$ (for $r \neq s$) leads to

$$d\omega_s^i = -\frac{\omega_s^i}{\sum_r \delta^{k_r^i - 1} \cdot \pi_r} \left[-\frac{h\pi_s}{1-\pi_s} \sum_{r \neq s} \delta^{k_r^i - 1} \cdot \pi_r + h\delta^{k_s^i - 1}\pi_s \right]$$

Thus

$$\frac{d\omega_s^i}{h} = -\omega_s^i \frac{1}{\sum_r \delta^{k_r^i - 1} \cdot \pi_r} \left[-\frac{\pi_s}{1-\pi_s} \sum_{r \neq s} \delta^{k_r^i - 1} \cdot \pi_r + \delta^{k_s^i - 1}\pi_s \right]$$

The parenthesis on the right hand side can be rearranged to yield

$$\frac{\pi_s}{1-\pi_s} \left[\delta^{k_s^i - 1}(1-\pi_s) - \sum_{r \neq s} \delta^{k_r^i - 1} \cdot \pi_r \right] = \frac{\pi_s}{1-\pi_s} \left[\delta^{k_s^i - 1} - \sum_r \delta^{k_r^i - 1} \cdot \pi_r \right]$$

where the sum is now over all states r. Inserting this term back into the equation above we get the result:

$$\frac{d\omega_s^i}{h} = -\omega_s^i \frac{\pi_s}{1-\pi_s}(\omega_s^i - 1)$$

■

Proposition 1 has the following corollary: let lottery L yield payoff x_m with total probability p_m, with $\sum_n p_n = 1$. Let $\mathbf{m}$ be the set of states where L pays x_m, and denote elements of $\mathbf{m}$ by $s_{\mathbf{m},j}$, where $j = 1, ..., |\mathbf{m}|$. Then $p_m = \sum_{s_{\mathbf{m},j} \in \mathbf{m}} \pi_{s_{\mathbf{m},j}} = p_m \sum_{s_{\mathbf{m},j} \in \mathbf{m}} \widetilde{\pi}_{s_{\mathbf{m},j}}$, where we write $\widetilde{\pi}_{s_{\mathbf{m},j}} = \pi_{s_{\mathbf{m},j}} / p_m$ (if L is being compared to another lottery L' and both lotteries are independent, then $\widetilde{\pi}_{s_{\mathbf{m},j}}$ is just the probability that L' gives payoff y_j where the set $\mathbf{m}$ is equal to the state (x_m, y_j)). Denote the salience distortion of p_m by

$$\omega_m = \frac{\sum_{s_{\mathbf{m},j} \in \mathbf{m}} \widetilde{\pi}_{s_{\mathbf{m},j}} \delta^{k_{s_{\mathbf{mj}}} - 1}}{\sum_{\mathbf{n}} \sum_{s_{\mathbf{n},j} \in \mathbf{n}} \widetilde{\pi}_{s_{\mathbf{n},j}} \delta^{k_{s_{\mathbf{nj}}} - 1} \cdot p_n}$$

Corollary 1 *If the probability p_m of payoff x_m is increased by $dp_m = hp_m$ and the probabilities of other states are reduced while keeping their odds constant, i.e. $dp_{\widetilde{m}} = -\frac{p_m}{1-p_m} hp_{\widetilde{m}}$ for all $\widetilde{m} \neq m$, then:*

$$\frac{d\omega_s}{h} = -\frac{p_m}{1-p_m} \cdot \omega_m \cdot (\omega_m - 1)$$

The proof of Corollary 1 is parallel to that of Proposition 1.

Lemma 1: *If the salience function is convex, then $r = v^{LT}(L_0) - v^{LT}(L_1)$ weakly decreases with x. Conversely, if the salience function is concave then r weakly increases with x.*

Proof. First note that, due to linear utility, the premium r is independent of x for a given salience ranking. In fact, for any salience ranking we have

$$\begin{aligned} v^{LT}(L_0) - v^{LT}(L_1) &= x - \frac{1}{p\delta^g + (1-p)\delta^l} \left[p\delta^g (x+g) + (1-p)\delta^l (x-l) \right] \\ &= \frac{1}{p\delta^g + (1-p)\delta^l} \left[l(1-p)\delta^l - gp\delta^g \right] \end{aligned}$$

Second, note that if $\sigma(x+g, x) - \sigma(x, x-l)$ increases in x, then the upside of the risky lottery L_1 becomes weakly more salient as x increases. In particular, if L_1's upside goes from being non salient (for low x) to being salient (for high x), the shift in r is negative:

$$\frac{l(1-p) - gp\delta}{p\delta + (1-p)} - \frac{l(1-p)\delta - gp}{p + (1-p)\delta} \propto -p(1-p)\left(1-\delta^2\right)[g+l]$$

which proves the claim when the salience function is convex (note that this shift goes to zero as δ approaches 1). The concave case is analogous. ■

B Preference Reversals under the Revealed Preference Approach

Section 5 showed that under the "valuation approach" our model explains when preference reversals occur in choice and pricing among lotteries of the form

$$L_1 = (x, p;\ 0, 1-p) \qquad\qquad L_2 = (2x, p/2;\ 0, 1-p/2).$$

We now show that the same patterns can arise when selling prices are computed under the revealed preference approach. To do so, recall that a local thinker with linear utility chooses the safer lottery L_1 in the choice set $\{L_1, L_2\}$ if and only if:

$$p \geq \frac{2(1-\delta)}{2-\delta-\delta^2}. \tag{34}$$

Therefore, it is sufficient to show that the local thinker may state a higher "revealed preference price " for the riskier lottery L_2 than for the safer lottery L_1 when (34) holds.

B.1 Reversal and Pricing in Isolation

To define the "revealed preference price" for a lottery $L = (y, q; 0, 1-q)$, consider the choice set $\{L, L_P\}$, where $L_P = (P, 1)$ is a lottery promising the sure amount P. Define the revealed preference price of L as the minimum P such that a local thinker weakly prefers the sure payoff L_P to L. When $y = x$ and $q = p$ we have that $L = L_1$ and P is the price of L_1. When instead $y = 2x$ and $q = p/2$ we have that $L = L_2$ and P is the price of L_2.

With linear utility and a given salience ranking, the certainty equivalent of a lottery is its expected value computed using the decision weights implied by salience. If the upside of lottery L is salient, then the decision weight attached to the lottery's upside is $\frac{q}{q+(1-q)\delta}$, and so $P = y \cdot \frac{q}{q+(1-q)\delta}$. If instead the downside of lottery L is salient, the decision weight attached to the lottery's upside is $\frac{q\delta}{q\delta+(1-q)}$, and so $P = y \cdot \frac{q\delta}{q\delta+(1-q)}$.

Since lotteries L_1 and L_2 have the same mean, the certainty equivalent of the risky lottery L_2 is higher than that of the safe lottery L_1 if L_2 has a salient upside (even if L_1's upside is also salient). This is because L_2 pays its salient upside with a smaller probability, and we know that smaller probabilities are subject to greatest distortions. By the same argument, the certainty equivalent of the risky lottery L_2 is lower than that of the safe lottery L_1 if the downside of L_2 is salient. As a consequence, the necessary and sufficient condition for the certainty equivalent of the risky lottery L_2 to be higher than that of the safe lottery L_1 (and thus for preference reversals to arise when (34) holds), is that the upside of L_2 is salient.

The upside of the general lottery L is salient when the lottery is contrasted with its price $P = y \cdot \frac{q}{q+(1-q)\delta}$ provided:

$$\sigma\left(y, y \cdot \frac{q}{q+(1-q)\delta}\right) > \sigma\left(0, y \cdot \frac{q}{q+(1-q)\delta}\right). \tag{35}$$

Condition (35) imposes that, at the high revealed preference price, the lottery's upside is indeed salient.[27]

For preference reversals to occur in the context of lotteries L_1 and L_2, it must then be the case that the above condition is satisfied for L_2, namely when $y = 2x$ and $q = p/2$. That is, it must be that:

$$\sigma\left(2x, xp\frac{2}{p+(2-p)\delta}\right) > \sigma\left(0, xp\frac{2}{p+(2-p)\delta}\right). \tag{37}$$

The mechanism for preference reversals encoded in (37) is the same as the one creating these reversals in the valuation approach of Section 5. When considered in isolation, the

[27] To fully characterize the revealed preference price P for L, note that the latter has a salient downside, and so $P = y\frac{\delta q}{q\delta+(1-q)}$ provided:

$$\sigma\left(y, y\frac{\delta q}{q\delta+(1-q)}\right) < \sigma\left(0, y\frac{\delta q}{q\delta+(1-q)}\right). \tag{36}$$

Conditions (35) and (36) are mutually exclusive, but for a given salience function there may be parameter values $(y, q, \ \delta)$ for which neither of them holds. This phenomenon is caused by jumps in the salience ranking, and disappears when probability distortions are a smooth function of salience. In this case, the minimum price P which is (weakly) preferred to the lottery is determined by $\sigma(y, P) = \sigma(0, P)$ (note that the certainty equivalent is not continuous, but is monotonic).

upside of the riskier lottery becomes salient, and so the lottery is overvalued relative to L_1. The only difference now is that condition (37) imposes an additional restriction on the expected value xp, namely it cannot be too high.

We can also use (37) to characterize a sufficient condition for preference reversals to arise. Since the lottery's price P increases in the probability p, condition (37) is less likely to hold as p increases. As a result, preference reversals can only occur when (37) is satisfied at the lowest possible probability $p = \frac{2(1-\delta)}{2-\delta-\delta^2}$ consistent with (34), namely when:

$$\sigma\left(2x, x\frac{2}{1+\delta+\delta^2}\right) > \sigma\left(0, x\frac{2}{1+\delta+\delta^2}\right). \tag{38}$$

It is only when (38) holds that it is possible to find $p \geq \frac{2(1-\delta)}{2-\delta-\delta^2}$ such that preference reversals occur. Note that the condition is more likely to hold when δ is high. Using the salience function of in the text (Equation (5)), the above condition becomes:

$$\frac{2x}{1+\delta+\delta^2} < \theta\left(\frac{1+\delta+\delta^2}{2} - 1\right).$$

The lotteries' payoff cannot be too large.

B.2 Preference Reversals in the Context of Choice

Consider now the case in which the certainty equivalents for the two lotteries are determined jointly, namely in the choice context. This is akin to presenting the local thinker with the choice set $\{L_1, L_2, L_{P_1}, L_{P_2}\}$, where $L_1 = (x, p; 0, 1-p)$, $L_2 = (2x, p/2; 0, 1-p/2)$ and $L_{P_i} = (P_i, 1)$, where P_i is the revealed preference price of lottery L_i. Now explicitly determining the prices is much more complicated because one needs to jointly determine two prices and two salience rankings. The key point, though, is that in equilbrium, the price of each lottery will be equal to the lottery's expected value calculated at the equilibrium salience ranking. This implies that when choosing among the lotteries the local thinker will value them at the resulting expected values. Accordingly, when pricing the lotteries the local thinker will state precisely the lotteries' perceived expected values. As a result, the local thinker's valuation of lotteries will be consistent with their pricing, and no preference reversals will occur.

References

Aizpurua, Jose, Jorge Nieto, and Jose Uriarte, "Choice Procedure Consistent with Similarity Relations," *Theory and Decision*, 29 (1990), 235–254.

Allais, Maurice, "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21 (1953), 503–546.

Barber, Brad, and Terrence Odean, "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *Review of Financial Studies*, 21 (2008), 785–818.

Bell, David, "Regret in Decision Making under Uncertainty," *Operations Research*, 30 (1982), 961–981.

Birnbaum, Michael, and Ulrich Schmidt, Allais Paradoxes can be Reversed by Presenting Choices in Canonical Split Form, Unpublished Manuscript, Kiel University, 2010.

Bostic, Raphael, Richard J. Herrnstein and R. Duncan Luce, "The Effect on the Preference-Reversal Phenomenon of Using Choice Indifferences," *Journal of Economic Behavior & Organization*, 13 (1990), 193–212.

Bordalo, Pedro, "Choice Set Effects and Salience Theory," Unpublished Manuscript, Harvard University, 2011.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," Unpublished Manuscript, Harvard University, 2011.

Camerer, Colin, "Choice under Risk and Uncertainty," in *Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth eds. (Princeton, N.J. : Princeton University Press, 1995).

Conlisk, John, "Three Variants on the Allais Example," *American Economic Review*, 79 (1989), 392–407.

Cox, James C., and Seth Epstein, "Preference Reversals without the Independent Axiom," *American Economic Review*, 79 (1989), 408–426.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99 (2009), 1145–1177.

Edwards, Ward, "Subjective Probabilities Inferred from Decisions," *Psychological Review*, 69 (1962), 109–135.

Fellner, William, "Distortion of Subjective Probabilities as a Reaction to Uncertainty," *Quarterly Journal of Economics*, 75 (1961), 670–689.

Fishburn, Peter, "Nontransitive Measurable Utility," *Journal Mathematical Psychology* 26 (1982), 31–67.

Friedman, Milton, and Leonard J. Savage, "The Utility Analysis of Choices involving Risk," *Journal of Political Economy*, 56 (1948), 279–304.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," NBER working paper No. 16911, 2011.

Gennaioli, Nicola, and Andrei Shleifer, "What Comes to Mind," *Quarterly Journal of Economics*, 125 (2010), 1399–1433

Grether, David, and Charles Plott, "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69 (1979), 623–638.

Harrison, Glenn, "Expected Utility Theory and the Experiments," *Empirical Economics*, 19 (1994), 223.

Holt, Charles, "Preference Reversals and the Independence Axiom," *American Economic Review*, 76 (1986), 508 – 515.

Kahneman, Daniel, "Maps of Bounded Rationality: Psychology for Behavioral Economics," *American Economic Review*, 93 (2003), 1449–1476.

Kahneman, Daniel, *Thinking Fast and Slow*, (New York: Farrar, Strauss, Giroux, 2011).

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

Kandel, Eric, James Schwartz, and Thomas Jessell, *Principles of Neural Science*, 4th ed. (Norwalk, Conn.: Appleton & Lange, 1991).

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

Koszegi, Botond, and Mathew Rabin, "Reference-Dependent Risk Attitudes," *American Economic Review*, 97 (2007), 1047–1073.

Koszegi, Botond, and Adam Szeidl, A Model of Focusing in Economic Choice, Unpublished Manuscript, University of California, Berkeley, 2011.

Laury, Susan, and Charles Holt, "Further Reflections on Prospect Theory," Andrew Young School of Policy Studies Research Paper Series No. 06-11, 2005.

Leland, Jonathan, "Generalized Similarity Judgments: an Alternative Explanation for Choice Anomalies," *Journal of Risk and Uncertainty*, 9 (1994), 151–172.

Lichtenstein, Sarah, and Paul Slovic, "Reversals of Preference Between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, 89 (1971), 46–55.

Loomes, Graham, and Robert Sugden, "Regret Theory: an Alternative Theory of Rational Choice under Uncertainty," *The Economic Journal*, 92 (1982), 805–824.

Loomes, Graham, and Robert Sugden, "A Rationale for Preference Reversal," *The American Economic Review*, 73 (1983), 428–432.

Luce, R. Duncan, *Utility of Gains and Losses: Measurement-Theoretic and Experimental Approaches*, Scientific Psychology Series (London: L. Erlbaum 2000).

McCoy, Allison, and Michael Platt, "Risk-Sensitive Neurons in Macaque Posterior Cingulate Cortex," *Nature Neuroscience*, 8 (2005), 1220–1227.

Manzini, Paola and Marco Mariotti, "Sequentially Rationalizable Choice," *American Economic Review*, 97 (2007), 1824–1839.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Ozbay, "Revealed Attention," *American Economic Review*, forthcoming, 2010.

Mullainathan, Sendhil, "Thinking through Categories," Unpublished Manuscript, Harvard University, 2002.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 66 (1998), 497–527.

Quiggin, John, "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization*, 3 (1982), 323–343.

Rabin, Mathew, "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68 (2000), 1281–1292.

Rubinstein, Ariel, "Similarity and Decision-Making under Risk," *Journal of Economic Theory*, 46 (1988): 145–153.

Savage, Leonard J., *Foundations of Statistics,* (New York: Dover Publications, 1954).

Schelling, Thomas, *The Strategy of Conflict,* (Cambridge, Mass.: Harvard University Press, 1960).

Schwartzstein, Joshua, "Selective Attention and Learning," Unpublished Manuscript, Harvard University, 2009.

Starmer, Chris, and Robert Sugden, "Testing Alternative Explanations of Cyclical Choices," *Economica*, 65 (1996), 347 – 361.

Taylor, Shelley, and Suzanne Thompson, "Stalking the Elusive Vividness Effect," *Psychological Review*, 89 (1982), 155–181.

Tversky, Amos, "Intransitivity of Preferences," *Psychological Review*, 76 (1969), 31–48.

Tversky, Amos, and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, 211 (1981), 453–458.

——, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (1992), 297–323.

Tversky, Amos, and Itamar Simonson, "Context-Dependent Preferences," *Management Science*, 39 (1993), 1179–1189.

Tversky, Amos, Paul Slovic, and Daniel Kahneman, "The Causes of Preference Reversals," *American Economic Review*, 80 (1990), 204–217.

Wu, George, and Richard Gonzalez, "Curvature of the Probability Weighting Function," *Management Science*, 42 (1996), 1676–1690.