Risk and Rationality: The Relative Importance of Probability Weighting and Choice Set Dependence

Adrian Bruhin* Maha Manai Luís Santos-Pinto

University of Lausanne
Faculty of Business and Economics (HEC Lausanne)

July 23, 2018

Abstract

We analyze the relative importance of probability weighting and choice set dependence in describing risky choices both non-parametrically and with a structural model. Our experimental design uses binary choices between lotteries that may trigger Allais Paradoxes. We change the choice set by manipulating the correlation structure of the lotteries' payoffs while keeping their marginal distributions constant. This allows us to discriminate between probability weighting and choice set dependence. There are three main results. First, probability weighting and choice set dependence both play a role in describing aggregate choices. Second, the structural model uncovers substantial individual heterogeneity which can be parsimoniously characterized by three types: 38% of subjects engage primarily in probability weighting, 34% are influenced predominantly by choice set dependence, and 28% are mostly rational. Third, the classification of subjects into types predicts preference reversals out-of-sample. These results may not only further our understanding of choice under risk but may also prove valuable for describing the behavior of consumers, investors, and judges.

KEYWORDS: Individual Choice under Risk, Choice Set Dependence, Probability Weighting, Latent Heterogeneity, Preference Reversals

JEL CLASSIFICATION: D81, C91, C49

Acknowledgments: We are grateful for insightful comments from the participants of the research seminars at Ludwig Maximillian University of Munich, NYU Abu Dhabi, University of Lausanne, and University of Zurich, as well as the participants of the Economic Science Association World Meeting 2018, and the Frontiers of Utility and Risk Conference 2018. All errors and omissions are solely our own. This research was supported by grant #152937 of the Swiss National Science Foundation (SNSF).

*Corresponding author: Bâtiment Internef 540, University of Lausanne, CH-1015 Lausanne, Switzerland; adrian.bruhin@unil.ch

1 Introduction

The past decades of mostly experimental economic research on choice under risk have revealed systematic violations of expected utility theory (EUT; von Neumann and Morgenstern, 1953). Important examples of EUT violations fall into two categories. First, as exposed in the famous Allais Paradoxes, most subjects tend to exhibit both both risk loving and risk averse behavior (Allais, 1953). This category of EUT violations contradicts EUT's independence axiom. Second, as demonstrated by Lichtenstein and Slovic (1971) and Lindman (1971), many subjects revert their preference when they have to choose between two lotteries or evaluate them in isolation. Cox and Epstein (1989) and Loomes et al. (1991) later showed experimentally that some forms of preference reversals contradict EUT's transitivity axiom. These and other systematic violations of EUT have spurred the development of various alternative decision theories.

A major class of alternative decision theories uses probability weighting to describe violations of the independence axiom. In these theories, subjects systematically overweight small probabilities and underweight large probabilities. Consequently, subjects may display risk loving behavior when buying a state lottery ticket and risk averse behavior when buying damage insurance, because they overweight the small probability of winning the state lottery and underweight the large probability of not suffering any damage. Prominent examples of this class of theories are Prospect Theory (Kahneman and Tversky, 1979), subsequently generalized to Cumulative Prospect Theory (CPT; Tversky and Kahneman, 1992), as well as Rank Dependent Utility (RDU; Quiggin, 1982).[1] These theories mainly differ in the way they account for probability weighting. For instance, RDU is silent about the origin of probability weighting, while in CPT, it directly results from reference-dependence and the Weber-Fechner law implying diminishing sensitivity away from reference points.

However, probability weighting cannot explain violations of the transitivity axiom. Subjects never revert their preference, since they always attach the same value to lotteries, regardless whether they have to choose among them or evaluate them in isolation.[2]

[1]When lottery payoffs are non-negative – as in this study – and subjects derive utility from lottery payoffs rather than absolute wealth levels, CPT and RDU coincide. Another example of a theory based on probability weighting is the model by Gul (1991) of disappointment aversion which belongs to the Chew-Deckel class of betweenness-respecting models (Deckel, 1986; Chew, 1989). For a detailed discussion see Fehr-Duda and Epper (2012).

[2]An extended version of CPT with an endogenous reference point allows for violations of transitivity

Another major class of decision theories postulates that the evaluation of lotteries is choice set dependent. This allows these theories to describe violations of the transitivity axiom and, under certain conditions, also of the independence axiom. Prominent members of this class are Salience Theory of Choice Under Risk (ST; Bordalo et al., 2012b) and Regret Theory (RT; Loomes and Sugden, 1982).[3] These theories have in common that, when subjects evaluate lotteries, they focus their limited attention on states of the world with large payoff differences between the alternatives. Hence, a lottery's value is choice set dependent as the weight attached to a state depends on the payoffs of the alternatives in that state. The main difference between these theories is how they operationalize choice set dependence. For example, ST respects diminishing sensitivity, as a given payoff difference renders a state less salient the further away it is from the reference point. In contrast, RT assumes that subjects use a convex regret function to evaluate lotteries with non-negative payoffs. Thus, they overweight states with payoff differences located further away from the reference point of zero – meaning that RT is at odds with diminishing sensitivity.[4]

Like probability weighting, choice set dependence can also explain why subjects sometimes display both risk loving and risk averse behavior. However, the intuition is different. Subjects tend to buy state lottery tickets because they overweight the state where they win the big prize due to the large payoff difference between winning the big prize and not buying the ticket. At the same time, they may buy damage insurance, because they overweight the state in which the damage occurs due to the large payoff difference between being insured and uninsured in that particular state.

These two major classes of decision theories often make similar predictions. Nevertheless, discriminating between them is important to better understand the behavior of various economic agents, such as investors, consumers, and judges. For example, in contrast to probability weighting, choice set dependence can naturally explain the counter-cyclicality of risk premia on financial markets (Bordalo et al., 2013a) and important behavioral phenomena

(Schmidt et al., 2008). However, when subjects consider lotteries with non-negative payoffs and derive utility from lottery payoffs rather than absolute wealth levels, the reference point is assumed to be exogenous and equal to zero (Tversky and Kahneman, 1992). In that case, CPT cannot explain preference reversals.

[3]Other examples of choice set dependent theories are by Rubinstein (1988); Aizpurua et al. (1990); Leland (1994); and Loomes (2010).

[4]We focus on ST in the present paper as the main example of a choice set dependent theory since it respects diminishing sensitivity and fits the aggregate choices in our dataset much better than RT (see Appendix A).

in consumer choices (Bordalo et al., 2012a, 2013b; Dertwinkel-Kalt et al., 2017) and judicial decisions (Bordalo et al., 2015). However, as it will become clear later, choice set dependence can describe violations of the independence axiom and, thus, the Allais Paradoxes only under some specific conditions. Hence, it is crucial to know the extent to which probability weighting and choice set dependence drive subjects' risky choices.

We address this question with a laboratory experiment which allows us to discriminate between probability weighting and choice set dependence while controlling for EUT. First, we provide non-parametric evidence at the aggregate level, i.e. at the level of a representative decision maker. Second, we account for heterogeneity in a parsimonious way by estimating a structural model which allows us to classify each subject into a type based on the decision theory that best describes her choices. Third, we perform out-of-sample predictions to assess the validity of this classification of subjects into types.

To discriminate between probability weighting and choice set dependence, the experiment uses a series of incentivized binary choices between lotteries that may trigger Allais Paradoxes. Every subject makes each binary choice twice. In one case, the two lotteries' payoffs are independent of each other, while in the other, they are perfectly correlated. Note that this manipulation of the correlation structure affects the joint payoff distribution of the two lotteries but not their marginal payoff distributions. Hence, if choices are driven by probability weighting, the predicted frequency of Allais Paradoxes is the same, as subjects evaluate each lottery in isolation and focus exclusively its marginal payoff distribution. However, if choices are driven by choice set dependence, the predicted frequency of Allais Paradoxes is different when payoffs are independent than when they are perfectly correlated due to the change in the joint payoff distribution.[5] As in Bordalo et al. (2012b), this design allows us to reliably discriminate between probability weighting and choice set dependence. Moreover, since EUT can never account for Allais Paradoxes, the design also enables us to control for EUT preferences.

To ensure that our results do not rely on a specific visual presentation of the binary choices, the experiment uses two presentation formats. Half of the subjects confront the "canonical presentation" while the other half confront the "states of the world presentation". In the canonical presentation, the two lotteries in a binary choice are represented

[5] As explained in detail in Section 3, when choice set dependence is the sole driver of risky choices, the predicted frequency of Allais Paradoxes is positive with independent payoffs and zero with perfectly correlated payoffs.

separately with distinct payoff distributions when payoffs are independent, and by their joint payoff distribution when payoffs are perfectly correlated. In contrast, in the states of the world presentation, the two lotteries are always represented by their joint payoff distribution, regardless whether payoffs are independent or perfectly correlated.[6] Ideally, our results should not depend on the presentation format.

To estimate the structural model and classify subjects into types, the lotteries' payoffs and probabilities vary systematically across the binary choices. Estimating the structural model and taking heterogeneity into account in a parsimonious way is important for two reasons. First, in order to make predictions about the subjects' choices in other risky situations one needs to know the decision theories and the corresponding parameters that mainly drive the subjects' behavior. Second, previous research uncovered substantial heterogeneity in risk attitudes (Hey and Orme, 1994; Harless and Camerer, 1994; Starmer, 2000), with a majority of non-EUT-types and a minority of EUT-types (Bruhin et al., 2010; Conte et al., 2011). This heterogeneity must be taken into account when testing the relative importance of different decision theories and making behavioral predictions – in particular in strategic settings where even small minorities can determine the aggregate outcome (Haltiwanger and Waldman, 1985, 1989; Fehr and Tyran, 2005).

Our structural model accounts for individual heterogeneity in a parsimonious way by using a finite mixture approach. That is, instead of estimating individual-specific parameters – which are typically noisy, hard to summarize in a concise way, and may suffer from small sample bias – the structural model assumes the population to be made up by three distinct types: EUT-types who are rational, CPT-types whose behavior is mostly driven by probability weighting, and ST-types whose behavior is predominantly driven by choice set dependence. Upon estimating the three types' relative sizes and their average type-specific parameters, we can classify every subject into the type that best fits her choices. This yields a parsimonious account of the relative importance of rational behavior, probability weighting, and choice set dependence. Furthermore, it also allows us to make type-specific predictions about the subjects' behavior in other domains of choice under risk.

For such behavioral predictions across domains to be meaningful, the subjects' classification of subjects into types and the estimated parameters need to reflect subjects' behavior – at least qualitatively – not only in-sample but also out-of-sample. To address this point,

[6]For screenshots illustrating the two presentation formats, see Figures 1 and 2 in Section 4.

the experiment exposes subjects to additional lotteries that may trigger preference reversals. Subjects always first choose between two of these additional lotteries and, later, evaluate each of them in isolation. Analyzing the frequency of preference reversals in these additional lotteries allows us to assess the validity of our classification of subjects into types in choices that were not used for estimating the structural model.

The experimental evidence gives rise to three main results. The first result summarizes the non-parametric evidence at the aggregate level; the second the insights gained from the structural model and the classification of subjects into types; and the third the out-of-sample predictions.

A non-parametric analysis of the aggregate choices provides the first main result. In the aggregate, EUT is clearly rejected, and both choice set dependence and probability weighting play a role. On the one hand, probability weighting plays a role, because the frequency of Allais Paradoxes exceeds the noise-level regardless whether the lotteries' payoffs are independent or perfectly correlated.[7] However, on the other hand, choice set dependence plays a role too, as Allais Paradoxes occur more frequently when lotteries' payoffs are independent than when they are perfectly correlated. This result holds under both presentation formats.

The structural model yields the second main result. There is vast heterogeneity in the subjects' choices and the population can be segregated into 38% CPT-types, 34% ST-types, and 28% EUT-types. However, while this classification indicates the best fitting decision theory for each type, an inspection of the subjects' average behavior per type reveals that both choice set dependence and probability weighting play some role across all types, although to a varying extent. Hence, the choices of all subjects seem to be driven by both probability weighting and choice set dependence to some degree, but their relative strength varies greatly across types.

The out-of-sample predictions about the preference reversals in the additional choices confirm the above finding and provide the third main result. Subjects classified as ST-types exhibit more preference reversals than those classified as EUT- and CPT-types. However, since the frequency of preference reversals exceeds the noise-level across all types, choice set dependence influences the choices of all three types. In conclusion, the classification of subjects into types passes this stringent out-of-sample test and remains qualitatively valid

[7]To determine the noise-level, we look at Allais Paradoxes going in the inverse direction, i.e. the direction that cannot be described by any non-EUT decision theory and, thus, is due to decision noise. See Figure 3 in Section 5 for details.

in choices that were not used for estimating the structural model.

Section 2 describes how these results contribute to the existing literature. Section 3 explains the strategy for discriminating between the different decision theories. Section 4 introduces the experimental design. Section 5 presents the non-parametric results at the aggregate level, while Section 6 discusses the structural model, its results, and the out-of-sample predictions. Finally, Section 7 concludes.

2 Related Literature

This section summarizes the related literature and highlights the paper's main contributions. The paper directly contributes to the empirical literature that aims at identifying the extent to which probability weighting and choice set dependence drive risky choices. On the one hand, there is considerable evidence suggesting that risk preferences depend on outcome probabilities irrespective of the choice set (for examples, see Kahneman and Tversky, 1979; Camerer and Ho, 1994; Loomes and Segal, 1994; Starmer, 2000; Fehr-Duda and Epper, 2012). On the other hand, the literature also has recognized that risky choices of many subjects depend on the choice set and that subjects sometimes revert their preferences (Lichtenstein and Slovic, 1971; Lindman, 1971; Grether and Plott, 1979; Pommerehne et al., 1982; Reilly, 1982; Cox and Epstein, 1989; Loomes et al., 1991). More recently, empirical tests of ST confirmed the role of choice set dependence in non-incentivized Mturk experiments (Bordalo et al., 2012b) and in two decisions each involving a choice between a lottery and a sure amount (Booth and Nolen, 2012).

Thus, the existing literature suggests that choice set dependence and probability weighting both influence risky choices. However, they have not been tested jointly in an incentivized experiment. Furthermore, it is unclear what is their relative importance, and whether choice set dependence and probability weighting each influence the behavior of all subjects to a varying extent or of just certain types of subjects. The present paper provides an answer to these questions by introducing an experiment that allows us not only to reliably discriminate between choice set dependence and probability weighting – both non-parametrically and with a structural model – but also to account for individual heterogeneity in a parsimonious way.

Moreover, the structural model adds to the literature that uses finite mixture models to classify subjects into types. This literature has mostly been focusing on discriminating rational from irrational behavior in decision making under risk (Bruhin et al., 2010; Fehr-

Duda et al., 2010; Conte et al., 2011)[8] and other complex decision situations (for examples see El-Gamal and Grether, 1995; Houser et al., 2004; Houser and Winter, 2004; Stahl and Wilson, 1995; Fischbacher et al., 2013). Our second main result enhances this strand of literature by showing that there is also substantial heterogeneity within the group of non-EUT subjects.

Uncovering this heterogeneity across non-EUT subjects not only contributes directly to our understanding of decision making under risk but also could prove to be fruitful in other domains as well. For instance, in deterministic consumer choice, there exist competing explanations for the famous endowment effect – i.e. the behavioral phenomenon that consumers tend to value goods higher as soon as they possess them (Samuelson and Zeckhauser, 1988; Knetsch, 1989; Kahneman et al., 1990; Isoni et al., 2011). One explanation of the endowment effect assumes loss aversion and an endogenous reference point, which shifts as soon as a subject obtains a good and expects to keep it (Kőszegi and Rabin, 2006). Another, competing explanation, is based on choice set dependence and has the following intuition: when the subject receives an endowment, she compares it to the status quo of having nothing which renders the good's best attribute salient and inflates its valuation (Bordalo et al., 2012a). Since our experimental design and structural model can isolate the group of subjects whose choices are mostly influenced by choice set dependence, they may offer an empirical way to study the relative importance of these competing explanations of the endowment effect.

Similarly, the experimental design and the structural model could also be used to study the links between limited attention and economic decisions. For instance, Kőszegi and Szeidl (2013) present a model in which limited attention and the focus on salient states affect intertemporal choice. Another model by Gabaix (2015) studies the role of limited attention on consumer demand and competitive equilibrium. Our methodology could provide a way to test the implications of these models, as it allows to discriminate subjects with limited attention from other behavioral and rational types.

3 Discriminating between Decision Theories

This section describes our strategy for discriminating between EUT, probability weighting – represented by CPT –, and choice set dependence – represented by ST. The strategy (i) relies

[8]Harrison and Rutström (2009) also apply finite mixture models in order to distinguish EUT from non-EUT behavior. However, they classify decisions instead of subjects.

on a series of binary choices between lotteries that may trigger the Allais Paradox and (ii) manipulates the choice set by making the lotteries' payoffs either independent or perfectly correlated.

We explain the strategy with the following binary choice, taken from Kahneman and Tversky (1979), between lotteries X and Y which may trigger the Common Consequence Allais Paradox.[9]

$$X = \begin{cases} 2500 & p_1 = 0.33 \\ z & p_2 = 0.66 \\ 0 & p_3 = 0.01 \end{cases} \qquad \text{vs.} \qquad Y = \begin{cases} 2400 & p_1 + p_3 = 0.34 \\ z & p_2 = 0.66 \end{cases}$$

Note that the two lotteries have a common consequence, i.e. a common payoff z which occurs with probability p_2 in both lotteries. In this example, the Common Consequence Allais Paradox refers to the robust empirical finding that if $z = 2400$, most subjects prefer Y over X, whereas if $z = 0$, most subjects prefer X over Y.

Next, we show that EUT can never describe the Allais Paradox, CPT can always describe it, and ST can only describe the Allais Paradox when the payoffs of the two lotteries are independent but not when they are perfectly correlated.

3.1 EUT

According to EUT, the decision maker evaluates any lottery L with non-negative payoffs $x = (x_1, \ldots, x_J)$ and associated probabilities $p = (p_1, \ldots, p_J)$ as

$$V^{EUT}(L) = \sum_{j=1}^{J} p_j \, v(x_j) \, ,$$

where v is an increasing utility function over monetary payoffs with $v(0) = 0$.[10] Note that the value $V^{EUT}(L)$ only depends on the attributes of lottery L and not on the attributes of the other lotteries in the choice set.

EUT cannot explain the Common Consequence Allais Paradox. In the above example, the decision maker evaluates lottery X as $V^{EUT}(X) = p_1 \, v(2500) + p_2 \, v(z) + p_3 \, v(0)$ and lottery Y as $V^{EUT}(Y) = (p_1 + p_3) \, v(2400) + p_2 \, v(z)$. When comparing the values of the two lotteries, $V^{EUT}(X)$ and $V^{EUT}(Y)$, the term involving the common consequence, $p_2 \, v(z)$,

[9]The analogous example for the Common Ratio Allais Paradox can be found in Appendix B.

[10]This assumes that subjects are interested in lottery payoffs and not final wealth states.

cancels out. Hence, the decision maker's choice between X and Y does not depend on the value of the common consequence.

3.2 CPT

According to CPT, the decision maker ranks the non-negative monetary payoffs of any lottery L such that $x_1 \geq \ldots \geq x_J$ and evaluates the lottery as

$$V^{CPT}(L) = \sum_{j=1}^{J} \pi_j^{CPT}(p)\, v(x_j)\,,$$

where π_j is the decision weight attached to the value of payoff x_j. As in EUT, the value $V^{CPT}(L)$ only depends on the attributes of lottery L, i.e. the decision maker evaluates the lottery in isolation. The decision weights are given by

$$\pi_j^{CPT}(p) = \begin{cases} w(p_1) & \text{for } \; j = 1 \\ w\left(\sum_{k=1}^{j} p_k\right) - w\left(\sum_{k=1}^{j-1} p_k\right) & \text{for } \; 2 \leq j \leq J \end{cases},$$

where p_k is payoff x_k's probability and w is the probability weighting function. The probability weighting function exhibits three properties (Prelec, 1998; Wakker, 2010; Fehr-Duda and Epper, 2012):

1. *Strictly increasing in probabilities with* $w(0) = 0$ *and* $w(1) = 1$. This property ensures that decision weights are non-negative and, together with $w(0) = 0$ and $w(1) = 1$, implies that decision weights sum to one.

2. *Inverse S-shape.* The probability weighting function is concave for small probabilities and convex for large probabilities. This ensures that the decision maker overweights small probabilities and underweights large probabilities. This is necessary for CPT to be able to explain the Common Consequence Allais Paradox, as shown further below.

3. *Subproportionality.* For the probabilities $1 \geq q > p > 0$ and the scaling factor $0 < \lambda < 1$ the inequality $\frac{w(q)}{w(p)} > \frac{w(\lambda q)}{w(\lambda p)}$ holds. Subproportionality is needed for CPT to be able to explain the Common Ratio Allais Paradox, as shown in Appendix B.

We now explain how CPT can describe the Common Consequence Allais Paradox in the choice between lotteries X and Y. When $z = 2400$, the choice is

$$X = \begin{cases} 2500 & p_1 = 0.33 \\ 2400 & p_2 = 0.66 \\ 0 & p_3 = 0.01 \end{cases} \qquad \text{vs.} \qquad Y = 2400\,.$$

The decision maker prefers lottery Y over X if

$$\begin{aligned} V^{CPT}(Y) &> V^{CPT}(X) \\ v(2400) &> \pi_1^{CPT} v(2500) + \pi_2^{CPT} v(2400) + \pi_3^{CPT} v(0) \\ v(2400) &> w(0.33)v(2500) + [w(0.99) - w(0.33)]\, v(2400) \\ &\quad + [1 - w(0.99)]\, v(0)\,. \end{aligned}$$

Intuitively, due to the decision maker's tendency to overestimate small probabilities and underestimate large probabilities, the decision weight $1 - w(0.99)$ attached to the lowest payoff of X is larger than its objective probability $p_3 = 0.01$, which renders X unattractive. Hence, when the common consequence is $z = 2400$, the decision maker has a tendency to prefer Y over X. In contrast, when $z = 0$, the choice is

$$X = \begin{cases} 2500 & p_1 = 0.33 \\ 0 & p_2 + p_3 = 0.67 \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 2400 & p_1 + p_3 = 0.34 \\ 0 & p_2 = 0.66 \end{cases}.$$

Now, the decision maker prefers lottery X over Y if

$$\begin{aligned} V^{CPT}(X) &> V^{CPT}(Y) \\ w(0.33)v(2500) + [1 - w(0.33)]\, v(0) &> w(0.34)v(2400) + [1 - w(0.34)]\, v(0)\,. \end{aligned}$$

Intuitively, the decision maker prefers lottery X over Y because the decision weights $w(0.33)$ and $w(0.34)$ are very close and, therefore, the decision is driven by the difference in utilities between $v(2500)$ and $v(2400)$ rather than the difference in probabilities. Hence, when the common consequence is $z = 0$, the decision maker no longer has a tendency to prefer Y over X.

3.3 ST

According to ST, due to cognitive limitations the decision maker is a local thinker who focuses her attention on some but not all states of the world. Salience shifts the focus of attention to states of the world in which one payoff stands out relative to the payoffs of the alternative. The decision maker overweights these salient states relative to the others. As the salience of a state directly depends on the payoffs of the alternative, a lottery's value is choice set dependent and – in contrast to EUT and CPT – lotteries are no longer evaluated in isolation.

Formally, if the decision maker has to choose between two lotteries L^1 and L^2, she ranks each possible state $s \in \{1, \ldots, S\}$ according to its salience $\sigma(x_s^1, x_s^2)$, where x_s^1 and x_s^2 are the payoffs of L^1 and L^2, respectively, in state s. The salience function σ satisfies three properties:

1. *Ordering.* For two states s and $\tilde{s}$ we have that if $[x_s^{\min}, x_s^{\max}]$ is a subset of $[x_{\tilde{s}}^{\min}, x_{\tilde{s}}^{\max}]$, then $\sigma(x_{\tilde{s}}^1, x_{\tilde{s}}^2) > \sigma(x_s^1, x_s^2)$. Ordering implies that states with bigger differences in payoffs tend to be more salient.

2. *Diminishing Sensitivity.* For any $\epsilon > 0$, then $\sigma(x_s^1, x_s^2) > \sigma(x_s^1 + \epsilon, x_s^2 + \epsilon)$. Diminishing sensitivity implies that for states with a given difference in payoffs, salience diminishes the further away from zero the difference in payoffs is.

3. *Symmetry:* $\sigma(x_s^1, x_s^2) = \sigma(x_s^2, x_s^1)$. Symmetry implies that for a given difference in payoffs, the salience of the state is the same regardless which of the two lotteries provides the higher or the lower payoff.

The decision weight of each state s depends on the state's salience-rank, $r_s \in \{1, \ldots, S\}$ with lower values being associated with higher salience:

$$\pi_s^{ST}(x^1, x^2) = p_s \, \frac{\delta^{r_s}}{\sum_{m \in S} \delta^{r_m} \, p_m} \,, \tag{1}$$

where p_s is the probability that state s is realized, and $\delta \leq 1$ is the decision maker's degree of local thinking. Note that for $\delta = 1$ the decision maker is rational and weights states by their objective probabilities, whereas for $\delta < 1$ the decision maker is a local thinker and overweights salient states. This yields the following values for lotteries L^1 and L^2:

$$V^{ST}(L^1 | \{L^1, L^2\}) = \sum_{s=1}^{S} \pi_s^{ST}(x^1, x^2) \, v(x_s^1)$$

and

$$V^{ST}(L^2 | \{L^1, L^2\}) = \sum_{s=1}^{S} \pi_s^{ST}(x^1, x^2) \, v(x_s^2) \,.$$

Note that the value of each lottery depends on all lotteries in the choice set $\{L^1, L^2\}$.

We now explain how ST can describe the Common Consequence Allais Paradox in the choice between lotteries X and Y when their payoffs are independent of each other. When $z = 2400$, there are three states of the world which rank in salience as follows:

$\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(2400, 2400)$. The decision maker prefers lottery Y over X if $V^{ST}(Y|\{X,Y\}) > V^{ST}(X|\{X,Y\})$, where

$$V^{ST}(Y|\{X,Y\}) = v(2400)\,,$$

and

$$\begin{aligned} V^{ST}(X|\{X,Y\}) &= \pi_2^{ST}(2500,2400)\,v(2500) + \pi_3^{ST}(2400,2400)\,v(2400) \\ &\quad +\pi_1^{ST}(0,2400)\,v(0)\,. \end{aligned}$$

Using $v(0) = 0$ and the decision weights given by equation (1), the condition for preferring Y over X becomes

$$\delta < \frac{0.01}{0.33}\,\frac{v(2400)}{v(2500) - v(2400)}\,. \tag{2}$$

Intuitively, lottery X provides the lowest payoff in the most salient state which makes lottery Y relatively attractive despite having a lower expected payoff. Hence, when the common consequence is $z = 2400$ and the degree of local thinking is severe enough, the decision maker prefers Y over X.

In contrast, when $z = 0$, there are four states of the world which rank in salience as follows: $\sigma(2500, 0) > \sigma(0, 2400) > \sigma(2500, 2400) > \sigma(0, 0)$. The decision maker prefers lottery X over Y if $V^{ST}(X|\{X,Y\}) > V^{ST}(Y|\{X,Y\})$, where

$$\begin{aligned} V^{ST}(X|\{X,Y\}) &= \left[\pi_1^{ST}(2500,0) + \pi_3^{ST}(2500,2400)\right]\,v(2500) \\ &\quad + \left[\pi_2^{ST}(0,2400) + \pi_4^{ST}(0,0)\right]\,v(0)\,, \end{aligned}$$

and

$$\begin{aligned} V^{ST}(Y|\{X,Y\}) &= \left[\pi_2^{ST}(0,2400) + \pi_3^{ST}(2500,2400)\right]\,v(2400) \\ &\quad + \left[\pi_1^{ST}(2500,0) + \pi_4^{ST}(0,0)\right]\,v(0)\,. \end{aligned}$$

Using $v(0) = 0$ and the decision weights given by equation (1), the decision maker prefers X over Y when

$$\begin{aligned} &(0.33)\,(0.66)\,v(2500) - \delta\,(0.67)\,(0.34)\,v(2400) \\ &\quad +\delta^2\,(0.33)\,(0.34)\,[v(2500) - v(2400)] \;>\; 0\,. \end{aligned} \tag{3}$$

Now, lottery X provides the highest payoff in the most salient state. Hence, when the common consequence is $z = 0$ and the degree of local thinking is severe enough, the decision maker prefers X over Y.

We now turn to the case in which the two lotteries' payoffs are perfectly correlated. In that case, ST can no longer describe the Common Consequence Allais Paradox. When the two lotteries' payoffs are perfectly correlated there are just the following three states of the world:

p_s	0.33	0.66	0.01
x_s	2500	z	0
y_s	2400	z	2400

The ranking in terms of salience of these three states, $\sigma(0, 2400) > \sigma(2500, 2400) > \sigma(z, z)$, is independent of the common consequence z. Hence, regardless of the common consequence, the decision maker has a tendency to prefer Y over X, and the Common Consequence Allais Paradox can no longer be described by ST when the lotteries' payoffs are perfectly correlated.

3.4 Summary

Table 1 summarizes the strategy for discriminating between EUT, probability weighting, and choice set dependence.

Table 1: When can the Allais Paradox occur?

	Lottery Payoffs	
	independent	correlated
Rationality: EUT	✗	✗
Probability Weighting: CPT	✓	✓
Choice Set Dependence: ST	✓	✗

EUT can never explain the Allais Paradox because the independence axiom implies that the preference functional is linear in probabilities and each lottery is evaluated independently of the other lotteries in the choice set. In contrast, probability weighting – represented by CPT – can explain the Allais paradox. This is because the decision weight of each payoff depends non-linearly on the marginal payoff distribution of the lottery under consideration which remains unchanged regardless whether the lotteries' payoffs are independent or perfectly correlated. Finally, choice set dependence – represented by ST – can explain the Allais paradox only when the lotteries' payoffs are independent but not when they are perfectly correlated. This is because decision weights depend on the joint payoff distribution of all lotteries in the choice set, which changes when we manipulate the correlation structure of the lotteries' payoffs.

4 Experimental Design

This section presents the experimental design which consists of two parts. In the main part, subjects make a series of binary choices between lotteries that may trigger the Common Consequence and the Common Ratio Allais Paradoxes. Based on these choices, we discriminate between rational behavior, probability weighting, as well as choice set dependence, and classify subjects into EUT-, CPT-, and ST-types, respectively. In the additional part, subjects make choices that could lead to preference reversals which allow us to validate the classification of subjects into types using out-of-sample-predictions.

4.1 Main Part

We now present the main part of the experiment. First, we explain how we constructed the series of binary choices. Subsequently, we describe the formats which we use to present the binary choices to the subjects.

4.1.1 Choices between Lotteries

Every subject goes through two blocks of binary choices between lotteries that may trigger the Allais Paradoxes. Both blocks feature the same binary choices, except that in one block the lotteries' payoffs are independent while in the other they are perfectly correlated. As described in the previous section, this allows us to discriminate non-parametrically between rational behavior, probability weighting, and choice set dependence by comparing within-subjects the frequency of Allais Paradoxes in the two blocks.

The binary choices within each block feature lotteries that vary systematically in payoffs and probabilities. This systematic variation not only allows us to estimate the parameters of a structural model for each decision theory but also ensures that our results are not driven by a particular set of lotteries.

The binary choices that may trigger the Common Consequence Allais Paradox are based on a $3 \times 3 \times 3$ design. The design uses the following three different payoff levels:

$$\text{Payoff Level 1:} \quad X = \begin{cases} 2500 & p_1 \\ z & p_2 \\ 0 & p_3 \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 2400 & p_1 + p_3 \\ z & p_2 \end{cases}$$

$$\text{Payoff Level 2:} \quad X = \begin{cases} 5000 & p_1 \\ z & p_2 \\ 0 & p_3 \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 4800 & p_1 + p_3 \\ z & p_2 \end{cases}$$

$$\text{Payoff Level 2:} \quad X = \begin{cases} 3000 & p_1 \\ z & p_2 \\ 500 & p_3 \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 2600 & p_1 + p_3 \\ z & p_2 \end{cases}$$

Varying the payoffs across these three levels while keeping probabilities constant identifies the curvature of the utility function, v. Similarly, the design features three different probability distributions, $p = (p_1, p_2, p_3)$, over the lotteries' payoffs:

$$\begin{aligned} &\text{Probability Distribution 1:} \quad p = (0.33, 0.66, 0.01) \\ &\text{Probability Distribution 2:} \quad p = (0.30, 0.65, 0.05) \\ &\text{Probability Distribution 3:} \quad p = (0.25, 0.60, 0.15) \end{aligned}$$

Varying the probability distributions while keeping the lotteries' payoffs constant identifies the shape of probability weighting function, w, in CPT and the degree of local thinking, δ, in ST. Finally, the design uses three different levels of the common consequence, z, to trigger the Common Consequence Allais Paradox:

1. $z = x_3$, i.e. the common consequence is equal to the lowest payoff of lottery X. In this case, lottery X and Y offer two payoffs each.

2. $z = y_1$, i.e. the common consequence is equal to the first payoff of lottery Y. In this case, lottery X offers three payoffs and lottery Y is a sure amount.

3. z is different from any other payoffs of the two lotteries but slightly below the first payoff of lottery Y.[11] In this case, lottery X offers three payoffs and lottery Y offers two payoffs.

The first two levels of the common consequence trigger the classical version of the Common Consequence Allais Paradox, as described in the previous section. We also include the third

[11] For Payoff Level 1: $z = 2000$; for Payoff Level 2: $z = 4000$; for Payoff Level 3: $z = 2000$.

level of the common consequence, since in combination with the fist level, it may trigger a more general version of the Common Consequence Allais Paradox in which the lottery Y does not degenerate into a sure amount.

The binary choices that may trigger the Common Ratio Allais Paradox are based on a similar $3 \times 3 \times 2$ design. The design uses different payoff and probability levels which are scaled up and down, respectively, to provoke the Common Ratio Allais Paradox. For details, please refer to Appendix C.

To avoid order effects, we randomize the order of the binary choices within each of the two blocks and counterbalance the order of the two blocks across subjects.

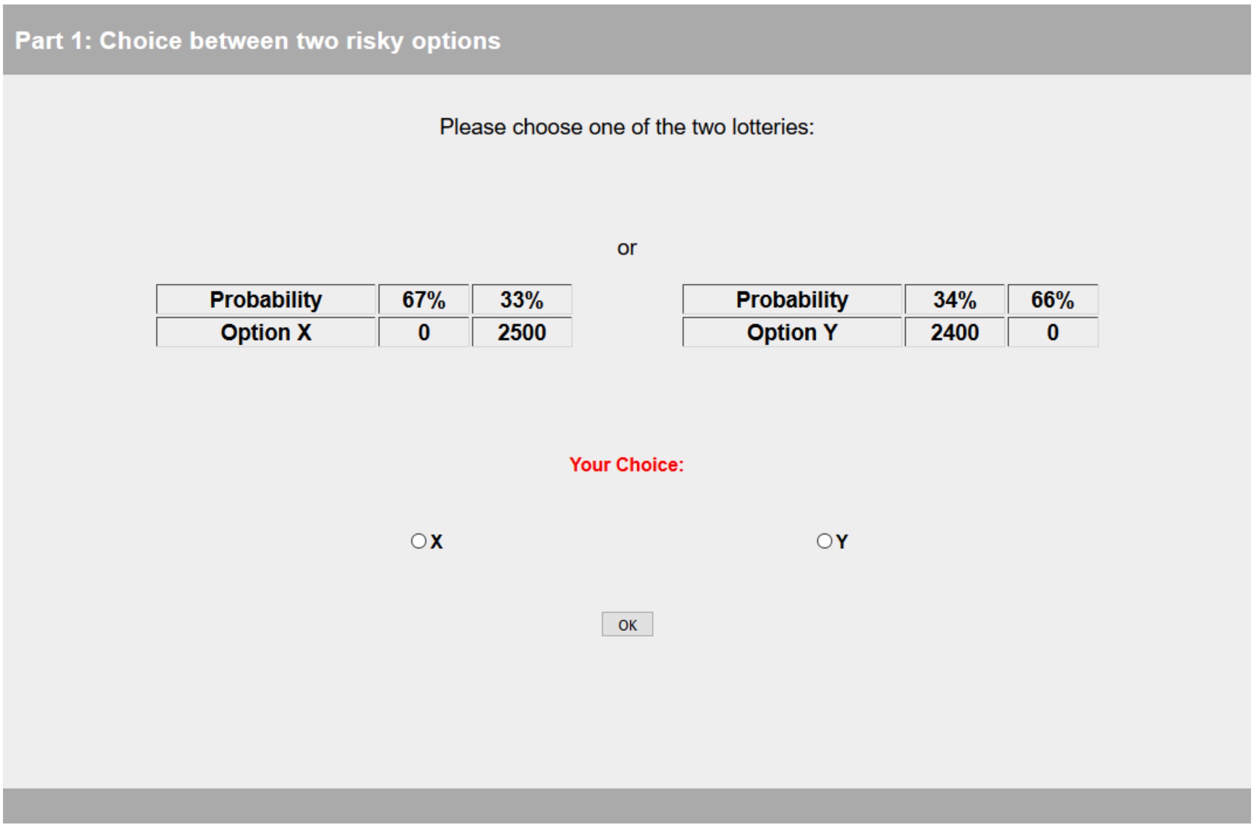
4.1.2 Presentation Format

We present the binary choices between lotteries in two formats: the "canonical presentation" and the "states of the world presentation". We apply a between-subjects-design, i.e. half of the subjects are exposed to the canonical presentation and the other half to the states of the world presentation.

The two presentation formats differ in the way they show the binary choices between lotteries with independent payoffs to the subjects. In the canonical presentation, as shown by the screenshot in Figure 1, the two lotteries X and Y are presented side by side as separate lotteries with independent payoff distributions. In the states of the world presentation, as shown by the screenshot in Figure 2, the lotteries are presented in a table displaying their joint payoff distribution. For binary choices between lotteries with correlated payoffs the two presentation formats are identical and display the two lotteries' joint payoff distribution.

The two presentation formats have distinct advantages and disadvantages. The main advantages of the canonical presentation are that it emphasizes the difference between lotteries with independent vs. correlated payoffs and that subjects are probably more used to the canonical presentation of lotteries with independent payoffs. However, the main disadvantage of the canonical presentation is that between the two blocks not only the correlation structure of the lotteries' payoffs changes but also their visual presentation. In contrast, the states of the world presentation keeps the visual presentation constant across the two blocks, but presents lotteries with independent payoffs in an unfamiliar way. Ideally, the two presentation formats should have no effect on the results.

Figure 1: Canonical Presentation of the Binary Choice between Two Lotteries with Independent Payoffs



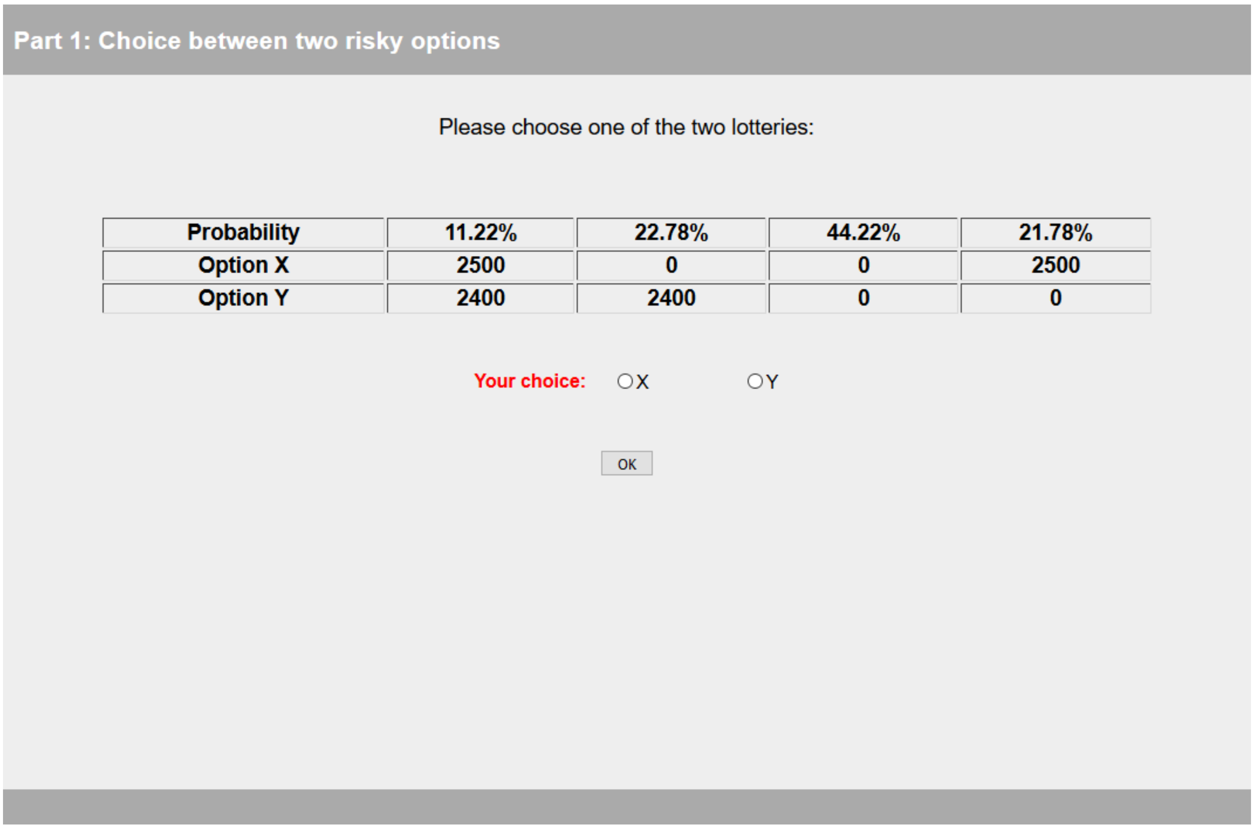
4.2 Additional Part

To validate the classification of subjects into types, we use out-of-sample predictions about the frequency of preference reversals. To establish whether a subject has a tendency to revert her preference, the main part of the experiment contains six binary choices between lotteries that are neither used for estimating the subject's preferences nor for classifying them into types. In the additional part of the experiment, the subject has to evaluate each of these additional lotteries in isolation by stating their certainty equivalent.

Each of the additional lotteries is presented in a choice menu in which the subject has to indicate whether she prefers the lottery or a certain payoff. The certain payoff increases from the lottery's lowest payoff, 0, to its highest payoff in 21 equal increments. The point where the subject switches form preferring the certain payoff to preferring the lottery allows us to approximate the certainty equivalent.[12]

[12]This method to trigger preference reversals does not use the Becker, De Groot, and Marschak mechanism

Figure 2: States of the World Presentation of the Binary Choice between Two Lotteries with Independent Payoffs



The order in which we elicit the certainty equivalents of the additional lotteries is randomized across subjects. Moreover, since the six binary choices between the additional lotteries appeared in the main part of the experiment, subjects should not recall the additional lotteries when stating their certainty equivalents. The six binary choices between the additional lotteries can be found in Appendix D.

By comparing the binary choices between the additional lotteries and their certainty equivalents, we can detect the number of preference reversals of every subject. Since there are six binary choices each subject can exhibit between 0 and 6 preference reversals.

(BDM; Becker et al., 1964) to elicit certainty equivalents and, therefore, avoids the problems pointed out by Karni and Safra (1987) and Segal (1988). Consequently, the preference reversals we observe in this part of the experiment represent violations of EUT's transitivity axiom. We did not impose a unique switch-point. 34 of 283 subjects (12.0%) switched more than once and, thus, did not reveal a unique certainty equivalent for at least one lottery. We excluded these subjects form the out-of-sample analysis of preference reversals in Section 6.3.

4.3 Number of Choices

Table 2 summarizes the number of choices in the two presentation formats. Subjects in the canonical presentation go through a total of 93 binary choices, while subjects in the states of the world presentation go through only 84 binary choices. The number of binary choices differs between the presentation formats since the 9 binary choices designed for triggering the Common Consequence Allais Paradox in which lottery X has three payoffs and lottery Y is a sure amount look identical regardless whether the lotteries' payoffs are independent or correlated. Moreover, we use 3 of the $3 \times 3 \times 2 = 18$ binary choices designed to trigger the Common Ratio Allais Paradox to make out-of-sample predictions about preference reversals. We leave these three binary choices out in the calculation of the frequencies of Allais Paradoxes and the structural estimations (see Appendices C and D). Therefore, the table shows these 3 binary choices among the total of 6 choices used for triggering preference reversals.

Regardless of the presentation format, each subject also evaluates 9 lotteries in isolation during the additional part of the experiment. This yields the certainty equivalents that we need to detect preference reversals and test our out-of-sample predictions.

4.4 Implementation in the Lab and Incentives

We conducted the experiment in the computer lab at the University of Lausanne between February and May 2015 using a web application based on PHP and MySQL. Most subjects were students of the University of Lausanne and the Ecole Polytechnique Federale de Lausanne, recruited via ORSEE (Greiner, 2015). The experiment consisted of 14 sessions with 283 students in total.[13]

At the beginning of the experiment, subjects received some general instructions informing them about the structure of the experiment, their anonymity, the show up fee, and the conversion rate of points into Swiss Francs.[14] At the beginning of each part, subjects received additional printed instructions. These additional instructions comprised the description of the choices made in that part, the description of the payment procedure for that part, and several comprehension questions which were controlled by the assistants before subjects could

[13]We also carried out 3 pilot sessions.

[14]Payoffs were shown in points. 100 points corresponded to one Swiss Franc. At the time of the experiment, 1 Swiss Franc corresponded to roughly 1.04 USD.

Table 2: Number of Binary Choices by Presentation Format and Type of Allais Paradox

	Canonical		
Allais Paradox	Independent Payoffs	Correlated Payoffs	Preference Reversal
Common Consequence	27	27	
Common Ratio[a]	15	18	
Total Binary Choices	42	45	6
	States of the World		
Allais Paradox	Independent Payoffs	Correlated Payoffs	Preference Reversal
Common Consequence	18[b]	27	
Common Ratio[a]	15	18	
Total	33	45	6

[a] Three of the $3\times3\times2 = 18$ binary choices to trigger the Common Ratio Allais Paradox were used to make out-of-sample predictions about preference reversals. These three binary choices were left out in the calculation of the frequencies of Allais Paradoxes and the structural estimations (see Appendices C and D).

[b] In the states of the world presentation, the nine binary choices where lottery X has three possible payoffs and lottery Y is a sure amount look identical regardless whether the lotteries' payoffs are independent or correlated. Since we did not want to present the same choices twice, subjects exposed to in the states of the world presentation had to go through nine binary choices less than those exposed to the canonical presentation.

begin. The additional instructions differed depending on whether a subject was exposed to the canonical presentation or the states of the world presentation. All instructions were written in French. English translations are available in the Online Appendix.

To incentivize subjects' choices in both parts of the experiment, we applied the prior incentive system (Johnson et al., 2014). This avoids violations of isolation, which may otherwise arise with a random lottery selection procedure, as pointed out by Holt (1986). In each part, every subject had to draw a sealed envelope from an urn before making any choices. The envelope contained one of the choices the subject was going to make in that part and which later was used for payment. At the very end of the experiment, the subject went to another room where she opened the envelopes together with an assistant, rolled some dice to determine the payoff of the chosen lotteries, and received her payment.

After making their choices, but before determining and receiving their payments, subjects filled in a demographic questionnaire, completed a short version of the Big 5 personality questionnaire, and a cognitive ability test with 12 questions based on Raven's matrices. The instructions were shown on screen at the beginning of each task. The cognitive ability test was also incentivized and subjects received 50 points per correct answer.[15]

Each subject was paid a show-up fee of 10 Swiss Francs. Total earnings varied between 12.00 and 142.50 Swiss Francs with a mean of 57.66 and a standard deviation of 26.39 Swiss Francs. Each session lasted for approximately 90 minutes.

5 Non-Parametric Results

In this section, we present the non-parametric results by analyzing the relative frequency of Allais Paradoxes. Figure 3 shows the average frequency of Allais Paradoxes relative to their maximum possible number between lotteries with independent and correlated payoffs.

First, we compare the relative frequency of Allais Paradoxes in the expected direction according to both probability weighting and choice set dependence (Panel (a)) to those in the inverse direction (Panel (b)). This comparison reveals that Allais Paradoxes in the expected direction are substantially and significantly more frequent than those in the inverse direction (t-tests: p-values < 0.001 in all pairwise comparisons). For example, for both presentation formats together, the average frequency of Allais Paradoxes in the expected direction is 28.2%

[15]We do not find any statistically significant relationship between these individual characteristics and the classification of subjects into types. Results are available on request.

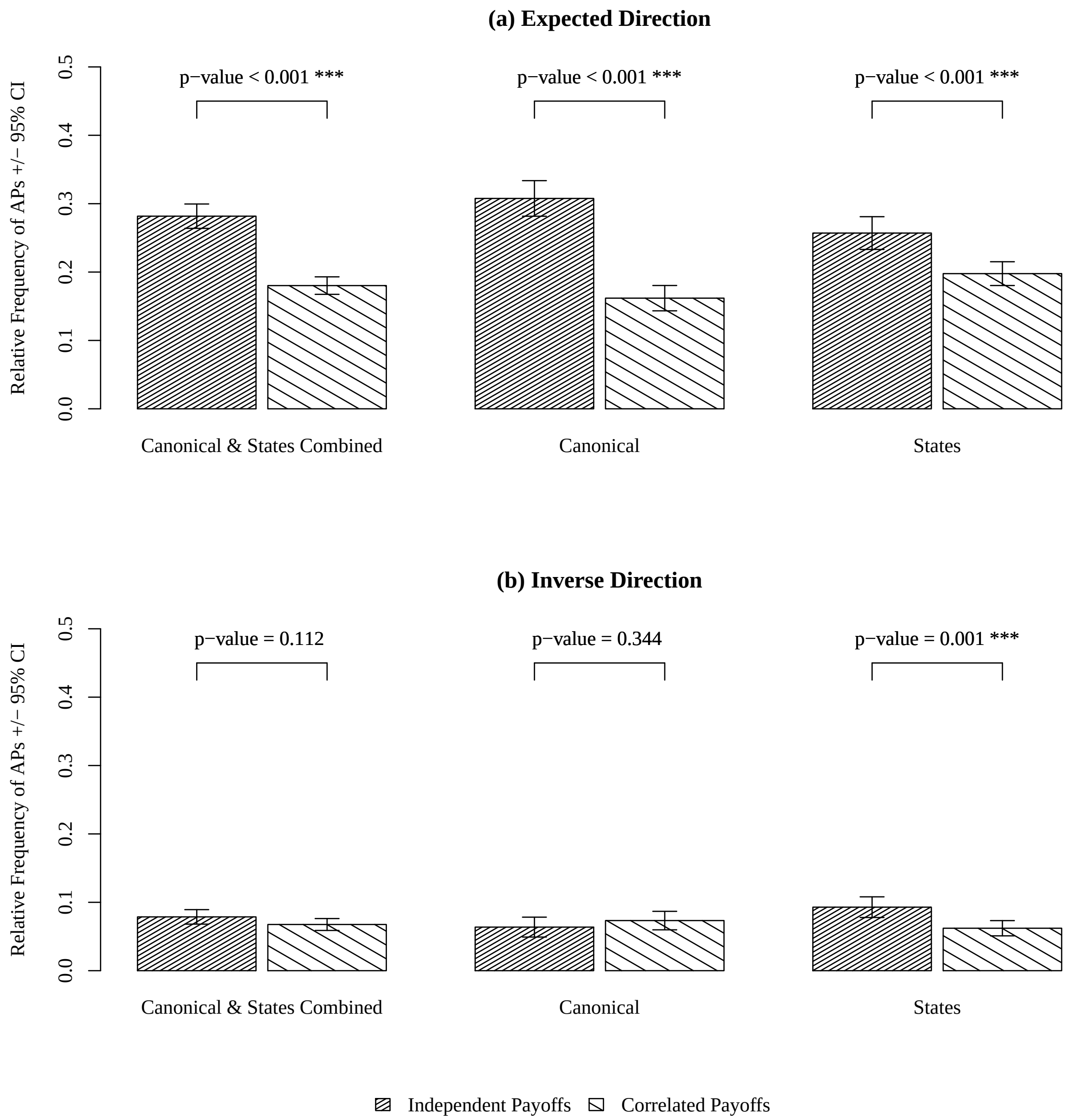
when lottery payoffs are independent and 18.0% when lottery payoffs are perfectly correlated. The corresponding frequencies of Allais Paradoxes in the inverse direction are only 7.8% and 6.8%, respectively. We interpret the Allais Paradoxes in the inverse direction as the result of decision noise. Since the relative frequency of Allais Paradoxes in the expected direction is significantly higher than in the inverse direction – regardless whether lottery payoffs are independent or perfectly correlated – we conclude that aggregate choices cannot be described by EUT plus decision noise.

Next, we focus on Allais Paradoxes in the expected direction and compare their frequencies when lottery payoffs are independent vs. perfectly correlated. Regardless of the presentation format, we find that Allais Paradoxes in the expected direction are always significantly more frequent when lottery payoffs are independent than when they are perfectly correlated (t-tests: p-values < 0.001 for both presentation formats separately as well as pooled).[16] This finding indicates that choice set dependence plays a role for describing aggregate choices. However, the frequency of Allais Paradoxes in the expected direction is significantly above the level explained by decision noise – even when lottery payoffs are perfectly correlated (t-tests: p-values < 0.001 for both presentation formats separately as well as pooled). This cannot be described exclusively by choice set dependence plus noise. Under choice set dependence plus noise, the frequency of Allais Paradoxes with perfectly correlated lottery payoffs in the expected direction should be driven exclusively by noise and, thus, match the corresponding frequency in the inverse direction. Since this is not the case (t-tests: p-values < 0.001 for both presentation formats separately as well as pooled), we conclude that probability weighting also plays a role in driving aggregate choices.

Figure 4 takes a more disaggregate look at the data and shows the distributions of the relative frequency of Allais Paradoxes in the expected direction separately for independent and perfectly correlated lottery payoffs. The shift between the two distributions confirms that subjects exhibit a higher frequency of Allais Paradoxes when lottery payoffs are independent than when they are perfectly correlated, implying that choice set dependence matters. However, in both cases the majority of subjects exhibits a substantial number of Allais Paradoxes, implying that probability weighting matters too. Taken together, this non-parametric evidence yields our first main result.

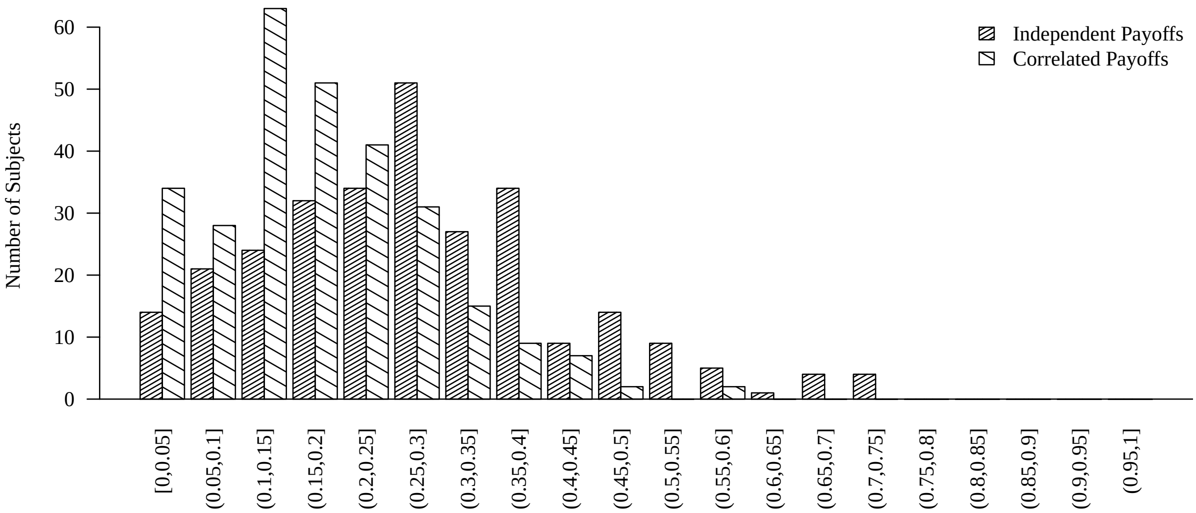
[16] When doing the same comparisons of the relative frequencies of Allais Paradoxes in the inverse direction, we find no systematic difference between independent and perfectly correlated payoffs. This reinforces our interpretation that the Allais Paradoxes in the inverse direction are the result of decision noise.

Figure 3: Relative Frequency of Allais Paradoxes



The figure shows the average frequency of Allais Paradoxes relative to their maximum possible number between lotteries with independent and perfectly correlated payoffs. Panel (a) depicts the relative frequency of Allais Paradoxes that go in the expected direction. Panel (b) shows the relative frequency of Allais Paradoxes that go in the inverse direction and probably reflect noise in the subjects' choices. The two bars on the left pool the choices from subjects exposed to the canonical presentation with those from subjects exposed to the states of the world presentation. The two bars in the middle and on the right separate the choices by presentation format.

Figure 4: Distribution of the Relative Frequency of Allais Paradoxes in the Expected Direction



The histograms show the distribution of the relative frequency of Allais Paradoxes in the expected direction for independent and perfectly correlated lottery payoffs. Choices from both presentation formats are pooled together.

Result 1 *For aggregate choices, EUT is clearly rejected and both choice set dependence as well as probability weighting play a role.*

In addition, when looking at the impact of the presentation format, we find a statistically significant but qualitatively unimportant effect. The frequency of Allais Paradoxes in the expected direction is significantly higher in the canonical presentation than in the states of the world presentation when lottery payoffs are independent (t-test: p-value = 0.005) and significantly lower when lottery payoffs are perfectly correlated (t-test: p-value = 0.006). However, even in the states of the world presentation the frequency of Allais Paradoxes remains significantly higher when lottery payoffs are independent than when they are correlated (t-test: p-value < 0.001). Thus, Result 1 is valid under both presentation formats. Consequently, from now on, we pool the choices from both presentation formats together.

6 Structural Model

In this section, we discuss the set up and the results of the structural model. It allows us to take individual heterogeneity into account in a parsimonious way and classify the subjects into distinct preference types. We also validate the classification of subjects into types using out-of-sample predictions.

6.1 Set Up

The structural model is based on a finite mixture model (see McLachlan and Peel, 2000, for an overview) and uses a random utility approach for discrete choices (McFadden, 1981). It discriminates between subjects whose preferences are rational and best described by EUT, subjects whose preferences display probability weighting and are best described by CPT, and subjects whose preferences display choice set dependence and are best described by ST. Controlling for the presence of rational subjects is important, as previous research revealed a minority of EUT-types in various student subject pools (Bruhin et al., 2010; Conte et al., 2011).

6.1.1 Random Utility Approach

Consider a subject $i \in \{1, \ldots, N\}$ whose preferences are best described by decision model M in $\mathcal{M} = \{EUT, CPT, ST\}$. She prefers lottery X_g over Y_g in binary choice $g \in \{1, \ldots, G\}$

when the random utility of choosing X_g, $V^M(X_g, \theta_M) + \epsilon_X$, is higher than the random utility of choosing Y_g, $V^M(Y_g, \theta_M) + \epsilon_Y$. The random errors ϵ_X and ϵ_Y are realizations of an extreme value 1 distribution with scale parameter $1/\sigma_M$, and the vector θ_M comprises decision model M's preference parameters. This implies that the probability of subject i choosing X_g, i.e. $C_{ig} = X$, is given by

$$
\begin{aligned}
P(C_{ig} = X; \theta_M, \sigma_M) &= Pr[V^M(X_g, \theta_M) - V^M(Y_g, \theta_M) \geq \epsilon_Y - \epsilon_X] \\
&= \frac{\exp[\sigma_M V^M(X_g, \theta_M)]}{\exp[\sigma_M V^M(X_g, \theta_M)] + \exp[\sigma_M V^M(Y_g, \theta_M)]} .
\end{aligned}
$$

Note that the parameter σ_M governs the choice sensitivity towards differences in deterministic value of the lotteries. If σ_M is 0, the subject chooses each lottery with probability 50% regardless of the deterministic value it provides. If σ_M is arbitrarily large, the probability of choosing the lottery with the higher deterministic value approaches 1.

Subject i's contribution to the density function of the random utility model corresponds to the product of the choice probabilities over all G binary decisions, i.e.

$$
f_M(C_i; \theta_M, \sigma_M) = \prod_{g=1}^{G} P(C_{ig} = X; \theta_M, \sigma_M)^{I(C_{ig}=X)} \, P(C_{ig} = Y; \theta_M, \sigma_M)^{1-I(C_{ig}=X)} ,
$$

where $I(C_{ig} = X)$ is 1 if subject i chooses lottery X_g and 0 otherwise.

6.1.2 Finite Mixture Model

Since risk preferences are heterogeneous, we do not directly observe which model best describes subject i's preferences. In other words, we do not know ex-ante whether subject i is an EUT-, CPT-, or ST-type. Hence, we have to weight i's type-specific density contributions by the corresponding ex-ante probabilities of type-membership, π_M, in order to obtain her contribution to the likelihood of the finite mixture model,

$$
\begin{aligned}
\ell(\Psi; C_i) &= \pi_{EUT} \, f_{EUT}(C_i; \theta_{EUT}, \sigma_{EUT}) + \pi_{CPT} \, f_{CPT}(C_i; \theta_{CPT}, \sigma_{CPT}) \\
&\quad + \pi_{ST} \, f_{ST}(C_i; \theta_{ST}, \sigma_{ST}) ,
\end{aligned}
$$

where the vector $\Psi = (\theta_{EUT}, \theta_{CPT}, \theta_{ST}, \sigma_{EUT}, \sigma_{CPT}, \sigma_{ST}, \pi_{EUT}, \pi_{CPT})$ comprises all parameters that need to be estimated, and $\pi_{ST} = 1 - \pi_{EUT} - \pi_{CPT}$.[17] Note that the ex-ante

[17]Note that i's likelihood contribution is highly non-linear. Maximizing the finite mixture model's likelihood is therefore not trivial and standard numerical maximization techniques, such as the BFGS algorithm,

probabilities of type-membership are the same across all subjects and correspond to the relative sizes of the types in the population.

Once we estimated the parameters of the finite mixture model, we can classify each subject into the type she most likely belongs to, given her choices and the the estimated parameters, $\hat{\Psi}$. To do so, we apply Bayes' rule and obtain subject i's individual ex-post probabilities of type-membership,

$$\tau_{iM} = \frac{\hat{\pi}_M \, f_M(C_i; \hat{\theta}_M, \hat{\sigma}_M)}{\sum_{m \in \mathcal{M}} \hat{\pi}_m \, f_m(C_i; \hat{\theta}_m, \hat{\sigma}_m)} \,. \tag{4}$$

Based on these individual ex-post probabilities of type-membership, we can also assess the ambiguity in the classification of subjects into types. If the finite mixture model classifies subjects cleanly into types, most τ_{iM} should be either very close to 0 or to 1. In contrast, if the finite mixture model fails to come up with a clean classification of subjects into distinct types, many τ_{iM} will be in the vicinity of 1/3.

6.1.3 Specification of Functional Forms

To keep the model parsimonious and yet flexible in fitting the data, we specify the following functional forms. In all three decision models, we use a power specification for the utility function v, i.e.

$$v(x) = \begin{cases} \frac{x^{1-\beta}}{1-\beta} & \text{for } \beta \neq 1 \\ \ln x & \text{for } \beta = 1 \end{cases} ,$$

which has a convenient interpretation, since β measures v's concavity. Moreover, this specification turned out to be a neat compromise between parsimony and goodness of fit (Stott, 2006). In CPT, we follow the proposal by Prelec (1998) and specify the probability weighting function as

$$w(p) = \exp(-(-\ln(p))^{\alpha}) \,,$$

where $0 < \alpha \leq 1$ measures the degree of likelihood sensitivity. When $\alpha = 1$, w is linear in probabilities. When α gets smaller, w becomes more inversely s-shaped. This specification

will usually fail to find its global maximum. We therefore apply the expectation maximization (EM) algorithm to obtain the model's maximum likelihood estimates $\hat{\Psi}$ (Dempster et al., 1977). The EM algorithm proceeds iteratively in two steps: In the E-step, it computes the individual ex-post probabilities of type-membership given the actual fit of the model (see equation (4)). In the subsequent M-step, it updates the fit of the model by using the previously computed ex-post probabilities to maximize each types' log likelihood contribution separately.

of the probability weighting function satisfies the three properties discussed in Section 3.2.[18] In ST, the decision weights depend on the degree of local thinking $0 < \delta \leq 1$ which we directly estimate using equation (1).[19]

6.2 Results

We now present and interpret the result of the structural model. Table 3 exhibits the type-specific parameter estimates of the finite mixture model. The results show that there is substantial heterogeneity in subjects' choices. The choices of 28.4% of subjects are best described by EUT, the choices of 37.9% are best described by CPT, and the choices of 33.7% are best described by ST. When classifying subjects into types using their ex-post probabilities of type-membership, we obtain a clean classification of subjects into 80 EUT-types, 108 CPT-types, and 95 ST-types.[20]

This classification confirms 1 obtained non-parametrically at the aggregate level. The choices of the majority of subjects is best described by either CPT or ST, while – consistent with previous evidence on risky choices of student subjects (Bruhin et al., 2010; Conte et al., 2011) – only a minority is best described by EUT.

On average, the 80 EUT-types display an almost linear utility function which makes them essentially risk neutral. Although the estimated concavity of $\hat{\beta} = 0.080$ is statistically significant, it is almost negligible in economic magnitude. Moreover, among the three types, the EUT-types exhibit the highest level of decision noise which translates into a relatively low estimated choice sensitivity.

The 108 CPT-types exhibit, on average, a concave utility function with $\hat{\beta} = 0.572$ and a low degree of likelihood sensitivity with $\hat{\alpha} = 0.469$. This confirms that the CPT-types' choices are strongly influenced by probability weighting. With these parameter estimates,

[18]We also tested the two-parameter version of Prelec's probability weighting function. However, as the second parameter measuring the function's net index of convexity is very close to 1, results remain virtually unchanged (see Appendix A). Hence, we opt for the one-parameter version to keep the total number of parameter the same for CPT and ST.

[19]In all of the binary choices we use for triggering the Allais Paradoxes, the salience ranking of the states of the world is fully determined by ordering, diminishing sensitivity, and symmetry (see Section 3.3 and the Online Appendix). Hence, we do not need to specify a particular salience function.

[20]Most of the ex-post probabilities of individual type-membership are either close to 0 or 1, confirming that almost all subjects can be unambiguously classified into one of these three types. Appendix E shows histograms with the ex-post probabilities of type-membership.

Table 3: Type-Specific Parameter Estimates of the Finite Mixture Model

Type-specific estimates	EUT	CPT	ST
Relative size (π)	0.284***	0.379***	0.337***
	(0.047)	(0.045)	(0.037)
Concavity of utility function (β)	0.080**	0.572***	0.870***
	(0.033)	(0.055)	(0.015)
Likelihood sensitivity (α)		0.469°°°	
		(0.026)	
Degree of local thinking (δ)			0.924°°°
			(0.013)
Choice sensitivity (σ)	0.010***	0.302***	2.756***
	(0.003)	(0.101)	(0.359)
Number of subjects[a]	80	108	95
Number of observations		23,316	
Log Likelihood		-11,458.71	
AIC		22,937.41	
BIC		23,017.98	

Subject cluster-robust standard errors are reported in parentheses. Significantly different from 0 (1) at the 1% level: *** (°°°).

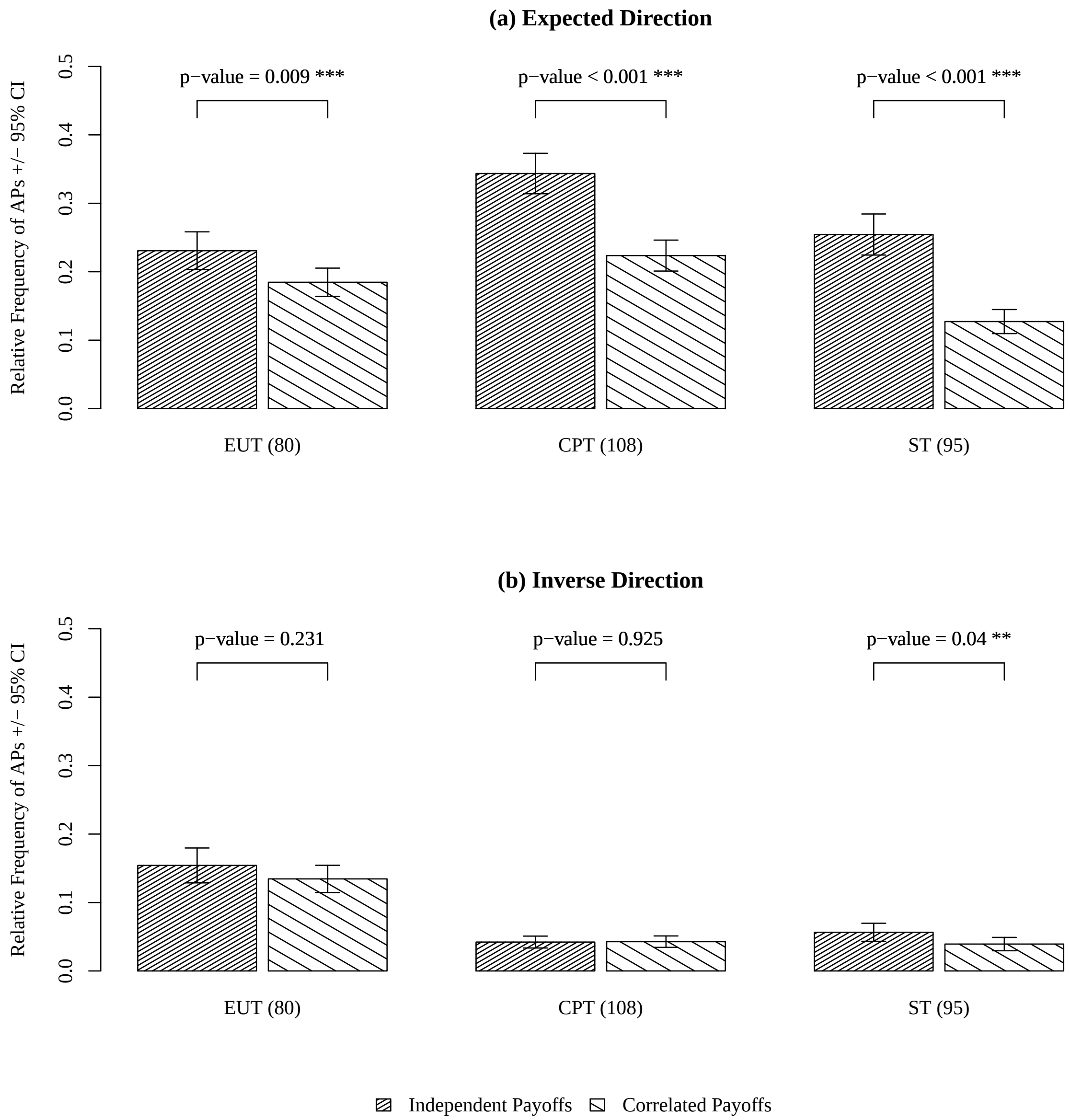
[a] Subjects are assigned to the best-fitting model according to their ex-post probabilities of type-membership (see Equation (4)).

the average CPT-type displays the Common Consequence Allais Paradox discussed in the motivating example in Section 3.

The 95 ST-types display, on average, a strongly concave utility function with $\hat{\beta} = 0.870$ and a weak but statistically significant degree of local thinking corresponding to $\hat{\delta} = 0.924$. Note that although the average ST-type's degree of local thinking appears to be low, she still exhibits the Common Consequence Allais Paradox discussed in the motivating example in Section 3. The reason is that with a strongly concave utility function, even a low degree of local thinking is sufficient to generate the Common Consequence Allais Paradox.[21]

[21] This is mainly due to Inequality (2), as the difference $v(2500) - v(2400)$ gets smaller. On the other hand, Inequality (3) is less affected by the concavity of the utility function and can still be satisfied with a small degree of local thinking.

Figure 5: Relative Frequency of Allais Paradoxes by Type



The figure shows the average frequency of Allais Paradoxes relative to their maximum possible number between lotteries with independent and perfectly correlated payoffs, separately for EUT-, CPT-, and ST-types. Panel (a) depicts the relative frequency of Allais Paradoxes that go in the expected direction. Panel (b) shows the relative frequency of Allais Paradoxes that go in the inverse direction and probably reflect noise in the subjects' choices. The numbers in parentheses indicate the number of subjects in each of the three types.

An interesting question that the finite mixture model's parameter estimates cannot directly address is whether probability weighting and salience exclusively drive the choices of the CPT- and ST-types, respectively, or whether they influence the choices of all types to a varying degree. To answer this question, we turn to Figure 5 which shows the relative frequency of Allais Paradoxes separately for EUT-, CPT-, and ST-types.

The relative frequency of Allais Paradoxes in the expected direction, shown in Panel (a), reveals the following. First, across all types, Allais Paradoxes are more frequent when lottery payoffs are independent than when they are perfectly correlated. This indicates that salience drives the choices not only of the ST-types – for whom the difference is most pronounced – but also of the CPT-types, and, to a smaller extent, even of the EUT-types. Second, all types exhibit a high relative frequency of Allais Paradoxes when lottery payoffs are perfectly correlated. This indicates that probability weighting drives the choices not only of the CPT-types – who display the highest relative frequency of Allais Paradoxes when lottery payoffs are correlated – but also of the ST- and EUT-types.

The relative frequency of Allais Paradoxes in the inverse direction, shown in Panel (b), shows that EUT-types make much noisier choices than CPT- and ST-types. This is consistent with the estimated choice sensitivity being much lower for the EUT-types. Moreover, it indicates that roughly two thirds of the EUT-types' Allais Paradoxes in the expected direction may be due to noise instead of salience or probability weighting.

Taken together, the structural estimations and the resulting classification of subjects into types yield our second main result.

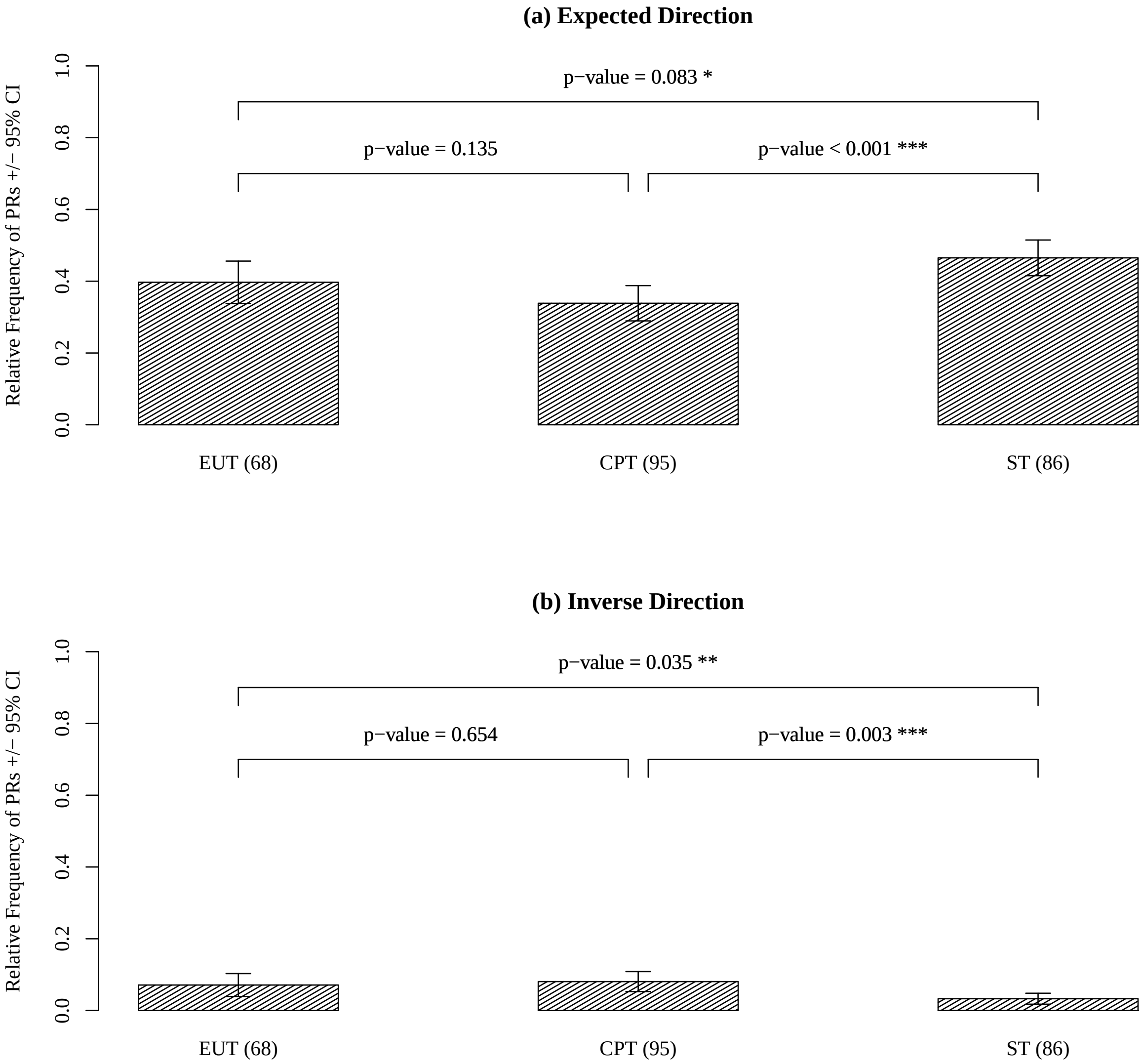
Result 2 *There is vast heterogeneity in the subjects's choices and the population can be segregated in a parsimonious way into 38% CPT-types, 34% ST-types, and 28% EUT-types. However, while this classification indicates the best fitting model for each type, both choice set dependence as well as probability weighting drive the choices of all types, although to a varying extent.*

6.3 Out-of-Sample Predictions

Next, we assess how well this parsimonious classification of subjects into types predicts the frequency of preference reversals out-of-sample, i.e. in the choices subjects made in additional part of the experiment described in Section 4.2.

We expect the ST-types to exhibit substantially more preference reversals than the EUT-

Figure 6: Relative Frequency of Preference Reversals by Type



The figure shows the average frequency of preference reversals by type relative to their maximum possible number in the choices of the additional part of the experiment (see Section 4.2). Panel (a) depicts the relative frequency of preference reversals that go in the expected direction. Panel (b) shows the relative frequency of preference reversals that go in the inverse direction and probably reflect noise in the subjects' choices. The numbers in parentheses indicate the number of subjects in each of the three types. 34 of the 283 subjects (12.0%) are excluded from the analysis because they exhibit more than one switch-point in at least one of the choice menus used for eliciting the certainty equivalents. Exhibiting more than one switch-point is independent of type-membership (χ^2-test of independence: p-value = 0.534).

and CPT-types, since their choices are mainly driven by choice set dependence. However, since choice set dependence also plays some role across in the EUT- and CPT-types, we expect to observe an above noise level of preference reversals in these two types as well.

Figure 6 shows the relative frequency of preference reversals by type and confirms this prediction. Panel (a) displays the preference reversals in the expected direction – i.e. those that can be explained with choice set dependence – while Panel (b) shows the preference reversals in the inverse direction – i.e. those that cannot be explained with choice set dependence and are most likely due to decision noise. The relative frequency of preference reversals in the expected direction is significantly higher for the ST-types than for both the EUT- and the CPT-types (t-tests: p-value = 0.083 for ST vs. EUT, and p-value < 0.001 for ST vs. CPT). The EUT- and CPT-types, on the other hand, exhibit a similar relative frequency of preference reversals in the expected direction (t-tests: p-value = 0.135). In addition, the relative frequency of preference reversals in the inverse direction is substantially lower than in the expected direction across all types, confirming that choice set dependence plays a role in the choices of all three types (t-tests: p-values < 0.001 across all types). In sum, this yields our third main result.

Result 3 *The out-of-sample predictions are qualitatively in line with Result 2, that is, subjects classified as ST-types exhibit more preference reversals than those classified as EUT- and CPT-types. However, since the frequency of preference reversals exceeds the noise level across all types, choice set dependence plays a role in driving the behavior of all three types.*

7 Conclusion

This paper discriminates between probability weighting and choice set dependence both non-parametrically and with a structural model. There are several main conclusions. First, for aggregate choices, both choice set dependence and probability weighting matter. Second, however, there is substantial individual heterogeneity which can be parsimoniously characterized by three types. 38% of subjects are CPT-types whose behavior is predominantly driven by probability weighting, while 34% of subjects are ST-types whose behavior is mainly driven by choice set dependence. The remaining 28% of subjects are EUT-types and their behavior is mostly rational. Finally, this classification of subjects is valid out-of-sample, as the subjects classified as ST-types exhibit significantly more preference reversals than their peers.

These conclusions are directly relevant for the literature that aims at identifying the main behavioral drivers of risky choices. This literature has so far treated probability weighting and choice set dependence as two mutually exclusive frameworks leading to two corresponding major classes of decision theories. Our results show, however, that both play a role for all subjects, although to a varying degree. Knowing about the relative importance of probability weighting and choice set dependence could thus inspire new decision theories taking both frameworks into account and lead to better predictions in various important domains of risk taking behavior, such as investment, asset pricing, insurance, and health behavior.

The conclusions also open up avenues for future research. First, our methodology could be used to study how the relative importance of probability weighting and choice set dependence varies with educational background, cognitive ability, and other socio economic characteristics in the general population. This could lead to new explanations for the observed variation in socio economic outcomes as the different types may fall pray to distinct behavioral traps during their lives. Second, while these results are valid out-of-sample within the domain of risky choices, it would also be interesting to know how far they extend to other domains in which choice set dependence plays a role too, such as consumer, voter, intertemporal, and judicial choices.

References

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

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

Becker, G. M., M. H. DeGroot, and J. Marschak (1964): "Measuring utility by a single-response sequential method," *Behavioral Science*, 9, 226–232.

Booth, A. and P. Nolen (2012): "Salience, risky choices and gender," *Economics Letters*, 117, 517–520.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012a): "Salience in Experimental Tests of the Endowment Effect," *American Economic Review: Papers & Proceedings*, 102, 47–52.

——— (2012b): "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics*, 1243–1285.

——— (2013a): "Salience and Asset Prices," *American Economic Review: Papers & Proceedings*, 103, 623–628.

——— (2013b): "Salience and Consumer Choice," *Journal of Political Economy*, 121, 803–843.

——— (2015): "Salience Theory of Judicial Decisions," *Journal of Legal Studies*, 44, s7–s33.

Bruhin, A., H. Fehr-Duda, and T. Epper (2010): "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 78, 1375–1412.

Camerer, C. F. and T.-H. Ho (1994): "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty*, 8, 167–196.

Chew, S. H. (1989): "Axiomatic utility theories with the betweenness property," *Annals of Operations Research*, 19, 273–298.

Conte, A., J. D. Hey, and P. G. Moffatt (2011): "Mixture Models of Choice Under Risk," *Journal of Econometrics*, 162, 79–88.

Cox, J. C. and S. Epstein (1989): "Preference Reversals Without the Independence Axiom," *American Economic Review*, 79, 408–426.

Deckel, E. (1986): "An axiomatic characterization of preferences under uncertainty: weakening the independence axiom," *Journal of Economic Theory*, 40, 304–318.

Dempster, A., N. Liard, and D. Rubin (1977): "Maximum Likelihood From Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society, Ser. B*, 39, 1–38.

Dertwinkel-Kalt, M., K. Köhler, M. R. J. Lange, and T. Wenzel (2017): "Demand Shifts Due to Salience Effects: Experimental Evidence," *Journal of the European Economic Association*, 15, 626–653.

El-Gamal, M. A. and D. M. Grether (1995): "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association*, 90, 1137–1145.

Fehr, E. and J.-R. Tyran (2005): "Individual Irrationality and Aggregate Outcomes," *Journal of Economic Perspectives*, 19, 43–66.

Fehr-Duda, H., A. Bruhin, T. Epper, and R. Schubert (2010): "Rationality on the rise: Why relative risk aversion increases with stake size," *Journal of Risk and Uncertainty*, 40, 147–180.

Fehr-Duda, H. and T. Epper (2012): "Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences," *Annual Review of Economics*, 4, 567–593.

Fischbacher, U., R. Hertwig, and A. Bruhin (2013): "How to model heterogeneity in costly punishment: insights from responders' response times," *Journal of Behavioral Decision Making*, 26, 462–476.

Gabaix, X. (2015): "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 129, 1661–1710.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1, 114–125.

Grether, D. M. and C. R. Plott (1979): "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69, 623–638.

Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59, 667–687.

Haltiwanger, J. C. and M. Waldman (1985): "Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity,," *American Economic Review*, 75, 326–340.

——— (1989): "Limited Rationality and Strategic Complements: The Implications for Macroeconomics," *Quarterly Journal of Economics*, 104, 463–483.

Harless, D. W. and C. F. Camerer (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62.

Harrison, G. and E. Rutström (2009): "Expected Utility Theory and Prospect Theory: One Wedding and a Decent Funeral," *Experimental Economics*, 12, 133–158.

Hey, J. D. and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62, 1291–1326. [1375,1383].

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

Houser, D., M. Keane, and K. McCabe (2004): "Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm," *Econometrica*, 72, 781–822.

Houser, D. and J. Winter (2004): "How Do Behavioral Assumptions Affect Structural Inference?" *Journal of Business and Economic Statistics*, 22, 64–79.

Isoni, A., G. Loomes, and R. Sugden (2011): "The Willingness to Pay- Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Comment," *American Economic Review*, 101, 991–1011.

Johnson, C., A. Baillon, H. Bleichrodt, Z. Li, D. van Dolder, and P. P. Wakker (2014): "Prince: An Improved Method For Measuring Incentivized Preferences," Working Paper.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1990): "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–292.

Karni, E. and Z. Safra (1987): ""Preference Reversal" and the Observability of Preferences by Experimental Methods," *Econometrica*, 55, 675–685.

Kőszegi, B. and M. Rabin (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121, 1133–1166.

Kőszegi, B. and A. Szeidl (2013): "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics*, 128, 53–104.

Knetsch, J. L. (1989): "The endowment effect and evidence of nonreversible indifference curves," *American Economic Review*, 79, 1277–1284.

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

Lichtenstein, S. and P. Slovic (1971): "Reversals of Preference between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, 89, 46–55.

Lindman, H. R. (1971): "Inconsistent Preferences Among Gambles," *Journal of Experimental Psychology*, 89, 390–397.

Loomes, G. (2010): "Modeling Choice and Valuation in Decision Experiments," *Psychological Review*, 117, 902–924.

Loomes, G. and U. Segal (1994): "Observing different orders of risk aversion," *Journal of Risk and Uncertainty*, 9, 239–256.

Loomes, G., C. Starmer, and R. Sugden (1991): "Observing Violations of Transitivity by Experimental Methods," *Econometrica*, 59, 425–439.

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

McFadden, D. (1981): *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, MA: MIT Press, chap. Econometric Models of Probabilistic Choice.

McLachlan, G. and D. Peel (2000): *Finite Mixture Models*, New York: Wiley Series in Probabilities and Statistics.

Pommerehne, W. W., F. Schneider, and P. Zweifel (1982): "Economic Theory of Choice and the Preference Reversal Phenomenon: A Reexamination," *American Economic Review*, 72, 569–574.

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

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

Reilly, R. J. (1982): "Preference Reversal: Further Evidence and Some Suggested Modifications in Experimental Design," *American Economic Review*, 72.

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

Samuelson, W. and R. Zeckhauser (1988): "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1, 7–59.

Schmidt, U., C. Starmer, and R. Sugden (2008): "Third-generation prospect theory," *Journal of Risk and Uncertainty*, 36, 203–223.

Segal, U. (1988): "Does the Preference Reversal Phenomenon Necessarily Contradict the Indepence Axiom?" *American Economic Review*, 78, 233–236.

Stahl, D. O. and P. W. Wilson (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10.

Starmer, C. (2000): "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice Under Risk," *Journal of Economic Literature*, 38, 332–382.

Stott, H. P. (2006): "Cumulative prospect theory's functional menagerie," *Journal of Risk and Uncertainty*.

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

von Neumann, J. and O. Morgenstern (1953): *Theory of Games and Economic Behavior*, Princeteon, NJ: Princeton University Press.

Wakker, P. P. (2010): *Prospect Theory*, Cambridge University Press.

A Structural Estimations at the Aggregate Level

Table 4: Structural Estimations at the Aggregate Level

Specification of Decision Theory	EUT	CPT	CPT2[a]	ST
Concavity of utility function (β)	0.125**	0.489***	0.503***	0.870***
	(0.010)	(0.045)	(0.038)	(0.012)
Likelihood sensitivity (α)		0.681°°°	0.692°°°	
		(0.027)	(0.030)	
Net index of convexity (γ)			0.962°	
			(0.020)	
Degree of local thinking (δ)				0.931°°°
				(0.008)
Choice sensitivity (σ)	0.020***	0.161***	0.186***	0.014***
	(0.001)	(0.044)	(0.041)	(0.001)
Number of subjects	283	283	283	283
Number of observations	23,316	23,316	23,316	23,316
Log Likelihood	-12,714.52	-12,386.13	-12,382.20	-12,650.83
AIC	25,433.03	24,778.25	24,772.39	25,307.65
BIC	25,449.15	24,802.42	24,804.62	25,331.82

Subject cluster-robust standard errors are reported in parentheses. Significantly different from 0 (1) at the 1% level: *** (°°°); at the 5% level: ** (°°); at the 10% level: * (°)

[a] CPT2 is a specification also based on Cumulative Prospect Theory but uses the more flexible, two-parameter version of the probability weighting function by Prelec (1998): $w(p) = \exp(-\gamma(-\ln(p))^{\alpha})$, where γ is the net index of concavity.

Table 4 reveals that, at the aggregate level, all decision models fit the subjects' choices considerably worse than the fintie mixture model (Table 3) which accounts for heterogeneity in a parsimonious way. Compared to the estimations at the aggregate level, the finite mixture model not only achieves a higher log likelihood but also lower values of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

Moreover, the alternative specification of Cumulative Prospect Theory, CPT2, using the more flexible, two-parameter version of Prelec's probability weighting function exhibits only a negligibly better fit than the baseline specification of CPT. This is because the estimated net index of concavity, $\hat{\gamma} = 0.962$, is very close to one. Thus, we opt for the baseline specification of CPT, as it exhibits the same number of parameters as ST and RT.

Table 5: Structural Estimations at the Aggregate Level (continued)

Specification of Decision Theory	RT[b]	RT2[c]
Concavity of utility function (β)	0.917**	0.575***
	(0.007)	(0.036)
Exponent of regret function (ζ)	0.477°°°	
	(0.018)	
Convexity of regret function (ξ)		0.008***
		(0.001)
Choice sensitivity (σ)	0.061***	0.628***
	(0.005)	(0.054)
Number of subjects	283	283
Number of observations	23,316	23,316
Log Likelihood	-13,452.20	-13,320.20
AIC	26,910.40	26,646.39
BIC	26,934.57	26,670.56

Subject cluster-robust standard errors are reported in parentheses. Significantly different from 0 (1) at the 1% level: *** (°°°); at the 5% level: ** (°°); at the 10% level: * (°)

[b] RT denotes a specification of Regret Theory with a power regret function: $r(x) = x^{\zeta}$ if $x \geq 0$, otherwise $r(x) = -(-x)^{\zeta}$.

[c] RT2 denotes a specification of Regret Theory with an exponential regret function: $r(x) = \exp(\xi\, x)$

Table 5 shows that Regret Theory fits aggregate choices only poorly. Regardless of the applied specification – RT or RT2 – it achieves a lower log likelihood and inferior values of the AIC and the BIC than any of the other decision theories reported in Table 4. Consequently, we opt for ST as our benchmark for choice set dependence.

B Common Ratio Allais Paradox

We now use an example of two lotteries, X and Y, that may induce the Common Ratio Allais Paradox:

$$X = \begin{cases} 6000 & p = \frac{1}{2}q \\ 0 & 1-p = 1-\frac{1}{2}q \end{cases} \qquad \text{vs.} \qquad Y = \begin{cases} 3000 & q \\ 0 & 1-q \end{cases}$$

In this example, the Common Ratio Allais Paradox refers to the empirical finding that if p is high most individuals prefer Y over X, whereas if p is scaled down by a factor $0 < \lambda < 1$

individuals prefer X over Y for a sufficiently small λ.

B.1 EUT

EUT cannot describe the Common Ratio Allais Paradox in the above example. The decision maker evaluates lottery X as $V^{EUT}(X) = p\,v(6000) + (1-p)\,v(0)$ and lottery Y as $V^{EUT}(Y) = 2p\,v(3000) + (1-2p)\,v(0)$. The decision maker chooses lottery X over Y if

$$\begin{aligned} V^{EUT}(X) &> V^{EUT}(Y) \\ p\,v(6000) &> 2p\,v(3000) - p\,v(0) \\ v(6000) &> 2v(3000) - v(0)\,. \end{aligned}$$

Hence, the choice does not depend on the value of the probability p.

B.2 CPT

CPT can describe the Common Ratio Allais Paradox in the above example. The decision maker prefers lottery Y over X if

$$\begin{aligned} V^{CPT}(Y) &> V^{CPT}(X) \\ w(q)\,v(3000) + [1 - w(q)]\,v(0) &> w(p)\,v(6000) + [1 - w(p)]\,v(0) \\ \frac{w(q)}{w(p)} &> \frac{v(6000) - v(0)}{v(3000) - v(0)}\,. \end{aligned}$$

Note that when p is scaled down by the factor λ, the right hand side of the above inequality remains unchanged, while the left hand side decreases due to the probability weighting function's subproportionality, i.e. $\frac{w(q)}{w(p)} > \frac{w(\lambda q)}{w(\lambda p)}$. Hence, for a sufficiently low λ the sign of the above inequality may change, and the decision maker prefers X to Y and exhibits the Common Ratio Allais Paradox.

B.2.1 ST

ST can describe the Common Ratio Allais Paradox in the above example when the two lotteries' payoffs are independent. In this case, there are four states of the world which rank in salience as follows: $\sigma(6000, 0) > \sigma(0, 3000) > \sigma(6000, 3000) > \sigma(0, 0)$. Hence, the decision maker evaluates lottery X as

$$
\begin{aligned}
V^{ST}(X|\{X,Y\}) &= \left[\pi_1^{ST}(6000,0)+\pi_3^{ST}(6000,3000)\right]\, v(6000) \\
&\quad + \left[\pi_2^{ST}(0,3000)+\pi_4^{ST}(0,0)\right]\, v(0)\,.
\end{aligned}
$$

and

$$
\begin{aligned}
V^{ST}(Y|\{X,Y\}) &= \left[\pi_2^{ST}(0,3000)+\pi_3^{ST}(6000,3000)\right]\, v(3000) \\
&\quad + \left[\pi_1^{ST}(6000,0)+\pi_4^{ST}(0,0)\right]\, v(0)\,.
\end{aligned}
$$

Using $v(0) = 0$ and the decision weights given by equation (1), the decision maker prefers Y over X when

$$
\begin{aligned}
v(3000)\,[\delta(1-p)q+\delta^2 pq] &> v(6000)\,[p(1-q)+\delta^2 pq] \\
v(3000)\,2\delta\,[1-p(1-\delta)] &> v(6000)\,[1-2p(1-\delta^2)] \\
\frac{1-p(1-\delta)}{1-2p(1-\delta^2)} &> \frac{v(6000)}{2\delta v(3000)}\,.
\end{aligned}
$$

Note that when p is scaled down, the right hand side of the above inequality remains unchanged, while the left hand side decreases. Hence, for a sufficiently low λ the sign of the above inequality may change, and the decision maker prefers X to Y and exhibits the Common Ratio Allais Paradox.

However, when the two lotteries are correlated, ST can no longer describe the Common Ration Allais Paradox. In this case, there are just three states of the world:

p_s	p	p	$1-2p$
x_s	6000	0	0
y_s	3000	3000	0

The ranking in terms of salience of these three states is as follows: $\sigma(0,3000) > \sigma(6000,3000) > \sigma(0,0)$. Hence, the decision maker evaluates lottery X as

$$
V^{ST}(X|X,Y) = \pi_2^{ST}(6000,3000)\, v(6000) + \left[\pi_1^{ST}(0,3000)+\pi_3^{ST}(0,0)\right]\, v(0)
$$

and evaluates lottery Y as

$$
V^{ST}(Y|X,Y) = \left[\pi_1^{ST}(0,3000)+\pi_2^{ST}(6000,3000)\right]\, v(3000) + \pi_3^{ST}(0,0)\, v(0)
$$

Using $v(0) = 0$ and the decision weights given by equation (1), the decision maker prefers X over Y when

$$\begin{aligned} v(6000)\,\delta p &> v(3000)\,(\delta p + p) \\ v(6000)\,\delta p &> v(3000)\,(\delta p + p) \\ \frac{v(6000)}{v(3000)} &> \frac{1+\delta}{\delta}\,. \end{aligned}$$

Hence, regardless of the value of p, the decision maker always prefers X over Y when the above inequality holds, and otherwise always prefers Y over X. Consequently, the decision maker never exhibits the Common Ratio Allais Paradox when the lotteries' payoffs are correlated.

C Choices to Trigger the Common Ratio Allais Paradox

The binary choices that may trigger the Common Ratio Allais Paradox are based on a subset of a $3 \times 3 \times 2$ design. The design uses the following three different payoff levels:

$$\text{Payoff Level 1:} \quad X = \begin{cases} 6000 & p = \frac{1}{2}q \\ 0 & 1-p = 1-\frac{1}{2}q \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 3000 & q \\ 0 & 1-q \end{cases}$$

$$\text{Payoff Level 2:} \quad X = \begin{cases} 5500 & p = \frac{1}{2}q \\ 500 & 1-p = 1-\frac{1}{2}q \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 3000 & q \\ 500 & 1-q \end{cases}$$

$$\text{Payoff Level 3:} \quad X = \begin{cases} 7000 & p = \frac{1}{2}q \\ 1000 & 1-p = 1-\frac{1}{2}q \end{cases} \quad \text{vs.} \quad Y = \begin{cases} 4000 & q \\ 1000 & 1-q \end{cases}$$

The design features three different probability levels $q \in \{0.90, 0.80, 0.70\}$. To trigger the Common Ratio Allais Paradox each of these three probability levels is scaled down: 0.90 is scaled down to 0.02, 0.80 to 0.10, and 0.70 to 0.20. From the resulting 18 binary choices this design generates, we exclude 3 binary choices which we use for triggering preference reversals and making out-of-sample predictions (see Appendix D).

D Choices to Trigger Preference Reversals

The six binary choices that may trigger preference reversals are based on the following lotteries $\tilde{X}$ and $\tilde{Y}$:

$$\text{Choice 1:} \quad \tilde{X} = \begin{cases} 400 & p = 0.96 \\ 0 & 1-p = 0.04 \end{cases} \quad \text{vs.} \quad \tilde{Y} = \begin{cases} 1600 & q = 0.24 \\ 0 & 1-q = 0.76 \end{cases}$$

$$\text{Choice 2:} \quad \tilde{X} = \begin{cases} 1600 & p = 0.24 \\ 0 & 1-p = 0.76 \end{cases} \quad \text{vs.} \quad \tilde{Y} = \begin{cases} 6400 & q = 0.06 \\ 0 & 1-q = 0.94 \end{cases}$$

$$\text{Choice 3:} \quad \tilde{X} = \begin{cases} 400 & p = 0.96 \\ 0 & 1-p = 0.04 \end{cases} \quad \text{vs.} \quad \tilde{Y} = \begin{cases} 6400 & q = 0.06 \\ 0 & 1-q = 0.94 \end{cases}$$

$$\text{Choice 4:} \quad \tilde{X} = \begin{cases} 3000 & p = 0.90 \\ 0 & 1-p = 0.10 \end{cases} \quad \text{vs.} \quad \tilde{Y} = \begin{cases} 6000 & q = 0.45 \\ 0 & 1-q = 0.55 \end{cases}$$

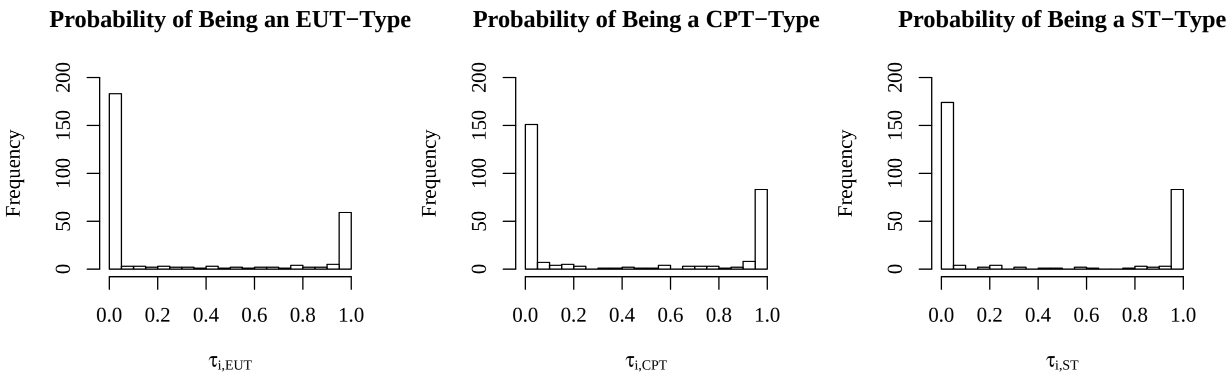
$$\text{Choice 5:} \quad \tilde{X} = \begin{cases} 3000 & p = 0.70 \\ 0 & 1-p = 0.30 \end{cases} \quad \text{vs.} \quad \tilde{Y} = \begin{cases} 6000 & q = 0.35 \\ 0 & 1-q = 0.65 \end{cases}$$

$$\text{Choice 6:} \quad \tilde{X} = \begin{cases} 3000 & p = 0.20 \\ 0 & 1-p = 0.80 \end{cases} \quad \text{vs.} \quad \tilde{Y} = \begin{cases} 6000 & q = 0.10 \\ 0 & 1-q = 0.90 \end{cases}$$

The first three binary choices are similar to the ones stated in Bordalo et al. (2012b). The last three binary choices are based on Payoff Level 1 of the $3 \times 3 \times 2$ design used for generating choices that may trigger the Common Ratio Allais Paradox (see Appendix C).

E Clean Classification of Subjects into Types

Figure 7: Distribution of Ex-Post Probabilities of Type-Membership



The figure shows the distribution of the subjects' individual ex post-probabilities of type-membership, τ_{iM}, according to Equation (4). The resulting classification of subjects into types is clean as for nearly all subjects these post-probabilities of type-membership are either close to 0 or 1.