Salience Theory of Choice Under Risk

Pedro Bordalo, Nicola Gennaioli, Andrei Shleifer*

April 2010

Abstract

We present a theory of choice among lotteries in which the decision maker's attention is drawn to (precisely defined) salient states of the world. This leads the decision maker to a context-dependent representation of lotteries in which true probabilities are replaced by probabilistic weights distorted in favor of salient payoffs. By endogenizing probability weights as a function of payoffs, our model delivers features of Prospect Theory without assumptions on the curvature of the utility function. The theory explains several famous paradoxes and experimental findings concerning choice under risk, but also makes novel predictions confirmed by new experiments.

---

*Harvard University, Universitat Pompeu Fabra and CREI, Harvard University. We are grateful to Nicholas Barberis, Tom Cunningham, Josh Schwartzstein, Jesse Shapiro and Tomasz Strzalecki for extremely helpful comments, to the Kauffman Foundation for support of this research, and to Allen Yang for excellent research assistance.

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 (Camerer 1995). A common feature of many of these violations is that experimental subjects change their choices after a change in the problem that should be irrelevant to an expected utility maximizer. For example, in problems such as the Allais (1953) paradox, experimental subjects change their preference between two lotteries when a common consequence is added to both. Experimental subjects also change their choices between two risky prospects when the same payoffs are presented as gains rather than as losses, the so-called reflection effect (Kahneman and Tversky, KT 1979). This influence of presentation or framing of a problem on decisions strikes at the very heart of the Expected Utility Theory.

Camerer (1995) reviews numerous attempts to amend the Expected Utility Theory by changing some of its axioms, but these attempts are not psychologically founded, and have not taken hold. The most successful psychological theory dealing with this evidence is KT's (1979) Prospect Theory. KT change three basic assumptions of the expected utility theory. First, they assume that subjects evaluate all risky choices relative to a reference point, so changes in that point can influence the perception of payoffs. Second, they assume different risk attitudes toward gains and losses, which helps them deal with reflection effects. Third, they assume that small probabilities get magnified by decision makers. Crucially, this magnification is a function of the probability itself, and not of payoffs. Together, these three assumptions enable KT to deal with a multitude of violations of EUT, but also to generate and test new predictions.

We propose an alternative model of choice under risk, which is centrally focused on how decision makers edit probabilities. Unlike in Prospect Theory, in our model the editing depends on the payoffs associated with each state of the world. In particular, we argue that, in comparing two lotteries, decision makers over-weight the probabilities of salient states, which are the states in which the contrast in the payoffs between these lotteries is the greatest. We formally define contrast and salience, and show how the decision maker renormalizes probabilities to put more weight on the more salient states. The new probability weights sum up to one. Through this process, the decision maker develops an endogenous, context-dependent, representation of each lottery. We do not need to make any non-standard assumptions about the value function. Once the representation of a risky choice transforms objective probabilities into probability weights, the decision maker is the standard expected utility maximizer. Indeed, we derive preferences toward risk from judgments about probabilities.

The psychological foundation of our theory is the idea of contrast. As Kahneman (2003) puts it, "changes and differences are more accessible to a decision maker than absolute values". The idea of contrast is best illustrated in the case of sensory information, where the relation between the magnitude of a stimulus and its perceived intensity can be examined quantitatively. To quote a familiar example, the brightness of an object depends not only on its luminosity (the objective measure of the amount of light it emits), but also on surrounding luminosities: the object will seem brighter when the surroundings are darker than when there are lighter. In a complex world, where many stimuli compete for attention, perception drives attention to some stimuli instead of others. We apply this idea to decision making, postulating that agents faced with a choice focus on those aspects that they perceive with greater intensity, namely those that exhibit the greater contrast among the alternatives. In lotteries, these are the states in which the payoffs of lotteries are most different.

At a broad level, our approach is similar to that pursued by Gennaioli and Shleifer (2010) in their study of the representativeness heuristic in probability judgments. The idea of both studies is that decision makers do not take into account fully all the information available to them, but rather over-emphasize the information their minds focus on[1]. Gennaioli and Shleifer (2010) call such decision makers local thinkers, because they neglect potentially important, but unrepresentative, data. Here, analogously, in evaluating lotteries, decision makers overemphasize the highest contrast states, and underemphasize the less distinctive ones. 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 differs from Expected Utility Theory only in the fact that true probabilities of various states are replaced by probability weights (adding up to one) based on well-defined salience of the payoffs in those states. But this modification yields important properties

---

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

that allow the model to account for a good deal of experimental evidence. First, it is the most basic property of the model that the over-weighting of probabilities is particularly severe when the salience and probability of states are very different. This implies, similarly to Prospect Theory, that small probabilities can be particularly over-weighted. However, in contrast to Prospect Theory, this would only occur if low probabilities are attached to highly salient states – those with extreme payoffs – but not otherwise. Second, a crucial property of the model is that salience is reflected from gains to losses. The most salient state in a choice between lotteries with positive payoffs remains the most salient state after all payoffs are reflected around zero. In particular, if the most salient state in the former choice has the highest payoffs, then the most salient state in the latter choice has the most negative payoffs. This induces a reflection effect: if a decision maker is risk-averse toward a lottery that only has gains as payoffs, he will be risk loving toward a lottery in which each payoff is replaced by its negative. We thus obtain risk-loving behavior toward losses in some situations without any assumptions on the value function.

Our model accounts for a broad range of available experimental findings, but also generates new predictions that go beyond existing theories of decision under risk. These predictions derive entirely from the assumption that payoffs, via the salience mechanism, shape probability assessments. We present experimental data addressing these predictions, and find a good deal of supporting evidence.

The paper proceeds as follows. In section 2, we begin with an experimental demonstration of the most basic prediction of the model, namely the switch from risk averse to risk-loving behavior toward gains as the payoffs – and their salience – change. We go through a special case of the model to analyze these experiments. In section 3, we present a more general formulation and derive the model's most fundamental properties, namely attitudes toward risk shaped endogenously by the salience of lottery payoffs and the reflection effect. In section 4, we use the model to analyze attitudes towards risk, and to describe in detail the forces that lead to risk averse and risk seeking behavior. In particular, we derive the building blocks of Prospect Theory from first principles, and provide experimental evidence for new predictions concerning risk attitudes. In Section 5 we analyse several famous violations of Expected Utility Theory, including framing effects and the Allais paradoxes. We illustrate our analysis by extending these paradoxes in several ways, including correlation effects. In Section 6, we extend our model to mixed lotteries and propose a finer notion of loss aversion, for which we also provide experimental evidence. We also study long-shot lotteries. Section 7 concludes.

2 A Simple Example

In this section, we present the results of two experiments, both of which involve a choice between a sure payoff and a lottery with two positive payoffs, one leading to a gain relative to the sure payoff, and one to a loss. We use these experiments to illustrate two central ideas of the 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 probability weights that shape choices. The same 78 subjects participated in the two experiments over the internet. In both experiments, subjects are asked which of the two options they prefer.

Experiment 1:

Option 1 : \$20 for sure
Option 2 : \$1 with probability 95% and \$381 with probability 5%.

Experiment 2:

Option 1 : \$320 for sure
Option 2 : \$301 with probability 95% and \$681 with probability 5%.

Three points about these experiments are noteworthy. First, the second experiment simply adds \$300 to all the options 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 very high payoff, and a high (95%) probability of a \$19 loss relative to the sure outcome.

Despite these similarities, we found that in Experiment 1 83% of the subjects chose the safe option 1, whereas in Experiment 2 70% of the same subjects chose the risky option 2. 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 Option 1 in the first experiment switched to Option 2 in the second.

Although in each experiment the two options offer the same expected value, the same subjects exhibit risk averse behavior in the first experiment, and risk loving behavior in the second. Standard expected utility theory typically assumes risk aversion, and so would have trouble accounting for Experiment 2. Prospect Theory would account for the two experiments by holding that small probabilities, in this case the 5% chance of a high outcome, are over-weighted by decision makers, which would lead them toward more risk loving behavior. However, the over-weighting of the 5% probability in Prospect Theory is independent of payoffs. To account for risk averse behavior in Experiment 1 and risk loving behavior in Experiment 2 requires a combination of the probability adjustment and declining absolute risk aversion. Indeed, the over-weighting of the 5% probability of the gain in the first experiment is a force against making a risk averse choice.[2]

Our explanation of these findings does not rely on any assumptions about the shape of the value function. It goes roughly as follows. In the first experiment, the payoff in the state of the bad outcome of a lottery, \$1, feels a lot lower than the sure payoff of \$20, and as a consequence subjects focus on this bad outcome. In the second experiment, the payoff in the state of the bad outcome of a lottery, \$301, does not feel nearly as bad compared to the sure payoff of \$320, and hence subjects focus on the good outcome of winning \$681. 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 sure payoff) pushes subjects toward risk-averse choice; the focus on gains pushes them toward risk-loving choice. We implement this idea by arguing that, in the first experiment, the salience of losses leads subjects to represent the problem as if the probability of a loss exceeded 95%. In contrast, in the second experiment, the salience of gains leads subjects to represent the

---

[2] In this analysis we assume that 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 preference reversal.

problem as if the probability of the gain exceeded 5%.

How do we implement this rough intuition formally? In section 3, we present a general model, but here consider just a special case helpful for analyzing the experiments. Suppose that utility is linear in payoffs. Suppose a decision maker compares a sure payoff x to a lottery that pays $x + g$ with probability p, and $x - l$ with probability $(1 - p)$. So, in our first experiment, $x = 20$, $g = 361$, $l = 19$, and $p = .05$. Assume as in the experiments that $pg = (1 - p)l$, so that the risk neutral expected utility maximizer is indifferent between the lottery and the safe choice x. For this choice between a safe payoff and a lottery, define the salience of the gain in the gamble as $(x + g)/x$ and the salience of the loss as $x/(x - l)$. The salience of each state is thus defined as the percentage difference, in that state, between the better and the worse outcome. (In Section 3 we present a more general formulation of salience.) We say that, in this decision problem, the gain state is more salient if $(x + g)/x > x/(x - l)$ and that the loss state is more salient if the inequality is reversed. The analogy here is to perception: the mind is drawn to salient states, which are the states with the strongest contrast between the good and the bad outcomes.

It is easy to check that in Experiment 1, the salience of the losses in the lottery, given by $20/1 = 20$, is greater than the salience of the gains, given by $381/20 = 19.05$. In contrast, in the second experiment, the salience of the losses, given by $320/301$, is smaller than the salience of the gains, given by $681/320$. Lottery losses are thus the salient state in the first experiment, while lottery gains are the salient state in the second experiment. In Section 3, we present a tentative formulation expressing the boost of the probability weight of the salient state relative to the objective probability in the represented decision problem. But the point is clear. The decision maker behaving as we described represents the lottery in the first experiment as less attractive than the objective odds indicate, and so, if his utility is linear in payoffs, chooses the safe option. In the second experiment, in contrast, the decision maker overstates the odds of a gain, represents the lottery as more attractive than it really is, and so chooses to take the risk.

Note a sharp difference of this model from Prospect Theory. In Prospect Theory, the decision maker would over-estimate low probabilities – those of gains – in both experiments. The payoffs do not influence the transformation of true probabilities into perceived probability weights. In our model, in contrast, what matters for the probability weights in the represented problem are the payoffs themselves, which, through their salience, influence which states the decision maker perceives to be likely. In the next several sections, we show that modification of probabilities based on salience makes many predictions similar to those of Prospect Theory, but also several distinctive ones.

3 The Model

An agent chooses between lotteries L_1 and L_2. The state space S includes all lotteries' payoff combinations occurring with positive probability. In a generic state $s \in S$ – which occurs with probability $\pi_s > 0$ – lottery L_1 pays x_s and L_2 pays y_s. We begin with lotteries that are either "positive", i.e. characterized by payoffs $x_s, y_s \geq 0$ for all s, or "negative", i.e. characterized by payoffs $x_s, y_s \leq 0$ for all s. We study "mixed" lotteries in Section 6.

Using this state space approach, we can consider both correlated and independent lotteries. If lotteries are independent $S = X \times Y$, where $X = (x_i)_{i=1,\dots,|X|}$ and $Y = (y_j)_{j=1,\dots,|Y|}$ are the sets of payoffs for L_1 and L_2 and each state $s = (x_i, y_j)$ occurs with probability $\pi_{i,j} = p_i \cdot q_j$, where p_i and q_j are the probabilities with which L_1 and L_2 pay x_i and y_j, respectively. The experiments of Section 2 are special cases of the choice between a lottery $L_1 = (\overline{x}, p; \underline{x}, 1-p)$ and a sure prospect $L_2 = (y, 1)$, where $\overline{x} > y > \underline{x}$. In this case, S consists of two elements: $(\overline{x}, y)$, the state in which L_1 "gains" over L_2, and $(\underline{x}, y)$, the state in which L_1 "loses" over L_2.

An expected utility maximizer with utility function $u : \mathbb{R} \rightarrow \mathbb{R}$ chooses L_1 over L_2 if and only if:

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

The agent trades off the state-by-state gains and losses of L_1 over L_2 by weighting them by their respective probabilities. For instance, the previous two outcomes lottery is preferred to the sure prospect when the gain in state $(\overline{x}, y)$, weighted by the probability of that gain, is larger than the loss in state $(\underline{x}, y)$, weighted by the probability of that loss.

The local thinker (LT) we consider instead over-weights in Equation (1) the most salient states $s \in S$ relative to the others. This process consists of an initial "representation" stage where states are ranked by salience and the probability π_s of each state is transformed into π_s^{LT}, and an "evaluation" stage where π_s^{LT} is used to trade-off lottery gains and losses just as 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 and symmetric salience function $\sigma(x, y)$ that satisfies four conditions:*
1) *Ordering: if $[x', y']$ is a subset of $[x, y]$, then $\sigma(x, y) > \sigma(x', y')$.*
2) *Diminishing sensitivity: if $x \geq y \geq \epsilon > 0$ then $\sigma(x + \epsilon, y) < \sigma(x, y - \epsilon)$.*
3) *Reflection: for any $x, y \geq 0$ and $x', y' \geq 0$, $\sigma(x, y) < \sigma(x', y')$ if and only if $\sigma(-x, -y) < \sigma(-x', -y')$.*
4) *Convexity: for any $\epsilon > 0$, $\lim\limits_{x \to +\infty} \frac{\sigma(x+\epsilon, x)}{\sigma(x, x-\epsilon)} = 1$.*

According to the ordering property, salience increases in the distance between payoffs. Diminshing sensitivity says that for positive states the salience of a state with a given difference between payoffs is higher at a lower average payoff level. Reflection says that the salience ordering of two states does not change when the sign of their payoffs is reversed. The consequence of properties 2) and 3) is that diminishing sensitivity works symmetrically with respect to negative payoffs, namely if $x, y > 0$ then $\sigma(-x - \epsilon, -y) < \sigma(-x, -y + \epsilon)$, so that in the negative domain the salience of a state with a given distance between payoffs falls as the average payoff becomes more negative. That is, for a given distance between payoffs, the salience of a state increases when the absolute value of the average payoff is lower. Salience is shaped by the magnitude rather than the sign of payoffs. Section 3.1 discusses the connection between these properties and the cognitive notion of salience. Finally, property 4) naturally implies that as the average payoff level in two states becomes very large, the effect of diminshing sensitivity becomes weak and the distance between payoffs becomes crucial to determining salience.

One example of a salience function satisfying the above properties is:

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

where $\theta \geq 0$ is a constant. We used this function in Section 2 with $\theta = 0$. In Equation

(2), $\sigma(x,y) \leq 1$ and θ captures the salience of zero payoffs. If $\theta = 0$, the salience of states with zero payoffs is the highest since $\sigma(x,0) = 1$ for every x. For $\theta > 0$, a state with a zero payoff can be less salient than a state with non-zero payoffs if the distance among payoffs in the latter state is sufficiently large. The salience function in (2) displays a particularly strong form of reflection because for this case $\sigma(x,y) = \sigma(-x,-y)$. We sometimes use this tractable salience function to illustrate the workings of our model, but our main results rely on the properties in Definition 1.

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

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

and $\delta \in (0,1]$. Thus, defining $\omega_s = \delta^{k_s - 1} / \left(\sum_r \delta^{k_r - 1} \cdot \pi_r\right)$, the probability weight on a generic state $s \in S$ is equal to:

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

Note that $\sum_s \pi_s^{LT} = 1$. 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 less salient states, so that state s is over-weighted (i.e. $\omega_s > 1$) if and only if $\delta^{k_s - 1} > \sum_r \delta^{k_r - 1} \cdot \pi_r$, namely when such state is sufficiently salient to be less discounted than the average. For a local thinker with $\delta < 1$, ω_s represents the extent of over-weighting of state s.

After this representation stage, the agent computes the expected utility of 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:

$$\sum_{s \in S} \delta^{k_s - 1} \pi_s \left[u(x_s) - u(y_s)\right] > 0. \tag{3}$$

which is clearly equivalent to $\sum_{s \in S} \pi_s^{LT} \left[u(x_s) - u(y_s) \right] > 0$. The case where $\delta \to 0$ describes the agent who only focuses on the most salient state and decides only based on payoffs in that state. More generally, when $\delta < 1$ the local thinker under-weights – compared to the expected utility maximizer – 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

A fundamental principle in human perception is that a sensorial stimulus gives rise to a subjective representation whose intensity is related the stimulus' magnitude but also depends on context (Kandel 1991). The salience function $\sigma(x, y)$ translates into the context of choice two key features of this perceptual principle. The first is the notion that the subjective intensity of a stimulus varies in its ratio relative to a reference level (Weber's law). This is captured by properties 1), 2) and 4) of Definition 1. The second feature is the idea that people focus on the magnitude of stimuli, rather than on their sign. This is captured by property 3), which posits that salience is influenced by the absolute level of payoffs. Just as we perceive relative changes in the temperature of hot or cold objects similarly, the salience ranking among states is preserved after the sign of payoffs is changed from positive to negative.[3],[4]

In biology, salience detection is viewed as a key attentional mechanism enabling organisms to focus their limited perceptual and cognitive resources on the most pertinent 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

---

[3] 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.

[4] One arguably important assumption of our model is that the salience ranking does not depend on a state's probability. In spite of this assumption, which allows to identify the pure effect of payoffs, a state's probability will affect the magnitude of salience-induced distortions, as the next section shows.

disproportionate weighting in subsequent judgments." We capture this idea by assuming in Definition 2 that the agent places greater weight on salient states by virtue of his limited cognitive resources, captured by the extent of "local thinking" $1/\delta$.[5] Such focus on salient states may in many cases facilitate an efficient use of limited cognitive resources, but can become problematic when the salience and probability of a state are far apart. The assumed exponential discounting of less salient states buys us analytical tractability, but the key properties of our model hold also if we assume that the weighting of state (x, y) is a continuous and increasing function of $\sigma(x, y)$.[6]

As we formally show in Section 4, our model can be viewed as providing a pshychological foundation for probability weighting. Following earlier work (Edwards 1962, Fellner 1961), Kahneman and Tversky (1979) explore the idea that agents distort the stated probabilities of lotteries. For a recent attempt to estimate the probability weighting function, see Wu and Gonzalez (1996). Unlike in our model, in these approaches the weighting function $\pi(p)$: i) is derived from revealed preferences rather than from psychological first principles, and ii) does not depend on lottery payoffs. Quiggin's (1982) rank-dependent expected utility and KT's (1992) Cumulative Prospect Theory constitute the first 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 Cumulative Prospect Theory. Compared to this work, our model captures the idea that weighting distortions do not depend only on the rank of a lottery's payoffs but also on the state space generated by the available lotteries.

Specifically, in our model the salience of a certain lottery outcome depends on the outcomes of the lotteries to which it is compared. As a consequence, in our model the agent's choice is sensitive to correlation between lotteries and to context more generally. Sections

---

[5] To reconcile the intuition from Weber's law with psychological research on salience, we think of a choice problem as a trade-off between two alternatives which differ along a set of states of the world. Our aim is to assign subjective intensities to these states. Thus, we are not assigning subjective intensities to the actual stimuli, such as monetary payoffs, but rather to the weight of each state in the trade-off.

[6] Under a general increasing "discounting" function $\delta\left[\sigma(x, y)\right]$, the cardinality properties of $\sigma(x, y)$ would generate additional effects beyond those we already find, without however changing our basic conclusions. The properties of Proposition 1 also hold under a general formula for inflating the odds of more salient states. The only significant restriction embodied in our $1/\delta$ expression is that the distortions in the odds of any two states do not depend on these state's underlying probabilities.

4.2 and 6.3 provide experimental support for these properties, which we show in Section 5.2 to play an important role in accounting for the Allais paradox.

Probability weighting is the only cognitive "friction" affecting the local thinker, who is endowed with a standard (concave or linear) utility function. This utility function can be defined over lottery payoffs or final wealth levels, but does not change curvature around a reference point. Although there is no explicit reference point in our model, our assumption that salience depends on the absolute size (rather than the sign) of lottery payoffs may be viewed as setting an implicit reference point of zero in the evaluation of payoffs. We make this assumption throughout the paper, but we do not exclude the possibility that agents may perceive payoffs relative to an endogenously changing reference point. In this respect, we view our approach as complementary to models of endogenous reference point formation such as Koszegi and Rabin (2007).

Our approach is also related to Rubinstein's (1988) model of similarity-based choice, in which agents simplify the choice among two lotteries by pruning the dimension (probability or payoff) along which lotteries are similar. We share with Rubinstein the general emphasis on a procedural approach to modelling boundedly rational choice and framing. However, the specific mechanisms for context dependence and the predictions of the two approaches are very different. Finally, we share the emphasis on the key role of lotteries' state space with Loomes and Sugden's Regret theory (1982), and more generally with Fishburn's formalism (1982). In these latter approaches, however, the state space (and correlation) of lotteries directly affects the agent utility rather than indirectly affecting choice through salience and probability weighting, as it does in our model.

3.2 Lottery Valuation and Risk Preferences

Consider again the choice between lotteries $L_\mathbf{1}$ and $L_\mathbf{2}$, with payoffs $x_\mathbf{s}$, $y_\mathbf{s}$ in state $s \in S$ which occurs with probability π_s. Given a local thinker's over-weighting ω_s of state s, it is easy to see that salience raises his evaluation of lottery L_1 relative to that of an expected utility maximizer (who has $\delta = 1$) if and only if:

$$\sum_{s \in S} \pi_s u(x_s)\left(\omega_s - 1\right) = \operatorname{cov}\left[u(x_s), \omega_s\right] > 0. \tag{4}$$

A similar expression holds for L_2. Salience induces over-valuation of a lottery if and only if that lottery's high payoff states tend to be more salient than average. As a consequence, the local thinker chooses L_1 over L_2 if and only if:

$$\sum_{s \in S} \pi_s u(x_s) - \sum_{s \in S} \pi_s u(y_s) > \operatorname{cov}\left[u(y_s) - u(x_s), \omega_s\right]. \tag{5}$$

When an expected utility maximizer chooses L_1, the local thinker might choose L_2 instead if the states where L_2 delivers higher payoffs than L_1 [i.e. where $u(y_s) - u(x_s) > 0$] are more salient.

Equation (5) illustrates that the local thinker's over-weighting of salient states plays a central role in his departures from expected utility theory. To account for these departures, 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:*

$$\frac{d\omega_s}{dh} = -\frac{\pi_s}{1-\pi_s}\omega_s\left(\omega_s - 1\right). \tag{6}$$

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

The over-weighting of the probability of a state is shaped by two factors: the state's true probability and the salience of its payoffs. Property 1) says that a marginal increase in the probability of a state reduces over-weighting ω_s if that state is over-weighted to begin with (i.e. $\omega_s > 1$); if instead the state is under-weighted (i.e. $\omega_s < 1$) an increase in that state's probability reduces such under-weighting (increases ω_s).[7] Put differently, low probability states are subject to the strongest distortions: they are severely over-weighted if salient and severely under-weighted otherwise. Intuitively, since large probabilities tend to be highly weighted regardless of their salience, low probability events are relatively much more under or over-weighted depending on their salience. In contrast to Kahneman and Tversky's (1979) assumption that low probability states are always over-weighted, our model implies that this is only the case if those states are associated with extreme, and thus highly salient, payoffs.

Consider now changes in payoffs increasing the salience of a state: namely, if $x_s > y_s$, consider increases in x_s or reductions in y_s. Property 2) says that higher salience necessarily increases the over-weighting of a state. When the difference between the payoffs of a state becomes sufficiently extreme that the agent pays more attention to that state, the probability weight placed by the agent on such state unambiguously increases.[8]

The broad message of Proposition 1 is that the strongest over- or under-weighting occurs when a low probability state is associated with an extreme payoff, or put differently, when the salience and probability of a lottery state are far apart. In the next two subsections, we explore the implications of this notion to investigate 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 concave utility $u(\cdot)$ chooses the

---

[7] Proposition 1 establishes this property by looking at changes in the probabilities that only affect the odds of state s with respect to all other states, but keeping the odds among the latter constant. The advantage is that, by focusing on this specific path of change, we can isolate the pure effect of a change in π_s which equally affects the odds of s relative to all other states.

[8] Property 1) above trivially translates into an effect on the weight placed on the specific outcome of one of two independent lotteries. In this case, $p_i^{LT} = p_i \sum_j q_j \omega_{ij} / \sum_{kj} p_k q_j \omega_{ij}$ which immediately implies that:

$$\frac{d}{dp_i}\left(\frac{p_i^{LT}}{p_i}\right) = -\frac{p_i}{1-p_i}\mathbb{E}_j\left(\omega_{i,j}\right)\left[\mathbb{E}_j\left(\omega_{i,j}\right) - 1\right]$$

where $\mathbb{E}_j\left(\omega_{i,j}\right)$ is the average overweighting of states where x_i is the payoff of lottery L_1.

Property 2) is not straightforward in this case, in the sense that a change in a payoff such as an increase in x_i would increase the salience of states where $x_i > y_j$ but decrease the salience of all other states, exerting a potentially ambiguous effect on p_i^{LT}/p_i.

lottery if and only if:

$$p \geq \widehat{p}(\overline{x}) \equiv \frac{u(y) - u(\underline{x})}{u(\overline{x}) - u(\underline{x})}, \tag{7}$$

namely when the lottery wins over the sure prospect with sufficiently high probability. Threshold $\widehat{p}(\overline{x})$ decreases 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 (7) with the salience distorted weight p^{LT}. To determine p^{LT}, note that the agent views the state $s_g = (\overline{x}, y)$ in which the lottery "gains" as more salient than the state $s_l = (\underline{x}, y)$ in which the lottery "loses" when $\sigma(\overline{x}, y) > \sigma(\underline{x}, y)$. Under Definition 1, for a given $(\underline{x}, y)$ there is a threshold $\overline{x}^*$ such that s_g is more salient than s_l provided $\overline{x} > \overline{x}^*$ (under the salience function of Equation (2) with $\theta \rightarrow 0$, $\overline{x}^* = y^2/\underline{x}$). As the agent inflates by $1/\delta$ the odds of the salient state, the weight of s_g is equal to:

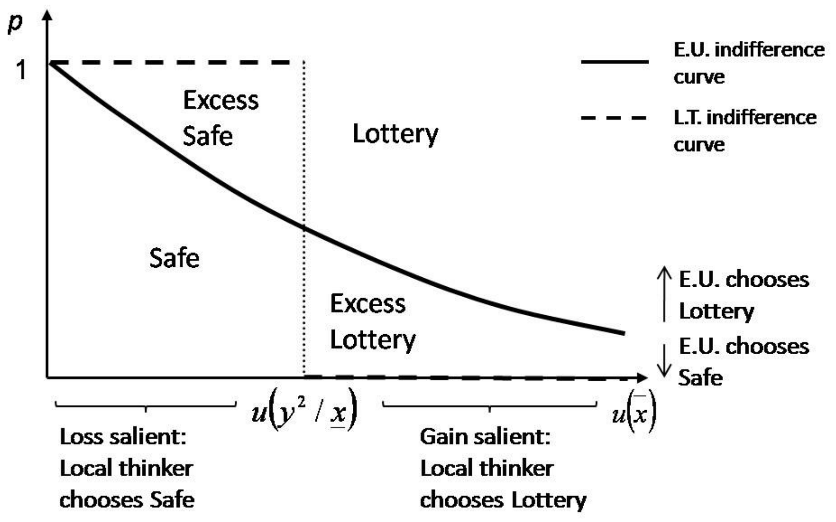
$$p^{LT} = \begin{cases} \frac{p\delta}{p\delta + (1-p)} & if \quad \overline{x} < \overline{x}^* \\ \frac{p}{p + \delta(1-p)} & if \quad \overline{x} > \overline{x}^* \end{cases}; \tag{8}$$

for simplicity, we leave aside the knife edge case $\overline{x} = \overline{x}^*$ (in which case $p^{LT} = p$). It is easy to see that, by Proposition 1, the over-weighting of state s_g at $\overline{x} > \overline{x}^*$ and its under-weighting at $\overline{x} < \overline{x}^*$ fall as the probability p of that state increases. Plugging p^{LT} into Equation (7) and rearranging terms, we find that the local thinker chooses the lottery over the sure prospect when:

$$L_1 \succ L_2 \Leftrightarrow \begin{cases} p > \widehat{p}_l^{LT}(\overline{x}) \equiv \frac{\widehat{p}(\overline{x})}{\widehat{p}(\overline{x}) + \delta[1 - \widehat{p}(\overline{x})]} & if \quad \overline{x} < \overline{x}^* \\ p > \widehat{p}_g^{LT}(\overline{x}) \equiv \frac{\delta \widehat{p}(\overline{x})}{\widehat{p}(\overline{x})\delta + [1 - \widehat{p}(\overline{x})]} & if \quad \overline{x} > \overline{x}^* \end{cases}; \tag{9}$$

where $\widehat{p}(\overline{x})$ is the expected utility maximizer's threshold from Equation (7), $\widehat{p}_l^{LT}(\overline{x})$ is the threshold for the case where the lottery's loss is salient, $\widehat{p}_g^{LT}(\overline{x})$ is the threshold for the case where the lottery's gain is salient. Note that $\widehat{p}_l^{LT}(\overline{x}) \geq \widehat{p}(\overline{x}) \geq \widehat{p}_g^{LT}(\overline{x})$. When the loss is salient, the local thinker rejects lotteries that the expected utility agent would take; the reverse is true when the gain is salient. To illustrate the role of salience in the starkest manner, Figure 2.a plots the expected utility maximizer's and the local thinker's choice as $\delta \rightarrow 0$:

Figure 1.a. Local Thinking vs. Expected Utility when $\delta \to 0$



The solid downward sloping curve separates the upper region in which the expected utility maximizer chooses the lottery from the lower region in which he chooses the sure prospect. The dashed curve separates these two regions for the local thinker. The curves (weakly) fall in $u(\overline{x})$: both the local thinker and the expected utility maximizer are more likely to choose the lottery as its upside increases.

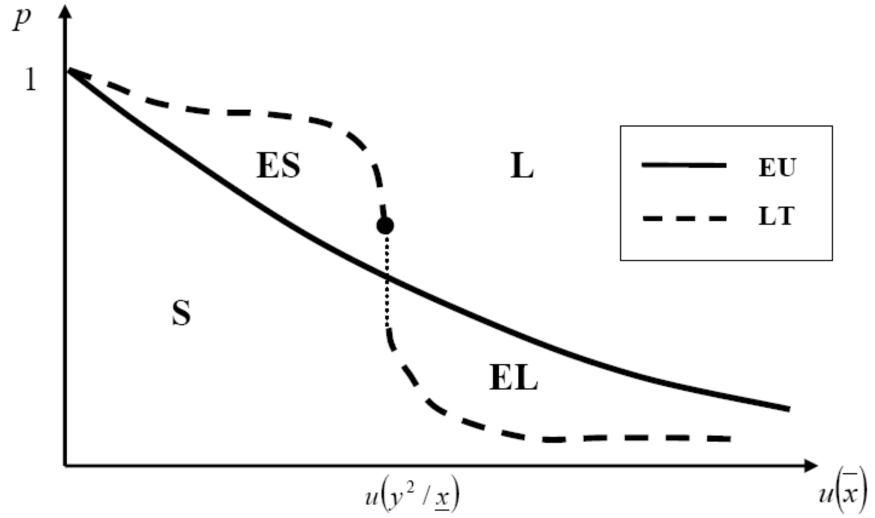
If the lottery's loss is salient, i.e. $u(\overline{x}) < u(\overline{x}^*)$, the local thinker chooses the sure prospect, if the gain is salient, i.e. $u(\overline{x}) > u(\overline{x}^*)$, he chooses the lottery. In both cases, the local thinker's choice is independent of the probability of the gain because when $\delta \to 0$ only the most salient state is considered regardless of p. In regions Safe and Lottery, these choices are the same as those of the expected utility maximizer. Departures from expected utility theory arise in regions Excess Safe and Excess Lottery. In the Excess Safe region, the local thinker displays "excess" risk aversion in the sense that he chooses the sure prospect when the expected utility maximizer would choose the lottery instead. In region Excess Lottery the local thinker displays excess risk seeking for the opposite reason.

When the loss is salient, excess risk aversion appears when the probability p of the gain is sufficiently high: for a low p not only the local thinker, but also the expected utility maximizer choose the sure prospect. When the gain is salient, excess risk taking appears when the probability of the gain is sufficiently low: when p is high not only the local thinker, but also the expected utility maximizer chooses the lottery. Departures from expected utility maximization thus occur when the agent focuses on a salient but relatively unlikely payoff, namely when the salience and probability of payoffs are far apart.

An increase in $u(\overline{x})$ that changes, via salience, the agent's focus from the lottery's loss to its gain can trigger a shift from Safe to Excess Lottery even if the lottery's expected utility is kept constant by reducing p (by moving in parallel slightly below the solid curve). This change, which is evidently irrelevant to an expected utility maximizer, triggers excess risk taking by the combined effect of Properties 1) and 2) in Proposition 1. Higher $u(\overline{x})$ boosts the salience of the lottery upside. This triggers an over-weighting of the upside that more than undoes the reduction in its objective probability p. Here the difference between the salience and probability of a payoff induces the local thinker to "over-react" to choice-irrelevant lottery changes.

When the extent of local thinking is not extreme, i.e. $\delta \in (0,1)$, the agent's choice pattern can be described by the figure below:

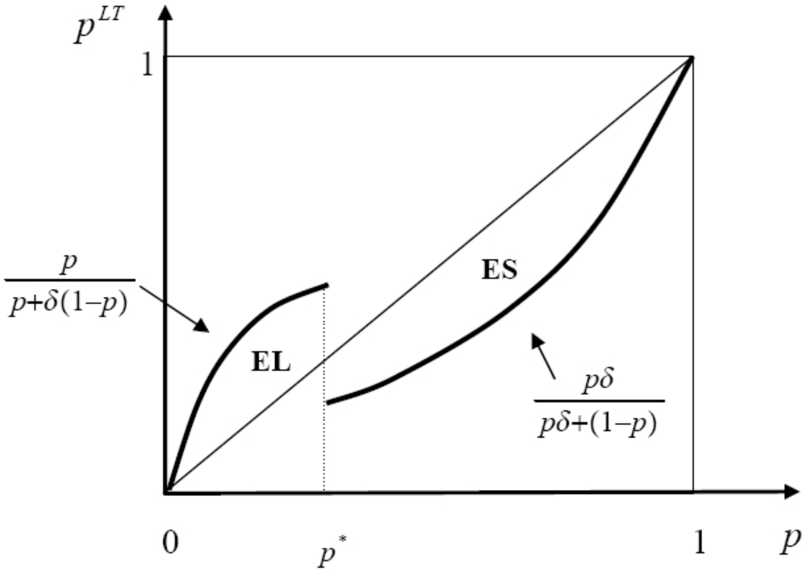
Figure 1.b. Local Thinking vs. Expected Utility for $\delta \in (0,1)$



The main qualitative properties of the case $\delta \to 0$ continue to hold, but the effect of salience is attenuated by the value of the actual probability of the lottery gain p, so that departures from expected utility are <u>not</u> most significant when p is either 1 or 0, but when p is moderately low/high. Generally speaking, however, departures from expected utility are stronger when the salient lottery outcome, the upside or downside, is relatively unlikely.

One implication of this logic is that our model can be used to recover a probability weighting function similar to Prospect Theory's inverse S-shaped π function (KT 1979). Several papers tried to measure π by revealed preference, including through determination of certainty equivalents of lotteries (Tversky and Kahneman 1992) or by more complex and robust methods (Wu and Gonzalez, 1996). For our purposes let L_1 be a mean preserving spread of L_2, that is $\overline{x} = \overline{x}(p) \equiv \underline{x} + (y - \underline{x})/p$. Intuitively, $\overline{x}(p)$ falls in p: for the lottery's expected value to stay constant at y as the probability of L_1's gain increases, the gain itself must decrease. Define the threshold p^* as the value of p such that $\overline{x}(p^*) = \overline{x}^*$. The agent then views the gain as being more salient than the loss provided $p < p^*$, while the reverse is true for $p > p^*$. A p is varied in this fashion, Equation (8) gives rise to the probability weighting function depicted in Figure 2:

Figure 2. Inverse S-shaped probability weighting function



This function captures the key features of the invese S-shaped function used by Kahneman and Tversky: over-weighting of low probabilities, under-weighting of high probabilities. In Figure 2 low probabilities are over-weighted because they occur in longshot lotteries that have a salient upside. High probabilities are under-weighted as they occur in lotteries with a small, non salient, upside. Just as in the Excess Safe and Excess Lottery regions of Figure 1.a, in ES and EL local thinkers respectively choose the lottery and the sure prospect, yielding the choice pattern observed in the experiments used to estimate the inverse S-shaped probability weighting function.

The inverse S-shaped probability weighting function is thus consistent with our result that departures from expected utility arise when the probability and salience of lottery outcomes are far apart. Crucially, however, in our approach such a weighting function does not capture subjects' universal perception of probabilities, but only their reaction to the salience of specific lottery payoffs. This difference is substantial. The most common estimates of the weighting function predict risk taking to occur in situations where the probability of a gain is lower than about 25-30%. Our model instead suggests that probability weighting may stifle risk taking even at small probabilities if gains are not salient and encourage risk taking at moderate probabilities and for moderate gains if the latter are salient relative to other lottery payoffs. This prediction, which we experimentally test in Section 4.2, indicates 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 the longshot bets typically stressed.

4.1 Reflection Effect

Kahneman and Tversky (1979) present some experiments showing that subjects' risk preferences shift from risk aversion to risk seeking as gains are reflected into losses. Our model delivers such shifts in risk attitudes without relying on the S-shaped value function of Prospect Theory. We in fact unify the shifts in risk taking typically attributed either to probability weighting or to the value function as the consequence of the changing salience of lottery payoffs.

Suppose that an agent has to choose between lottery $L_1 = (x_i, p_i)$ and the sure prospect $L_2 = (y, 1)$, both of which are defined in the gain domain (i.e. $x_i > 0$ for all i, $y > 0$) and where L_1 is a mean preserving spread of L_2, namely $E(x_i) = y$. Denote by i the state in which the lottery pays x_i. The local thinker then chooses L_2 over L_1 if and only if:

$$u(y) - \sum_i p_i u(x_i) > \text{cov}\left[u(x_i), \omega_i\right]. \tag{10}$$

Suppose now that lotteries L_1, L_2 are reflected into lotteries over losses $L'_1 = (-x_i, p_i)_i$ and $L'_2 = (-y, 1)$. Crucially, the reflection of payoffs does not alter the salience ranking of states. That is, if in the choice between L_1 and L_2 the salience ranking of a lottery payoff x_i is k_i, in the choice between L'_1 and L'_2 the salience ranking of state $-x_i$ is also k_i. Formally, this is due to the fact that for $x, y > 0$ the salience ranking of states is preserved under reflection. This implies that the agent chooses L'_2 over L'_1 when:

$$u(-y) - \sum_i p_i u(-x_i) > \text{cov}\left[u(-x_i), \omega_i\right] \tag{11}$$

with the same ω_i. To gauge the difference between Equations (10) and (11), suppose that in the gain domain lower-than average payoffs are more salient, i.e. $\text{cov}[u(x_i), \omega_i] < 0$. This relaxes inequality (10), creating a force toward risk aversion. On the other hand, the fact that $\text{cov}[u(x_i), \omega_i] < 0$ is likely to imply that in Equation (11) $\text{cov}[u(-x_i), \omega_i] > 0$, which means that as payoffs are reflected salience becomes a force toward risk seeking! Intuitively, since the salience ranking is invariant to reflection, when a low payoff is salient in the gain domain the reflected loss is salient under reflection. The former effect induces a preference for the sure prospect, the latter effect goes in the reverse direction. A symmetric argument holds when $\text{cov}[u(x_i), \omega_i] > 0$, in which case risk seeking for gains is reflected into risk aversion for losses.

Of course, the curvature of $u(\cdot)$ may prevent the sign of $\text{cov}[u(x_i), \omega_i]$ from switching under the reflection of payoffs, and more generally prevent a shift in risk attitudes to occur as a mere consequence of payoffs' reflection. The linear utility case nicely illustrates the force toward a reflection effect in our model:

Proposition 2 *Let L_1, L_2 be lotteries over gains, and L'_1, L'_2 be the reflected lotteries over losses. If $u(\cdot)$ is linear, the agent chooses L_2 over L_1 if and only if he chooses L'_1 over L'_2.*

In the linear case risk attitudes shift perfectly as gains are reflected into losses. To explain the "reflection effect", KT propose a value function that is: i) defined over lottery payoffs or generally increments relative to a reference point, and ii) concave for gains, convex for losses. As in Prospect Theory, in our model lottery payoffs play a crucial role because these payoffs, not final wealth levels, determine the salience of lottery outcomes. Unlike in Prospect Theory, though, the curvature of the utility function does not change and reflection affects risk attitudes by shaping the weighting of probabilities. We come back to this important difference in Section 5.1 when we discuss the experimental evidence on the framing effects created by reflection.[9]

4.2 Evidence on Shifts in Risk Attitudes

In this subsection, we explore and test experimentally the relationship between payoff salience and shifts in risk attitudes implied by our model (and briefly considered in Section 2). We consider choices in which the lottery has the same expected value as the sure prospect, i.e. $(\overline{x} - y) = (y - \underline{x})(1 - p)/p$. To highlight the role of payoff salience, we consider the case where $u(\cdot)$ is linear, so that risk attitudes are not due to the curvature of $u(\cdot)$. In this case, for a local thinker Equation (7) becomes:

$$p^{LT} \gtrless p, \tag{12}$$

so the agent is risk loving when the gain is more salient than the loss and risk averse otherwise. To further streamline the model, we consider the salience function of Equation (2). In this case, since the lottery's expected value is equal to y, we can calculate that the gain is more salient than the loss whenever:

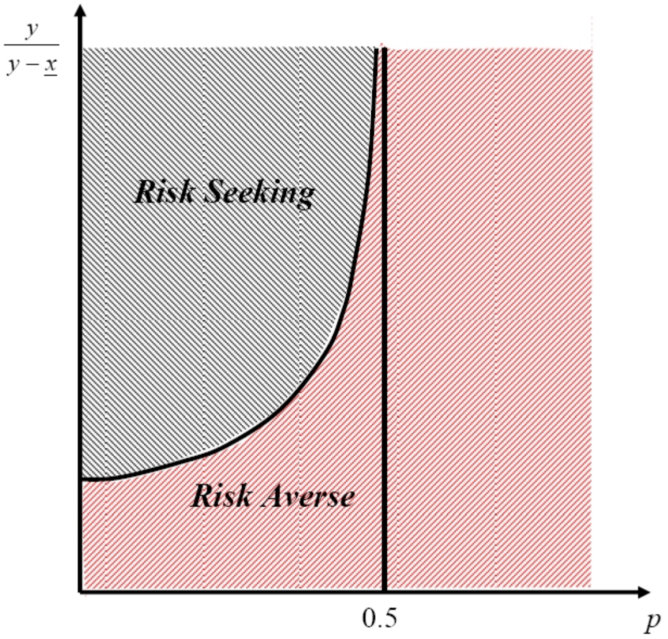
$$(y + \frac{\theta}{2})(1 - 2p) > (y - \underline{x})(1 - p), \tag{13}$$

---

[9] Note two distinctive implications of our model with respect to the occurrence of the reflection effect. First, reflection of payoff gains into losses does not necessarily increase preference for risk, and may sometimes reduce it, as when the agent is risk seeking for payoff gains (see the example of Section 3.2.2). Relatedly, if the agent's utility $u(\cdot)$ is concave, reflection of payoffs does not necessarily reverse risk preferences as in the linear case of Proposition 2; the shift from risk taking to risk aversion may occur at different points in the loss and gain domains. In this sense, the principle that larger reductions of wealth are relatively more painful – a cornerstone of neoclassical utility theory – constitutes in our model a natural and independent force against reflection.

which uniquely identifies the parameter values for which the agent is risk seeking. Note that the condition can only be met when $p < 1/2$, for in this case the lottery gain is larger than the loss, which is a necessary condition for the gain to be more salient. In fact, as Equation (12) shows, in the linear case it is the presence of overweighting,not its magnitude that determines risk preferences. The risk attitudes implied by Equation (13) are graphed below:

Figure 3. Shifts in risk attitudes

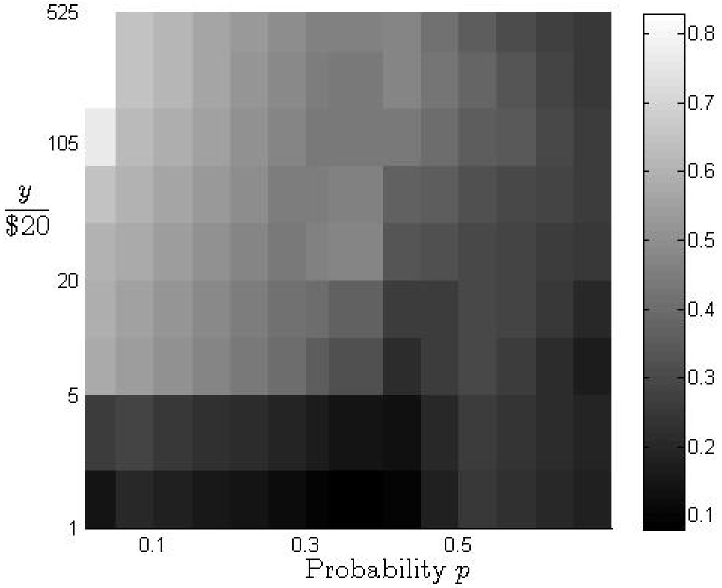


In Figure 3, two patterns stand out. First, as in Section 2, for a fixed probability $p < 1/2$ an increase in the lottery's expected value y that keeps its absolute loss $y - \underline{x}$ constant eventually induces risk seeking. As all payoffs become large, absolute payoff differences become crucial for determining salience, favoring risk-taking because at $p < 1/2$ the lottery's absolute gain is larger than its loss. This implication, which follows from the convexity property in Definition 1, implies that risk seeking can occur also for moderate probabilities provided y is sufficently high.

Second, for a given expected value y, as p increases towards $1/2$ the difference between the gain $\overline{x} - y$ and the loss $y - \underline{x}$ falls, rendering the gain less salient. In this example, a higher p reduces risk seeking by affecting salience, not over-weighting. Risk taking does not occur for $p \geq 1/2$ because when the gain is smaller than the loss the latter is always more salient.[10]

To test these predictions, we gave experimental subjects a series of binary choices between a lottery $L_1(p) = (\overline{x}, p; \underline{x}, 1-p)$ and a sure prospect $L_2 = (y, 1)$. In all experiments, L_1 is a mean preserving spread of L_2, i.e. $p(\overline{x} - y) = (1-p)(y - \underline{x})$; we set the downside $(y - \underline{x})$ of the lottery at \$20, yielding an upside $(\overline{x} - y)$ of $\$20 \cdot (1-p)/p$. We varied the expected value y in $\{\$20, \$100, \$400, \$2100, \$10500\}$ and the probability p in $\{.01, .05, .2, .33, .4, .5, .67\}$. For each of these 35 points, 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 Figure 4

Figure 4. Proportion Choosing $L_1(p)$



---

[10] Although Figure 3 is obtained under the salience function of Equation (2), its main qualitative features are guaranteed by the 4 properties of Definition 1. In particular, defining $\overline{x}(p, y)$ as the value of the lottery's upside at which the lottery has the same expected value of the sure prospect, Equation (12) holds and the local thinker becomes risk seeking provided $\sigma(\overline{x}(p, y), y) > \sigma(\underline{x}, y)$. 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 more salient and the agent is risk averse. This is surely the case for $p \geq 1/2$. The convexity property instead implies that as y increases then, for fixed $p < 1/2$, the uspide eventually becomes more salient, generating the qualitative patterns of Figure 3.

The axes are the same as those in Figure 3: the x-axis shows the probability of the lottery gain, and the y-axis shows the ratio of that gain to $20. The patterns are qualitatively consistent with our predictions. For a given expected value y, the proportion of risk takers falls as the probability p of the gain increases. At the same time, for a given probability p of the gain, the proportion of risk takers increases with the expected value y. The effect is very significant: at $p = 0.05$ a large majority of subjects (80%) were risk averse when $y = \$20$, but as y increases to $2100 a large majority (65%) becomes risk seeking. Finally, and again consistent with our prediction, Figure 4 shows 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 40%, we observe a 5-fold rise in risky choices as y increases. These patterns are broadly consistent with the predictions of our model.

To compare our predictions with those of Prospect Theory, in the Appendix we calibrate standard functional forms for the value function and probability weighting function and simulate the above experiments using a single-agent stochastic choice model (Camerer and Ho 1994). Under the most commonly used specification of the value function $v(x) = x^{\alpha}$ with $\alpha \sim 0.88$, Prospect Theory predicts strong risk seeking for $p \leq 0.2$ and much more risk averse behavior for $p > 0.2$. Given the agent's mild risk aversion, preferences toward risk are driven by the inverse S-shaped probability weighting function, inducing risk taking if and only if p is small. These predictions do not account for subjects' conservative behavior at low expected values y even if p is low and more generally for the strong increase in risk taking occurring as the lottery's expected value y goes up (for fixed $y - \underline{x}$) even for moderate probabilities $p > 0.2$. Our model can parsimoniously reconcile these apparently contradictory patterns of risk taking by highlighting the impact of the lottery's expected value on the salience of its upside and thus on probability weighting.

5 Local Thinking and Experimental Anomalies

Experimental evidence has identified some robust forms of "invariance failures," namely preference reversals occurring across lottery choice problems that would be perceived as equivalent by an expected utility maximizer. We now show how our model can provide a unified explanation of these "anomalies" by stressing how the equivalence among these probems is broken down by the changing salience of lottery payoffs.

5.1 Framing of Lottery Outcomes

KT (1981) exploited the reflection effect to show how choice among lotteries can be affected by changing the presentation of lottery outcomes. Some experimental 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 over the sure gain. 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 over the sure loss. 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. Prospect Theory views this framing effect as the combination of two forces. First, the agent automatically adjusts his reference point to include the initial lump-sum payoff. Second, the concavity of the value function for gains and its convexity for losses leads to risk aversion in L_1 vs. L_2 but to risk seeking in L_1' vs. L_2'.

Our explanation for this framing effect does not depend on the curvature of the value function but emphasizes the idea that reflecting payoffs from gains to losses affects probability weighting. This is because salience in our model depends on the payoffs of the alternative lotteries as they are presented to the agent, and not on wealth changes that the agent may experience regardless of the lottery chosen. As a result, a reduction in lottery payoffs accompanied by a contextual increase in the agent's baseline wealth affects the salience of lottery outcomes and choice even if the change is neutral in terms of final wealth states.

To see how this works, note that in choosing between L_1 and L_2, the lottery's downside is more salient than its upside because by diminishing sensitivity $\sigma(0, 500) > \sigma(1000, 500)$. As a consequence, the zero outcome is over-weighted, implying that an agent with initial wealth w and linear utility chooses L_2 over L_1 if and only if:

$$(w + 1000 + 500) > \frac{1}{1+\delta}(w + 1000 + 0) + \frac{\delta}{1+\delta}(w + 1000 + 1000). \tag{14}$$

Note that the agent integrates the lump sum amount of 1000 into his initial wealth. The above condition is satisfied for any $\delta < 1$ and thus the agent is averse to risk because the lottery and the sure prospect have the same expected value.

In the choice between L_1' and L_2', the most salient lottery payoff is still the one where the lottery pays zero because, by diminishing sensitivity in the loss domain, $\sigma(0, -500) > \sigma(-1000, -500)$. The zero outcome is thus over-weighted in this case as well, implying that the agent chooses L_1' over L_2' if and only if:

$$(w + 2000 - 500) < \frac{\delta}{1+\delta}(w + 2000 - 1000) + \frac{1}{1+\delta}(w + 2000 - 0). \tag{15}$$

Once more, the initial amount of 2000 sheckels is integrated into the agent's wealth. Now the focus on zero induces the agent to over-weight the gain of lottery L_1', namely the zero loss state. As a consequence, for any $\delta < 1$ the agent is risk seeking and chooses L_1' over L_2', reproducing the reversal documented in experiments.

Presenting the lottery as payoff gains rather than losses boosts the salience of the highest lottery outcomes, inducing a shift from risk aversion to risk seeking. This effect is due to the impact of lottery payoffs on salience and does not require the agent to change his reference consumption level as a function of the initial lump sum payment.

The same intuition can account for KT's (1981) famous Public Health Dilemma, which describes the outbreak of a disease that is expected to kill 600 people. One group of subjects was asked to choose between a program saving 200 people for sure and a program saving 600 people with probability 1/3 and none with probability 2/3. A second group of subjects was asked to choose between a program under which 400 people die for sure and a program under which 0 people die with probability 1/3 and 600 die with probability 2/3. In the former version of the problem, 75% of respondents chose the program with a sure outcome; in the latter version, only 22% of respondents chose the program with a sure outcome. Our model can account for this shift in risk preferences by simply noting that when the choice is framed in terms of lives saved, the local thinker specifies the payoffs as gains, when the choice is framed in terms of lives lost, the local thinker specifies the payoffs as losses. In other words, the change in language from people saved to people lost plays here the same role as the initial monetary payoff in the previous example: it allows to reflect the consequences of the health programs. Because salience is shaped by the magnitude of those consequences rather than by the final states, in the "saved" frame the most salient outcome is the one where nobody is saved and in the "die" frame the most salient outcome is the one where nobody dies, triggering preference reversal.

In the remainder of this Section, we show how this role of payoffs in shaping probability weighting can unify the explanation of framing effects arising from the manipulation of lottery outcomes with the explanation of other failures of invariance.

5.2 Local Thinking and the 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:

$$L_1^z = (2500,\ 0.33;\quad 0,\ 0.01;\quad z,\ 0.66),\quad L_2^z = (2400,\ 0.34;\quad z,\ 0.66) \tag{16}$$

for different values of the payoff z. By the independence axiom, an expected utility maximizer should not change his choice with the "common consequence" z, which cancels out in the comparison between L_1^z and L_2^z. In reality, experiments run for the values $z = 2400$ and $z = 0$ reveal that $L_2^{2400} = (2400,\ 1)$ is preferred to $L_1^{2400} = (2500,\ 0.33;\ \ 0,\ 0.01;\ \ 2400,\ 0.66)$ while $L_1^0 = (2500,\ 0.33;\ \ 0,\ 0.67)$ is preferred to $L_2^0 = (2400,\ 0.34;\ \ 0, 0.66)$. In violation of the independence axiom, z affects the experimental subjects' choices.

KT (1979) explain the switch from L_1^0 to L_2^{2400} by assuming subcertainty of the probability weighting function, $\pi(0.66) + \pi(0.34) < 1$. As a result, the common consequence $z = 2400$ is under-weighted in lottery L_1^{2400} relative to L_2^{2400}, favoring the latter choice (the choice of the riskier L_1^0 over L_2^0 only requires risk aversion to be small)[11].

---

[11] In Cumulative Prospect Theory (Tversky and Kahneman, 1992), probability weights sum up to one.

Our theory does not feature subcertainty, but explains the Allais paradox by arguing that the common consequence z alters the salience and weighting of lottery outcomes. To see this, consider the choice between L_1^{2400} and L_2^{2400}. In this case the most salient state is the one where L_1^{2400} pays zero, and the riskless L_2^{2400} lottery pays 2400, because:

$$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400). \tag{17}$$

Since the local thinker then over-weights the probability of obtaining zero in L_1^{2400}, for a sufficiently low δ he prefers L_2^{2400} to L_1^{2400}. The risky lottery L_1^{2400} is not chosen 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, most important, both lotteries have the same downside of 0. As a consequence, the most salient state is determined by the lotteries' upsides because:

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

The agent over-weights 2500 in L_1^0 relative to 2400 in L_2^0, so that for δ small the agent chooses the riskier lottery L_1^0.

In sum, as we formally prove in the Appendix, for low δ the agent exhibits the Allais paradox. The independence axiom is violated here because the common consequence z shapes the salience of lottery outcomes. When $z = 2400$ lottery L_2^z is perfecly safe, making the zero outcome "stand out" in L_1^z. When $z = 0$ lottery L_2^z is risky, making the 2500 outcome "stand out" in L_1^z. This marked shift in the salience of the downside versus the upside of lottery L_1^z explains the reversal in choices as z changes.

Another version of the Allais paradox involves the "common ratio" effect, which occurs in the choice between lotteries:

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

However, the paradoxical choice pattern is driven, as in Prospect Theory, by the fact that the common consequence is more valuable when it is associated with a sure prospect than when it associated with a risky prospect.

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:

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

Crucially, 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 value of p. Parameter α here identifies the "common ratio" between probabilities q and p.

In contrast to 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 $L_2^{0.002} = (3000,\ 0.002;\ \ 0,\ 0.998)$. KT (1979) explain this evidence by assuming that the probability weighting function grows slower than linearly for small probabilities; hence, $\pi(\alpha p)/\pi(p)$ is relatively high at low p, inducing the choice of L_1^q when $p = 0.002$.

In our theory, the common ratio effect arises because reductions in p endogenously increase the agent's focus on the salient upside of lottery L_1^q. To see this, note that the salience ordering among lottery states is $s_1 = (6000, 0)$, $s_2 = (0, \alpha \cdot 6000)$, $s_3 = (6000, \alpha \cdot 6000)$, $s_4 = (0, 0)$. With a common ratio $q/p = \alpha$, this implies that the local thinker evaluates the odds with which the risky lottery L_1^q pays out relative to the safe one L_2^p as:

$$\frac{q^{LT}}{p^{LT}} = \alpha \cdot \frac{(1-p) + p\delta^2}{(1-\alpha p)\delta + \alpha p \delta^2}. \tag{21}$$

The agent's choice then follows the criterion in Equation (20) with the only difference that the local thinker replaces the true ratio $q/p = \alpha$ with the distorted ratio of Equation (21).

For $\delta < 1$ the odds ratio q^{LT}/p^{LT} in Equation (21) is over-estimated when p is low, inducing agents, for given α, to choose the riskier lottery L_1^q, just as the experimental evidence indicates. In particular, q^{LT}/p^{LT} tends to α/δ as $p \to 0$, which can be very large if the extent of local thinking is severe. At small p, both lotteries pay out with low probability and the agent becomes increasingly focused on the most salient upside of the riskier lottery. As in Proposition 1, when p falls the probability of salient events is more over-weighted, triggering the observed shift towards risk taking.[12]

5.2.1 Sensitivity to Correlations and the Independence Axiom

The general theme of the previous analysis is that in our model violations of the independence axiom are due to the impact of the lotteries' state space on salience. By changing the state space, the common consequence z alters the salience ranking of lottery outcomes. In the common ratio effect, reductions in p boost the impact of the salience ranking embodied in the state space. To further examine the role of the state space, we now introduce a correlation among the above lotteries. Correlation changes the state space *but not* a lottery's distribution over final outcomes, so it may affect choice under local thinking but not under Expected Utility Theory or Prospect Theory.

Consider first the following well known correlated version of the common ratio effect (KT 1979). Subjects are given a choice problem 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, depending on what the subjects chose initially, either the lottery L_1^q or L_2^p from Equation (19) is played where $p = 1$ and $\alpha = 0.5$. This composite gamble corresponds to a choice between L_1^q and L_2^p, with $p = 0.25$ and $\alpha = 0.5$. When the lotteries are correlated individuals tend to choose the safer lottery L_2^p just as when, in single stage gambles, the probability that such lottery pays off is large, i.e. $p = 1$. When instead people are directly presented with the single stage gambles L_1^q and L_2^p with $p = 1/4$, they tend to choose L_1^q, just as in the low probability case of the common ratio effect. In short, when the probability with which both gambles pay out is reduced by adding a correlated state in which both lotteries pay zero, the common ratio effect disappears.

In explaining this anomalous behavior, KT informally argue that individuals "edit out" the common consequence when the two lotteries are presented as correlated, thereby choosing

---

[12] When p is large, the choice of the safer lottery L_2^p can be explained by the concavity of $u(\cdot)$. Concavity of $u(\cdot)$ is however not necessary for the agent to choose L_2^p for large p. As $p \to 1$ the local thinker's odds ratio is smaller than the common ratio α, leading even a risk neutral agent to choose L_2^p. The intuition is that as $p \to 1$ the salient state s_1 in which L_1^q wins and L_2^p loses is highly unlikely to occur. Thus, lottery evaluation is shaped by the more likely state s_2 in which L_1^q loses and L_2^p wins, inducing a choice of the latter.

as if $p = 1$. Our model yields such editing out as a consequence of the low salience of the correlated state in which both lotteries pay 0. To see this formally, note that for $\alpha = 0.5$ the state space of the composite gamble is as follows:

salience ranking	1	2	3	4	4
payoff of L_1^q	6000	0	6000	0	0
payoff of L_2^p	0	3000	3000	0	0

In the rightmost state the gamble ends and both lotteries pay zero; the first four states are associated with the continuation of the gamble and thus with realized lottery payoffs. The salience of the state in which the gamble ends is the same as the salience of the state in which both lotteries are played and pay zero. As a consequence, the probabilities of these states must be weighted equally.

It is then easy to see that for any probability $r < 1$ of the game ending, the odds ratio q^{LT}/p^{LT} assessed by the agent in this composite gamble is equal to that of Equation (21). In other words, in the composite gamble the local thinker disregards the correlated state and its probability r, choosing between the second stage lotteries L_1^q and L_2^p as if these were single stage lotteries. This is what experimental subjects do. Our model reconciles the subjects' choice in single and two stage lotteries: in the single stage case, the rescaling of probabilities matters because it boosts the distortions caused by the given salience ranking of states; in the two stage case, the rescaling adds a new correlated state which is not salient. Rescaling does not matter now because agents disregard the new state when making their choices.

One way to view this result is that the introduction of the correlated state induces subjects to notice that rescaling of p changes the probability with which the lotteries pay zero in a *neutral* way, leading them to appreciate the normative prescription of the expected utility model. The role of correlation is even starker in the common consequence effect. Consider the following correlated version of the lotteries L_1^z and L_2^z in Equation (16):

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

,

This version makes clear that the two lotteries pay the common consequence z in the same state. For a local thinker, not only is the state in which both lotteries pay z the least salient one, but the value of z does not affect the choice at all because states where the lotteries pay the same amount drop from Equation (3). As a result, our model – but not Prospect Theory – implies that the Allais paradox should not occur when L_1^z and L_2^z are presented in the correlated form as above.

We tested experimentally this prediction by presenting experimental subjects lotteries L_1^z and L_2^z in the above correlation structure 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. 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.[13]

---

[13] To further probe into the possibility for our model to jointly explain this evidence on the role of correlation and the basic violation of the Allais paradox, we ran an experiment where subjects were explicitly presented the lotteries of Equation (16) with $z = 2400$ as uncorrelated, with a state space consisting of the four possible states. As the Appendix shows, 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).

One way to summarize this evidence is that local thinkers want to act according to the independence axiom (as many experimental subjects indeed profess to do), but are only able to do so when its application is made evident by the correlation structure. When this is so, the Allais paradox disappears. The Appendix provides additional evidence on correlation.

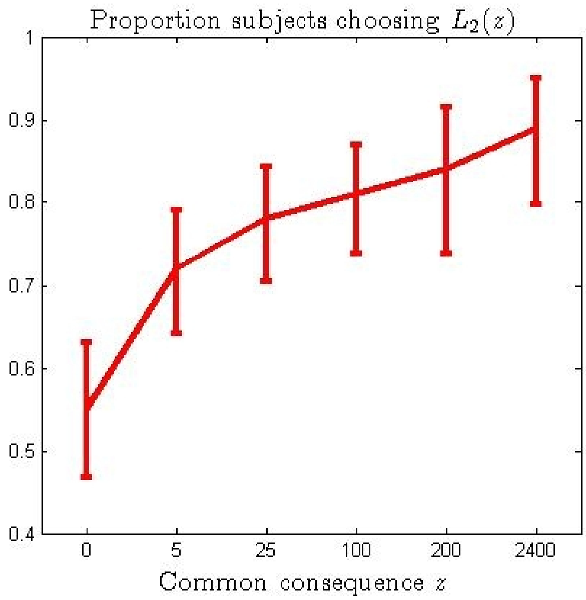
5.2.2 Payoffs, Salience and the Common Consequence Effect

In this section, we use the Allais paradox to illustrate how in our model lottery payoffs exerts a stark effect on choice by shaping the salience of lottery outcomes. Consider once more the "common consequence lotteries" L_1^z and L_2^z of Equation (16). Here we explained the shift toward risk taking occurring when z is reduced from 2400 to 0 as being triggered by the fact that at the specific value $z = 0$ the lotteries have the same downside. This implies that the upside 2500 of the risky lottery L_1^0 becomes salient, enhancing risk taking in our theory.

This intuition suggests that the shift toward risk taking may not occur even if z is reduced a lot below 2400 but not all the way down to 0. In fact, at any $z > 0$ the risky lottery L_2^z has still a lower downside than L_1^z, which stifles risk taking.

To test this possibility, we let subjects choose between L_1^z and L_2^z for $z =$ \$0, \$5, \$25, \$100, \$200, and \$2400. As expected, as soon as $z > 0$ as many as 70% of subects still prefer the safe lottery L_2^z even if z is as low as \$5. Interestingly, this boost in the preference for the safer lottery L_2^z as z increase from 0 to \$5 comes from subjects who had chosen L_1^z when $z = 0$. The pattern of responses for $z = 100$ already approaches that reported by Kahneman and Tversky for $z = 1000$. Figure 4 below illustrates these findings.

Figure 4. Proportion choosing L_2^z



The error bars represent a 95% confidence interval of the probability of choosing L_2^z. Given a value of $z = 5$, a Prospect Theory agent chooses the safer lottery L_2^z when:

$$L_2^z - L_1^z = V(5)[1 - \pi(.66) - \pi(.34)] - \pi(.33)V(2500) + \pi(.34)V(2400) \geq 0 \qquad (22)$$

In the appendix we evaluate the above expression by calibrating standard functional forms and using a single-agent stochastic choice model (Camerer and Ho 1994). The key parameter determining whether condition (22) holds is the curvature of $V(\cdot) = z^\alpha$. At the commonly used value of $\alpha \sim 0.88$, condition (22) does not hold. Intuitively, in this case the agent is nearly risk neutral so increasing z from 0 to 5 has very little impact on the attractiveness of the safer lottery (recall that the riskier lottery is chosen for $z = 0$).

Our model stresses instead that the strong sensitivity of subjects' choice to a change from $z = 0$ to $z = 5$ may be due not to the curvature of the utility function but rather to the impact of z on salience. If $\sigma(0, 2400) > \sigma(5, 2500)$, then also for $z = 5$ the local thinker continues to focus on the lower downside of the riskier lottery L_1^z, triggering the choice of the safer lottery L_2^z.[14]

---

[14] Using the salience function of Equation (2), condition $\sigma(0, 2400) > \sigma(5, 2500)$ is satisfied if θ is sufficiently

Overall, this section suggests that our focus on the salience of lottery states does not only account for the Allais paradoxes by endogenizing probability weighting but is also consistent with new experimental results that test the sensitivity of these paradoxes to changes in the presentation of lotteries. Such changes would be of little relevance in models in which the weighting of probabilities is exogenously given and does not depend on lottery states.

6 Further predictions

6.1 Mixed Lotteries

We now analyze mixed lotteries, namely those which involve both positive and negative payoffs. To do so, we extend the notion of salience to mixed states of the world. The ordering property of salience $\sigma(x, y)$ naturally extends to mixed states. A tension potentially arises between diminishing sensitivity and reflection. To see this, note that diminishing sensitivity and reflection imply that the salience of state $(z, z+\epsilon)$ where $\epsilon > 0$, decreases in z when $z > 0$ but increases in z when $z < -\epsilon$. In the former case, the state $(z, z+\varepsilon)$ is positive and by diminishing sensitivity σ increases as z falls to zero. In the latter case, $(z, z+\varepsilon)$ is negative and by diminishing sensitivity σ increases as z increases to zero. In the former case, state $(z, z+\epsilon)$ is positive and diminishing sensitivity implies that $\sigma(z, z+\epsilon)$ falls as as z increases because the state gets further away from zero. In the latter case, state $(z, z+\epsilon)$ is negative and diminishing sensitivity implies that an increase in z now raises $\sigma(z, z+\epsilon)$ because it moves the state closer to zero. In other words, diminishing sensitivity in the positive and negative domain implies that the effect of an increase in z on salience must change sign when $z \in [-\epsilon, 0]$. This is precisely the case when the state is mixed. We account for this switch in sign in the mixed domain by extending Definition 1 as follows:

Definition 3 ***Definition 4*** *The salience of a generic state of the world $s = (x, y)$ is measured by a continuous and symmetric salience function $\sigma(x, y)$ that satisfies four conditions:* 1) *Ordering: if $[x', y']$ is a subset of $[x, y]$ then $\sigma(x, y) > \sigma(x', y')$.*

---

small. Indeed, the lower is θ the more salient are states where one payoff is zero. In Section 6.2 we back up a value of θ from experimental evidence on longshot lotteries and verify that at such value condition $\sigma(0, 2400) > \sigma(5, 2500)$ is indeed satisfied.

2) *Diminishing sensitivity: for any* x, y *such that* $x > y$ *and for any* $\epsilon > 0$ *such that* $x + y - \epsilon > 0$, $\sigma(x + \epsilon, y) < \sigma(x, y - \epsilon)$.

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) *Loss salience: for any* $x \geq 0$, *if* $y > 0$ *then* $\sigma(-x - y, x) > \sigma(-x, x + y)$

5) *Convexity: for every* $\epsilon > 0$, $\lim\limits_{x \to +\infty} \frac{\sigma(x+\epsilon, x)}{\sigma(x, x-\epsilon)} = 1$.

The convexity property 5) does not change from Definition 1 because it only applies to states that are non-mixed. Property 2) extends diminishing sensitivity to states whose average payoff is non-negative, which we call *positive mixed states.* One implication of property 2) is that, in the domain of positive mixed states and for a given distance between payoffs, maximum salience is attained by "spreading" such distance symmetrically around zero. Formally, in the class of positive mixed states with payoff distance $|z - w|$, maximum salience is is attained in the limit at state $(|z - w|/2, -|z - w|/2)$. This property is reasonable because it implies that just as a given payoff distance in the positive domain is more salient the closer its average payoff is to 0, the same holds for mixed states whose payoffs lie around zero.

Property 3) extends reflection to positive mixed states. Together with property 2), it yields diminishing sensitivity for *negative mixed states* whose average payoff is negative. In other words, in the domain of states with negative average payoff the salience of states with a given distance between payoffs is larger the closer the payoffs are to zero and is highest for the state whose payoffs are symmetric around zero.

In sum, properties 2) and 3) extend quite naturally to mixed lotteries the idea of diminishing sensitivity, namely that the salience of a given payoff difference increases as the absolute level of payoffs becomes closer to zero. Once more, this captures the crucial fact that salience depends on the magnitude of payoffs, and not on their signs.

This discussion leaves unspecified whether the salience of a state with given distance between payoffs is higher when that state is positive or negative mixed. Property 4) above addresses this issue by ranking the impact of increases or decreases in the average payoff around zero. In particular, it states that increasing losses in absolute terms leads to a larger increase in salience that increasing the gains. The intuition for this condition is that ceteris paribus losses tend to be more salient than gains, and one of its implications is to reproduce some of Kahneman and Tversky's (1979) experimental findings on loss aversion.

Condition 4) is however weaker than loss aversion as it does not prevent a local thinker with linear utility from sometimes taking risk around zero. To see this, suppose that a local thinker with linear utility chooses between a mixed lottery $L_1 = (x, 1/2; -y, 1/2)$, where $-y < 0 < x$, and the sure prospect $L_2 = (z, 1)$ with $z = (x - y)/2$. As the local thinker trades off the state where the lottery loses $s_l = (-y, z)$ against the state where the lottery gains $s_g = (z, x)$, he is swayed by the salience of each. Definition 3 then implies that the local thinker's choice depends on the values of x and y.

Proposition 3 *If the mixed lottery L_2 has a non-negative expected value $(x \geq y)$, a risk neutral local thinker displays risk aversion $(L_2 \succ L_1)$. If L_2 has a sufficiently negative expected value $(0 < x < y/3)$ a risk neutral local thinker displays risk seeking behavior $(L_2 \prec L_1)$. The local thinker switches preferences within the range $x \in (y/3, y)$.*

In other words, the local thinker is risk averse and thus rejects the bet when the sure prospect z is positive, whereas he is risk seeking and thus takes the bet when the sure prospect z is (sufficiently) negative. Mere loss aversion would instead imply that for sufficiently small stakes the agent will be risk averse regardless of the sign of the sure prospect z. The key point here is that our model generates this risk seeking because a negative z renders the (mixed) state in which the lottery gains more salient, thereby inducing risk seeking. This is a consequence of the diminishing sensitivity property, whereby mixed states tend to be more salient because they are characterized by a lower average payoff. A positive z then triggers risk averse behavior by focusing the agent's attention on the (mixed) state where the lottery loses. This effect highlights the role of context dependent evaluation underlying our model: a sure loss makes subjects focus on the lottery gain, whereas a sure gain makes subjects focus on the lottery loss. This basic intuition then generalizes to the case of generic mixed lotteries.[15]

---

[15] In particular, a local thinker is risk averse regarding mixed lotteries with zero expected value, $x = y$. The analysis is readily extended to choices between two mixed lotteries with zero expected value. Let $L_x = (-x, 1/2; x, 1/2)$; then for a risk neutral local thinker, $L_x \succ L_y$ if and only if $x < y$. In fact, in this case the two most salient states are $s = (-y, x)$ and $s' = (y, -x)$. These states are obtained from the symmetric state $s' = (-x, x)$ by subtracting $y - x$ to the downside or adding it to the upside respectively; if follows

Note incidentally that to the extent that aversion to mixed lotteries with positive expected values follows from salience, it does not constrain the utility function. In particular, a local thinker with linear utility may be averse to mixed lotteries with positive expected value at all levels of wealth. Even though a local thinker is at heart an expected utility maximizer, he is immune to Rabin's critique (Rabin, 2000).

To test the prediction of risk seeking behavior regarding mixed lotteries when compared to a sure loss, we let subjects choose between $L_1^0 = (-\$40, \frac{1}{2}; \$40, \frac{1}{2})$ and $L_2^0 = (0, 1)$ and also between $L_1^{-20} = (-\$60, \frac{1}{2}; \$20, \frac{1}{2})$ and $L_2^{-20} = (-\$20, 1)$. The first problem involves a symmetric mixed lottery and the agent chooses it against a sure prospect of zero. In the second problem the agent chooses between a lottery that loses 20 on average and a sure loss of -20. The proportion of subjects choosing the two prospects in the two problems was:

	L_1^{-20}	L_2^{-20}
L_1^0	19%	12%
L_2^0	37%	32%

As predicted, subjects were mostly risk averse for Problem 1 (69%) and mostly risk seeking for Problem 1' (58%). Moreover, a significant proportion of subjects who were risk averse in Problem 1 became risk seeking in Problem 1' (37%).[16]

6.2 Long-Shot Lotteries

Long shots lotteries refer to lottery tickets in the everyday meaning of the term, namely a lottery with a low expected value and a minuscule probability of a very large gain. The theoretical challenge raised by these lotteries is that people demand these lotteries when they have very large payoffs but small expected values (e.g. jackpots in the millions of dollars and per ticket cost of \$10 to \$20). For instance, in Prospect Theory if agents take the longshot lottery at a low expected value, they should do so even if the expected value is high (modulo risk aversion, see Appendix).

---

from the property of loss salience that $\sigma(-y, x) > \sigma(y, -x)$. Therefore, a local thinker focuses on the state where L_1 loses and thus becomes risk averse.

[16] In Appendix 2C, we find that according to Kahneman and Tversky's callibration, a Prospect Theory agent would need a risk aversion coefficient of 0.9 to make this switch. We find this value to be too high to be plausible.

To see the implications of local thinking, consider the choice between the following lotteries:

$$L = (y/p, p; 0, 1-p) \ \ vs \ \ S = (y, 1),$$

where y is the lottery's expected value.[17]

In chosing between L and S, the lottery gain is more salient than the loss and the agent therefore chooses L when $\sigma(y/p, y) > \sigma(y, 0)$, which using the salience function in Equation (2) boils down to:

$$y < \theta \frac{1-2p}{2p}. \tag{23}$$

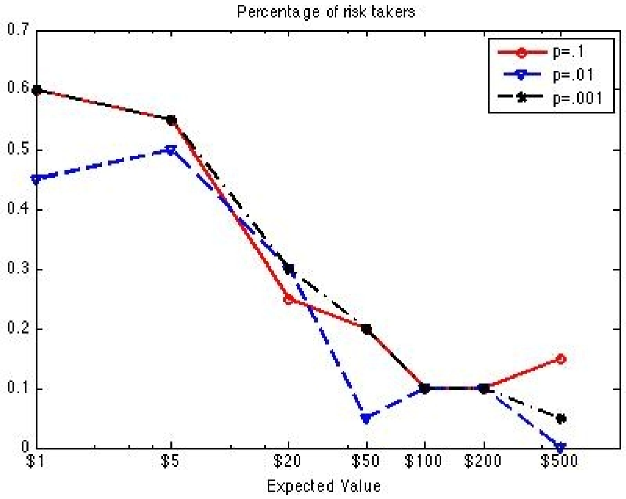
The local thinker chooses L when the lottery's expected value is sufficiently small. Intuitively, as the lottery's expected value (and thus cost) increases, the state of losing the lottery becomes more and more salient, discouraging risk taking. Note the fundamental difference from Section 4: in that context, an increase in the lottery's expected value fostered risk taking because there the loss under the lottery $y - \underline{\mathbf{x}}$ was held constant. Here instead the loss under the lottery increases in its expected value, which explains why higher y discourages risk taking. In this sense, our model places a structure on the demand for longshot lotteries that – enriched with a supply side – can help shed light the market for longshot lotteries in the real world.

To test this prediction, we conducted an online survey where subjects were asked to choose between L and S for expected values $y \in \{1, 5, 20, 50, 100, 200, 500\}$ and probabilities $p \in \{10^{-3}, 10^{-2}, 10^{-1}\}$. The proportion of risk seeking choices is shown in Figure 5.

---

[17] This can be viewed as the choice between buying a fair lottery ticket and saving (and thus gaining) the ticket price y for sure. To capture more accurately the choice faced in the context of state lotteries, the same choice can be modelled as occurring between lotteries $L' = (x, p; -y, 1-p)$ and $S' = (0, 1)$. In this formulation y is the cost of the lottery ticket and S' is the option of doing nothing. Nothing substantial changes under this alternative formulation.

Figure 5. Proportion taking the longshot lottery



This finding confirms our prediction that subjects take longshot lotteries only at low expected values: risk seeking has already dropped dramatically at $y = 20$. We can use this finding to back out a value for θ. In our experiments subjects transition from being risk seeking to being risk averse when the lottery cost is around $y \sim \$15$ and the prize is as low as $y/p \sim \$1500$. If the agent is risk neutral, by plugging these numbers into Equation (23) we obtain $\theta \sim 0.3$, which is consistent with all of our previous findings.

6.3 Context Dependent Evaluation

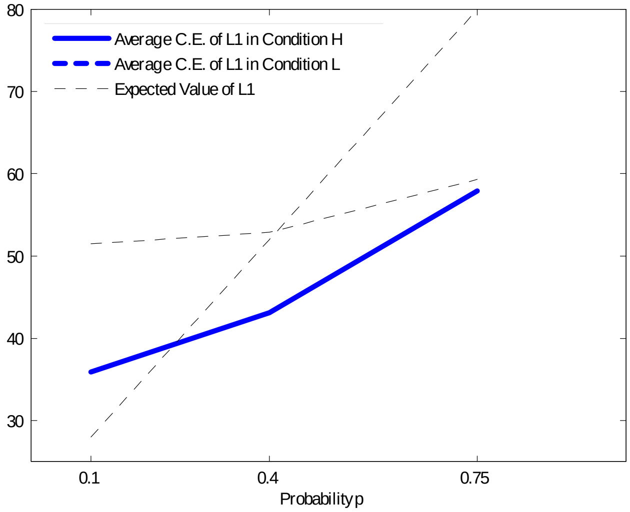
We finally address experimentally one of the basic implications of our model: that a local thinker's evaluation of a lottery depends on its alternative. To test for such context dependence, we gave subjects a series of choices between a lottery $L = (x, p; y, 1-p)$ and a sure prospect $S = (z, 1)$, where $0 < x < z < y$. We ask subjects to first choose between L and S and then to provide their certainty equivalent for L. Our identifying assumption is that such certainty equivalent is computed based on the latter's valuation of L in the comparison with S.

$$\mathrm{CE}(L\,|S\,) = u^{-1}\left[p^{LT}u(x) + \left(1 - p^{LT}\right)u(y)\right]. \tag{24}$$

In our model the above certainty equivalent should fall in the sure prospect z because a higher value of the latter reduces the salience of the lottery upside $\bar{x}$ while increasing the salience of the lottery downside $\underline{x}$. Under Expected Utility (or Prospect Theory), the lottery's certainty equivalent should be independent of the alternative z.

Figure 6 plots the certainty equivalent reported for $L = (\$20, p; \$100, 1-p)$ at probability levels $p \in \{.1, .4, .75\}$ in two conditions: i) $z = \$90$ (which we call condition C_h), and ii) $z = \$25$ (which we call condition C_l):

Figure 6. Average evaluation of lotteries



Two properties stand out. First, the certainty equivalent proposed by the subjects increases with the expected value of the lottery in both conditions C_h and C_l. Second, consistent with our prediction, the certainty equivalent in C_h is always lower than that in C_l. That is, when the lottery is compared to a higher (lower) sure prospect, agents attach a lower (higher) value to it. This is consistent with the idea that a higher sure prospect renders the lottery downside more salient, reducing the agent's evaluation of L.[18]

---

[18] Note that these results are inconsistent with the mere possibility that the sure prospect may act as an anchor. If this were the case, agents' evaluation of the lottery should go up rather than down with the value of z.

Appendix 1. Proofs

Appendix 2. Experimental Evidence and Calibrations

2.A General Experimental Strategy

All experimental results were obtained from surveys conducted on Amazon Mechanical Turk, an online marketplace service hosted by Amazon.com. MTurk allows *requesters* to post tasks that *workers* can complete in exchange for compensation. Typical tasks include data management (e.g. finding the best category for a product), content management (e.g. tagging contents with keywords), and consumer surveys.

We posted a series of surveys on decision-making under risk. Surveys consisted of several multiple choice questions (typically between 4 and 6) which were either choice problems or elicitations of value. The surveys are available upon request. Throughout the surveys, we presented choice problems in the traditional form (KT79). A representative choice problem (which we use below) is for instance:

Problem 1: Choose between:

Lottery A: gives \$2500 with 33% chance and \$0 otherwise, Lottery B: gives \$2400 with 34% chance and \$0 otherwise[19]

For each survey, the sample size was between 75 and 100 subjects. As requesters, we have no information about workers who complete our surveys, except for their worker ID. Using this information, we find that over 1100 different subjects participated in our surveys, of whom over 60% participated only once. We required two conditions for participation: i) that workers be living in the U.S. (so that subjects had as much as possible a similar understanding of the questions involved) and ii) that workers' reputation index be above 0.96 (see discussion below on incentives). We did not collect demographic information on our subjects. However, other surveys on MTurk workers demographics have found that U.S. workers are predominantly female (60%), their age distribution is similar to the general population's but somewhat younger on average (ages range from 16 to 60, and 45% are above 30 years old), have higher education level than the general population (about half have a bachelor degree or higher) and report an average household income level of $\$40,000$.

---

[19] Independent lotteries were always presented side by side.

Though this data is self reported, it indicates that the pool of workers is very representative of the internet-using population, and reasonably representative of the general population.

Given the anonymous survey setting, we used several approaches to ensure high quality data. Monetary incentives were not feasible due to the large volume of surveys and the range of lottery payoffs involved. Moreover, the evidence on the impact of monetary incentives suggests at most a quantitative correction to levels of risk aversion, but no qualitative impact on our results. To understand the workers' motivation, note that workers choose a task in terms of its compensation and interest to them[20] (there are thousands of tasks to choose from at any given moment). Once they choose a task, they have a strong incentive to perform, as it can affect their reputation index: in fact, requesters have the option to accept or reject a worker's task, and the index captures the percentage of a worker's tasks which were accepted. Although we did not reject answers (because no flagrant poor performance was evident) we systematically discarded surveys completed in a very short time, under 45 secs, as these surveys may have been answered without due attention. We included all other surveys in the analysis.

To test for effects of ordering and presentation of the questions on the results, we repeated a few surveys changing the ordering and the presentation. We also repeated a few surveys in identical form to check for consistency of preferences across subjects and across time. Preferences were largely robust to such manipulations. Finally, we occasionally introduced test questions, such as a choice between a two-outcome lottery and a sure prospect lower than the lottery's downside. The rate of "wrong" anwers was always negligible.

2.B Additional Experimental Evidence

2.B.1 Assumption of Independence:

In our model, the state space of a choice problem plays a central role. Consider for instance the choice between Lottery A, which gives \$2500 with 33% chance and \$0 otherwise, and Lottery B, which gives \$2400 with 34% chance and \$0 otherwise. We assume that

---

[20] Workers seemed to take a personal interest in our surveys, often providing feedback and justifying their choices.

subjects interpret the above problem as a choice between two independent lotteries. In other words, consider the following Problem 1', which presents a state space that makes the absence of correlation in Problem 1 explicit:

Problem 1'	21.78%	22.78%	11.22%	44.22%
Lottery A'	\$2500	\$0	\$2500	\$0
Lottery B'	\$0	\$2400	\$2400	\$0

The preferences in *Problem 1* and *Problem 1'* were indistinguishable: Lottery A [49%] and Lottery A' [51%] (N=75 for each problem). This supports our assumption that subjects' default representation of choice problems is that lotteries are independent.

2.B.2 Further tests of correlation

In Section 5 we examined the role of correlation in the context of the Allais paradoxes. Here, we test the impact of correlations on choice problems that, unlike those of the Allais paradox, do not feature a common consequence state. Consider the following two choice problems:

Problem 2:
Lottery 1: \$40 with p=67% and \$120 with p=33%, or
Lottery 2: \$30 with p=50% and \$110 with p=50%,

and

Problem 2'	1/6	1/3	1/2
Lottery 1':	\$40	\$120	\$40
Lottery 2':	\$110	\$110	\$30

The options in Problem 2 and Problem 2' are identical up to correlation in payoffs. The two lotteries have similar expected values, E.V.(Lottery 1)= \$66.4 and E.V.(Lottery 2)= \$70.

The joint preferences were:

Prob2\Prob2'	Lottery 1'	Lottery 2'
Lottery 1	11%	28%
Lottery 2	17%	44%

As predicted, there is a significant shift towards lottery 2', which wins in the most salient and least likely state (even though lottery 1 wins in 2 out of 3 states).

2.B.3 Salience of mixed states

Recall from Definition 3 that salience satisfies diminishing sensitivity in mixed states. For negative mixed states, this property implies that for any $\varepsilon > 0$

$$\sigma(-x, x) > \sigma(-x - \varepsilon, x - \varepsilon)$$

where $x - \varepsilon$ is positive. This is the property for which we had the least direct (from research on perception) and indirect (from lottery choice) evidence. To test this property, we ran – under the hypothesis that our model is correct – the following question, where the lotteries' payoffs depended upon the flip of a coin:

Problem 3	Heads	Tails	
Lottery 1:	get \$0	lose \$5	[73.3%]
Lottery 2:	lose \$10	get \$5	

In the first problem, subjects must trade off two negative mixed states in which payoffs have the same distance and occurring with equal probability. Note that subjects disproportionately choose Lottery 1, which is consistent with them putting more weight on the state where payoffs are on average closer to zero, consistent with the notion of diminishing sensitivity postulated in Definition 3.

We then tested the property of Loss Salience, namely that salience decreases faster for positive mixed states than for negative mixed states:

Problem 3'	Heads	Tails	
Lottery 1':	get $5	lose $5	[75.7%]
Lottery 2':	lose $15	get $15	

The results confirm that subjects put more weight on the negative mixed state (Heads) than on the positive mixed state (Tails). This prediction is consistent also with classical loss aversion, but it is interesting to note that, in the context of our model, it does not imply loss aversion, as discussed in Section 6.

2.B.4 A two-stage common consequence experiment:

In our account of the "isolation effect" of Section 5, we ran the correlated version of the common consequence Allais paradox without assuming that the common consequence occurred in an initial stage of a two stage game. In fact, we wanted to illustrate that our results on correlation do not necessarily rely on compounded lotteries. Here however we check whether the Allais paradox disappears when the common consequence z occurs in the initial stage of a two stage problem. To do so, we ran the following choice problem:

Problem 4: Suppose you are presented with a two-stage game. In the first stage, you have a 66% chance of getting $0 and ending the game, and a 34% chance of going to the second stage. In the second stage, you will play one of the following lotteries:

Lottery 1, you gain	$2500 with 97% chance	Lottery 2, you gain	$2400 for sure
	$0 with 3% chance		

.

Problem 4' differed from *Problem 4* in that the first stage offered a 66% chance of getting $2400, instead of $0. The results were:

Prob4\Prob4'	Lottery 1'	Lottery 2'
Lottery 1	6.7%	2.7%
Lottery 2	6.7%	83.9%

.

Preferences were essentially the same in both Problems, as expected from local thinking and in accordance with the independence axiom. Furthermore, the overall choice patterns coincided with those observed for the one-stage problem with common consequence $z = \$2400$. This emphasizes the point that it is the salience of the zero payoff that drives choices in this problem.

2.C Calibrations

In this section we analyze the experimental results presented in the paper in terms of calibrations of our model and of Prospect Theory. We will use the following standard functional forms for (Cumulative) Prospect Theory's value function v and probability weighting function π (Tversky and Kahneman, 1992) :

$$v(x) = \begin{cases} x^{\alpha} & \text{for } x \geq 0 \\ -\lambda(-x)^{\alpha} & \text{for } x < 0 \end{cases}, \qquad \pi(p) = \frac{p^{\gamma_{KT}}}{(p^{\gamma_{KT}} + (1-p)^{\gamma_{KT}})^{1/\gamma_{KT}}}$$

An agent characterized by v has constant relative risk aversion with coefficient $1 - \alpha > 0$ (implying decreasing absolute risk aversion). The parameter $\lambda > 1$ captures the fact that losses are weighted more heavily than gains and is a measure of loss aversion. This functional form is standard in applications of Prospect Theory (e.g. Benartzi and Thaler, 1995). However, its empirical validity is undermined by the fact that estimations of α from different sets of experiments yield very different values: Tversky and Kahneman calibrate $\alpha \sim 0.88$, $\lambda \sim 2.25$ using choices between two outcome lotteries and sure prospects, but Wu and Gonzalez (1996) get $\alpha \sim 0.5$ and $\alpha \sim 0.37$ from more complex choice problems. This indicates that an important part of the choice process is not being included in the calibration. It would be interesting to understand whether and how Salience can account for these more complex choice problems.

The probability weighting function is calibrated with $\gamma_{KT} \sim 0.61$ (TK1992). Interestingly, the calibration of π is much more robust than that of v: more careful calibrations based on different data yield similar values (Wu and Gonzalez (1996) estimate $\gamma_{KT} \sim 0.71$).

In turn, Prelec's (1998) simplest representation of the weighting function is

$$\pi(p) = e^{-(-\ln p)^{\gamma_P}}$$

with $\gamma_P \sim 0.65$, which is numerically very similar to the original (TK92) calibration.

2.C.1 Attitudes towards Risk

To analyze our experimental results in light of these calibrations, we use a single-agent stochastic choice model (Camerer and Ho 1994), namely we consider that all our subjects are described by the same calibrated model and that their choices follow a stochastic process described by $P(L_1 \succ L_2) = 1/(1 + e^{u(L_2)-u(L_1)})$. This process introduces no extra degrees of freedom and ensures that $P(L_1 \succ L_2) = 0.5$ if and only if $u(L_1) = u(L_2)$, and also $P(L_1 \succ L_2) = 1$ if and only if $u(L_1) >> u(L_2)$.

Recall that subjects were given the choice between a sure prospect y and a mean-preserving spread lottery $(\overline{x}, p; \underline{x}, 1-p)$, where $\overline{x} > y > \underline{x} > 0$. According to (cumulative) Prospect Theory, the lottery is evaluated as

$$V(L) = \pi(p)v(\overline{x}) + (1 - \pi(p))v(\underline{x})$$

Consider now the evidence on risk attitudes:

Proportion of Risk-Seeking Subjects

\$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

The evidence is characterized by two important points: i) for any $p \leq 0.5$, subjects are risk averse for small y, and become progressively more risk seeking as payoff level increases (even for p as large as 0.4 there is a five-fold increase in risk seeking preferences as y increases) and ii) as the loss becomes larger than the gain (i.e. as p goes above 0.5), risk seeking drops dramatically for all levels of y. These results suggest that probability distortions are driven by payoff salience. However, can a combination of risk aversion and probability weighting reproduce these results?

Prospect Theory

$\alpha = 0.6, \gamma_{KT} = 0.61$

\$10500	0.78	0.62	0.52	0.50	0.50	0.49	0.48
\$2100	0.89	0.70	0.54	0.50	0.49	0.48	0.47
\$400	0.94	0.80	0.57	0.49	0.48	0.45	0.44
\$100	0.93	0.82	0.57	0.47	0.44	0.41	0.38
\$20	0.32	0.22	0.13	0.11	0.10	0.10	0.11
	0.01	0.05	0.2	0.33	0.4	0.5	0.67

$\alpha = 0.88, \gamma_{KT} = 0.61$

\$10500	1	0.99	0.85	0.52	0.39	0.28	0.21
\$2100	1	1	0.89	0.52	0.37	0.25	0.16
\$400	1	1	0.92	0.51	0.34	0.20	0.12
\$100	1	1	0.93	0.47	0.28	0.15	0.08
\$20	1	1	0.74	0.17	0.08	0.04	0.02
	0.01	0.05	0.2	0.33	0.4	0.5	0.67

The predictions of Prospect Theory depend on the values of α and γ_{KT}. If α is large (e.g. $\alpha = 0.88$ as suggested by (TK1992)), then the value function is almost linear and preferences in this experiment depend mainly on the weighting of probabilities. The pattern obtained (table on the right) is risk seeking for any p such that $\pi(p) > p$ and risk aversion otherwise, largely independently of y (the exception is small y where agents are mostly risk averse). If instead α is small (e.g. $\alpha = 0.37$ as suggested by (WG96)), then choices are driven mainly by risk preferences and agents are essentially risk neutral for any $x > 100$, independently of p. The intermediate $\alpha = 0.6$ (table on the left) ensures that preferences depend both on payoff level x and on its variance (through p). However, these patterns are significantly different from the experimental results, since for small p, agents are not progressively more risk seeking; instead they are extremely risk seeking (for $x \geq 100$) and then progressively more risk neutral.

Prospect Theory predicts a drop in risk seeking when $\pi(p) < p$, i.e. when $p > 0.35$, which corresponds broadly to what we observe. However, two comments are in order.

First, in its original formulation the probability weighting function was meant to overweight small probabilities (in KT, 1979 this seems to mean $p < 0.1$). In that sense, our prediction of risk seeking for large probabilities $p \sim 0.35$ is entirely original. However, the subsequent calibration of π (e.g. KT, 1992) was derived from experiments similar to these and naturally the calibrated parameters must reflect the revealed preferences.

Second, the results show that risk aversion is reasonably constant for any x for large p (as predicted by our model). Instead, in Prospect Theory the interaction between v and π implies that risk seeking increases in x even for large p. Decreasing γ_{KT} just increases risk seeking on the left side of the table ($p < 0.4$) and decreases it on the right.

Let us now consider the predictions of our model:

Salience Theory ($\delta = 0.7, \alpha = 0.88$)

\$10500	0.79	0.79	0.75	0.71	0.69	0.34	0.39
\$2100	0.79	0.83	0.79	0.75	0.72	0.31	0.37
\$400	0.68	0.82	0.82	0.78	0.76	0.27	0.34
\$100	0.38	0.72	0.81	0.79	0.77	0.22	0.30
\$20	0.001	0.002	0.013	0.03	0.04	0.07	0.16
	0.01	0.05	0.2	0.33	0.4	0.5	0.67

The predictions of Salience Theory are in broad agreement with the experimental results. Even for the simplest specification with a linear utility and constant discount factor, Salience Theory reproduces the overall pattern of shifts towards risk seeking for $p \leq 0.5$, and shifts towards risk aversion for $p > 0.5$. To get a closer fit with the data, one can either add some risk aversion or make the discount factor dependent on salience. In the table above we added a small amount of risk aversion, to match the calibration of Prospect Theory. The fit is not perfect, as it predicts too much risk seeking for intermediate probabilities, $p = 0.2$ to $p = 0.4$, but the overall pattern is clear.

2.C.2 Allais Paradoxes

2.C.2.a Common consequence paradox:

Recall from Section 5.2 that the "common consequence" Allais paradox compares the choices

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

for $z = 0$ and $z = 2400$. When $z = 0$, the salience ranking of the states of the world is

$$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400).$$

Given this ranking, a local thinker with concave utility $u(\cdot)$ chooses the lottery L_1^z when

$$[u(2500) - u(2400)]\,(0.33)\,\delta + [u(0) - u(2400)]\,(0.01) > 0, \tag{25}$$

namely when the utility gain afforded by the lottery when it pays 2500 is sufficiently discounted relative to the most salient state, namely the utility loss borne by the agent when the lottery pays zero. If instead $z = 2400$, the salience ranking is

$$\sigma(2500, 0) > \sigma(0, 2400) > \sigma(2500, 2400) > \sigma(0, 0).$$

Given this ranking, the same agent chooses L_1^z when:

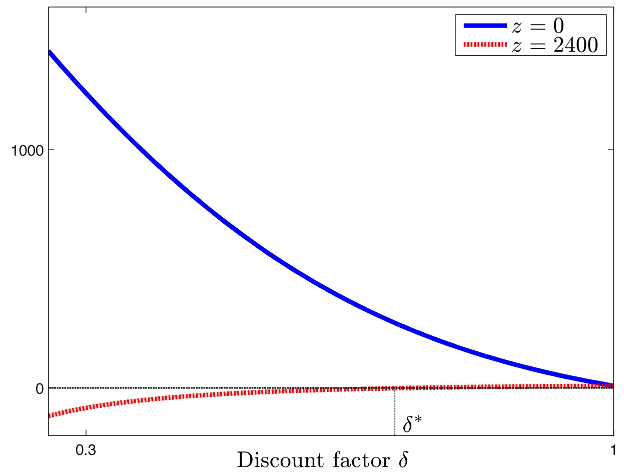
$$[u(2500) - u(2400)]\,(0.33)\,\delta + [u(0) - u(2400)]\,(0.01) \cdot R(\delta) > 0 \tag{26}$$

where $R(\delta) = \frac{[1-0.34(1-\delta)]\delta - 34(1-0.34)(1-\delta)}{1-(0.34)\left(1-\delta^2\right)} \cdot \delta$. For any utility funcion $u(\cdot)$, in Equation (26) the utility loss $[u(0) - u(2400)]$ is heavily discounted relative to Equation (25) because the factor $R(\delta)$ in Equation (26) is smaller than one for every $\delta < 1$. As a result, when $\delta < 1$, Equation (26) is more likely to hold than Equation (25). This implies not only that for a local thinker the common consequence z is not neutral, but also that – as in the Allais paradox – L_1^z is much more likely to be chosen when $z = 0$ than when $z = 2400$.[21] To see this graphically, we compute the difference between the local thinker's expected utility of L_1^z and his expected utility of L_2^z for the two cases $z = 0, 2400$ and we plot them as a function of δ (assuming for simplicity linear utility):

---

[21] Note that Equation (26) is very likely to hold for every $u(\cdot)$, as the multiplicative term $R(\delta)$ is negative as soon as $\delta < 0.95$.

Figure 7. Allais Paradox: Local thinker's evaluation of $\mathbf{L_1^z - L_2^z}$



The agent chooses L_1^z if the difference between the expected utilities is positive, and chooses L_2^z otherwise. When $z = 0$ (the solid line) the agent always chooses the lottery with the highest upside L_1^z. When $z = 2400$ (the dotted line), the agent chooses the lottery with the highest upside L_1^{2400} only if local thinking is not severe $\delta > \delta^*$, where $\delta^*$is the level at which both lotteries have the same expected value. When local thinking is severe (i.e. $\delta < \delta^*$), the local thinker chooses the lottery with smaller downside L_2^{2400}, thereby reproducing the Allais paradox. As $\delta \to 1$, the model converges to expected utility and the value of z becomes irrelevant.

2.C.2.b Robustness:

We now examine the evidence concerning the robustness of the Allais common consequence paradox. Recall that the agent choses between

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

for $z = 0, 5, 25, 100, 200, 2400$. The Prospect Theory agent evaluates the difference between the prospects as

$$\begin{aligned} & u(L_2(z)) - u(L_1(z)) = v(z) + \pi(.34)\left[v(2400) - v(z)\right] - \pi(.33)v(2500) - \pi(.66)v(z) = \\ = \; & v(z)[1 - \pi(.66) - \pi(.34)] - \pi(.33)v(2500) + \pi(.34)v(2400) \end{aligned}$$

In the above expression, for subjects to choose L_1 over L_2 when $z = 0$ [i.e. $L_2(0) - L_1(0) < 0$] it must be that $\pi(.34)V(2400) - \pi(.33)V(2500) < 0$. As z increases, since the weighting function is assumed to be subadditive [i.e. $1 > \pi(.66) + \pi(.34)$], the leftmost term in the above expression becomes larger and larger, eventually rendering the entire expression positive for $z = 2400$. One would however require a quite steep value function $V(\cdot)$ for the switch in the sign of the above expression to already occur at $z = 5$, especially because the extent of subadditivity of $\pi(\cdot)$ is typically small. Using the above calibrations for the value function v and the probability weighting function π, we get

$$u(L_2(z)) - u(L_1(z)) = 0.08 \cdot z^\alpha - 0.348 \cdot 2500^\alpha + 0.355 \cdot 2400^\alpha$$

Preferences are then modeled as $P(L_1 \succ L_2) = 1/(1 + e^{u(L_2) - u(L_1)})$. Varying the risk aversion coefficient α has two effects: increasing α increases the impact of z on the demand for the safer option L_1 as $z \to 2400$, but it also diminishes that impact when z is small. Instead, if α is small, increasing z from 0 to 5 significantly increases demand for L_1, but for large z the agent is indifferent between L_1 and L_2. This trade-off is inconsistent with the experimental data, which shows both a significant increase in demand for L_1 for $z = 5$ relative to $z = 0$, and also a nearly universal demand for L_1 as $z \to 2400$. Both of these effects are predicted by Salience Theory.

The conclusion of this section on calibration is that Salience Theory can account for all experimental evidence on the Allais paradoxes as long as the degree of local thinking δ is small enough. Salience Theory is also quantitatively compatible with the experiments on risk attitudes for intermediate values of δ.

2.D Mixed Lotteries

We now turn to the evidence on risk seeking behavior for mixed lotteries. Recall from Section 6.1 that subjects chose between $L_1^0 = (-\$40, \frac{1}{2};\ \$40, \frac{1}{2})$ and $L_2^0 = (0, 1)$ and also between $L_1^{-20} = (-\$60, \frac{1}{2}; \$20, \frac{1}{2})$ and $L_2^{-20} = (-\$20, 1)$. The joint distribution of preferences was

	L_1^{-20}	L_2^{-20}
L_1^0	19%	12%
L_2^0	37%	32%

As highlighted in the paper, over a third (37%) of subjects shift towards risk seeking in the second problem (compared to 12% who shift in the opposite direction). How does Prospect Theory account for these results? To estimate the joint distribution, we can use the same single-agent stochastic choice model with the assumption that the two choices are independent, $P(L_1^0 \succ L_2^0, L_1^{-20} \succ L_2^{-20}) = P(L_1^0 \succ L_2^0) \cdot P(L_1^{-20} \succ L_2^{-20})$. We can then fit the risk aversion parameter α to the results above. For instance, $\alpha = 0.1$ leads to

	L_1^{-20}	L_2^{-20}
L_1^0	16%	10%
L_2^0	45%	29%

Intuitively, Prospect Theory can only account for the results if, in the choice between L_1^{-20} and L_2^{-20}, the convexity of the value function for losses outweights loss aversion. But a very convex value function for losses (i.e. a very low α) also implies a very concave value function for gains, with a risk aversion coefficient of the order of 0.9, an unreasonably large value. Prospect Theory may account for these results by allowing different curvatures for the value function for gains and losses; but this would not ease the constrains on the convexity of the value function for losses.

References

Allais, M., 1953, “Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’Ecole Americaine”, Econometrica 21, 503- 546.

Benartzi, S. and Thaler, R. 1995. “Myopic Loss Aversion and the Equity Premium Puzzle", Quarterly Journal of Economics 110, 73

Camerer, C. 1995. “Choice under Risk and Uncertainty”, in Handbook of Experimental Economics, JH Kagel and AE Roth eds.

Camerer, C. and Ho, T. 1995. “Violations of the betweenness axiom and nonlinearity in probability", Journal of Risk and Uncertainty 8, 167

Edwards, W. 1962. “Subjective probabilities inferred from decisions", Psychological Review 69, 109

Fellner, W. 1961. “Distortion of subjective probabilities as a reaction to uncertainty", Quarterly Journal of Economics 75, 670

Gennaioli, N., Shleifer, A. 2010. “What comes to mind”, Quarterly Journal of Economics, forthcoming.

Kahneman, D. and Tversky, A.. 1979. “Prospect theory: an analysis of decision under risk”, Econometrica, vol 47, issue 2, 263

Kahneman, D. and Tversky, A.. 1981. “The framing of decisions and the psychology of choice", Science 211, 453

Kahneman, D. 2003. “Maps of bounded rationality: Psychology for Behavioral Economics”, American Economic Review 93, 1449-1476

Kandel, E., Schwartz, J., Jessell, T. 1991. “Principles of Neural Science”, McGraw-Hill/Appleton

Koszegi, B., Rabin, M. 2007. “Reference-dependent risk attitudes”, American Economic Review, 97(4), 1047

Loomes, G. and Sugden, R. 1982. “Regret theory: An alternative theory of rational choice under uncertainty”, The Economic Journal, 805

McCoy, A. and Platt, M. 2005. “Risk-sensitive neurons in macaque posterior cingulate cortex”, Nature Neuroscience, 8(9), 1220

Mullainathan, S. 2002. “Thinking through Categories”, NBER working paper.

Prelec, D. 1998. “The probability weighting function”, Econometrica 66(3), 497

Rabin, M. 2000. “Risk aversion and expected-utility theory: A calibration theorem”, Econometrica 68(5), 1281

Rubinstein, A. 1988 “Similarity and decision-making under risk", Journal of Economic Theory 46, 145-153

Schwartzstein, J. 2009. “Selective Attention and Learning”, working paper.

Taylor, SE., Thompson, SC. 1982. “Stalking the elusive vividness effect", Psychological Review 89, 155

Tversky, A. 1969. “Intransitivity of Preferences”, Psychological Review, 76, 31-48

Tversky, A. and Kahneman, D.. 1981. “The framing of decisions and the psychology of choice”, Science, 211(4481), 453-458

Tversky, A. and Kahneman, D.. 1992. “Advances in prospect theory: Cumulative representation of uncertainty”, Journal of Risk and Uncertainty, 5(4), 297-323.

Wu, G. and Gonzalez, R. 1996. “Curvature of the probability weighting function”, Management Science, 42(12), 1676