NBER WORKING PAPER SERIES

SALIENCE THEORY OF CHOICE UNDER RISK

Pedro Bordalo
Nicola Gennaioli
Andrei Shleifer

Working Paper 16387
http://www.nber.org/papers/w16387

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2010

We are grateful to Nicholas Barberis, Gary Becker, Colin Camerer, John Campbell, Tom Cunningham, Xavier Gabaix, Ming Huang, Jonathan Ingersoll, Emir Kamenica, Botond Koszegi, David Laibson, Drazen Prelec, Matthew Rabin, Josh Schwartzstein, Jesse Shapiro, Tomasz Strzalecki, Richard Thaler, Georg Weiszacker and George Wu 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Salience Theory of Choice Under Risk
Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer
NBER Working Paper No. 16387
September 2010
JEL No. D03,D81

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 endogenizing 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.

Pedro Bordalo
Department of Economics
Harvard University
Littauer Center
Cambridge, MA 02138
bordalop@nber.org

Nicola Gennaioli
CREI
Universitat Pompeu Fabra
Ramon Trias Fargas 25-27
08005 Barcelona (Spain)
ngennaioli@crei.cat

Andrei Shleifer
Department of Economics
Harvard University
Littauer Center M-9
Cambridge, MA 02138
and NBER
ashleifer@harvard.edu

An online appendix is available at:
http://www.nber.org/data-appendix/w16387

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 Expected Utility Theory to deal with these findings by changing some of its axioms, 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." In line with this idea, in our model the decision maker focuses on salient payoffs. He is then risk seeking when a lottery's upside is salient and risk averse when its downside is salient.

To formalize this idea in a choice between two lotteries, we define a state of the world to be salient if, roughly speaking, the distance between the lotteries' payoffs in that state is large. We thus follow Kahneman (2003), who writes that "changes and differences are more accessible to a decision maker than absolute values". The model then describes how decision makers replace the objective probabilities they face with decision weights that increase in the salience of states. Through this process, the decision maker develops a context-dependent representation of each lottery. Aside from replacing objective probabilities with decision weights, the agent's utility is standard.[1]

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.[2] 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 is the standard economic agent.

Our model leads to an understanding of what encourages and discourages risk seeking, but also to an explanation of the Allais paradoxes. The strongest departures from Expected Utility Theory 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 draws attention to the highest lottery payoffs, and risk averse when the common consequence is large and draws attention to the lowest payoffs. We explore this prediction 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.

[1]In most of the paper, we assume a linear utility function. However, this function does not deal with the phenomenon of loss aversion, i.e. the extreme risk aversion with respect to small positive expected value bets. To deal with this phenomenon, we modify preferences around zero along the lines of Kahneman and Tversky (1979) in Section 6.2.

[2]Other models in the same spirit are Mullainathan (2002) and Schwartzstein (2009).

It is useful to compare our model to the gold standard of existing theories of choice under risk, Kahneman and Tversky's (KT, 1979) Prospect Theory. Prospect Theory 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 endogenously derived 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, that 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 present an experiment illustrating the switch from risk averse to risk-loving behavior as lottery payoffs, and their salience, change. In section 3, we present a general model and derive its most fundamental properties, namely how changes in the structure of lotteries affect the endogenous decision weights. In section 4, we use the model to analyze in detail the forces that lead to risk averse and risk seeking behavior. In particular, we derive some of the building blocks of Prospect Theory from first principles, and provide experimental evidence for new predictions concerning risk attitudes. In Section 5 we focus on the two principal phenomena our model accounts for: Allais paradoxes and preference reversals. We illustrate our analysis by extending these paradoxes in several ways, including correlation effects. In Section 6, we consider framing effects and mixed lotteries. Section 7 concludes.

2 A Simple Example

In this section, we present the results of two experiments illustrating two central intuitions behind our model: how the contrast between payoffs in different states makes some states more salient to the decision maker than others, and how the relative salience of different states leads to a transformation of objective probabilities into decision weights, shaping risk attitudes. The procedures for all experiments in the paper are described in Appendix 2. The two experiments are:

Experiment 1: Choose between the two options:

$$L_1 = \begin{cases} \$1 \quad \text{with probability } 95\% \\ \$381 \text{ with probability } 5\% \end{cases}, \qquad L_2 = \{\$20 \;\; \text{for sure.}$$

Experiment 2: Choose between the two options:

$$L_1 = \begin{cases} \$301 \text{ with probability } 95\% \\ \$681 \quad \text{with probability } 5\% \end{cases}, \qquad L_2 = \{\$320 \;\; \text{for sure.}$$

Three points are noteworthy. First, the second experiment simply adds \$300 to all the payoffs in the first. Second, in both experiments the two options have the same expected payoffs. Third, in both experiments there is the same relatively small (5%) probability of a high payoff, and a high (95%) probability of a \$19 loss relative to the sure outcome.

The same 120 subjects participated in the two experiments over the internet. In Experiment 1, 83% of the subjects chose the safe option L_2, whereas in Experiment 2 67% of the same subjects chose the risky option L_1. The difference between the two experiments in the probability of choosing the safe option is highly statistically significant. In fact, over half the subjects who chose L_2 in the first experiment switched to L_1 in the second.

Although in each experiment the two options offer the same expected value, the same subjects are risk averse in the first experiment and risk loving in the second. Expected Utility Theory typically assumes risk aversion, and so would have trouble accounting for Experiment 2. Prospect Theory (both in its standard and cumulative versions) holds that the small 5% probability of the high outcome is over-weighted by decision makers, creating a force toward risk loving behavior in both experiments. To account for risk averse behavior in Experiment 1 and risk loving behavior in Experiment 2, Prospect Theory requires a combination of probability weighting and declining absolute risk aversion in the value function.[3]

Our explanation of these findings does not rely on the shape of the value function. It goes roughly as follows. In the first experiment, in the state where the lottery loses relative

[3]This is only true if the reference point of a Prospect Theory agent is the status quo. If instead the reference point is the sure prospect, then both problems are identical and Prospect Theory cannot account for the switch from risk aversion in Experiment 1 to risk seeking in Experiment 2.

to the sure payoff, the lottery's payoff \$1 feels a lot lower than the sure payoff of \$20: this bad outcome is more salient than the possibility of winning \$381, and subjects focus on the former when making their decision. In the second experiment, the lottery's payoff in the state where it loses, \$301, does not feel nearly as bad compared to the sure payoff of \$320: the outcome of winning \$681 is thus more salient and subjects focus on it when making their decision. Put differently, the risk of losses is more salient than the opportunity for gains in the first experiment, and vice versa in the second. The focus on losses (relative to the safe payoff) then pushes subjects toward risk-averse choice in Experiment 1, while the focus on gains pushes them toward risk-loving choice in Experiment 2. The analogy here is to sensory perception: the salient states are those featuring the strongest difference between lottery payoffs, and the decision maker's mind is drawn to the salient state when making a decision.

How do we implement this rough intuition formally? In section 3 we present a general model, but here consider a special case helpful for analyzing this experiment. Define the salience of a state of the world as the percentage difference between the higher and the lower payoff in that state. In Experiment 1 the salience of the state where the lottery loses is equal to $20/1 = 20$, the salience of the state where the lottery wins is given by $381/20 = 19.05$. In contrast, in Experiment 2 the salience of the state where the lottery loses is given by $320/301 = 1.06$, while the salience of the state where the lottery wins is given by $681/320 = 2.13$. In line with our previous intuition, in Experiment 1 the lottery loss is more salient than the gain, and vice versa in Experiment 2.

We formally assume that the agent overweights the probability of the more salient state, underweighting the other state's probability. Thus, in Experiment 1 subjects under-weight the probability of a 5% gain, whereas in Experiment 2 they over-weight the probability of a 5% gain. The agent computes the lottery's expected utility using these transformed decision weights. An agent with a linear utility will then value the lottery less than its expected value in Experiment 1, behaving in a risk averse manner, while he will value the lottery more than its expected value in Experiment 2, behaving in a risk loving manner, just as the experimental results indicate. By focusing the agent on different lottery states, extreme payoffs shape risk preferences. In the next sections we show that our emphasis on the salience of lottery payoffs gives an intuitive account of risk attitudes and yields novel predictions on how the context in which a choice is presented should affect behavior.

3 The Model

An agent chooses between two lotteries L_1 and L_2. The state space S includes all lotteries' payoff combinations occurring with positive probability. In a generic state $s = (x_s, y_s) \in S$, which occurs with probability $\pi_s > 0$, lottery L_1 pays x_s and L_2 pays y_s. If lotteries are statistically independent, then $S = X \times Y$, where $X = (x_i)_{i=1,\ldots,|X|}$ and $Y = (y_j)_{j=1,\ldots,|Y|}$ are the sets of payoffs for L_1 and L_2 , respectively, and state $s_{ij} = (x_i, y_j)$ occurs with probability $\pi_{i,j} = p_i \cdot q_j$, where $p_i, q_j > 0$ are the probabilities with which L_1 and L_2 pay x_i and y_j, respectively. The experiments of Section 2 involve the choice between the lottery $L_1 = (\overline{x}, p;\ \underline{x}, 1-p)$ and the sure prospect $L_2 = (y, 1)$, where $\overline{x} > y > \underline{x}$. Here S consists of the state $(\overline{x}, y)$ where L_1 "gains" over L_2, and the state $(\underline{x}, y)$ where L_1 "loses" over L_2.

The agent evaluates lottery payoffs using a value function $v : \mathbb{R} \to \mathbb{R}$. Absent probability distortions, the agent chooses lottery L_1 over L_2 if and only if:

$$\sum_{s \in S} \pi_s \left[v(x_s) - v(y_s)\right] > 0. \tag{1}$$

Throughout most of the paper, we illustrate the mechanism that generates risk preferences in our model by assuming a linear value function v, which allows us to deal with cases involving lotteries that have all positive or all negative payoffs. In section 6.2, when we focus on mixed lotteries, we consider a piece-wise linear value function featuring loss aversion, as in Kahneman and Tversky (1979). [4] With a linear value function, expression (1) says that the agent trades off the state-by-state gains and losses of L_1 relative to L_2 by weighting them by their respective probabilities. For instance, the two outcomes lottery is preferred to the sure prospect when its expected gain $p(\overline{x} - y)$ in state $(\overline{x}, y)$ is larger than its expected loss

[4] Our approach is complementary to Koszegi and Rabin (2006), who build a model of reference point formation. Koszegi and Rabin (2007) use this model to explain shifts in attitudes towards risky lotteries in the real world. Our focus on lottery choice in the lab allows us to focus on the role of salience, holding reference points constant. Our model of probability weights can also be studied using a two part utility function of the kind introduced by Koszegi and Rabin (2006) which combines a standard concave utility function with a reference dependent value function featuring loss aversion around status quo wealth.

$(1-p)(y-\underline{x})$ in state $(\underline{x}, y)$.

The local thinker (LT) departs from Equation (1) by overweighting the most salient states $s \in S$ relative to the others. This process consists of two stages: first, the states of the world are ranked by salience and the probability π_s of each state is transformed into decision weights π_s^{LT}; second, the weights π_s^{LT} replace the probabilities π_s in Equation (1). To analyze the first stage, we offer a formal definition of salience:

Definition 1 *The salience of a generic state of the world $s = (x, y)$ is measured by a continuous, bounded, and symmetric salience function $\sigma(x, y)$ that satisfies four conditions:*
1) *Ordering: if $[x', y']$ is a subset of $[x, y]$ then $\sigma(x, y) > \sigma(x', y')$. For any x, $\lim_{y\to\infty} \sigma(x, y) = \sup_{x',y'} \sigma(x', y')$.*
2) *Diminishing sensitivity: for any x, y such that $x+y > 0$ and for any $\epsilon > 0$, $\sigma(x+\epsilon, y+\epsilon) < \sigma(x, y)$.*
3) *Reflection: for any x, y, x', y' such that $x + y > 0$ and $x' + y' > 0$, $\sigma(x, y) < \sigma(x', y')$ if and only if $\sigma(-x, -y) < \sigma(-x', -y')$.*
4) *Convexity: for any $x, y, \epsilon > 0$, the ratio $\frac{\sigma(x,x+y)}{\sigma(x+\epsilon,x+\epsilon+y)}$ decreases in x and tends to 1 as $x \to \infty$.*

To illustrate these four properties, consider the following specific salience function:

$$\sigma(x, y) = \frac{|x-y|}{|x| + |y| + \theta}, \tag{2}$$

where $\theta > 0$ is a constant and $\sigma(x, y) \in [0, 1]$. According to the ordering property 1, salience increases in the distance between payoffs. In Equation (2), this is captured by the numerator $|x-y|$. According to diminishing sensitivity and reflection (properties 2 and 3), salience decreases as a state's average payoff gets further away from zero in both the positive and negative domains. This captures the intuitive idea that salience is shaped by the magnitude rather than the sign of payoffs. To illustrate diminishing sensitivity and reflection, Figure 1 plots the salience function of Equation (2) keeping the distance between payoffs constant at $|x-y| = 1$. Note that in (2) diminishing sensitivity is captured by the term $|x| + |y|$ in the denominator, and reflection takes the strong form $\sigma(x, y) = \sigma(-x, -y)$, as illustrated by the symmetry of salience in Figure 1. Finally, according to the convexity property 4, salience falls at a decreasing rate as payoffs become very large in absolute value. This property limits the effect of diminishing sensitivity, implying that at large absolute payoff values the distance between payoffs is the principal determinant of salience. As shown by Figure 1, in Equation (2) this effect is captured by the convexity of salience as a function of $x + y$. Section 3.1 discusses the connection between these properties and the cognitive notion of salience.

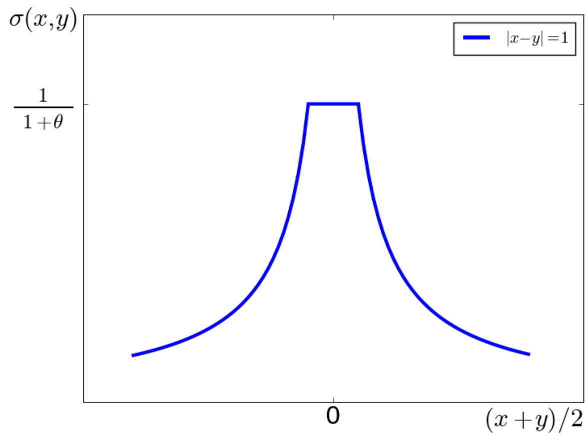

Figure 1: A simple salience function, Eq. (2)

Although our main results rely only on the properties in Definition 1, we sometimes use the tractable salience function (2) to illustrate the workings of our model. One feature of this functional form is that it parsimoniously parameterizes, via θ, the relative strength of ordering and diminishing sensitivity. If $\theta = 0$, diminishing sensitivity is strong in the sense that any state with a zero payoff has maximal salience, $\sigma(0, x) = 1$, regardless of the distance between 0 and the payoff x. This is the case we used in Section 2, since for $\theta = 0$ salience is an increasing function of the ratio $\max\{|x|, |y|\}/\min\{|x|, |y|\}$. If instead $\theta > 0$, even a state with a zero payoff can be less salient than a state with positive payoffs provided the distance between payoffs in the latter state is sufficiently large. In this sense, when $\theta > 0$ ordering can overcome diminishing sensitivity.

For a given salience function $\sigma(x, y)$, the local thinker ranks the states and distorts their probability weights as follows:

Definition 2 *Given any two states* $s, \widetilde{s} \in S$, $s \neq \widetilde{s}$, *we say that* s *is more salient than* $\widetilde{s}$ *if*

$\sigma(x_s, y_s) > \sigma(x_{\tilde{s}}, y_{\tilde{s}})$. Index $k_s \in \{1, ..., |S|\}$ denotes the salience ranking of a generic $s \in S$, with lower k_s indicating higher salience. States with the same salience are given the same ranking. For any two states $s, \tilde{s} \in S$ having $k_s \geq k_{\tilde{s}}$, the local thinker distorts the odds $\pi_{\tilde{s}}/\pi_s$ into $\pi_{\tilde{s}}^{LT}/\pi_s^{LT}$, where:

$$\frac{\pi_{\tilde{s}}^{LT}}{\pi_s^{LT}} = \left(\frac{1}{\delta}\right)^{k_s - k_{\tilde{s}}} \frac{\pi_{\tilde{s}}}{\pi_s}, \tag{3}$$

with $\delta \in (0, 1]$, and by assumption $\sum_s \pi_s^{LT} = 1$. Defining $\omega_s = \delta^{k_s - 1} / \left(\sum_r \delta^{k_r - 1} \cdot \pi_r\right)$, the decision weight attached by a local thinker to a generic state $s \in S$ is equal to:

$$\pi_s^{LT} = \pi_s \cdot \omega_s. \tag{4}$$

The parameter δ captures the (inverse of the) extent of the local thinker's focus on salient states. With $\delta < 1$, the local thinker discounts the probability of the less salient states. In this case, the state s is overweighted if and only if $\omega_s > 1$, which occurs if and only if s is sufficiently salient to be less discounted than the average ($\delta^{k_s - 1} > \sum_r \delta^{k_r - 1} \cdot \pi_r$). For a local thinker with $\delta < 1$, ω_s represents the extent of overweighting of state s. In the case of independent lotteries, Equation (4) implies that the weight placed on each specific lottery outcome is equal to $p_i^{LT} = p_i \sum_j q_j \omega_{ij} / \sum_{kj} p_k q_j \omega_{ij}$.

The agent evaluates lotteries L_1 and L_2 by using the weights π_s^{LT} rather than the objective probabilities π_s. He chooses L_1 whenever:

$$\sum_{s \in S} \delta^{k_s - 1} \pi_s \left[v(x_s) - v(y_s)\right] > 0, \tag{5}$$

which is equivalent to $\sum_{s \in S} \pi_s^{LT} \left[v(x_s) - v(y_s)\right] > 0$. The case where $\delta \to 0$ describes the agent who only focuses on the most salient state and decides based solely on payoffs in that state. In general, for $\delta < 1$ the local thinker under-weights – compared with true probabilities – the gains and losses of L_1 relative to L_2 that occur in low salience states. As $\delta \to 1$, our model converges to Expected Utility Theory.

3.1 Discussion of Assumptions and Setup

Our formalization of salience relies on the basic principle of human perception that a sensorial stimulus gives rise to a subjective representation whose intensity increases in the stimulus' magnitude but also depends on context (Kandel et al, 1991). The function $\sigma(x, y)$ maps this idea to lottery states: the strength of the stimulus is the payoff difference within a state, and salience captures the subjective intensity with which a state is perceived. This subjective intensity decreases with the distance of the state's payoffs from the status quo of zero, which is our measure of context. As in Weber's law of diminishing sensitivity, whereby a change in luminosity is perceived less intensely if it occurs at a higher luminosity level, our agent perceives less intensely payoff differences occurring at high (absolute) payoff levels.[5]

We model the agent's overweighting of salient states by assuming that the probability of the kth most salient state is multiplied by the $(k-1)$th power of the "local thinking" parameter δ, which may reflect an individual's ability to pay attention to multiple aspects of the problem, cognitive load, or simply intelligence. This rank-based discounting buys us analytical tractability, but our main results also hold if the probability of each state (x, y) is multiplied by an increasing function $\delta\left[\sigma(x, y)\right]$, for instance $\delta^{-\sigma(x,y)}$; obviously, in this case, the cardinal properties of $\sigma(x, y)$ yield additional effects beyond those we already find. The only significant restriction embodied in our formalization is that the distortions in the odds of any two states do not depend on these states' underlying probabilities. The salience function in Equation (2) and the decision weights in Definition 2 provide a tractable and parsimonious version of our model, characterized only by the two parameters (θ, δ). This allows us to derive predictions concerning a large variety of experimental lottery choice problems and to find ranges for θ and δ that are consistent with the observed choice patterns.

To compare our model to Prospect Theory, recall that the latter is based on four modifications of Expected Utility Theory. First, lotteries are evaluated without integrating them into

[5] There is some neurobiological evidence connecting visual perception to risk taking behavior. 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)." As we will see, a similar mechanism underlies risky choices in our model.

the decision maker's wealth (narrow framing). Second, people have a stable value function that is concave in gains and convex in losses. Third, people exhibit loss aversion, the kink in utility around the reference point. Fourth, and crucially to the comparison, KT follow the earlier work of Edwards (1962) and Fellner (1961) in assuming a stable probability weighting function, which centrally features overweighting of small probabilities but does not depend on lottery payoffs. It is the various properties of this function, such as sub-additivity, that allow KT to account for risk loving behavior and the Allais paradoxes.[6]

Our framework adopts the narrow framing assumption. In Section 6.2, we also incorporate loss aversion; although our model does not deal centrally with mixed lotteries, loss aversion seems to be a natural way to explain risk aversion toward small bets, such as the refusal to take a 50-50 chance of winning $10 and losing $5. We do not need to assume that utility is concave for gains and convex for losses; we show in Section 6.1 how our model generates these features of preferences without additional assumptions. Most importantly, we adopt KT's general idea of probability weighting but replace their fixed probability weighting function with context-dependent decision weights based on the salience of payoffs.

Quiggin's (1982) rank-dependent expected utility and Tversky and Kahneman's (1992) Cumulative Prospect Theory (CPT) constitute early attempts to study the impact of the rank order of a lottery's payoffs on probability weighting. Prelec (1998) axiomatizes a set of theories of choice based on probability weighting, which include CPT. 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: the lottery upside may still be underweighted if the payoff associated with it is not sufficiently large. As we show in Section 4, this feature is crucial to explaining shifts in risk attitudes.

Second, and more important, in our model probability weights depend on the choice context, namely on the available alternatives as they are presented to the agent. Our model thus provides a tractable setup to study context dependence. In Section 5 we exploit this feature to shed light on the psychological forces behind the Allais paradoxes and preference reversals. We are not the first to propose a model of context dependent choice among lotteries. Rubinstein (1988) builds a model of similarity-based preferences, in which agents

[6] For a recent attempt to estimate the probability weighting function, see Wu and Gonzalez (1996).

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 very different from Rubinstein's, even though we share the general idea that the common ratio Allais paradox (see Section 5) is due to subjects' focus on lottery payoffs. Another theory of context dependent choice is Regret Theory (Loomes and Sugden 1982, Bell 1982, Fishburn 1982). In this theory, the alternative lotteries directly affect the agent's utility by adding a regret/rejoice term to a standard utility function. In our model, instead, context affects decisions by shaping the salience of states and decision weights. By adopting a traditional utility theory perspective, Regret Theory cannot capture framing effects.

Context-dependence in our model captures the influence of payoff salience on decisions and enables us to address several long-standing puzzles in decision theory. The theory could perhaps be generalized to take into account determinants of salience other than payoff values, such as prior experiences, which might matter in some situations. Our emphasis on the state space as a source of context dependence does not lead to accurate predictions when lotteries are presented in such a way that the state space is neglected. For example, consider a choice problem in which 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, while the other lottery has payoffs 2,3,4,5,6,1. In our model, the state in which the first lottery pays 6 and the second pays 1 would appear most salient to the agent, leading him to prefer the first lottery. But of course a moment's thought would lead him to realize that the second lottery is just a rearrangement of the first. What is salient here is that the two lotteries are the same. Our model needs to be extended to understand how people represent different problems to deal with this situation.

3.2 Lottery Valuation and Risk Preferences

Let us go back to the choice between L_1 and L_2. Given over-weighting ω_s of state $s = (x_s, y_s)$, one can see that he chooses L_1 over L_2 if and only if:

$$\sum_{s \in S} \pi_s v(x_s) - \sum_{s \in S} \pi_s v(y_s) > \text{cov}\left[v(y_s) - v(x_s), \omega_s\right], \tag{6}$$

where the left-hand side measures the difference between the valuation of L_1 and L_2 computed by an agent who does not distort probabilities. Equation (6) says that probability weighting tilts the choice in favor of L_2 when the right hand side of Equation (6) is positive, namely when the states where L_2 delivers higher payoffs than L_1 [i.e. where $v(y_s) - v(x_s) > 0$] are more salient, and thus have higher ω_s.

Equation (6) illustrates that the local thinker's over-weighting of salient states plays a central role in his choices. To explore this role, we begin with the following result:

Proposition 1 *The local thinker's over-weighting ω_s has two properties:*

1) *Change in a state's probability. 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:*

$$\frac{d\omega_s}{h} = -\frac{\pi_s}{1-\pi_s} \cdot \omega_s \cdot (\omega_s - 1) . \tag{7}$$

2) *Change in a state's payoffs. A change in the payoffs x_s, y_s of a state s which, by increasing (only) that state's salience $\sigma(x_s, y_s)$, improves its ranking from k_s to $k_s - 1$ increases this state's over-weighting ω_s.*

The over-weighting of the probability of a state depends on two factors: the state's true probability and the salience of its payoffs. According to property 1, an increase in a state's probability π_s reduces the weight ω_s if the state is over-weighted (i.e. if $\omega_s > 1$), while it increases ω_s otherwise. Put differently, low probability states are subject to the strongest distortions: they are severely over-weighted if salient and severely under-weighted otherwise. Intuitively, a given distortion caused by salience is relatively more severe if the state's true probability is low.

Property 2 says that changing the payoffs of a state s in a way that increases its salience ranking (without affecting the salience of other states), unambiguously increases the decision weight placed by the local thinker on this state, for any given objective probability π_s.

This proposition stands in contrast with KT's (1979,1992) assumption that high rank, low probability payoffs are always overweighted. In our model states are overweighted if and only if they are salient, regardless of their probability. Probabilities do however affect the extent of distortion: the most severe over-weighting occurs when salient states are unlikely. We next investigate the implications of this notion for the local thinker's preferences toward risk, shedding further light on the experimental findings of Section 2.

4 Salience and Attitudes Towards Risk

Consider the choice between a lottery $L_1 = (\overline{x}, p; \underline{x}, 1-p)$ and a sure prospect $L_2 = (y, 1)$, where $\overline{x} > y > \underline{x} > 0$. An expected utility maximizer with linear utility chooses the lottery when the latter's expected value is larger than y, namely when the lottery wins with sufficiently high probability:

$$p \geq \frac{y - \underline{x}}{\overline{x} - \underline{x}}. \tag{8}$$

The right hand side falls in $\overline{x}$ because the lottery becomes more attractive at any given p as long as its upside goes up.

A local thinker's choice is described by replacing p in Equation (8) with the salience distorted weight p^{LT}. To determine this weight, recall that the lottery's "gain" state $s_g = (\overline{x}, y)$ is more salient than the "loss" state $s_l = (\underline{x}, y)$ when $\sigma(\overline{x}, y) > \sigma(\underline{x}, y)$. Definition 1 ensures that there is a threshold payoff $\overline{x}^*$ such that the gain state is more salient than the loss state if and only if $\overline{x} > \overline{x}^*$. As the agent inflates by $1/\delta$ the odds of the salient state, the weight of the gain state s_g is equal to:

$$p^{LT} = \begin{cases} \frac{p\delta}{p\delta + (1-p)} & if \quad \overline{x} < \overline{x}^* \\ \frac{p}{p + \delta(1-p)} & if \quad \overline{x} > \overline{x}^* \end{cases} ; \tag{9}$$

in the knife edge case $\overline{x} = \overline{x}^*$ we have $p^{LT} = p$. If $\overline{x} > \overline{x}^*$, the lottery gain is over-weighted, i.e. $\omega_{s_g} = 1/[p + \delta(1-p)] > 1$. If instead $\overline{x} < \overline{x}^*$, the lottery gain is under-weighted, i.e. $\omega_{s_g} = \delta/[p\delta + (1-p)] < 1$. By Proposition 1, the overweighting of the gain state when $\overline{x} > \overline{x}^*$ and its underweighting when $\overline{x} < \overline{x}^*$ fall in the probability p of the state.

To illustrate the role of salience in the starkest manner, notice that as local thinking becomes severe, i.e. as $\delta \to 0$, then by Equation (9) the agent forms his decision by focusing solely on the most salient state. Figure 2 compares the choice of a local thinker and an expected utility maximizer in this case:

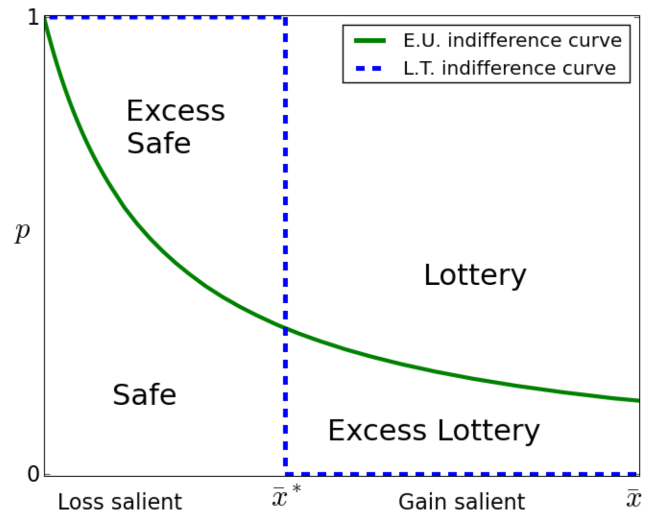

Figure 2: Local Thinking vs. Expected Utility when $\delta \to 0$

Above the solid downward sloping curve, the expected utility maximizer chooses the lottery; below it he chooses the sure prospect. The dashed curve separates these two regions for the local thinker: if $\overline{x} < \overline{x}^*$, the local thinker focuses on the lottery loss and chooses the sure prospect, if $\overline{x} > \overline{x}^*$ he focuses on the gain and chooses the lottery.

In Figure 2, two patterns stand out. First, by boosting the salience of the lottery upside, a higher $\overline{x}$ induces the local thinker to switch from risk aversion to risk seeking. Second, departures from Expected Utility Theory emerge when the salient state is relatively unlikely. If the gain is both salient and likely, we are in the Lottery region where both the expected utility maximizer and the local thinker take the risk; if the loss is both salient and likely, we are in the Safe region where both agents choose the sure prospect. In contrast, in the Excess Lottery region not only is the lottery gain salient, but its probability p is quite low. It is precisely in this region that the local thinker chooses the lottery even if the the Expected Utility maximizer does not. Similarly, in the Excess Safe region not only is the lottery loss salient, but its probability $1-p$ is quite low. As a consequence, the local thinker refuses the lottery even if the the Expected Utility maximizer would take it.

4.1 Payoffs, Probabilities and Risk Attitudes

In Figure 2 an independent change in $\overline{x}$ affects not only the salience of the lottery's upside but also the premium that the lottery pays relative to the sure prospect (namely the difference in expected values). For $\delta > 0$, both effects change the local thinker's risk taking.[7]

To see the "pure" impact of probabilities and payoffs on the local thinker's risk attitudes, consider the case in which the premium is zero, namely where the lottery has the same expected value as the sure prospect, i.e. $(\overline{x} - y) = (y - \underline{x})(1 - p)/p$. This is akin to moving along the solid downward sloping curve of Figure 2. In this case, from Equation (8) a local thinker chooses the lottery if and only if:

$$p^{LT} \geq p, \tag{10}$$

so the agent is risk loving when the gain is more salient than the loss and is risk averse otherwise. Now risk attitudes are independent of $\delta < 1$: due to the agent's underlying risk neutrality, the slighest over or under-weighting of the lottery gain induces a strict preference for the lottery or sure prospect, respectively.

To obtain a sharp prediction, consider the salience function of Equation (2). In this case, the lottery gain is more salient than the loss whenever:

$$\left(y + \frac{\theta}{2}\right)(1 - 2p) > (y - \underline{x})(1 - p), \tag{11}$$

which uniquely identifies the parameter values for which the agent is risk seeking. To shed light on the experiments of Section 2, we now investigate how the agent's risk attitudes vary holding the lottery loss $(y - \underline{x})$ constant at some value l (in Section 2, $l = 19$). The risk attitudes implied by Equation (11) are graphed in Figure 3:

Two patterns stand out. First, as in Section 2, for a fixed $p < 1/2$, an increase in the

[7] This is shown by replacing Equation (9) into Equation (8): when $\delta > 0$ the local thinker chooses the sure prospect when $\overline{x} < \overline{x}^*$ and $p \geq \frac{y - \underline{x}}{\delta(\overline{x} - \underline{x}) + (y - \underline{x})(1 - \delta)}$, while he chooses the lottery when $\overline{x} > \overline{x}^*$ and $p \geq \frac{(y - \underline{x})\delta}{(\overline{x} - y) + (y - \underline{x})\delta}$. The main features of the local thinker's choice track those represented in Figure 2, but now the local thinker's risk attitudes become sensitive to the lottery premium. For instance, as the lottery upside grows closer to $\overline{x}^*$ in the Excess Safe region then – for a given probability p – the local thinker eventually starts taking the lottery because its premium has gone up.

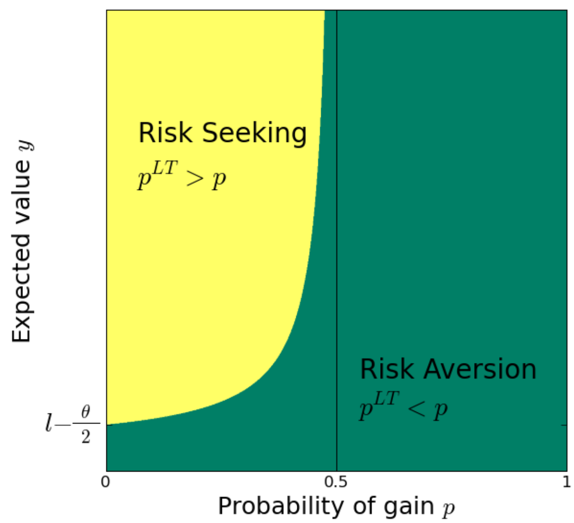

Figure 3: Shifts in risk attitudes

lottery's expected value y fosters risk seeking. When the expected value y is low, the lottery's downside is close to zero and diminishing sensitivity implies that the loss is salient, inducing risk aversion. As y becomes larger, diminishing sensitivity becomes weaker and absolute payoff differences become crucial for determining salience. Since for $p < 1/2$ the lottery gain is larger than the loss, for y sufficiently large this effect leads to risk taking.[8]

Second, for a given expected value y, a higher probability p of the gain reduces risk seeking. Intuitively, as p increases the lottery's upside must fall for the expected value to stay constant. As a consequence, the lottery gain becomes less salient, inducing risk aversion. Risk taking never occurs when $p \geq 1/2$, since in this case the gain is smaller than the loss and thus the latter is always more salient by diminishing sensitivity.

One way to present this logic is to derive the probability weighting function generated by our model. Fix a value of y in Figure 3 and increase the probability p along the horizontal axis. Figure 4 below shows the decision weight p^{LT} along this path as determined

[8]The main features of Figure 3 are obtained under the properties in Definition 1. Define $l = y - \underline{x}$ as the fixed downside of the lottery, and $\overline{x}(p, y)$ as the upside at which the lottery's expected value is equal to the sure prospect. Then the local thinker is risk seeking provided $\sigma(\overline{x}(p, y), y) > \sigma(y - l, y)$. Since $\overline{x}(p, y)$ decreases in p, by the ordering property we know that, holding y constant, for p sufficiently large the lottery downside becomes salient and the agent is risk averse. This is surely the case for $p \geq 1/2$. On the other hand, since $\sigma(y - l, y)$ decreases in y the convexity property implies that, for fixed $p < 1/2$, as y increases the upside eventually becomes salient, generating the patterns in Figure 3.

by our model, where p^* is the threshold at which the agent switches from risk seeking to

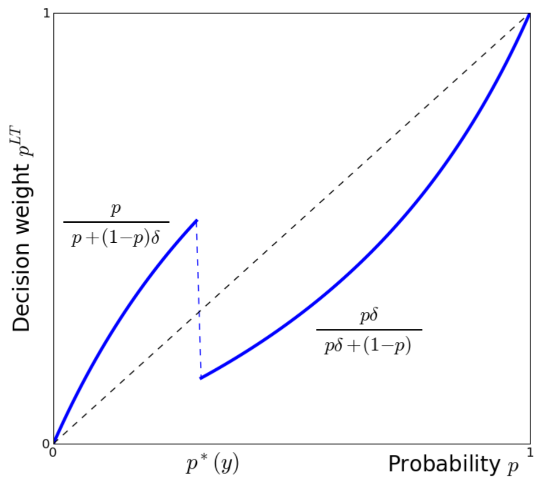

Figure 4: Context dependent probability weighting function

risk aversion in Figure 3. Remarkably, our model endogenously recovers the key features of Prospect Theory's inverse S-shaped $\pi(p)$ function (KT 1979): over-weighting of low probabilities, under-weighting of high probabilities. In Figure 4, 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.

Crucially, however, in our model the specific weighting function of Figure 4 does not capture subjects' universal perception of probabilities, because the weighting function is *context dependent.* At high expected values y, the probability weighting function shifts up, which expands the range of overweighting, leading to risk seeking even at moderate probabilities; at low expected values y, the probability weighting function shifts down, which expands the range of under-weighting, leading to risk aversion even at low probabilities. More generally, decision weights in our model are shaped by the way lottery choice is presented. Section 5 uses this idea to unify observed patterns of violation of the independence axiom and preference reversals.

These results indicate that the salience of particular states can be a powerful force inducing individuals to gamble or engage in risky behavior in conditions that are far more common than those characterizing longshot bets. More generally, changes in lottery stakes or in probabilities that are irrelevant to an expected utility maximizer can, by shaping salience, induce a local thinker to "over-react" in the direction of risk seeking or risk aversion.

4.2 Experimental Evidence on Risk Attitudes

We tested the predictions of Figure 3 by giving experimental subjects a series of binary choices between a lottery $L_1(p) = (\overline{x}, p; \underline{x}, 1-p)$ and a sure prospect $L_2 = (y, 1)$ where L_1 is a mean preserving spread of L_2. We set the downside of the lottery at $l = \$20$, yielding an upside $(\overline{x} - y)$ of $\$20 \cdot (1-p)/p$. We varied y in $\{\$20, \$100, \$400, \$2100, \$10500\}$ and p 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 p or y constant. The observed proportion of subjects choosing the lottery for every combination (y, p) is reported in Table 1; for comparison with the predictions of Figure 3, the results are shown graphically in Figure 5.

\$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
	0.01	0.05	0.2	0.33	0.4	0.5	0.67

Table 1: Proportion of Risk-Seeking Subjects

The patterns are qualitatively consistent with the predictions of Figure 3. For a given expected value y, the proportion of risk takers falls as the probability p of the gain increases; for a given probability $p < 0.5$ of the gain, the proportion of risk takers increases with the expected value y. The effect is statistically significant: at $p = 0.05$ a large majority of subjects (80%) are risk averse when $y = \$20$, but as y increases to \$2100 a large majority (65%) becomes risk seeking. Finally, there is a large drop in risk taking as p crosses 0.5. Note that the increase in y raises the proportion of risk takers from less than 10% to 50% even for moderate probabilities in the range $(0.2, 0.35)$. Even for probabilities as high as

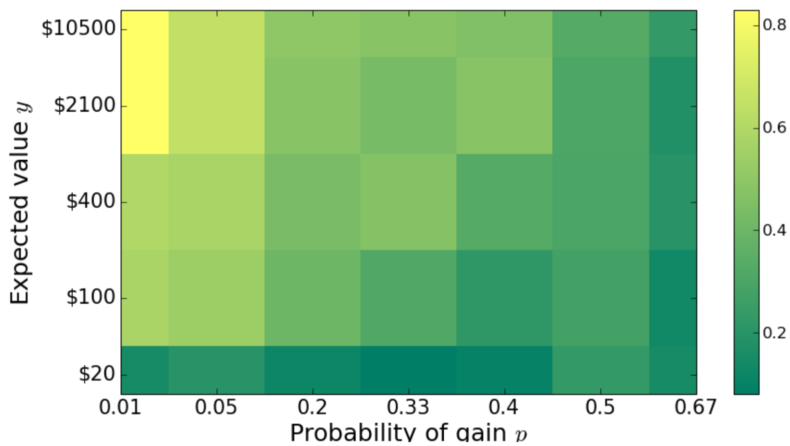

Figure 5: Proportion of Risk-Seeking Subjects

40%, we observe a 5-fold rise in risky choices as y increases. These patterns are broadly consistent with the predictions of our model. As hinted in Section 2, the weighing function of Prospect Theory and CPT can explain why risk seeking prevails at low p, but not the shift from risk aversion to risk seeking as y rises. To explain this feature, both theories need a concave value function characterized by strong diminishing returns.[9]

We further explore our model's predictions by considering the choice between a sure prospect of gaining y and a longshot lottery L that promises, for the same expected value y, a large positive payoff with small probability and zero with large probability. Equation (11) implies that for these lotteries, where $\underline{x} = 0$, risk taking occurs when y is sufficiently small:

$$y < \theta \frac{1 - 2p}{2p}. \tag{12}$$

Now a high expected value y discourages risk seeking. Note the difference with Figure 5, where higher y fosters risk taking: in that case the amount of the lottery's loss $(y - \underline{x})$ stays constant so that by diminishing sensitivity the salience of the loss state decreases as all payoffs rise. In the case of longshot lotteries instead, a higher y unambiguously boosts the salience $\sigma(0, y)$ of the loss while it exerts two conflicting effects on the salience of the gain,

[9] In the Appendix we explore different calibrations of Prospect Theory, and in particular find that these results cannot be explained with Tversky and Kahneman's (1992) baseline calibration of the value function for two outcome lotteries.

$\sigma(y, y/p)$, for it increases the payoff difference but also raises the average payoff level. Using the salience function of Equation (2), these effects imply that an increase in y discourages risk taking by making more salient the state in which the longshot lottery loses.

To test this prediction, we asked subjects to choose between the longshot lottery and the sure prospect for $y \in \{1, 5, 20, 50, 100, 200\}$ and $p = 0.01$. As predicted by our model, subjects take longshot lotteries only at low expected values and risk seeking already drops dramatically between $y = \$5$ and $y = \$20$. Interestingly, Prospect Theory and CPT cannot yield this pattern under a standard value function $v(x) = x^\alpha$, for in this case risk attitudes over longshot lotteries only depend on probability weighting, and not on payoffs.

In the Appendix, we compute values for our parameters θ and δ that are consistent with the evidence on risk preferences of this section, namely $\delta \sim 0.7$ and $\theta \sim 0.1$. Although these values are not a formal calibration, we employ them as a useful reference for discussing Allais paradoxes in the next section.

5 Local Thinking and Context Dependence

We have seen that the context dependent probability weighting function produced by our model gives an intuitive account of risk attitudes. We now use such context dependence to shed further light on the Allais paradoxes and on preference reversals.

5.1 The "Common Consequence" Allais Paradox

The Allais paradoxes (1953) are the best known and most discussed instances of failure of the independence axiom. Kahneman and Tversky's (1979) version of the "common consequence" paradox compares the choices:

$$L_1^z = (2500,\ 0.33;\quad 0,\ 0.01;\quad z,\ 0.66),\quad L_2^z = (2400,\ 0.34;\quad z,\ 0.66) \tag{13}$$

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, for the latter cancels out in the comparison between L_1^z and L_2^z.

In reality, experiments reveal that for $z = 2400$ most subjects are risk averse, preferring $L_2^{2400} = (2400,\ 1)$ to $L_1^{2400} = (2500,\ 0.33;\ \ 0,\ 0.01;\ \ 2400,\ 0.66)$. When instead $z = 0$, most subjects are risk seeking, preferring $L_1^0 = (2500,\ 0.33;\ \ 0,\ 0.67)$ to $L_2^0 = (2400,\ 0.34;\ \ 0, 0.66)$. In violation of the independence axiom, z affects the experimental subjects' choices.

Prospect Theory and CPT (KT 1979 and TK 1992) explain the switch from L_2^{2400} to L_1^0 by the so called "certainty effect", namely the idea that adding a downside risk to the sure prospect L_2^{2400} undermines agents' valuation much more than adding the same downside risk to the already risky lottery L_1^{2400}. This effect is directly built into the probability weighting function $\pi(p)$ by the assumption of subcertainty, e.g. $\pi(0.34) - \pi(0) < 1 - \pi(0.66)$.[10]

Our model endogenizes this feature of decision weights, and thus explains the Allais paradox, by arguing that the common consequence z alters the salience of lottery outcomes. To see this, consider the choice between L_1^{2400} and L_2^{2400}. Here there are three states of the world and the most salient state is one where L_1^{2400} pays zero and the riskless lottery L_2^{2400} pays 2400 because:

$$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400). \tag{14}$$

The inequalities follow from diminishing sensitivity and ordering, respectively, and can be verified for the case of the salience function in Equation (2). By Equation (5), a local thinker then prefers the riskless lottery L_2^{2400} provided:

$$(0.01) \cdot 2400 - \delta \cdot (0.33) \cdot 100 > 0, \tag{15}$$

which holds for $\delta < 0.73$. Although the risky lottery L_1^{2400} has a higher expected value, it is not chosen when local thinking is sufficiently 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 there are four states of the world, whose salience ranking is:

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

The first inequality follows from ordering, and the second from diminishing sensitivity. By

[10] In CPT 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.

Equation (5), a local thinker prefers the risky lottery L_1^0 provided:

$$(0.33) \cdot (0.66) \cdot 2500 - \delta \cdot (0.67) \cdot (0.34) \cdot 2400 + \delta^2 \cdot (0.33) \cdot (0.34) \cdot 100 > 0 \tag{17}$$

which is verified for all δ, so that any local thinker chooses the risky lottery L_1^0 because its upside is very salient.

In sum, when $\delta < 0.73$ – which holds in the parameterization $\delta = 0.7$, $\theta = 0.1$ – a local thinker exhibits the Allais paradox. 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 agent 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, just as the downside of the risky lottery. As a result, the lotteries' upsides are now crucial to determining salience. This induces the agent to overweight the larger upside of L_1^0, triggering risk seeking. The salience of payoffs endogeneizes the "certainty effect" as a form of context dependence: when the same downside risk is added to the lotteries, the sure prospect is particularly damaged because the common downside payoff induces the agent to focus on the larger upside of the risky lottery, leading to risk seeking choices.

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 (13):

Probability	0.01	0.33	0.66
payoff of L_1^z	0	2500	z
payoff of L_2^z	2400	2400	z

(18)

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 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 (5), so that the value of z should not affect the choice at all. 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 above.

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	8%	8%
L_2^0	13%	71%

The vast majority of subjects do not reverse their preferences (79% of subjects lay on the 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 (18). Among the few subjects reversing their preference, no clear pattern is detectable. 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 by stressing that when the lotteries pay z in the same state, this state is not salient and is disregarded in evaluation.[11]

5.2 The "Common Ratio" Allais Paradox

The "common ratio" effect occurs in the choice between lotteries:

$$L_1^q = (6000,\ q; \quad 0,\ 1-q), \quad L_2^p = (\alpha \cdot 6000,\ p; \quad 0,\ 1-p), \tag{19}$$

where $\alpha < 1$. By the independence axiom, an expected utility maximizer with utility function $u(\cdot)$ chooses L_2^p over L_1^q when:

$$\frac{q}{p} \cdot u(6000) + u(0) \left(1 - \frac{q}{p}\right) \leq u(\alpha \cdot 6000). \tag{20}$$

[11] 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 Appendix shows, our results are robust to all these variations. We also ran an experiment where subjects were explicitly presented the lotteries of Equation (13) 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 an agent 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 Appendix).

The choice should not vary as long as the odds q/p with which L_1^q pays out relative to L_2^p are kept constant. A stark case arises when $q/p = \alpha$; now the two lotteries have the same expected value and a risk averse expected utility maximizer always prefers L_2^p to L_1^q for any p. Parameter α identifies the "common ratio" between q and p at different levels of p.

It is well known (KT 1979) that, contrary to the Expected Utility Theory, the choices of experimental subjects depend on the value of p: for fixed $q/p = \alpha = 0.5$, when $p = 0.9$ 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 $p = 0.002$, 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 goes down 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 $\pi(p)$ grows slower than linearly for small p; hence, $\pi(\alpha p)/\pi(p)$ is relatively high at low p, inducing the choice of L_1^q when $p = 0.002$.

Consider the choice between L_1^q and L_2^p in our model. For $\alpha = 1/2$ there are four states of the world, and the salience ranking among them is

$$\sigma(6000, 0) > \sigma(0, 3000) > \sigma(6000, 3000) > \sigma(0, 0), \tag{21}$$

as implied by ordering and diminishing sensitivity. This implies that the local thinker evaluates the odds with which the riskier lottery L_1^q pays out relative to the safer one L_2^p as:

$$\frac{q^{LT}}{p^{LT}} = \frac{q}{p} \cdot \frac{(1-p) + p\delta^2}{(1-q)\delta + q\delta^2}. \tag{22}$$

The local thinker then follows the criterion in Equation (20) where the true ratio $q/p = 1/2$ is replaced by the distorted ratio of Equation (22). With a linear utility, the local thinker selects the safer lottery L_2^p if $q^{LT}/p^{LT} \leq 1/2$ and the riskier lottery L_1^q otherwise. This implies that the local thinker chooses the safer lottery when:

$$p \geq \frac{2(1-\delta)}{2 - \delta - \delta^2}. \tag{23}$$

As in the common ratio effect, the local thinker is risk averse when p 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 $L_1^{0.01}$ just as experimental subjects do. Our model therefore generates the common ratio effect, also under the parameterization $\delta = 0.7$, $\theta = 0.1$.

The intuition for this result is that there is a tradeoff between the salience and likelihood of states in our model. The upside of the riskier lottery L_1^q is salient at every p, creating a force toward risk seeking. When p is large, though, obtaining a positive payoff under the safer lottery is much more likely than under the riskier one. If δ is not too low, this implies that the decision weight attached to the second most salient state in ranking (21) is higher than the one attached to the first most salient state, so that the agent is risk averse. When instead p is low, the safer lottery pays out with just marginally higher probability. Now decision weights are mostly shaped by the salience of payoffs, inducing the local thinker to over-weight the higher upside of L_1, which triggers risk seeking.

Crucially, experimental evidence shows that this common ratio effect is also not robust to the introduction of correlation among lotteries. KT (1979) gave subjects a choice between $L_1^{0.5}$ or L_2^1 from Equation (19), in a game with two stages. 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 chosen lottery is played out. Most subjects here are risk averse, choosing the safer lottery L_2^1. The key point is that the final distribution of payoffs in this two stage gamble corresponds to a choice between $L_1^{0.125}$ and $L_2^{0.25}$. Crucially, when directly presented with the single stage choice between $L_1^{0.125}$ and $L_2^{0.25}$, most subjects are risk seeking and choose $L_1^{0.125}$, just as in the low probability case of the common ratio effect. This is not so when the probabilities are reduced by adding a correlated state in which both lotteries pay zero. In this case, there is no violation of the independence axiom.

In explaining this behavior, KT informally argue that individuals "edit out" the common consequence when the two lotteries are presented as correlated, thereby choosing as if $p = 1$. Our model yields such editing as a consequence of the low salience and cancellation of the correlated state in which both lotteries pay 0. To see this, note that adding a correlated state where both lotteries pay 0 neither affects the salience ranking in Equation (21), in which the correlated state is just a replica of $(0, 0)$, nor – more importantly – the odds ratios between the first three most salient states. As a consequence, for any probability

$r < 1$ of the correlated $(0, 0)$ state, the odds ratio q^{LT}/p^{LT} assessed by the agent is equal to that in Equation (22) where the correlated state is absent. In other words, the local thinker chooses as if he disregards the correlated state and its probability r, namely as if he is choosing among the single stage lotteries $L_1^{0.5}$ and L_2^1. This is what experimental subjects do. We thus reconcile the subjects' choice in single and two stage lotteries: in the single stage case, the rescaling of probabilities affects choice by changing the weights attached to salient states. In the two stage case, rescaling occurs by adding a correlated non-salient state which is disregarded in choice, leaving the weighting of salient states unaffected.

In sum, our model explains the Allais paradoxes as the product of a specific form of context dependence working though the salience of lottery payoffs. Adding a common payoff to all lotteries changes risk preferences by changing the salience of lotteries' upsides or downsides. Rescaling the lotteries' probabilities shapes the importance of salience vs. likelihood in determining decision weights, which also affects choice. Crucially, the presence of context dependence implies that risk attitudes depend on how the lotteries are correlated. Adding a common payoff or rescaling probabilities by introducing in the lotteries a non-salient payoff state does not affect choice: it is too enticing for subjects to disregard this non-salient state and to abide by the independence axiom.

5.3 Preference Reversals

Context dependence in our model can also explain the phenomenon of preference reversal described by Lichtenstein and Slovic (1971). They showed that subjects may systematically prefer a safer lottery L_p to a riskier lottery $L_\$$ and yet have a higher minimum selling price for the riskier lottery $L_\$$ (we follow the conventional notation, whereby L_p has a high probability of a low payoff, and $L_\$$ a low probability of a high payoff). Subjects' preferences are thus revealed by choice to be the opposite of preferences revealed by pricing. This phenomenon, confirmed also by Grether and Plott (1979) and Tversky et al. (1990), is at odds with both standard Expected Utility Theory and Prospect Theory, leading to claims that choosing and pricing follow two fundamentally different principles.

To study preference reversals in our model, we must specify how a local thinker prices a lottery. As in Expected Utility Theory, we define the minimum selling price to be the price at which the local thinker is indifferent between playing the lottery and receiving that price. For a local thinker, this price is found by replacing the lottery's probabilities with decision weights. Maintaining our interpretation of decision weights as spontaneously arising from the agent's perception of the alternatives, it is natural to assume that the lotteries being considered jointly determine the weights used by a local thinker to compute the selling price of each of them. Formally, in the context of choosing between lotteries L_1 and L_2, a local thinker with a value function $v(\cdot)$ prices L_1 at:

$$P(L_1\,|L_2) = v^{-1}\left[\sum p_s^{LT} v(x_s)\right], \tag{24}$$

where p_s^{LT} are the decision weights in the choice of L_1 vs. L_2. When the agent is asked to price a lottery in isolation, we assume that he evaluates it in the context of a choice between the lottery and the status quo of not having it $L_0 \equiv (0,1)$, i.e. of getting zero for sure. The price $P(L_1|L_2)$ is then the (decision weights-adjusted) certainty equivalent of L_1 in the context of the comparison with L_2; with a linear value function, this corresponds to the expected value of L_1 as perceived by the local thinker.

Consider now how preference reversals arise in our model. In the experiments, subjects are first asked to price in isolation the following two lotteries:

$$L_p = \begin{cases} x, & p \\ 0, & 1-p \end{cases}, \qquad L_\$ = \begin{cases} 2x, & p/2 \\ 0, & 1-p/2 \end{cases} \tag{25}$$

and subsequently to choose among them. The lotteries in Equation (25) are of the kind used to test for the common ratio effect (where $\alpha = 1/2$) and typically have $p > 3/4$. Equation (23) thus says that, when asked to choose, a local thinker is risk averse and prefers L_p to $L_\$$, just like most experimental subjects do.

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

$$P(L_p\,|L_0) = x \cdot \tfrac{p}{p+(1-p)\delta}\ , \quad P(L_\$\,|L_0) = 2x \cdot \tfrac{p/2}{p/2+(1-p/2)\delta}. \tag{26}$$

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

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

Both lotteries are priced above their expected value, but $L_\$$ is more overpriced than L_p because it pays a higher gain with a smaller probability, and from Proposition 1 we know that lower probabilities are relatively more distorted.[12]

As a consequence, while in a choice context the local thinker prefers the safer lottery L_p, in isolation he prices the risky lottery $L_\$$ higher, thus exhibiting 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 in the sense that the alternatives of choice are different in the two cases.

Another way to describe our account of preference reversals is to note that these arise through violations of "procedural invariance", defined by Tversky et al. (1990) as situations in which a subject prices a lottery above its expected value, $P(L|L_0) > EV(L)$, and yet prefers the expected value to the lottery, $L \prec (EV(L), 1)$. In our model, such violations directly follow from context dependent pricing. Tversky et al (1990) show that the vast majority of observed reversals follow from the violations of procedural invariance, as predicted by our model. Regret Theory can produce preference reversals by a distinct mechanism, intransitivities in choice, but does not violate procedural invariance.

Our context-based explanation suggests that reversals between choice and pricing should only occur when pricing takes place in isolation but not if agents price lotteries in the choice context itself. To test this hypothesis, subjects were given a choice between lotteries, $L_\$ = (16, 0.31;\ 0, 0.69)$ and $L_p = (4, 0.97;\ 0, 0.03)$. Subjects also stated their certainty equivalents for the two lotteries, in isolation and in the context of choice.[13] Our prediction

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

[13] In our experimental design, each subject priced each lottery only once, and different lotteries were priced

of context dependent pricing would then be verified if preference reversal occurs between choice and pricing in isolation, but not between choice and pricing in the choice context.

We find that, despite considerable variation in subjects' evaluations (which is a general feature of such elicitations[14]) the results are consistent with our predictions. First, among the subjects who chose L_p over $L_\$$, L_p was on average priced lower than $L_\$$ in isolation:

$$P(L_p^{\text{iso}}) = 4.6 < P(L_\$^{\text{iso}}) = 5.2 \ .$$

Thus, there is a preference reversal between choice and average pricing in isolation. 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_p would exhibit the standard preference reversals (see Appendix 2).[15]

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_p higher, on average, than the alternative risky lottery $L_\$$:

$$P(L_p|L_\$) = 4.3 > P(L_\$|L_p) = 4.1$$

As predicted by our model, in the context of choice, the average pricing is consistent with choice.[16] One might object that this agreement may be caused by the subjects' wish to be coherent when they price just after a choice. However, recall that each subject priced only one of the two lotteries in the choice context. Moreover, we ran another version of the survey

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 Appendix 2 for more details).

[14]See Grether and Plott (1979), Bostic et al (1990), Tversky et al (1990).

[15]The average prices above imply that some subjects priced the safer lottery L_p 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 Appendix 2 we consider truncations of the data that filters out such overpricing.

[16]In our data, the distribution for $P(L_p|L_\$)$ does not dominate that for $P(L_\$|L_p)$. 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 Appendix 2 we look in a more detailed way at the manifestation and significance of fact ii) in light on Tversky's et al. (1990) analysis of preference reversals.

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_p in isolation, but similarly to L_p under comparison. 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.

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

6 Other applications and extensions

6.1 Reflection and Framing Effects

KT (1979) show that experimental subjects shift from risk aversion to risk seeking as gains are reflected into losses. Our model yields such 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 $L_1 = (x_i, p_i)$ and sure prospect $L_2 = (y, 1)$, both of which are defined over gains (i.e. $x_i > 0$, $y > 0$) and where $E(L_1) = y$. By Equation (6), a local thinker with linear utility is risk averse, choosing L_2 over L_1 when:

$$y - \sum_i p_i x_i = 0 > \text{cov}\left[x_i, \omega_i\right], \tag{27}$$

namely when the agent overweights lower payoffs. Suppose now that L_1 and L_2 are reflected into lotteries over losses $L_1' = (-x_i, p_i)_i$ and $L_2' = (-y, 1)$. Property 3) in Definition 1 then implies that such reflection does not alter the salience ranking. That is, if the ranking of state (x_i, y) is k_i, then the ranking of state $(-x_i, -y)$ is also k_i. As a result, the same agent is risk seeking, choosing L_2' over L_1' when:

$$0 < \text{cov}\left[-x_i, \omega_i\right] = -\text{cov}\left[x_i, \omega_i\right]. \tag{28}$$

Risk seeking prevails when small losses (i.e. low values of x_i) are more salient than larger ones. Note that Equation (27) is the same as Equation (28), implying that a local thinker with linear utility is risk averse for gains if and only if he is risk seeking for losses. Thus, our model yields the fourfold pattern of risk preferences[17] without assuming, as Prospect Theory does, a value function that is concave for gains and convex for losses. In our model, reflection of payoffs generates shifts in risk attitudes by inducing the agent to shift his attention from the lottery upside to its downside and vice versa.

KT (1981) exploited the reflection effect to investigate how choice among lotteries can be shaped by framing. Some subjects were first asked to imagine that they had been given a lump sum of 1000 Israeli shekels, and then to choose between $L_1 = (1000,\ 0.5; 0,\ 0.5)$ and $L_2 = (500,\ 1)$; in this experiment, 16% percent of the subjects chose the lottery. A different group of subjects was instead asked to imagine that they had been given 2000 shekels at the outset, and then to choose between the reflected version of the previous lotteries, $L_1' = (-1000,\ 0.5; 0,\ 0.5)$ and $L_2' = (-500,\ 1)$; in this case, as many as 69% of subjects chose the lottery. The two problems are identical in terms of final monetary payoffs, so an expected utility maximizer would always choose either the lottery or the sure prospect but never switch between them as gains are reflected into losses. In Prospect Theory, this framing effect arises because the agent adjusts his reference point to include the initial lump-sum payoff and then the S-shaped value function leads to risk aversion in L_1 vs. L_2 but to risk seeking in L_1' vs. L_2'.

In our model instead, this framing effect arises because salience depends on lottery payoffs as presented (not on final wealth states). Thus, even if the agent's utility is linear, risk attitudes change because the reflection of lottery payoffs shifts the agent's focus from the lottery downside to its upside. This intuition can also account for KT's (1981) famous Public Health Dilemma, which describes the outbreak of a disease that is expected to kill 600 people. When the choice between medical responses is framed in terms of lives saved (resp. lost), the local thinker specifies the payoffs as gains (resp. losses). Thus, in the "saved" frame the most salient outcome is the one where nobody is saved, leading to a risk-averse choice,

[17]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 et al (1990).

while in the "die" frame the most salient outcome is the one where nobody dies, triggering a risk-seeking choice.

6.2 Mixed Lotteries

We now 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. Under the salience function of Equation (2), which features $\sigma(x, y) = \sigma(-x, -y)$ 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 agent rejects L_S when his dislike for losses more than compensates for his focus on the lottery gain, i.e. $\lambda > 2/\delta$.[18] In lotteries where the negative downside is larger than the

[18] 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 (2) 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

positive upside, salience and loss aversion go in the same direction in triggering risk aversion.

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:

5) *Loss salience: for every state* (x, y) *such that* $x + y > 0$ *we have that*

$$\sigma(-x, -y) > \sigma(x, y).$$

This condition relaxes the symmetry around zero of the salience function of Equation (2) 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.[19]

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

for any degree of loss aversion $\lambda > 1$, the Local Thinker prefers the safer lottery L_2.

[19] If we endow the local thinker with a standard utility function, instead of a value function, then in the absence of property 5) the utility function would be subject to Rabin's critique (Rabin, 2000) in the domain of mixed lotteries (but not in the domain of positive lotteries). Adding property 5) would entail that aversion to mixed lotteries with positive payoffs follows from salience, and not from underlying preferences. Thus, even though such a local thinker is at heart an expected utility maximizer, he is immune to Rabin's critique.

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 paper leaves some important issues open. One of them is how agents choose between more than two lotteries. We focused on the two-lottery problems because most of the experimental evidence on paradoxes and shifts in risk attitudes deals with problems of this sort. Although there is no reason why it should not be possible to extend our formalism to account for choice among more than 2 lotteries, the N-lottery case does not directly follow from our current model because to define salience we need to specify what lottery states an agent is drawn to when he considers a larger choice set. We plan to pursue this extension in future work, which may be useful for important applications (such as those to finance).

More broadly, our specification of contrast as a driver of salience could be useful for thinking about a variety of economic situations. For instance, in 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). The same might be true when investors choose stocks, entrepreneurs choose projects, or voters choose political candidates. Applied to these cases, 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.

References

Allais, Maurice. 1953. "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'Ecole Americaine."*Econometrica*, 21(4): 503–546.

Barber, Brad, and Terrence Odean. 2008. "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors.", *Review of Financial Studies*, 21(2): 785–818.

Barberis, Nicholas, and M Huang. 2008. "Stocks as Lotteries: the implications of Probability Weighting for Security Prices.", *American Economic Review*, 98(5): 2066–2100.

Bell, David. 1982. "Regret in Decision Making under Uncertainty.", *Operations Research*, 30(5): 961–981.

Bostic, Raphael, R.J. Herrnstein and R. Duncan Luce. 1990. "The Effect on the Preference-Reversal Phenomenon of Using Choice Indifferences.", *Journal of Economic Behavior & Organization*, 13(2): 193–212.

Camerer, Colin. 1995. "Choice under Risk and Uncertainty."In *Handbook of Experimental Economics*, ed. John H. Kagel and Alan E. Roth, 587–667. Princeton, N.J. : Princeton University Press.

Camerer, Colin, and Teck-Hua Ho. 1994. "Violations of the betweenness axiom and nonlinearity in probability.", *Journal of Risk and Uncertainty*, 8(2): 167–196.

Chetty, Raj, Adam Looney and Kory Kroft. 2009. "Salience and taxation: Theory and Evidence.", *American Economic Review*, 99(4): 1145–1177.

Edwards, Ward. 1962. "Subjective probabilities inferred from decisions."*Psychological Review*, 69(2): 109–135.

Fellner, William. 1961. "Distortion of subjective probabilities as a reaction to uncertainty" *Quarterly Journal of Economics*, 75(4): 670–689.

Fishburn, Peter. 1982, "Nontransitive Measurable Utility.", *Journal Mathematical Psychology* 26(1): 31–67.

Friedman, Milton, and L. J. Savage. 1948. "The utility analysis of choices involving risk.", *Journal of Political Economy*, 56 (4): 279–304.

Gennaioli, Nicola, and Andrei Shleifer. Forthcoming. "What comes to mind." *Quarterly Journal of Economics.*

Grether, David, and Charles Plott. 1979. "Economic theory of choice and the preference reversal phenomenon." *American Economic Review*, 69(4): 623–638.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: an analysis of decision under risk." *Econometrica*, 47(2): 263–292.

Kahneman, Daniel. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93(5): 1449–1476.

Kandel, Eric, James Schwartz and Thomas Jessell. 1991. "Principles of Neural Science." 4th ed. Norwalk, Conn.: Appleton & Lange, c1991.

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

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

Holt, Charles, and Susan Laury. 2002. "Risk aversion and incentive effects." *American Economic Review*, 92(5): 1644–1655.

Lichtenstein, Sarah, and Paul Slovic. 1971. "Reversals of preference between bids and choices in gambling decisions." *Journal of Experimental Psychology*, 89(1): 46–55.

Loomes, Graham, and Robert Sugden. 1982. "Regret theory: An alternative theory of rational choice under uncertainty." *The Economic Journal*, 92(368): 805–824.

McCoy, Allison, and Michael Platt. 2005. "Risk-sensitive neurons in macaque posterior cingulate cortex." *Nature Neuroscience*, 8(9): 1220–1227.

Mullainathan, Sendhil. 2002. "Thinking through Categories." *mimeo* Harvard University.

Prelec, Drazen. 1998. "The probability weighting function." *Econometrica*, 66(3): 497–527.

Quiggin, John. 1982. "A Theory of Anticipated Utility " *Journal of Economic Behavior & Organization*, 3(4): 323–343.

Rabin, Mathew. 2000. "Risk aversion and expected-utility theory: A calibration theorem." *Econometrica*, 68(5): 1281–1292.

Rubinstein, Ariel. 1988. "Similarity and decision-making under risk." *Journal of Economic Theory*, 46(1): 145–153.

Schelling, Thomas. 1960. "The Strategy of Conflict ", Cambridge, Mass.: Harvard University Press.

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

Taylor, Shelley, and Suzanne Thompson. 1982. "Stalking the elusive vividness effect." *Psychological Review*, 89(2): 155–181.

Tversky, Amos and Daniel Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science*, 211(4481): 453–458.

Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, 5(4): 297–323.

Tversky, Amos, and Itamar Simonson. 1993. "Context-dependent preferences." *Management Science*, 39(10): 1179–1189.

Tversky, Amos, P Slovic and Daniel Kahneman. 1990. "The Causes of Preference Reversals." *American Economic Review*, 80(1): 204–217.

Wu, George, and Richard Gonzalez. 1996. "Curvature of the probability weighting function." *Management Science*, 42(12): 1676–1690.