NBER WORKING PAPER SERIES

SALIENCE THEORY OF CHOICE UNDER RISK

Pedro Bordalo
Nicola Gennaioli
Andrei Shleifer

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

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

We are grateful to Nicholas Barberis, Gary Becker, Colin Camerer, John Campbell, Tom Cunningham, Xavier Gabaix, Morgan Grossman-McKee, Ming Huang, Jonathan Ingersoll, Emir Kamenica, Daniel Kahneman, Botond Koszegi, David Laibson, Pepe Montiel Olea, Drazen Prelec, Matthew Rabin, Josh Schwartzstein, Jesse Shapiro, Tomasz Strzalecki, Richard Thaler, Georg Weiszacker and George Wu for extremely helpful comments, and to Allen Yang for excellent research assistance. Gennaioli thanks the Spanish Ministerio de Ciencia y Tecnologia (ECO 2008-01666 and Ramon y Cajal grants), the Barcelona GSE Research Network, and the Generalitat de Catalunya for financial support. Shleifer thanks the Kauffman Foundation for research support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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

ABSTRACT

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

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

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

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

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

1 Introduction

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

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

We propose a new psychologically founded model of choice under risk, which naturally exhibits the systematic instability of risk preferences and accounts for the puzzles. In this model, risk attitudes are driven by the salience of different lottery payoffs. Psychologists view salience detection as a key attentional mechanism enabling humans to focus their limited cognitive resources on a relevant subset of the available sensory data. As Taylor and Thompson (1982) put it: "Salience refers to the phenomenon that when one's attention is differentially directed to one portion of the environment rather than to others, the information contained in that portion will receive disproportionate weighting in subsequent judgments." In line with this idea, in our model the decision maker focuses on salient payoffs. He is then risk seeking when a lottery's upside is salient and risk averse when its downside is salient.

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

At a broad level, our approach is similar to that pursued by Gennaioli and Shleifer (2010) in their study of the representativeness heuristic in probability judgments. The idea of both studies is that decision makers do not take into account fully all the information available to them, but rather over-emphasize the information their minds focus on.[2] Gennaioli and Shleifer (2010) call such decision makers local thinkers, because they neglect potentially important but unrepresentative data. Here, analogously, in evaluating lotteries, decision makers overweight states that draw their attention and neglect states that do not. We continue to refer to such decision makers as local thinkers. In both models, the limiting case in which all information is processed correctly is the standard economic agent.

Our model leads to an understanding of what encourages and discourages risk seeking, but also to an explanation of the Allais paradoxes. The strongest departures from Expected Utility Theory in our model occur in the presence of extreme payoffs, particularly when these occur with a low probability. Due to this property, our model predicts that subjects in the Allais experiments are risk loving when the common consequence is small and attention is drawn to the highest lottery payoffs, and risk averse when the common consequence is large and attention is drawn to the lowest payoffs. We explore the model's predictions by describing, and then experimentally testing, how Allais paradoxes can be turned on and off. We also show that preference reversals can be seen as a consequence of lottery evaluation in different contexts (that affect salience), rather than the result of a fundamental difference between pricing and choosing. The model thus provides a unified explanation of risk preferences and invariance violations based on a psychologically motivated mechanism

[1]In most of the paper, we assume a linear utility function. However, this functional form does not deal with the phenomenon of loss aversion, i.e. the extreme risk aversion with respect to small positive expected value bets. To deal with this phenomenon, we modify preferences around zero along the lines of Kahneman and Tversky (1979) in Section 7.3.

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

of salience.

It is useful to compare our model to the gold standard of existing theories of choice under risk, Kahneman and Tversky's (KT, 1979) Prospect Theory. Prospect Theory incorporates the assumption that the probability weights people use to make choices are different from objective probabilities. But the idea that these weights depend on the actual payoffs and their salience is new here. In some situations, our endogenously derived decision weights look very similar to KT's, but in other situations – for instance when small probabilities are *not* attached to salient payoffs or when lotteries are correlated – they are very different. We conduct multiple experiments, both of simple risk attitudes and of Allais paradoxes with correlated states, that distinguish our predictions from KT's, and uniformly find strong support for our model of probability weighting.

The paper proceeds as follows. In Section 2, we present an experiment illustrating the switch from risk averse to risk-loving behavior as lottery payoffs, and their salience, change. In Section 3, we present a salience-based model of choice among two lotteries, and show how changes in the structure of lotteries affect the endogenous decision weights. In Section 4, we use this model to study risk attitudes, derive from first principles Prospect Theory's weighting function for a class of choice problems where it should apply, and provide experimental evidence for our predictions. In Section 5 we show that our model accounts for Allais paradoxes and preference reversals. We obtain new predictions concerning these paradoxes, and test them. In Section 6, we extend the model to choice among many lotteries. We then introduce salience into a standard asset pricing model, which may shed light on some empirical puzzles in finance, such as the growth-value anomaly. In Section 7, we address framing effects, failures of transitivity and mixed lotteries. Section 8 concludes.

2 A Simple Example

We begin by presenting the results of two experiments illustrating two central intuitions behind our model: how the contrast between payoffs in different states makes some states more salient to the decision maker than others, and how this process shapes risk attitudes. The procedures for all experiments in the paper are described in the Appendix 2 (Supplementary

Material). The two experiments are:

Experiment 1: Choose between the two options:

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

Experiment 2: Choose between the two options:

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

Three points are noteworthy. First, Experiment 2 simply adds \$300 to all the payoffs in Experiment 1. Second, in both experiments L_1 and L_2 have the same expected payoffs. Third, in both experiments lottery L_1 has the same relatively small (5%) probability of a high payoff, and a high (95%) probability of a \$19 loss relative to the sure outcome.

The same 120 subjects participated in the two experiments over the internet. In Experiment 1, 83% of the subjects chose the safe option L_2, whereas in Experiment 2, 67% of the same subjects chose the risky option L_1. Thus, there is a statistically significant switch from a large majority of risk averse choices to a large majority of risk seeking choices. In fact, over half the subjects who chose L_2 in the first experiment switched to L_1 in the second.

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

Our explanation of these findings does not rely on the shape of the value function. It

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

goes roughly as follows. In Experiment 1, in the state where the lottery loses relative to the sure payoff, the lottery's payoff of \$1 feels a lot lower than the sure payoff of \$20. Because this downside is more salient than winning \$381, the subjects focus on it when making their decisions. This focus triggers the risk averse choice. In Experiment 2, the lottery's payoff in the bad outcome state, \$301, does not appear nearly as bad compared to the sure payoff of \$320. The upside of winning \$681 is more salient and subjects focus on it when making their decisions. This focus triggers the risk seeking choice. The analogy here is to sensory perception: a lottery's salient payoffs are those which differ most strongly from the payoffs of alternative lotteries, and the decision maker's mind focuses on salient payoffs when making a choice. We now describe a model that formalizes this intuition.

3 The Model

A choice problem is described by: i) a set of states of the world S, where each state $s \in S$ occurs with objective and known probability π_s such that $\sum_{s\in S} \pi_s = 1$, and ii) a choice set $\{L_1, L_2\}$, where the L_i are risky prospects that yield monetary payoffs x_s^i in each state s. For convenience, we refer to L_i as lotteries.[4] Here we focus on choice between two lotteries, leaving the general case of choice among $N > 2$ lotteries to Section 6.

The agent uses a value function[5] v to evaluate lottery payoffs relative to the reference point of zero.[6] Absent distortions in decision weights, the agent evaluates L_i as:

$$V(L_i) = \sum_{s\in S} \pi_s v(x_s^i). \tag{1}$$

[4]Formally, L_i are acts, or random variables, defined over the choice problem's probability space (S, F_S, π), where S is assumed to be finite and F_S is its canonical σ-algebra. However, as we will see in Equation (7), the decision maker's choice depends only on the L_i's joint distribution over payoffs and not on the exact structure of the state space. Thus we use the term lotteries, in a slight abuse of nomenclature relative to the usual definition of lotteries as probability distributions over payoffs.

[5]Throughout most of the paper, we illustrate the mechanism generating risk preferences in our model by assuming a linear value function v. In section 7.3, when we focus on mixed lotteries, we consider a piece-wise linear value function featuring loss aversion, as in Kahneman and Tversky (1979).

[6]This is a form of narrow framing, also used in Prospect Theory. Koszegi and Rabin (2006, 2007) build a model of reference point formation and use it to explain shifts in risk attitudes in the real world. We instead study risk attitudes in the lab holding reference points constant. These approaches are complementary, as one could combine our model of decision weights with Koszegi and Rabin's two part value function.

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

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

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

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

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

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

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

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

Section 3.1 discusses the connection between these properties and the cognitive notion of salience. To illustrate Definition 1, consider the salience function:

$$\sigma(x_s^i, x_s^{-i}) = \frac{|x_s^i - x_s^{-i}|}{|x_s^i| + |x_s^{-i}| + \theta}. \tag{2}$$

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

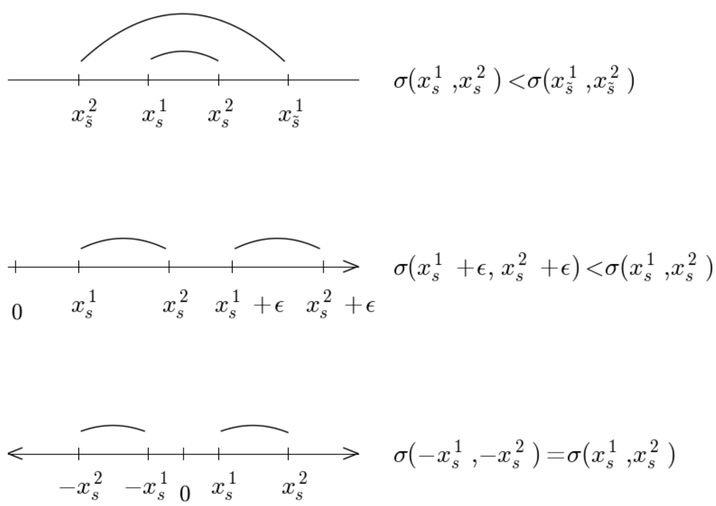

Figure 1: Properties of a salience function, Eq. (2)

The specification (2) exhibits two additional properties. The first is "symmetry", namely $\sigma(x_s^1, x_s^2) = \sigma(x_s^2, x_s^1)$, which is a natural property in the case of two lotteries but which is dropped in the $N > 2$ lottery case. The second property of (2) is "convexity": salience falls at a decreasing rate as payoffs become larger in absolute value. This latter property limits the extent of diminishing sensitivity, implying that at large absolute payoff values the distance between payoffs (the numerator) becomes the principal determinant of salience.[7] Our main results rely only on the properties in Definition 1, but we often use the tractable functional form (2) to illustrate our model. The example of Section 2 follows from (2) evaluated at $\theta \simeq 0$. In this example, there are two states of the world: one in which the lottery yields its upside, the other in which it yields its downside.[8] In Experiment 1 the state $(1, 20)$ where

[7]The convexity of (2) formally means that $\left|\sigma(x_s^i + \epsilon, x_s^{-i} + \epsilon) - \sigma(x_s^i, x_s^{-i})\right|$ (weakly) decreases as the magnitude of payoffs (x_s^i, x_s^{-i}) goes up. Parameter θ in (2) captures the relative strength of ordering (the numerator) vs. diminishing sensitivity (the denominator). If $\theta = 0$, diminishing sensitivity is strong because any state with a zero payoff has maximal salience: $\sigma(0, x) = 1$ regardless of the value of x. When $\theta > 0$, even a state with a zero payoff can be not very salient if x is small.

[8]In this example, constructing the state space from the alternatives of choice is straightforward. Section

the lottery yields its downside of \$1 is more salient than the state $(381, 20)$ where the lottery yields its upside \$381, since $\sigma(1, 20) \simeq \frac{20-1}{21} > \sigma(381, 20) \simeq \frac{381-20}{401}$. However, in Experiment 2, where payoffs have been shifted up, the lottery's upside \$681 is more salient than its downside \$301, $\sigma(301, 320) \simeq \frac{320-301}{621} < \sigma(681, 320) \simeq \frac{681-320}{1001}$.

3.1 Salience, Decision Weights and Risk Attitudes

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

Definition 2 *Given states $s, \widetilde{s} \in S$, we say that for lottery L_i state s is more salient than $\tilde{s}$ if $\sigma(x_s^i, x_s^{-i}) > \sigma(x_{\widetilde{s}}^i, x_{\widetilde{s}}^{-i})$. Let $k_s^i \in \{1, ..., |S|\}$ be the salience ranking of state s for L_i, with lower k_s^i indicating higher salience. States with the same salience obtain the same ranking. Then, if s is more salient than $\widetilde{s}$, namely if $k_s^i < k_{\widetilde{s}}^i$, the local thinker transforms the odds $\pi_{\widetilde{s}}/\pi_s$ of $\widetilde{s}$ relative to s into the odds $\pi_{\widetilde{s}}^i/\pi_s^i$, given by:*

$$\frac{\pi_{\widetilde{s}}^i}{\pi_s^i} = \delta^{k_{\widetilde{s}}^i - k_s^i} \cdot \frac{\pi_{\widetilde{s}}}{\pi_s} \tag{3}$$

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

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

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

3.2 describes how the state space S is derived in more complex cases.

As show in Appendix 1, Definition 2 implies that the extent of overweighting also depends on objective probabilities:

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

$$\frac{\mathrm{d}\omega_s^i}{h} = -\frac{\pi_s}{1-\pi_s} \cdot \omega_s^i \cdot \left(\omega_s^i - 1\right). \tag{5}$$

This result states that an increase in a state's probability π_s reduces the distortion of the decision weight in that state by driving ω_s^i closer to 1. That is, low probability states are subject to the strongest distortions: they are severely over-weighted if salient and severely under-weighted otherwise. This stands in marked contrast to KT's (1979,1992) assumption that low probability, high rank payoffs are always overweighted. In our model, payoffs are overweighted if and only if they are salient, regardless of probability. On the other hand, by Proposition 1 our model also implies that the largest distortions of choice occur when salient payoffs are relatively unlikely. This property plays a key role for explaining some important findings such as the common ratio Allais Paradox in Section 5.1.

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

$$V^{LT}(L_i) = \sum_{s \in S} \pi_s^i v(x_s^i) = \sum_{s \in S} \pi_s \omega_s^i v(x_s^i). \tag{6}$$

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

Since salience is defined on the state space S, one may wonder whether splitting states, or generally considering a different state space compatible with the lotteries' payoff distributions, may affect the local thinker's evaluation (6). Denote by $S_{\mathbf{x}}$ the set of states in S where the lotteries yield the same payoff combination $\mathbf{x}$, formally $S_{\mathbf{x}} \equiv \{s \in S \,|\mathbf{x}_s = \mathbf{x}\}$. Clearly, $S = \cup_{\mathbf{x} \in X} S_{\mathbf{x}}$ where X denotes the set of distinct payoff combinations occurring in S. By Definition 1, all states s in $S_{\mathbf{x}}$ are equally salient for either lottery, and thus have the same value of ω_s^i, which for simplicity we denote $\omega_{\mathbf{x}}^i$. Using (4) we can rewrite $V^{LT}(L_i)$ in (6) as:

$$V^{LT}(L_i) = \sum_{\mathbf{x}\in X}\left(\sum_{s\in S_{\mathbf{x}}}\pi_s\right)\omega_{\mathbf{x}}^i v(x_{\mathbf{x}}^i), \tag{7}$$

where $x_{\mathbf{x}}^i$ denotes L_i's payoff in $\mathbf{x}$. Equation (7) says that the state space only influences evaluation through the total probability of each distinct payoff combination $\mathbf{x}$, namely $\pi_{\mathbf{x}} = \sum_{s\in S_{\mathbf{x}}}\pi_s$. This is because salience $\sigma(.,)$ depends on payoffs, and not on the probabilities of different states. Hence, splitting a given probability $\pi_{\mathbf{x}}$ across different sets of states does not affect evaluation (or choice) in our model. There is therefore no loss in generality from viewing S as the "minimal" state space X identified by the set of distinct payoff combinations that occur with positive probability. In the remainder of the paper, we keep the notation of Equation (6), with the understanding that S is this "minimal" state space.

In a choice between two lotteries, Equation (6) implies that - due to the symmetry of the salience function (i.e. $k_s^1 = k_s^2$ for all s) - the local thinker prefers L_1 to L_2 if and only if:

$$\sum_{s\in S}\delta^{k_s}\pi_s\left[v(x_s^1) - v(x_s^2)\right] > 0. \tag{8}$$

For $\delta = 1$, the agent's decision coincides with that of an Expected Utility maximizer having the same value function $v(.)$. For $\delta < 1$, local thinking favors L_1 when it pays more than L_2 in the more salient (and thus less discounted) states.

3.2 Discussion of Assumptions and Setup

We now discuss our formalization of salience and of the state space, the key ingredients of our approach.

Salience and Decision Weights

In human perception, a sensorial stimulus gives rise to a subjective representation whose intensity increases in the stimulus' magnitude but also depends on context (Kandel et al, 1991). In our model, the strength of the stimulus is the payoff difference among lotteries in a given state and the salience function $\sigma(.,.)$ captures the subjective intensity with which this stimulus is perceived. Through diminishing sensitivity and reflection, this subjective intensity decreases with the distance of the state's payoffs from the status quo of zero, which is our measure of context. As in Weber's law of diminishing sensitivity, whereby a change in luminosity is perceived less intensely if it occurs at a higher luminosity level, the local thinker perceives less intensely payoff differences occurring at high (absolute) payoff levels.[9]

Consistent with psychology of attention, we assume that the agent evaluates options by focusing on (weighting more) their most salient states. The "local thinking" parameter $1/\delta$ captures the agent's focus on salient states, proxying for his ability to pay attention to multiple aspects, cognitive load, or simply intelligence. Our assumption of rank-based discounting buys us analytical tractability, but our main results also hold if the distortion of the odds in (3) is a smooth function of salience differences, for instance $\delta^{\left[\sigma(x_s^i, x_s^{-i}) - \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i})\right]}$. The main restriction embodied in our model is that this function does not depend on a state's probability. The salience function in Equation (2) provides a tractable benchmark characterized by only two parameters (θ, δ). This allows us to look for ranges of θ and δ that are consistent with the observed choice patterns.

The State Space

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

In our approach, the state space S and the states' objective probabilities are a given of the choice problem.[10] In the lab, specifying a state space for a choice problem is straightforward when the feasible payoff combinations – and their probabilities – are available, for instance

[9]Neurobiological evidence connects visual perception to risk taking. 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)."

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

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

Our emphasis on the state space as a source of context dependence does not lead to accurate predictions when lotteries are presented in a way that induces the decision maker to neglect the state space. For example, suppose that the payoffs of two lotteries are determined by the roll of the same dice. One lottery pays 1,2,3,4,5,6, according to the dice's face; the other lottery pays 2,3,4,5,6,1. The state in which the first lottery pays 6 and the second pays 1 may appear most salient to the decision maker, leading him to prefer the first lottery. But of course, a moment's thought would lead him to realize that the lotteries are just rearrangements of each other, and recognize them as identical. In the following, we assume that, before evaluating lotteries, the decision maker edits the choice set by discarding lotteries that are mere permutations of other lotteries. We also assume (see Section 6) that he discards dominated lotteries from the choice set. Such editing is plausibly related to salience itself: in these cases, before comparing payoffs, what is salient to the decision maker are the properties of permutation or dominance of certain lotteries. To focus our study on the salience of lottery payoffs, we do not formally model this editing process. However, endogenizing the choice set is an important direction for future work. In a similar spirit, the model could be generalized to take into account determinants of salience other than payoff values, such as

[11] In Appendix 2 (Supplementary Material) we provide experimental evidence consistent with this assumption, as well as details on the information given in the experimental surveys.

prior experiences and details of presentation, or even color of font. These may matter in some situations but are not considered here.

Our theory of decision weights can be viewed as a way to endogeneize the probability weighting function introduced by Edwards (1962), Fellner (1961) and later used by KT in Prospect Theory. The various properties of this probability weighting function, such as overweighting of small probabilities and subadditivity, allow KT to account for risk loving behavior and the Allais paradoxes. Quiggin's (1982) rank-dependent expected utility and Tversky and Kahneman's (1992) Cumulative Prospect Theory (CPT) develop weigthing functions in which the rank order of a lottery's payoffs affects probability weighting.[12]

Our theory exhibits two sharp differences from these works. First, in our model the magnitude of payoffs, not only their rank, determines salience and probability weights: the lottery upside may still be underweighted if the payoff associated with it is not sufficiently large. As we show in Section 4, this feature is crucial to explaining shifts in risk attitudes. Second, and more important, in our model decision weights depend on the choice context, namely on the available alternatives as they are presented to the agent. In Section 5 we exploit this feature to shed light on the psychological forces behind the Allais paradoxes and preference reversals. We are not the first to propose a model of context dependent choice among lotteries. Rubinstein (1988), followed by Aizpurua et al (1990) and Leland (1994), builds a model of similarity-based preferences, in which agents simplify the choice among two lotteries by pruning the dimension (probability or payoff, if any), along which lotteries are similar. The working and predictions of our model are different from Rubinstein's, even though we share the idea that the common ratio Allais paradox (see Section 5.1.2) is due to subjects' focus on lottery payoffs. In Regret Theory (Loomes and Sugden 1982, Bell 1982, Fishburn 1982), the choice set directly affects the agent's utility via a regret/rejoice term added to a standard utility function. In our model, instead, context affects decisions by shaping the salience of payoffs and decision weights. By adopting a traditional utility theory perspective, Regret Theory cannot capture framing effects and violations of procedural invariance (Tversky et al. 1990).

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

We now show that our model provides an intuitive explanation for several well known anomalies of choice under risk and deliver new predictions, which we experimentally test. Section 6 then shows how our model can be used in relevant economic applications.

4 Salience and Attitudes Towards Risk

Consider the choice between a lottery $L_1 = (x_h^1, \pi_h;\, x_l^1, 1-\pi_h)$ and a sure prospect $L_2 = (x, 1)$ that have the same mean, namely $\mathbb{E}_s(x_s^1) = x$. Here we assume that all payoffs are positive, and leave issues related to loss aversion to Section 7.3. This setup is often used by experimenters to elicit risk attitudes, and illustrates in the starkest manner how salience shapes risk attitudes. In state $s_h = (x_h^1, x)$ the lottery gains relative to the sure prospect, while in state $s_l = (x_l^1, x)$ it loses. Since $\mathbb{E}_s(x_s^1) = x$, it is easy to see that Equation (8) implies that for any $\delta < 1$, a local thinker with linear utility chooses the lottery if and only if the gain state s_h is more salient than the loss state s_l, i.e. when $\sigma(x_h^1, x) > \sigma(x_l^1, x)$. Indeed, in this case $\pi_h^1 > \pi_h$ and the local thinker perceives the expected value of L_1 to be above that of L_2, behaving in a risk seeking manner. Using the salience function in Equation (2), this occurs when:

$$\left(x + \frac{\theta}{2}\right)(1 - 2\pi_h) > (x - x_l^1)(1 - \pi_h), \tag{9}$$

which uniquely identifies the parameter values for which the agent is risk seeking. Holding the lottery loss $(x - x_l^1)$ constant at some value $\widetilde{l}$ (as in the experiments of Section 2, where $\widetilde{l} = 19$), the risk attitudes implied by Equation (9) are pictured in Figure 2. Recall that $x > \tilde{l}$ so that $x_l^1 > 0$. For convenience, we set $\theta/\tilde{l} \simeq 0$.

Two patterns stand out. First, as in Section 2, for a fixed $\pi_h < 1/2$, a higher expected value x fosters risk seeking by inducing a vertical move from the grey to the white region. When x is low, the lottery's downside x_l^1 is close to zero. By diminishing sensitivity, the loss is salient, inducing risk aversion. As x becomes large, the effect of diminishing sensitivity weakens, due to the convexity of the salience function in (2). Since for $\pi_h < 1/2$ the lottery gain is larger than the loss, it eventually becomes salient, inducing risk seeking.[13]

[13] Besides the properties of Definition 1, to obtain Figure 2 it suffices for the salience function to be convex. Indeed, define $x_h^1(\pi_h, x)$ as the upside at which the lottery's expected value is equal to the sure prospect.

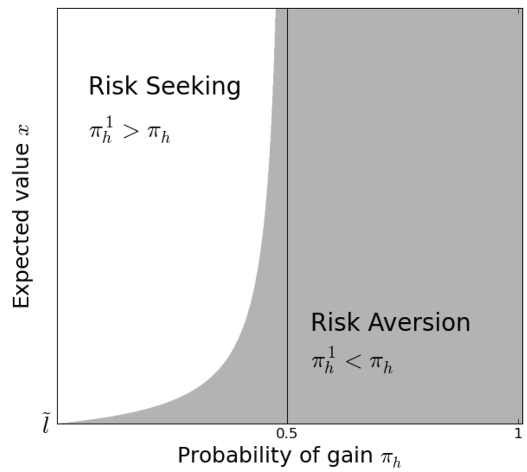

Figure 2: Shifts in risk attitudes

Second, for a given expected value x, a higher probability π_h of the gain reduces risk seeking by inducing a horizontal move from the white to the grey region of Figure 2. As π_h increases, the lottery's upside must fall for the expected value of L_1 to stay constant. As a consequence, the lottery gain becomes less salient, inducing risk aversion. Risk seeking never occurs when $\pi_h \geq 1/2$: now the gain is weakly smaller than the loss in absolute terms. By diminishing sensitivity, the loss is more salient.

Remarkably, in this context our model of decision weights recovers the key features of Prospect Theory's inverse S-shaped probability weighting function (KT 1979): over-weighting of low probabilities, and under-weighting of high probabilities. To see how, fix a value of $x > \widetilde{l}$ in Figure 2 and increase the probability π_h along the horizontal axis. Figure 3 shows the decision weight π_h^1 along this path, where $\pi_h^*(x)$ is the threshold at which the agent switches from risk seeking to risk aversion in Figure 2. Low probabilities are over-weighted because they are associated with salient upsides of longshot lotteries. High probabilities are under-weighted as they occur in lotteries with a small, non salient, upside.

Note however that in our model the weighting function is context dependent. In contrast

The local thinker is risk seeking if $\sigma(x_h^1(\pi_h, x), x) > \sigma(x - \tilde{l}, x)$. Since $x_h^1(\pi_h, x)$ falls in π_h, ordering implies that for π_h sufficiently large $x - l$ becomes salient and the agent is risk averse. This is surely the case for $\pi_h \geq 1/2$. On the other hand, since by convexity $\sigma(x - \tilde{l}, x)$ decreases in x, as x becomes large the upside eventually becomes salient, yielding Figure 2.

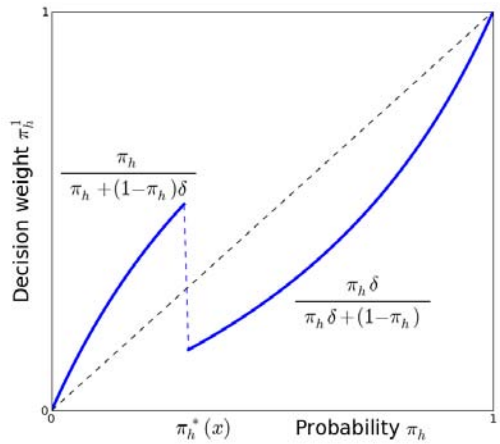

Figure 3: Context dependent probability weighting function

to Prospect Theory, risk seeking behavior is no longer only associated with a low probability of a gain. At high expected values x, the threshold $\pi_h^*(x)$ approaches 1/2 so risk seeking occurs even at moderate probabilities. At low x the threshold is low, so risk aversion occurs even at low probabilities. The salience of particular states can induce risk seeking behavior in conditions that are far more common than those characterizing longshot bets.

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

The patterns are qualitatively consistent with the predictions of Figure 2. For a given expected value x, the proportion of risk takers falls as π_h increases; for a given $\pi_h < 0.5$, the proportion of risk takers increases with the expected value x. The effect is statistically significant: at $\pi_h = 0.05$ a large majority of subjects (80%) are risk averse when $x = \$20$, but as x increases to \$2100 a large majority (65%) becomes risk seeking. Finally, there is a large

Table I: Proportion of Risk-Seeking Subjects

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

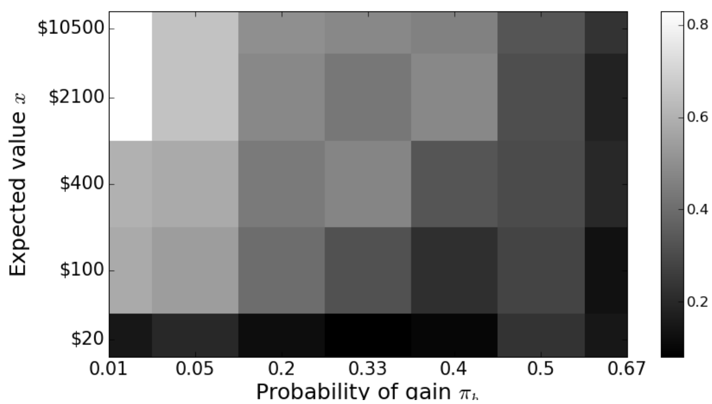

Figure 4: Proportion of Risk-Seeking Subjects

drop in risk taking as π_h crosses 0.5. Note that the increase in x raises the proportion of risk takers from around 10% to 50% even for moderate probabilities in the range $(0.2, 0.4)$. These patterns are broadly consistent with the predictions of our model. The weighing function of Prospect Theory and CPT can explain why risk seeking prevails at low π_h, but not the shift from risk aversion to risk seeking as x rises. To explain this finding, both theories need a concave value function characterized by strong diminishing returns.[14]

In Appendix 2 (Supplementary Material) we show that parameter values $\delta \sim 0.7$ and

[14]In Appendix 2 we provide further support for these claims by showing that standard calibrations of Prospect Theory cannot explain our experimental findings. For example, the calibration in KT(92) features the value function $v(x) = x^{0.88}$, which is insufficiently concave. Appendix 2 performs additional experiments on longshot lotteries whose results are also consistent with out model but inconsistent with Prospect Theory under standard calibrations of the value function.

$\theta \sim 0.1$ are consistent with the above evidence on risk preferences, as well as with risk preferences concerning longshot lotteries. These values are not a formal calibration, but we employ them as a useful reference for discussing Allais paradoxes in the next section.

5 Local Thinking and Context Dependence

We now illustrate the distinctive implications of our model regarding the role of context dependence in the Allais paradoxes and in preference reversals.

5.1 The Allais Paradoxes

5.1.1 The "common consequence " Allais Paradox

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

$$L_1^z = (2500,\ 0.33;\quad 0,\ 0.01;\quad z,\ 0.66),\quad L_2^z = (2400,\ 0.34;\quad z,\ 0.66) \tag{10}$$

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

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

Prospect Theory and CPT (KT 1979 and TK 1992) explain the switch from L_2^{2400} to L_1^0 by the so called "certainty effect", the idea that adding a downside risk to the sure prospect L_2^{2400} undermines agents' valuation much more than adding the same downside risk to the already risky lottery L_1^{2400}. This effect is directly built into the probability weighting function $\pi(p)$ by the assumption of subcertainty, e.g. $\pi(0.34) - \pi(0) < 1 - \pi(0.66)$.[15]

[15] In CPT the mathematical condition on probability weights is slightly different but carries the same

Our model endogenizes this feature of decision weights, and thus explains the Allais paradox, because the common consequence z alters the salience of lottery outcomes. To see this, consider the choice between L_1^{2400} and L_2^{2400}. The minimal state space is $S = \{(2500, 2400), (0, 2400), (2400, 2400)\}$ so there are three states of the world and the most salient state is one where the risky lottery L_1^{2400} pays zero because:

$$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400). \tag{11}$$

The inequalities follow from diminishing sensitivity and ordering, respectively, and can be easily verified for the case of the salience function in Equation (2). By Equation (8), a local thinker then prefers the riskless lottery L_2^{2400} provided:

$$-(0.01) \cdot 2400 + \delta \cdot (0.33) \cdot 100 < 0, \tag{12}$$

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

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

$$\sigma(2500, 0) > \sigma(0, 2400) > \sigma(2500, 2400) > \sigma(0, 0). \tag{13}$$

The first inequality follows from ordering, and the second from diminishing sensitivity. By Equation (8), a local thinker prefers the risky lottery L_1^0 provided:

$$(0.33) \cdot (0.66) \cdot 2500 - \delta \cdot (0.67) \cdot (0.34) \cdot 2400 + \delta^2 \cdot (0.33) \cdot (0.34) \cdot 100 > 0 \tag{14}$$

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

intuition: the common consequence is more valuable when associated with a sure rather than a risky prospect.

In sum, when $\delta < 0.73$ – which holds in the parameterization $\delta = 0.7$, $\theta = 0.1$ – a local thinker exhibits the Allais paradox. It is worth spelling out the exact intuition for this result. When $z = 2400$, the lottery L_2^{2400} is safe, whereas the lottery L_1^{2400} has a salient downside of zero. The agent focuses on this downside, leading to risk aversion. When instead $z = 0$, the downside payoff of the safer lottery L_2^0 is also 0. As a result, the lotteries' upsides are now crucial to determining salience. This induces the agent to overweight the larger upside of L_1^0, triggering risk seeking. The salience of payoffs endogeneizes the "certainty effect" as a form of context dependence: when the same downside risk is added to the lotteries, the sure prospect is particularly hurt because the common downside payoff induces the agent to focus on the larger upside of the risky lottery, leading to risk seeking choices.

This role of context dependence invites the following test. Suppose that subjects are presented the following correlated version of the lotteries L_1^z and L_2^z in Equation (10):

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

(15)

where the table specifies the possible joint payoff outcomes of the two lotteries and their respective probabilities. Correlation changes the state space *but not* a lottery's distribution over final outcomes, so it does not affect choice under either Expected Utility Theory or Prospect Theory. Critically, this is not true for a local thinker: the context of this correlated version makes clear that the state in which both lotteries pay z is the least salient one, and also that it drops from evaluation in Equation (8), so that the value of z should not affect the choice at all. That is, in our model – but not in Prospect Theory – the Allais paradox should not occur when L_1^z and L_2^z are presented in the correlated form as in (15).

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

	L_1^{2400}	L_2^{2400}
L_1^0	7%	9%
L_2^0	11%	73%

The vast majority of subjects do not reverse their preferences (80% of choices lie on the NW-SE diagonal), and most of them are risk averse, which in our model is also consistent with the fact that $(0, 2400)$ is the most salient state in the correlated choice problem (15). Among the few subjects reversing their preference, no clear pattern is detectable. This contrasts with the fact that our experimental subjects exhibit the Allais paradox when lotteries are presented in an uncorrelated form (see Appendix 2, Supplementary Material). Thus, when the lotteries pay the common consequence in the same state, choice is invariant to z and the Allais paradox disappears. Our model accounts for this fact because, as the common consequence z is made evident by correlation, it becomes non-salient. As a result, subjects prune it and choose based on the remaining payoffs.[16]

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

[16]We tested the robustness of the correlation result by changing the choice problem in several ways: 1) we framed the correlations verbally (e.g. described how the throw of a common die determined both lotteries' payoffs), 2) we repeated the experiment with uncertain real world events, instead of lotteries, and 3) we varied the ordering of questions, the number of filler questions, and payoffs. As the Appendix shows, our results are robust to all these variations. We also ran an experiment where subjects were explicitly presented the lotteries of Equation (10) with $z = 2400$ as uncorrelated, with a state space consisting of the four possible states. The choice pattern exhibited by subjects is: i) very similar to the one exhibited when the state space is not explicitly presented, validating our basic assumption that an agent assumes the lotteries to be uncorrelated when this is not specified otherwise, and ii) very different from the choice pattern exhibited under correlation (with 35% of subjects changing their choice as predicted by our model, see Appendix 2).

5.1.2 The "common ratio" Allais Paradox

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

$$L_1^{\pi'} = (6000,\ \pi';\quad 0,\ 1-\pi'),\quad L_2^{\pi} = (\alpha \cdot 6000,\ \pi;\quad 0,\ 1-\pi), \tag{16}$$

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

$$v(\alpha \cdot 6000) \geq \frac{\pi'}{\pi} \cdot v(6000) + v(0)\left(1 - \frac{\pi'}{\pi}\right). \tag{17}$$

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

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

Consider the choice between $L_1^{\pi'}$ and L_2^{π} in our model. For $\alpha = 1/2$ there are four states of the world, $S = \{(6000, 3000), (0, 3000), (6000, 0), (0, 0)\}$, and the salience ranking among them is

$$\sigma(6000, 0) > \sigma(0, 3000) > \sigma(6000, 3000) > \sigma(0, 0), \tag{18}$$

as implied by ordering and diminishing sensitivity. It is convenient to express the agent's decision as a function of the transformed probabilities of the lottery outcomes (as opposed to those of states of the world).[17] Denoting these transformed probabilities by $\widehat{\pi}'$ and $\widehat{\pi}$, we find that the local thinker evaluates the odds with which the riskier lottery $L_1^{\pi'}$ pays out relative to the safer one L_2^{π} as:

$$\frac{\hat{\pi}'}{\hat{\pi}} = \frac{\pi'}{\pi} \cdot \frac{(1-p) + p\delta^2}{(1-\pi')\delta + \pi'^2}. \tag{19}$$

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

$$\pi \geq \frac{2(1-\delta)}{2-\delta-\delta^2}. \tag{20}$$

As in the common ratio effect, the local thinker is risk averse when π is sufficiently high and risk seeking otherwise. In particular, for $\delta \in (0.22, 1)$, the local thinker switches from $L_2^{0.9}$ to $L_1^{0.01}$ just as experimental subjects do. The parameterization $\delta = 0.7$, $\theta = 0.1$ is thus consistent also with the common ratio effect.

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

Experimental evidence shows that this common ratio effect is also not robust to the introduction of correlation. KT (1979) asked subjects to choose between two lotteries of the type (18) in a two-stage game where in the first stage there is a 75% probability of the game ending without any winnings and a 25% change of going to stage two. In stage two, the lottery chosen at the outset is played out. The presence of the first stage is equivalent to

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

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

In explaining this behavior, KT informally argue that individuals “edit out” the correlated first stage state where both lotteries pay zero. Our model yields this editing as a consequence of the low salience and cancellation of such state. Adding a correlated state where both lotteries pay 0 neither affects the salience ranking in Equation (18) nor – more importantly – the odds ratios between states. As a result, the local thinker chooses as if he disregards the correlated state and its probability. This is what experimental subjects do.

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

5.2 Preference Reversals

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

To study preference reversals in our model, consider how a local thinker prices a lottery. As in Expected Utility Theory, the minimum selling price is the amount of money at which the local thinker is indifferent between receiving that amount and playing the lottery, namely the lottery's certainty equivalent. For a local thinker, this price is found by replacing the lottery's probabilities with decision weights, which in turn depend on the agent's perception of the choice alternatives. Formally, when choosing between lotteries L_1 and L_2, a local thinker with a value function $v(\cdot)$ prices L_1 at:

$$P(L_1 \,|L_2\,) = v^{-1}\left[\sum_{s\in S}\pi_s^1 v(x_s^1)\right], \tag{21}$$

where π_s^1 is the decision weights of state s for lottery L_1 in the context of its choice against lottery L_2. With a linear value function, the price $P(L_1|L_2)$ is the expected value of L_1 as perceived by the local thinker. If the agent is asked to price a lottery in isolation, we naturally assume that he evaluates it in the context of a choice between the lottery and the status quo of not having it $L_0 \equiv (0,1)$, i.e. of getting zero for sure.

Consider now preference reversals in our model. In the experiments, subjects are first asked to price in isolation, and then to choose among, the following two independent lotteries:

$$L_\$ = \begin{cases} x, & \text{with probability} \quad \pi' \\ 0, & \text{with probability} \quad 1-\pi' \end{cases}, \qquad L_\pi = \begin{cases} \alpha x, & \text{with probability} \quad \pi \\ 0, & \text{with probability} \quad 1-\pi \end{cases}, \tag{22}$$

where typically $\pi'/\pi = \alpha = 1/2$, as in the common ratio experiments. We know from (20) that, with linear utility, the local thinker selects the safer lottery L_π when $\pi > 2(1-\delta)/(2-\delta-\delta^2)$. In the literature, we typically have $\pi > 3/4$, so this constraint holds for any $\delta \geq 2/3$. Thus, when asked to choose, a local thinker having linear utility and $\delta = 0.7$ is risk averse and prefers L_π to $L_\$$, just as most experimental subjects do.

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

$$P(L_\pi\,|L_0\,) = \tfrac{x}{2} \cdot \tfrac{\pi}{\pi+(1-\pi)\delta}\ , \quad P(L_\$\,|L_0\,) = x \cdot \tfrac{\pi/2}{\pi/2+(1-\pi/2)\delta}. \tag{23}$$

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

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

Both lotteries are priced above their expected value, but $L_\$$ is more overpriced than L_π because it pays a higher gain with a smaller probability, and from Proposition 1 we know that lower probabilities are relatively more distorted.[18]

Thus, while in a choice context the local thinker prefers the safer lottery L_π, in isolation he prices the risky lottery $L_\$$ higher, exhibiting a preference reversal. Crucially, this behavior is not due to the fact that choosing and pricing are different operations. In fact, in our model choosing and pricing are the same operation, as in standard economic theory. Preference reversals occur because, unlike in standard theory, evaluation in our model is context dependent. Pricing and choosing occur in different contexts because the alternatives of choice are different in the two cases. One noteworthy feature of our model is that it generates preference reversals through violations of "procedural invariance", defined by Tversky et al. (1990) as situations in which a subject prices a lottery above its expected value, $P(L_1|L_0) > \mathbb{E}(x_s^1)$, and yet prefers the expected value to the lottery, $L_1 \prec (\mathbb{E}(x_s^1), 1)$. Tversky et al (1990) show that the vast majority of observed reversals follow from the violations of procedural invariance, as predicted by our model. Regret Theory can produce preference reversals by a distinct mechanism, intransitivities in choice, but does not violate procedural invariance.

One distinctive implication of our context-based explanation is that reversals between choice and pricing should only occur when pricing takes place in isolation but not if agents price lotteries in the choice context itself. We tested this hypothesis by giving subjects a choice between lotteries $L_\$ = (16, 0.31;\ 0, 0.69)$ and $L_\pi = (4, 0.97;\ 0, 0.03)$, which Tversky et

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

al (1990) found to lead to a high rate of preference reversals. Subjects stated their certainty equivalents for the two lotteries, in isolation and in the context of choice.[19] Our model then predicts that preference reversal should occur between choice and pricing in isolation, but not between choice and pricing in the choice context.

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

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

Thus, our subjects pool exhibits the standard preference reversal between choice and average pricing in isolation.[20]

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

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

As predicted by our model, in the choice context the average price ranking is consistent with choice.[21] One may object that this agreement is caused by the subjects' wish to be coherent

[19]In our experimental design, each subject priced each lottery only once, and different lotteries were priced in different contexts. This design ensures that subjects do not deform their prices to be consistent with their choices; however, it also implies that preference reversals are not observed within-subject but only at the level of price distributions across subject groups (see Appendix 2 for more details).

[20]This reversal holds not only with respect to average prices but also for the distribution of prices we observe. Assuming that subjects draw evaluations randomly from the price distributions, we estimate that around 54% of the subjects who choose L_π would exhibit the standard preference reversals (see Appendix 2). The average prices above imply that some subjects priced the safer lottery L_π above its highest payoff. Such overpricing can occur even in a laboratory setting and with incentives schemes (Grether and Plott 1974, Bostic et al 1990), perhaps due to misunderstanding of the pricing task. In Appendix 2 we consider truncations of the data that filters out such overpricing.

[21]In our data, the distribution for $P(L_\pi|L_\$)$ does not dominate that for $P(L_\$|L_\pi)$. This is due to the fact that: i) on average subjects attribute similar values to both lotteries in the choice context, and ii) there is substantial variability in choice (and thus in pricing), as about half the subjects chose each lottery. In Appendix 2 we look in a more detailed way at the manifestation and significance of fact ii) in light on Tversky's et al. (1990) analysis of preference reversals.

when they price just after a choice. However, each subject priced only one of the lotteries in the choice context.[22] It appears to be the act of comparing the lotteries that drives their evaluation during choice, and not (only) an adjustment of value subsequent to choice.

Another potential objection is that our experiments do not elicit true selling prices. It is well known that it is difficult to design price elicitation mechanisms for subjects who violate the independence axiom of Expected Utility Theory. To avoid these problems, Cox and Epstein (1989) study preference reversals by only eliciting the *ranking* of selling prices across lotteries. In their experiments, Cox and Epstein directly compared lotteries to each other, so their procedure can be viewed as eliciting evaluations in the context of choice. They find some evidence of preference reversals, but crucially they show that these reversals are equally likely in both directions (from risk averse choice to risk seeking pricing, and from risk seeking choice to risk averse pricing). Symmetric reversal patterns are typically attributed to arbitrary fluctuations in evaluation, see Bostic et al (1990). Thus we interpret Cox and Epstein's results as consistent with our predictions that systematic preference reversals subside when prices are elicited in a choice context.

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

6 Choice Among Many Lotteries, with an Application to Asset Pricing

6.1 Setup and Definitions

We now extend our model to choice among $N \geq 2$ of lotteries. Before doing so, note that salience in a general choice problem cannot be inferred from that of pairwise comparisons, since salience – and thus evaluation – will generally change for each pairwise comparison. Pairwise intransitive preferences may even arise in some cases, as we show in Section 7.

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

To model choice from an arbitrary set of alternatives, which is particularly useful for economic applications, we need to generalize the notion of salience to a general choice set. Suppose that the local thinker is faced with a state space S and a choice set $\aleph = \{L_1, \ldots, L_N\}$ of lotteries defined over S, as in Section 3. Let $x_s = (x_s^1, \ldots, x_s^N)$ be the vector of payoffs delivered in a generic state s, and denote by $x_s^{-i} = \{x_s^j\}_{i \neq j}$ the vector of payoffs excluding x_s^i. The salience of state s for lottery L_i is then captured by a function $\widehat{\sigma}(x_s^i, x_s^{-i})$ which contrasts L_i's payoff x_s^i in s with all other payoffs $x_{\mathbf{s}}^{-i}$ in the same state. Let $x_s^{-i} + \epsilon$ denote the vector with elements $\{x_s^j + \epsilon\}_{j \neq i}$. In line with Definition 1, we impose the following properties:

Definition 3 *Given a state space S and a choice set $\aleph$, the salience of state s for lottery L_i is given by a continuous and bounded function $\widehat{\sigma}(x_s^i, \mathbf{x}_s^{-i})$ that satisfies three conditions:*
1) *Ordering: if $x_s^i = \max \mathbf{x}_s$, then for any $\epsilon, \epsilon' \geq 0$ (with at least one strict inequality):*

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

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

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

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

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

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

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

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

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

where $\sigma(.,.)$ is the salience function employed in the two lottery case of Section 3, and $\mathbf{f}(\mathbf{x}_s^{-i}) : R^{N-1} \rightarrow R$ is a function of the residual vector $\mathbf{x}_s^{-i}$. One intuitive specification which, together with Definition 1, fulfills Definition 3 is:

$$\mathbf{f}(\mathbf{x}_s^{-i}) = \frac{1}{N-1} \sum_{j \neq i} x_s^j, \tag{24}$$

in which case the salience of a state for a lottery depends on the contrast between the lottery's payoff and the average of the other lotteries' payoffs in s. Even if $\sigma(.,.)$ is symmetric, salience is in general not symmetric in the sense that the same state may have different salience for different lotteries. For instance, a state where lottery L_i's payoff is very different from the payoffs of all the other L_j's (which in turn are similar to each other) is very salient for L_i but not salient for the L_j's. However, the same state may be very salient for all lotteries, if for example the lotteries' payoffs in that state are equally divided into two very different values.

Given a lottery specific salience ranking k_s^i based on the salience function $\widehat{\sigma}$, each state is assigned a decision weight π_s^i according to Equation (4), and a value $V^{LT}(L_i)$ is computed for each lottery L_i according to Equation (6).

One important new effect arises in our model of choice among $N > 2$ lotteries. Specifically, the preference ranking among any two lotteries depends not only on the contrast between their payoffs but also on the remaining alternatives. The choice set is a source of context effects. To see how this can arise, consider the pairwise choice between a sure prospect $L_1 = (x, 1)$ and a risky lottery $L_2 = (x_h, \pi_h;\ x_l, 1 - \pi_h)$ where $x_h > x > x_l > 0$, $\pi_h < 1/2$ and $\mathbb{E}(x_s) = x$. Using the salience function (2) with $\theta = 0$, state $s_l = (x, x_l)$ is salient when:

$$x_h \cdot x_l < x^2. \tag{25}$$

In this case, which we assume throughout, the lottery's downside is salient and the sure prospect L_1 is chosen from the choice set $\{L_1, L_2\}$.

Suppose now that we add to this choice set another (correlated) risky lottery $L_3 = (x + y, \pi_h;\ x - y, 1 - \pi_h)$ whose payoffs are spread by an amount $y > 0$ relative to the sure prospect L_1. Before examining the impact of L_3 on the preference between L_1 and L_2, note that in a pairwise comparison L_1 is always preferred to L_3. This is because by diminishing sensitivity the spread y is more salient as a loss than as a gain.

Consider now the agent's choice from the enlarged choice set $\{L_1, L_2, L_3\}$. If the axiom of independence of irrelevant alternatives holds, the agent's preference of L_1 over L_2 should not be affected by the presence of L_3. In our model, however, this does not need to be so. To see this, note that when evaluating L_2 in the enlarged choice set, the salience of the state s_h is now computed by contrasting the lottery's payoff x_h with the average payoff $x + y/2$ of L_1 and L_3 in the same state s_h, while the salience of the state s_l in which L_2 loses now contrasts the lottery's payoff x_l with $x - y/2$. Consider the most interesting case where lottery L_2 is still risky relative to the average lottery in the choice set, in the sense that $x_h > x + y/2$ and $x_l < x - y/2$. It is then easy to show that the loss state s_l remains salient for L_2 provided:

$$x_h \cdot x_l < x^2 - \frac{y^2}{4}. \tag{26}$$

Thus, in the range $x_h x_l \in (x^2 - y^2/4, x^2)$ the local thinker prefers L_1 to L_2 in a pairwise comparison but prefers L_2 to L_1 when L_3 is added to the choice set. Intuitively, when L_3 is added, the downside x_l of the original risky lottery L_2 looks less bad, relative to the alternatives. Although L_2's upside x_h also looks less good, by the convexity property the former effect is stronger. In other words, the inclusion of an "unfavourable" risky lottery L_3 boosts the risk tolerance of the agent by making the risky lottery L_2 look better by comparison. This can induce the agent to switch his choice to L_2 over L_1. The general analysis of these so-called "decoy" effects is beyond the scope of this paper, and can be found in Bordalo (2011). An interesting implication of the model is that such shifts in salience (and resulting violations of the axiom of independence of irrelevant alternatives) should subside when the choice set is large (i.e. N is large).

Such choice set effects imply that manifestly dominated lotteries may affect salience and thus evaluation. As shown in Bordalo (2011), one can even construct fine-tuned examples in which a dominated lottery is overvalued relatively to, and even chosen over, a dominating lottery. In reality, however, the agent quickly realizes that the dominating option is more valuable: what is salient then is the dominance relation between lotteries. To deal with this feature while keeping the model tractable, we have assumed in Section 3.2 that the agent edits the choice set by immediately identifying, and discarding, dominated lotteries. (This is reminiscent of the editing out of dominated lotteries in Prospect Theory.)

6.2 An application of salience to asset pricing

In this subsection, we show how our model of salience can be included in a very standard asset pricing model in a way that accounts for some puzzling existing evidence. An investor living for two periods $t = 0, 1$ decides at $t = 0$ how to allocate his wealth w between current and future consumption by investing in a measure 1 of assets, each available in net supply of 1. The investor values consumption according to a concave utility function $u(c)$ ($u' > 0$, $u'' \leq 0$) and there is no time discounting. Each unit of asset $i \in [0, 1]$ costs p_i units of current consumption and yields at $t = 1$ a dividend x_s^i in state $s \in S$. We denote by $x_s = \int_i x_s^i \cdot di$ the aggregate payoff of all assets in $s \in S$. In line with Equation (24), we define the salience of a state s for a generic asset i as a function of the asset's payoff x_s^i in that state and the average market payoff x_s in the same state, denoting it by:

$$\sigma(x_s^i, x_s), \tag{27}$$

where $\sigma(.,.)$ satisfies Definition 1. Since each asset captures an infinitesimal market share, all assets are compared to the same market benchmark x_s. Equation (27) implies that a state is salient for an asset if in that state the asset's payoff "stands out" relative to the market payoff. We then adopt the following definition:

Definition 4 *At prices $(p_i)_{i \in [0,1]}$ an equilibrium portfolio for a local thinker consists of a measure one of asset holdings $(\alpha_i)_{i \in [0,1]}$ such that:*
1) *The portfolio $(\alpha_i)_{i \in [0,1]}$ is feasible, namely $\int_i \alpha_i \cdot p_i \cdot di \leq w$, and*
2) *The portfolio $(\alpha_i)_{i \in [0,1]}$ is preferred to any portfolio obtained from it by a small deviation along the holding α_j of asset j, for any $j \in [0, 1]$. The deviation for each asset j is evaluated in light of that asset's salience weighting.*

Definition 4 extends the standard theory of portfolio choice to the case of a local thinker. Just as in our prior analysis the local thinker evaluated the expected gain (relative to the status quo) of accepting a lottery in light of the lottery's salient states, condition 2) states that when deciding of whether to buy an extra unit of an asset the investor evaluates the incremental gain so obtained in light of that asset's salience ranking. Additionally, just as the salience of a lottery is shaped by the choice set, according to Equation (27) the salience of an asset is determined by its comparison with the other assets in the market (and not with the investor's status quo portfolio). The market is the key source of context dependence in our specification.

To see what Definition 4 implies for the investor's optimal choice, consider a specific asset $j \in [0,1]$ and let $(\alpha'_i)_{i\in[0,1]}$ be a portfolio which coincides with $(\alpha_i)_{i\in[0,1]}$ along the holdings of every asset $i \neq j$, but is not constrained in its holding of asset j. By condition 2), $(\alpha_i)_{i\in[0,1]}$ is an equilibrium portfolio provided:

$$\alpha_j = \arg \max_{\alpha'_j | \alpha'_i = \alpha_i} \left\{ u \left(w - \int_{i\in[0,1]} \alpha'_i p_i di \right) + \mathbb{E}_s \left[\omega_{s,j} \cdot u \left(\int_{i\in[0,1]} \alpha'_i x^i_s di \right) \right] \right\}, \text{ for all } j \in [0,1] \,. \tag{28}$$

This is the standard utility maximization condition except for the fact that a deviation along the holding of each asset j is evaluated using the asset-specific decision weights $\omega_{s,j}$. From (28), the first order condition for the investor is:

$$p_i \cdot u'(c_0) = \mathbb{E}\left[\omega^i_s \cdot x^i_s \cdot u'(c_{1,s})\right], \text{ for all } i \in [0,1] \,, \tag{29}$$

where $c_0 = w - \int \alpha_i p_i$ and $c_{1,s} = \int \alpha_i \cdot x^i_s$ are the consumption levels at each time and in each state. When deciding whether to buy an extra unit of asset i, the investor realizes that the cost of doing so is the utility value of the consumption p_i forsaken at $t = 0$, while the expected benefit is the utility value of earning x^i_s in each future state s. Equation (29) departs from the standard Euler equation because the benefit x^i_s of the asset is weighted by its salience ω^i_s. When thinking about an asset, the investor is drawn to over-value its salient payoffs and under-value its non-salient payoffs. The investor will then seek to buy a higher quantity of an asset whose upside is salient, especially if such upside occurs in a state where the marginal utility of consumption $u'(c_{1,s})$ is high.

Consider the implications of this model for asset prices. In general equilibrium, the investor must hold the market portfolio, namely $\alpha_i = 1$ for all i, $c_{1,s} = x_s$ for all s, and

$c_0 = w - \int_i p_i \cdot di > 0$. Suppose that at $t = 1$ there is no aggregate risk so that the average payoff is $x_s = x$ for all s (and thus $c_{1,s} = c_1 = x$ for all s). Then, for any utility function $u(c)$, the investor described by (29) is willing to hold the market portfolio provided:

$$p_i = \frac{\mathbb{E}\left[x_s^i\right]}{R} + \frac{\text{cov}\left[\omega_s^i, x_s^i\right]}{R}, \text{ for all } \; i \in [0,1]\,, \tag{30}$$

where $R = u'(c_0)/u'(c_1)$ is the return on the riskless asset. The first term on the right hand side of (30) is the price prevailing when the agent is fully rational (i.e. when $\delta = 1$): in the absence of aggregate risk, the investor is risk neutral at the margin, valuing each asset at its expected discounted dividend. Indeed, for a diversified investor, standard risk aversion is second order, leaving prices unaffected.

Relative to an Expected Utility maximizer, the local thinker over or under values an asset by the second term on the right hand side, which increases in the covariance between the asset's payoffs and their salience. Specifically, the local thinker overvalues an asset – exhibiting risk seeking behavior – when the asset's highest payoffs are salient, while he undervalues an asset, exhibiting risk averse behavior, when the asset's lowest payoffs are salient. By shaping the agent's focus on specific asset payoffs, salience creates a *first order source* of risk attitudes.

To illustrate how this mechanism works, consider the well-known empirical finding in the cross-section of stock returns, namely the fact that value stocks – those with low stock market values relative to measures of "fundamentals" such as assets or earnings – earn higher average returns than growth stocks, those with high market values relative to measures of fundamentals (Fama and French, 1993, Lakonishok, Shleifer and Vishny, 1994). Consider two stocks g and v characterized by the dividend structure:

$$x_s^g = \begin{cases} x^g & s \in S^g \\ l & s \in S/S^g \end{cases} \qquad x_s^v = \begin{cases} x^v & s \in S^v \\ l & s \in S/S^v \end{cases}$$

where $x^g > x^v > x > l$ (where x is the average market payoff) and $\Pr(s \in S^g) = \pi_g < \Pr(s \in S^v) = \pi_v$. Asset g is a "growth stock", delivering a large above market payoff x^g with small probability π_g and a below market payoff l with a high probability $1-\pi_g$. Asset v is a "value stock", yielding a small above market payoff x^v with high probability π_v and a below market payoff l with low probability $1-\pi_v$. Suppose now that the salience of different states satisfies:

$$\sigma(x^g, x) > \sigma(l, x) > \sigma(x^v, x). \tag{31}$$

That is, the upside of the growth stock stands out relative to the market, while for the value stock the downside stands out. Using the salience function of Equation (2), this condition is met when $x^g l > x^2 > x^v l$, which requires the upside x^v of the more stable value stock to be sufficiently close to the market average x relative to the upside x^g of the growth stock. In this case, the investor thinks of the growth stock as an opportunity to obtain a large windfall while he magnifies the downside risk of the value stock. This agent (partly) neglects the fact that the growth stock has an objectively higher probability of a low payoff because its upside x^g is sufficiently salient to catch the agent's attention. Equation (30) then yields equilibrium prices given by:

$$p_g = \frac{\mathbb{E}\left[x_s^g\right]}{R} + \frac{\pi_g(1-\pi_g)(1-\delta)}{R} \cdot \left(x^g - l\right), \tag{32}$$

$$p_v = \frac{\mathbb{E}\left[x_s^v\right]}{R} - \frac{\pi_v(1-\pi_v)(1-\delta)}{R} \cdot \left(x^v - l\right). \tag{33}$$

The growth stock is over-valued and the value stock is under-valued, the more so the lower is δ.

This implication of the model is consistent with the empirical evidence we already mentioned, but it goes further than that. Fama and French (1992,1993) have conjectured that the reason that value stocks earn higher average returns is that they are disproportionately exposed to a separate risk factor, which they referred to as distress risk. Subsequent research, however, has not been able to find evidence that value stocks are particularly risky (Lakonishok et al, 1994). Furthermore, Campbell, Hilscher and Szilagyi (2008) find that stocks of companies vulnerable to the risk of bankruptcy earn if anything lower average returns, contradicting the Fama-French view that "value" reflects bankruptcy risk. Our model of salience might help explain what is going on. It suggests that while value stocks are not fundamentally riskier, the possibility of their bankruptcy (or a very low payoff) is salient to the investors, and as a consequence value stocks are underpriced. The model thus puts together the Fama-French idea that investors fear bankruptcy of value stocks with the empirical observation that this possibility is salient and thus exaggerated, so value stocks are indeed underpriced.

More generally, this example shows that the extent to which certain asset payoffs "stand out" relative to the market may cause – through salience – distortions in the perception of asset specific risks and thus of asset prices, for instance helping to explain why right-skewed assets tend to be overvalued. This principle may also imply that, precisely by reducing right-skewness, a diversified basket of stocks could also be relatively undervalued. We leave further analysis of the impact of salience on asset prices to future work.

7 Other applications and extensions

7.1 Reflection and Framing Effects

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

$$V^{LT}(L_1) = \sum_{s \in S} \pi_s \omega_s^1 x_s^1 = \mathbb{E}(x_s^1) + \text{cov}[\omega_s^1, x_s^1] \tag{34}$$

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

$$\text{cov}\left[\omega_s^1, -x_s^1\right] = -\text{cov}\left[\omega_s^1, x_s^1\right] > 0, \tag{35}$$

which is fulfilled if and only if the agent was originally risk averse. Intuitively, a salient downside inducing risk aversion in the gain domain becomes a salient upside inducing risk seeking in the loss domain. Our model thus yields the fourfold pattern of risk preferences[23] without assuming, as Prospect Theory does, a value function that is concave for gains and convex for losses. Reflection of payoffs generates shifts in risk attitudes by inducing the agent to shift his attention from the lottery upside to its downside and vice versa. The same logic shows our model can account for the KT's (1981) famous framing experiments and the Public Health Dilemma, even with a linear value function.

7.2 Intransitivity of pairwise preferences

Intransitivities may arise in our model, but their occurrence rests on a delicate balance between probabilities, payoffs and degree of local thinking δ. In certain classes of cases, such a balance does not exists. For example, intransitivities never occur in choices among independent lotteries sharing the same support with fewer than 4 outcomes. To illustrate how intransitive preferences may arise in our model, consider the following three lotteries:

$$L_\pi = \begin{cases} \alpha x, & \pi \\ 0, & 1-\pi \end{cases}, \quad L_\$ = \begin{cases} x, & \alpha\pi \\ 0, & 1-\alpha\pi \end{cases}, \quad L_s = (y, 1), \tag{36}$$

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

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

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

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

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

7.3 Mixed Lotteries

We now apply our model to mixed lotteries, those involving both positive and negative payoffs. To this end, we come back to the KT (1979) piecewise linear value function exhibiting loss aversion, for loss aversion provides an intuitive explanation for risk aversion with respect to small mixed bets. Using the salience function of Equation (2), for which $\sigma(x, y) = \sigma(-x, -y)$ for all x, y, all risk aversion for lotteries symmetric around zero is due to loss aversion. For non-symmetric lotteries, salience and loss aversion interact to determine risk preferences. To see this, consider Samuelson's wager, namely the choice between the lotteries:

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

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

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

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

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

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

The agent rejects L_S when his dislike for losses more than compensates for his focus on the lottery gain, i.e. $\lambda > 2/\delta$.[25] In lotteries where the negative downside is larger than the positive upside, salience and loss aversion go in the same direction in triggering risk aversion.

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

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

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

This condition relaxes the symmetry around zero of the salience function of Equation (2) represented in Figure 1, postulating that departures from zero are more salient in the negative than in the positive direction. In this specification, local thinking can itself be a force towards risk aversion for mixed lotteries, complementing loss aversion. In particular, if losses are sufficiently more salient than gains, one can account for Samuelson's wager based on salience alone (and linear utility): if $\sigma(-100, 0) > \sigma(200, 0)$, a local thinker with linear utility rejects Samuelson's bet as long as $200 \cdot \frac{\delta}{1+\delta} - 100 \cdot \frac{1}{1+\delta} < 0$, or $\delta < 1/2$. A specification where risk aversion for mixed lotteries arises via the salience of lottery payoffs may give distinctive implications from standard loss aversion, but we do not investigate this possibility here.[26]

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

[26] If we endow the local thinker with a standard utility function, instead of a value function, then in the absence of property 5) the utility function would be subject to Rabin's critique (Rabin, 2000) in the domain

8 Conclusion

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

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

Our specification of contrast as a driver of salience could be useful for thinking about a variety of economic situations. We have discussed an application to asset pricing and the growth/value anomaly, but other misperceptions in finance might also be influenced by

of mixed lotteries (but not in the domain of positive lotteries). Adding property 5) would entail that aversion to mixed lotteries with positive payoffs follows from salience, and not from underlying preferences. Thus, even though such a local thinker is at heart an expected utility maximizer, he is immune to Rabin's critique.

salience. Salience may also affect consumer behavior: when considering which of different brands to buy, a consumer might focus on the attributes where the potential brands are most different, neglecting the others (see Tversky and Simonson 1993). Bordalo (2011) and Koszegi and Szeidl (2011) use a version of our model of salience to investigate this issue. In many applications, the key idea of our approach is that mental frames, rather than being fixed in the mind of the consumer, investor, or voter, are endogenous to the contrasting features of the alternatives of choice. This notion could perhaps provide a way to study how context shapes preferences in many social domains.

Appendix 1.

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

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

Proof. By definition,

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

Therefore,

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

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

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

Thus

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

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

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

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

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

■

References

Aizpurua, Jose, Jorge Nieto and Jose Uriarte, "Choice procedure consistent with similarity relations." *Theory and Decision*, 29 (1990): 235–254.

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

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

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

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

Birnbaum, Michael and Ulrich Schmidt, "Allais paradoxes can be reversed by presenting choices in canonical split form ", mimeo, Kiel University 2010.

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

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

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

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

Campbell, John, Jens Hilscher and Jan Szilagyi, "In search of distress risk.", *Journal of Finance*, 63(2008): 2899–2939.

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

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

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

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

Fama, Eugene F. and Kenneth R. French, "The cross section of expected stock returns." *Journal of Finance*, 47(1992): 427–465.

Fama, Eugene F. and Kenneth R. French, "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33 (1993): 3–56.

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

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

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

Gabaix, X., "A Sparsity-Based Model of Bounded Rationality." NBER working paper 16911, 2011.

Gennaioli, Nicola, and Andrei Shleifer, "What comes to mind." *Quarterly Journal of Economics*. 125 (2010): 1399–1433

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

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

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

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

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

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

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

Koszegi, Botond, and Adam Szeidl, "A model of focusing in economic choice."mimeo, University of California, Berkeley, 2011.

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

Lakonishok, Josef, Andrei Shleifer and Robert W. Vishny, "Contrarian investment, extrapolation, and risk." *Journal of Finance*, 49 (1994): 1541–1578.

Leland, Jonathan, "Generalized similarity judgments: an alternative explanation for choice anomalies." *Journal of Risk and Uncertainty*, 9 (1994): 151–172.

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

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

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

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

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

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

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

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

Savage, Leonard J., “Foundations of Statistics ”, (2nd revised edition, 1972). New York: Dover Publications.

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

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

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

Tversky, Amos. 1969. “Intransitivity of Preferences.” *Psychological Review*, 76(1): 31–48.

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

——, “Advances in Prospect Theory: Cumulative representation of uncertainty.” *Journal of Risk and Uncertainty*, 5 (1992): 297–323.

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

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

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