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)

January 11, 2019

Abstract

The existing literature on choice under risk suggests that probability weighting and choice set dependence both influence risky choices. However, they have not been tested jointly. We design an incentivized laboratory experiment to assess the relative importance of probability weighting and choice set dependence both non-parametrically and with a structural model. Our design uses binary choices between lotteries that may trigger Allais Paradoxes. To reliably discriminate between probability weighting and choice set dependence, we manipulate the lotteries' correlation structure while keeping their marginal distributions constant. The non-parametric analysis reveals that probability weighting and choice set dependence jointly play a role in describing aggregate choices. To take potential heterogeneity into account parsimoniously, we estimate a structural model based on a finite mixture approach. The model classifies subjects into three distinct types: a Cumulative Prospect Theory (CPT) type whose choices are primarily driven by probability weighting, a Salience Theory (ST) type whose choices are predominantly driven by choice set dependence, and an Expected Utility Theory (EUT) type. The structural model uncovers substantial heterogeneity in risk preferences: 38% of subjects are CPT-types, 34% are ST-types, and 28% are EUT-types. This classification of subjects into types also predicts preference reversals out-of-sample. Overall, these results show that probability weighting and choice set dependence play a similarly important role in describing risky choices. Beyond the domain of choice under risk, they may also help to improve our understanding of consumer, investor, and judicial choices.

KEYWORDS: Choice under Risk, Choice Set Dependence, Probability Weighting, Salience Theory, 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 economic research on choice under risk have revealed systematic violations of expected utility theory (EUT; von Neumann and Morgenstern, 1953). As exposed in the famous Allais Paradoxes, subjects frequently exhibit both risk loving and risk averse behavior (Allais, 1953). For example, outside the laboratory, many individuals display risk loving behavior when buying state lottery tickets and risk averse behavior when buying damage insurance (Garrett and Sobel, 1999; Cicchetti and Dubin, 1994; Forrest et al., 2002; Sydnor, 2010). Showing such risk loving and risk averse behavior at the same time violates EUT's independence axiom. Moreover, as demonstrated by Lichtenstein and Slovic (1971) and Lindman (1971), subjects often revert their choice when they have to choose between two lotteries or evaluate them in isolation. Some of these preference reversals violate EUT's transitivity axiom (Cox and Epstein, 1989; Loomes et al., 1991). These and other systematic violations of EUT have spurred the development of various alternative decision theories which fit into two major classes.

The first major class of decision theories uses probability weighting to describe why individuals may behave in a risk loving and risk averse manner at the same time. The most prominent example is Prospect Theory (Kahneman and Tversky, 1979), subsequently generalized to Cumulative Prospect Theory (CPT; Tversky and Kahneman, 1992), which is the best-fitting model for aggregate choices in this class (Starmer, 2000; Wakker, 2010).[1] According to CPT, individuals systematically overweight small probabilities and underweight large probabilities. Consequently, they display risk loving behavior when buying a state lottery ticket because they overestimate the small probability of winning and risk averse behavior when buying damage insurance because they underweight the large probability of not suffering any damage. However, CPT cannot explain preference reversals. Individuals never revert their choice, since they always attach the same value to lotteries, regardless whether they have to choose among them or evaluate them in isolation.[2]

The other major class of decision theories postulates that the evaluation of lotteries is

[1] Another example in this class of decision theories is Rank Dependent Utility (RDU; Quiggin, 1982). In our paper, RDU and CPT formally coincide, as we exclusively use lotteries with non-negative payoffs.

[2] When subjects consider lotteries with non-negative payoffs and derive utility from lottery payoffs rather than absolute wealth levels, then the reference point is equal to zero (Tversky and Kahneman, 1992). In this case, CPT cannot explain preference reversals. However, an extended version of CPT assuming an endogenous reference point can generate preference reversals (Schmidt et al., 2008).

choice set dependent. These theories are able to describe preference reversals as they allow for violations of the transitivity axiom. Prominent members of this class are Salience Theory (ST; Bordalo et al., 2012b) and Regret Theory (RT; Loomes and Sugden, 1982).[3] We focus on ST in this paper because it is becoming the main contender to CPT as the most descriptive theory of choice under risk. According to ST, individuals 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. ST can also explain why individuals often display both risk loving and risk averse behavior. However, the intuition is different than in CPT. Individuals buy state lottery tickets because they overweight the state where they win the big prize due to the large payoff difference between buying the ticket and winning versus not buying the ticket. At the same time, they 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, there are important differences. Besides its ability to describe preference reversals, ST can also naturally explain several behavioral phenomena in consumer choice – such as the endowment effect – (Bordalo et al., 2012a, 2013b; Dertwinkel-Kalt et al., 2017), the counter-cyclicality of risk premia (Bordalo et al., 2013a), and how legally irrelevant information affects judicial decisions (Bordalo et al., 2015). However, in contrast to CPT, whether ST can describe the Allais Paradox or not depends on the choice set, in particular on the lotteries' correlation structure. Hence, to better understand and predict the behavior of consumers, investors, and judges it is crucial to know the relative importance of probability weighting and choice set dependence. However, to the best of our knowledge, their relative importance has not yet been tested jointly.

We address this question with an experiment which allows us to discriminate between probability weighting and choice set dependence while controlling for EUT. The experiment uses a series of incentivized binary choices between lotteries that may trigger Allais Para-

[3]The main difference between ST and RT is how they operationalize choice set dependence. ST focusses on payoff differences while RT focusses on utility differences. Moreover, ST respects diminishing sensitivity as the salience function is concave while RT is at odds with diminishing sensitivity as the regret function is convex. Other examples of choice set dependent theories are by Rubinstein (1988); Aizpurua et al. (1990); Leland (1994); and Loomes (2010).

doxes. Every subject faces the lotteries of each binary choice twice. In one case, the lotteries' payoffs are independent of each other, while in the other, they are perfectly correlated. This manipulation of the correlation structure affects the joint payoff distribution of the lotteries but leaves their marginal payoff distributions unchanged. If risk preferences are driven by probability weighting, the predicted frequency of Allais Paradoxes is the same, as subjects evaluate each lottery in isolation and focus exclusively on its marginal payoff distribution. Hence, CPT can explain the Allais Paradox regardless of whether payoffs are independent or perfectly correlated. However, if risk preferences are driven by choice set dependence, the predicted frequency of Allais Paradoxes differs between the case with independent payoffs and the case with perfectly correlated payoffs due to the change in the joint payoff distribution. As explained in detail in Section 2, when choice set dependence drives risk preferences, the predicted frequency of Allais Paradoxes is positive with independent payoffs and zero with perfectly correlated payoffs. Thus, ST cannot explain the Allais Paradox when payoffs are perfectly correlated. Since EUT can never account for Allais Paradoxes, the design also enables us to control for EUT preferences.

Moreover, 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 presented 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 presented by their joint payoff distribution, regardless whether payoffs are independent or perfectly correlated.[4]

To obtain our first main result, we analyze the importance of probability weighting and choice set dependence non-parametrically at the aggregate level, i.e., at the level of a representative decision maker. In the aggregate, EUT is rejected, and both choice set dependence and probability weighting play a role. Probability weighting plays a role, because the frequency of Allais Paradoxes exceeds the noise-level regardless whether lotteries' payoffs are independent or perfectly correlated.[5] However, choice set dependence plays a role too, be-

[4]For screenshots illustrating the two presentation formats, see Figures 1 and 2 in Section 3.

[5]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 4 for details.

cause Allais Paradoxes occur more frequently when lotteries' payoffs are independent than when they are perfectly correlated. This result holds under both presentation formats, that is, under the canonical presentation as well as the states of the world presentation and does not depend on specific functional forms.

As a next step, we estimate a structural model which offers two conceptual advantages. First, it allows us to take individual heterogeneity into account. This is important because it is unclear whether probability weighting and choice set dependence each influence the behavior of all subjects to the same extent or whether the population consists of distinct preference types. Furthermore, previous research uncovered substantial heterogeneity in risk preferences (Hey and Orme, 1994; Harless and Camerer, 1994; Starmer, 2000), which may be characterized by a majority of non-EUT-types and a minority of EUT-types (Bruhin et al., 2010; Conte et al., 2011). One should take this heterogeneity into account when testing the relative importance of different decision theories or when 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). Second, the structural model yields estimated preference parameters which can be used to calibrate theoretical models or to predict behavior in various contexts.

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 and may suffer from small sample bias – it assumes that there are three distinct preference types: CPT-types whose behavior is mostly driven by probability weighting, ST-types whose behavior is primarily driven by choice set dependence, and EUT-types. By estimating the three types' relative sizes and their average type-specific parameters, the structural model uncovers the relative importance of EUT-behavior and of the two most prominent non-EUT theories of choice under risk. Moreover, it also provides a classification of every subject into the type that best fits her choices.

Another feature of the experimental design benefitting the structural model is that it does not require us to impose a particular salience function. More specifically, the binary choices in our experiment allow us to reliably discriminate between probability weighting and choice set dependence as long as subjects exhibit a salience function which satisfies the three general properties of ordering, diminishing sensitivity, and symmetry (Bordalo et al., 2012b).

The structural model yields the second main result. There is vast heterogeneity in the subjects' risk preferences and the population consists of 38% CPT-types, 34% ST-types, and 28% EUT-types. This result shows that probability weighting and choice set dependence play a similarly important role in describing the non-EUT-types' choices.

Finally, we assess whether structural model's classification of subjects into types has predictive power out-of-sample. This is an important question, since if we want to use the structural model to predict behavior in other contexts, its classification of subjects into types needs to be valid not only for the choices used for estimating the model but also for the choices in these other contexts.

To address this question, 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. By analyzing the frequency of preference reversals in these additional lotteries, we can assess the validity of our classification of subjects into types for choices that we did not use for estimating the structural model.

The out-of-sample predictions about the frequency of preference reversals in the additional lotteries provide the third main result. Subjects classified as ST-types exhibit more preference reversals than those classified as CPT- and EUT-types, confirming that the ST-types' choices are indeed mostly driven by choice set dependence. In conclusion, the classification of subjects into types passes this stringent out-of-sample test and remains valid for choices that we did not use to estimate the structural model.

The paper directly contributes to the empirical literature that tests the performance of probability weighting and choice set dependence at explaining risky choices. On the one hand, there is considerable evidence suggesting that risky choices 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 recognizes that risky choices depend on the choice set, and that many subjects sometimes revert their choices (Lichtenstein and Slovic, 1971; Lindman, 1971; Grether and Plott, 1979; Pommerehne et al., 1982; Reilly, 1982; Cox and Epstein, 1989; Loomes et al., 1991). In particular, 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). More recently, Frydman and Mormann (2018) find that Allais Paradoxes occur less frequently when lotteries' payoffs are perfectly correlated and that the evaluation of lotteries changes if they add an additional "phantom lottery" which subjects can see but not choose. Thus, the existing literature suggests that probability weighting and choice set dependence both influence risky choices.

However, the relative importance of probability weighting and choice set dependence has not been tested jointly. Furthermore, it is unclear whether probability weighting and choice set dependence each influence the behavior of all subjects to the same extent or whether there are distinct preference types among the non-EUT subjects. The present paper provides an answer to these questions by introducing an incentivized experiment which reliably discriminates between choice set dependence and probability weighting, by featuring a parsimonious structural model that takes individual heterogeneity into account, and by assessing whether the results are valid out-of-sample.

The paper also adds to the literature that uses finite mixture models to classify subjects into types. This literature has mostly been focused on discriminating EUT from non-EUT preferences in decision making under risk (Bruhin et al., 2010; Fehr-Duda et al., 2010; Conte et al., 2011).[6] These studies label the non-EUT subjects as CPT-types because they were not designed to discriminate between CPT and ST. The second main result enhances this strand of literature by uncovering additional heterogeneity within the group of non-EUT subjects.

Knowing about this additional heterogeneity within the group of non-EUT individuals could be insightful also in domains other than individual decision making under risk. For instance, in deterministic consumer choice, it may help to identify the relative importance of the competing explanations for the famous endowment effect – i.e., the 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 an individual obtains a good and expects to keep it (Kőszegi and Rabin, 2006). Another explanation is choice set dependence which has the following intuition: when the individual receives an endowment, she compares it to the status quo of having nothing which

[6]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. Other studies have also used finite mixture models to analyze strategic decision making in various domains (for examples see El-Gamal and Grether, 1995; Houser et al., 2004; Houser and Winter, 2004; Stahl and Wilson, 1995; Fischbacher et al., 2013; Bruhin et al., forthcoming)

renders the good's best attribute salient and inflates its valuation (Bordalo et al., 2012a). Since our experimental design and our structural model can isolate the group of subjects whose choices are mostly influenced by choice set dependence, they may offer a way to study it relative importance for explaining the endowment effect. More precisely, one could investigate whether subjects labeled as ST-types based on our experiment and structural model are more prone to exhibit the endowment effect than the other types.

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 ST-types with limited attention from other types.

The paper has the following structure. Section 2 explains the strategy for discriminating between the different decision theories. Section 3 introduces the experimental design. Section 4 presents the non-parametric results at the aggregate level, while Section 5 discusses the structural model, its results, and the out-of-sample predictions. Finally, Section 6 concludes.

2 Discriminating between Decision Theories

This section describes our empirical strategy for discriminating between EUT, probability weighting, and choice set dependence. We focus on the two most prominent behavioral theories, i.e., we represent probability weighting by CPT and choice set dependence by ST. The empirical strategy (i) relies on a series of binary choices between lotteries that may trigger Common Consequence and Common Ratio Allais Paradoxes 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 between lotteries X and Y, taken from Kahneman and Tversky (1979), which may trigger the Common Consequence Allais Paradox.[7]

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

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

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 individuals prefer Y over X, whereas if $z = 0$, most individuals 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.

2.1 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 $v(0) = 0$.[8] 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 since, when comparing the values of the two lotteries $V^{EUT}(X)$ and $V^{EUT}(Y)$, the term involving the common consequence, $p_2 \, v(z)$, cancels out. Hence, the decision maker's choice between X and Y does not depend on the value of the common consequence.

2.2 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

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

lottery in isolation. The decision weights are given by

$$\pi_j^{CPT}(p) = \begin{cases} w(p_1) - w(0) & \text{for } \ j = 1 \\ w\left(\sum_{k=1}^{j} p_k\right) - w\left(\sum_{k=1}^{j-1} p_k\right) & \text{for } \ 2 \leq j \leq J-1 \\ w(1) - w\left(\sum_{k=1}^{J-1} p_k\right) & \text{for } \ j = J \end{cases},$$

where p_k is payoff x_k's probability and w is the probability weighting function. Typically, the probability weighting function in CPT exhibits three properties (Kahneman and Tversky, 1979; Prelec, 1998; Wakker, 2010; Fehr-Duda and Epper, 2012):

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

2. *Inverse S-shape.* The probability weighting function is concave for small probabilities and convex for large probabilities. This ensures 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 explained 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 A.

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

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

In this case, the decision maker tends to prefer Y over X. Due to the decision maker's tendency to overestimate small probabilities and underestimate large probabilities, the decision weight attached to the lowest payoff of X, $1 - w(0.99)$, is larger than its objective probability $p_3 = 0.01$, which renders X unattractive.

In contrast, when $z = 0$, the choice is

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

In this case, the decision maker tends to prefer X over Y. Now, the decision weights of the two lotteries' highest payoffs, $w(0.33)$ 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.

In sum, CPT can always explain the Allais Paradox because the decision weights depend 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.

2.3 ST

According to ST, cognitive limitations cause the decision maker to be 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 $\epsilon > 0$, $\sigma(x_s^1, x_s^2) > \sigma(x_s^1 + \epsilon, x_s^2 + \epsilon)$. 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 permutations of payoffs between lotteries leave the salience of a state unchanged.

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. For $\delta = 1$ the decision maker 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:

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

and

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

Note that the value of each lottery depends on both 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 $V^{ST}(Y) > V^{ST}(X)$, where

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

and

$$V^{ST}(X) = \pi_2^{ST}(2500, 2400)\, v(2500) + \pi_3^{ST}(2400, 2400)\, v(2400) + \pi_1^{ST}(0, 2400)\, v(0)\,.$$

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 $V^{ST}(X) > V^{ST}(Y)$, where

$$V^{ST}(X) = \left[\pi_1^{ST}(2500, 0) + \pi_3^{ST}(2500, 2400)\right]\, v(2500) + \left[\pi_2^{ST}(0, 2400) + \pi_4^{ST}(0, 0)\right]\, v(0)\,,$$

and

$$V^{ST}(Y) = \left[\pi_2^{ST}(0,2400) + \pi_3^{ST}(2500,2400)\right]\, v(2400) + \left[\pi_1^{ST}(2500,0) + \pi_4^{ST}(0,0)\right]\, v(0)\,.$$

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

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

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.

In sum, 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 the two lotteries in the choice set, which changes when we manipulate the correlation structure of the lotteries' payoffs.

2.4 Empirical Strategy

Table 1 summarizes the empirical strategy to discriminate between EUT, probability weighting, and choice set dependence.

Table 1: When can the Allais Paradox occur?

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

EUT can never explain the Allais Paradox. In contrast, probability weighting – represented by CPT – can explain the Allais paradox regardless wether 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.

3 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 EUT-preferences, 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 with out-of-sample-predictions.

3.1 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 two formats which we use to present the binary choices to the subjects.

3.1.1 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 lotteries, 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 EUT-preferences, probability weighting, and choice set dependence by comparing the frequency of Allais Paradoxes in the two blocks within-subjects.

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 each decision theory in the structural model 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:

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

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

$$\text{Payoff Level 3:} \qquad X = \begin{cases} 3000 & p_1 \\ z & p_2 \\ 500 & p_3 \end{cases} \qquad \text{vs.} \qquad Y = \begin{cases} 2600 & p_1 + p_3 \\ z & p_2 \end{cases}$$

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.[9] 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 level of the common consequence, since in combination with the first 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 B.

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.

3.1.2 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 and expose half of the subjects to the canonical presentation and the other half to the states of the world presentation.

The two formats differ in the way they present 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 perfectly correlated payoffs the two presentation formats are identical and display the two lotteries' joint payoff distribution.

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

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

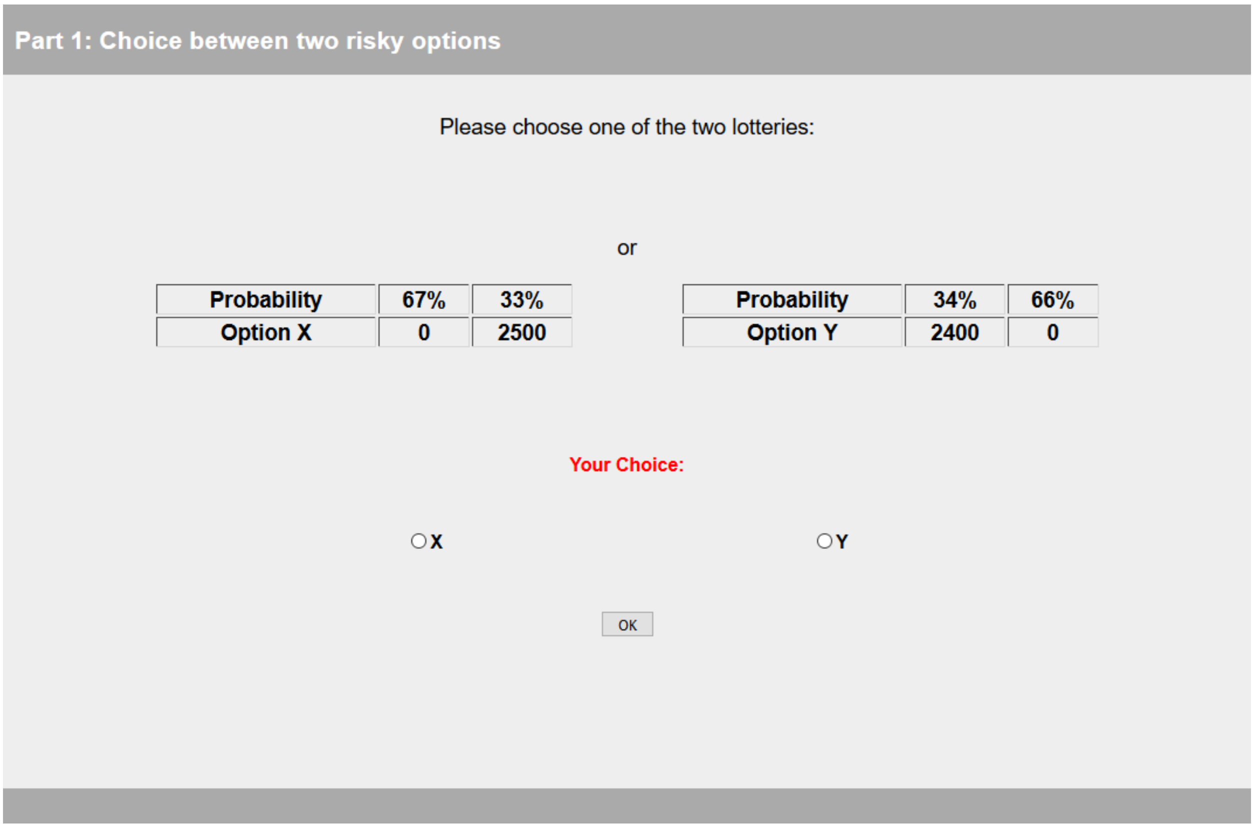
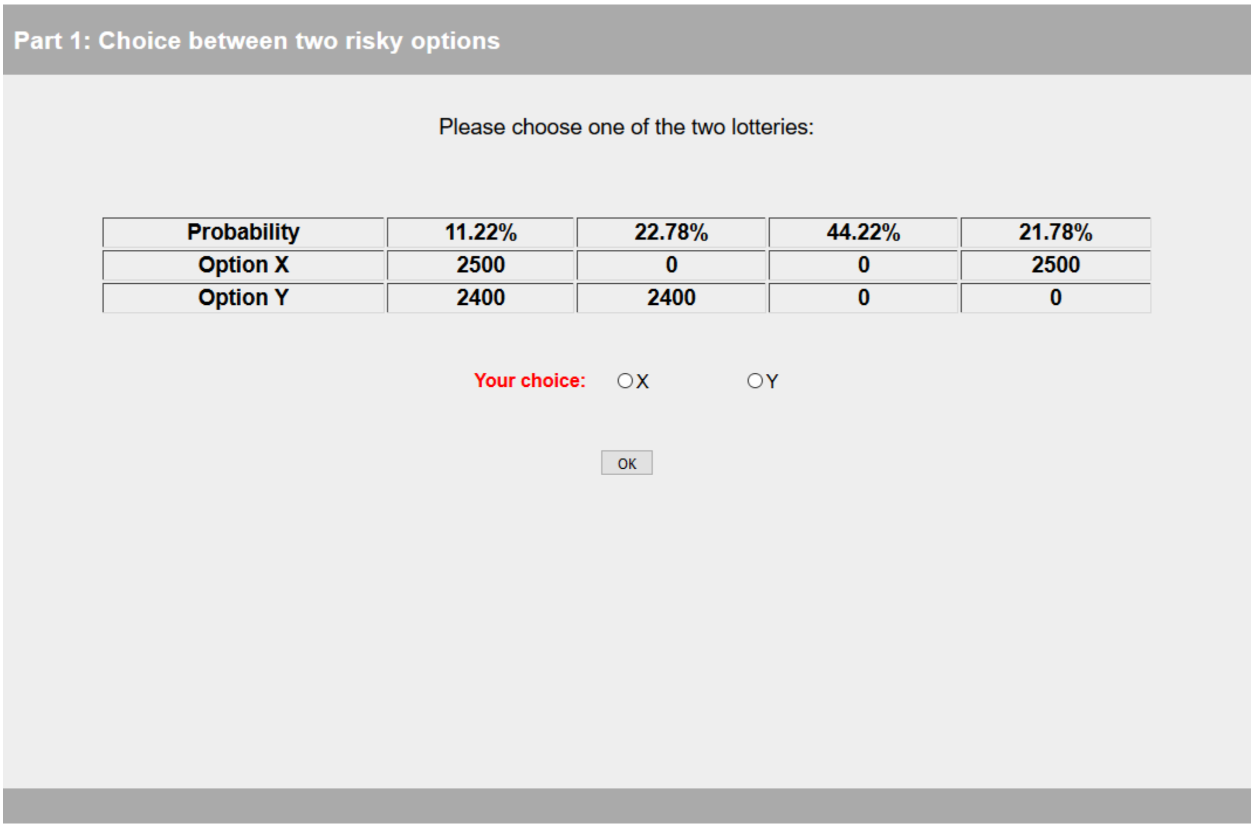


Figure 2: States of the World Presentation of the Binary Choice between Two Lotteries with Independent Payoffs



Part 1: Choice between two risky options

Please choose one of the two lotteries:

Probability	11.22%	22.78%	44.22%	21.78%
Option X	2500	0	0	2500
Option Y	2400	2400	0	0

Your choice: ○X ○Y

OK

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. perfectly 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, our results should remain valid under both presentation formats.

3.2 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 choice, the main part of the experiment contains six binary choices between lotteries that we neither use 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.

We present each of the additional lotteries 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.[10]

We randomize the order in which we elicit the certainty equivalents of the additional lotteries 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 C.

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

[10]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 dropped these subject from the out-of-sample analysis shown in Section 5.3. However, exhibiting more than one switch-point is independent of these subjects' type-membership (χ^2-test of independence: p-value = 0.534).

are six binary choices, each subject can exhibit between 0 and 6 preference reversals.

3.3 Number of Choices

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 perfectly correlated. Table 3 in Appendix D decomposes the number of choices in each presentation format. Regardless of the presentation format, each subject also evaluates 9 lotteries in isolation during the additional part of the experiment.

3.4 Implementation in the Lab and Incentives

We conducted the experiment in the computer lab at the University of Lausanne (LABEX) 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 subjects in total.

At the beginning of the experiment, subjects received 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.[11] 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 the answers to which the assistants verified before subjects could 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 incentive system, as pointed out by Holt (1986). In each part, every subject had to draw a sealed envelope from an urn before making any choices.

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

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.[12]

Each subject received 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 approximately 90 minutes.

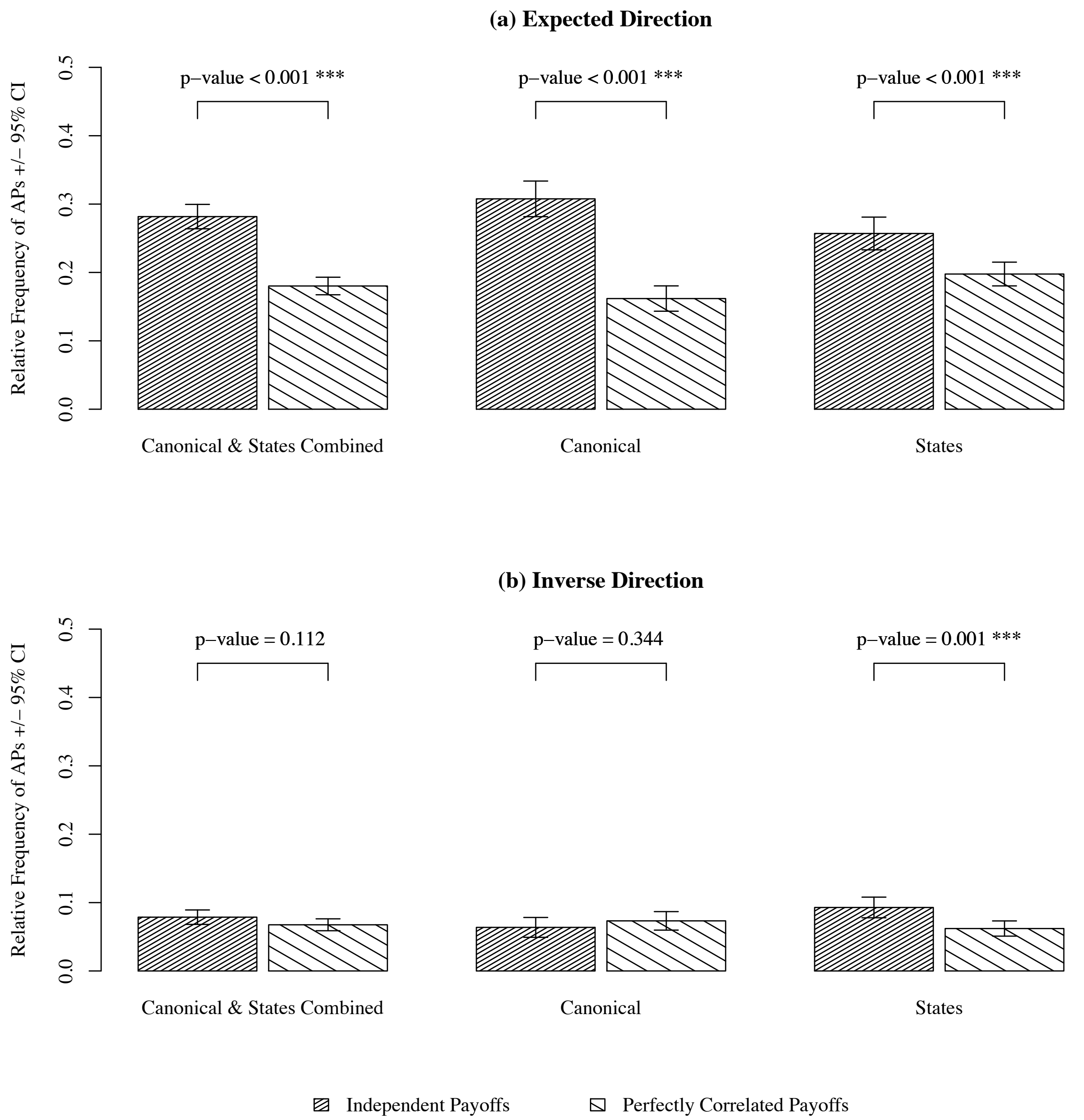
4 Non-Parametric Results

In this section, we present the non-parametric results by analyzing the relative frequency of Allais Paradoxes at the aggregate level. Figure 3 shows the average frequency of Allais Paradoxes relative to their maximum possible number separately for lotteries with independent and perfectly correlated payoffs. Panel (a) exhibits the frequency of Allais Paradoxes in the expected direction, that is, Allais Paradoxes in the direction predicted by CPT and ST. Panel (b) exhibits the frequency of Allais paradoxes in the inverse direction. As neither theory can describe these Allais Paradoxes in the inverse direction, we interpret them as the result of decision noise. This interpretation is in line with the literature which acknowledges the existence and relevance of decision noise (e.g. Hey, 2005).

We start by summarizing the systematic patterns in the frequency of Allais Paradoxes before discussing whether they can be described by EUT, CPT, and ST. There are three systematic patterns. First, Allais Paradoxes are substantially more frequent in the expected than in the inverse direction (t-tests: p-values < 0.001 in all pairwise comparisons). For example, for both presentation formats combined, the frequency of Allais Paradoxes in the expected direction is 28.2% with independent payoffs and 18.0% with perfectly correlated

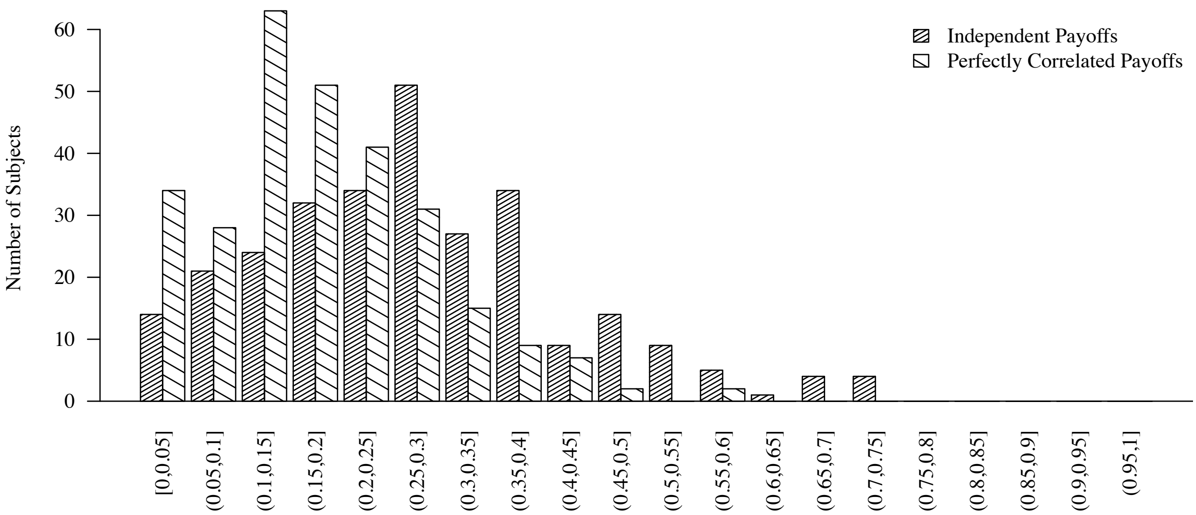
[12] We do not find any statistically significant relationship between these individual characteristics and the classification of subjects into types. Results are available on request.

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 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.

payoffs. The corresponding frequencies in the inverse direction are 7.8% and 6.8%.[13] Second, Allais Paradoxes in the expected direction are significantly more frequent with independent than with perfectly correlated payoffs (t-tests: p-values < 0.001 in all pairwise comparisons). Third, Allais Paradoxes in the inverse direction are about as frequent with independent as with perfectly correlated payoffs – the differences are small in magnitude and only significant for the states of the world presentation (t-tests: p-value $= 0.112$ for both presentations combined, p-value $= 0.344$ for the canonical presentation, and p-value $= 0.001$ for the states of the world presentation). This third pattern confirms that the Allais Paradoxes in the inverse direction probably result from decision noise.

We now asses which of the three theories is able to describe the above patterns in the

[13]These frequencies are close to those found by Huck and Müller (2012) who analyzed the frequency of Allais Paradoxes both in the lab and in the Dutch population. In the lab, they found the frequency of Allais Paradoxes to be 13.0% in the expected direction and 2.7% in the inverse direction. In the Dutch population, the frequencies are 21.7% in the expected and 9% in the inverse direction.

frequency of Allais Paradoxes. EUT fails to describe the patterns as it never predicts any Allais Paradoxes and, thus, their frequency should never exceed the noise-level. CPT and ST can each describe some but not all of the patterns. While CPT can describe that Allais Paradoxes are more frequent in the expected than in the inverse direction, it cannot describe that Allais Paradoxes in the expected direction are more frequent with independent than with perfectly correlated payoffs. ST, on the other hand, can describe that Allais Paradoxes in the expected direction are more frequent with independent than with perfectly correlated payoffs. However, it fails to describe that with perfectly correlated payoffs, Allais Paradoxes are more frequent in the expected than in the inverse direction.

In conclusion, none of the three theories alone can can explain the systematic patterns in the frequency of Allais Paradoxes at the aggregate level. However, CPT and ST seem both to play a role as each of them is able to describe some of the patterns.

Figure 4 confirms this conclusion by taking a more disaggregate look at the data. It shows the distributions of the relative frequency of Allais Paradoxes in the expected direction separately for independent and perfectly correlated lottery payoffs. In both cases, the majority of subjects exhibits a substantial number of Allais Paradoxes, implying that CPT matters. However, 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 ST matters too. Taken together, this non-parametric evidence yields our first main result.

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

This first result suggests that there may be considerable heterogeneity in subjects' risk preferences. In particular, the choices of some subjects may be predominantly influenced by probability weighting whereas the choices of others may be primarily driven by choice set dependence. We examine this possibility with the structural model presented in the next section. Moreover, from now on, we pool the choices from the canonical presentation and the states of the world presentation, since there are no economically relevant differences between these two presentation formats (for details, see Appendix E).

5 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. Later, we also validate the classification of subjects into types using out-of-sample predictions.

5.1 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 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 EUT subjects is important, as the behavior of a minority of our subjects may still be best described by EUT, as previously found by other studies (Bruhin et al., 2010; Conte et al., 2011).

5.1.1 Random Utility Approach

The random utility approach allows the structural model to explicitly take decision noise into account. Consider a subject $i \in \{1, \ldots, N\}$ whose preferences are best described by decision model M in the set of decision models $\mathcal{M} = \{EUT, CPT, ST\}$. 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

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

The parameter σ_M governs the choice sensitivity with respect to differences in the lotteries' deterministic value. 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.,

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

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

5.1.2 Finite Mixture Model

Since risk preferences may be 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 $\pi_{ST} = 1 - \pi_{EUT} - \pi_{CPT}$.[14] Note that the ex-ante 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

[14] Since i's likelihood contribution is highly non-linear, we 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.

subjects cleanly into types, most τ_{iM} should be either 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.

5.1.3 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 likelihood sensitivity and reflects the shape of the probability weighting function. When $\alpha = 1$, w is linear in probabilities. When α gets smaller, w becomes more inversely S-shaped. This specification of the probability weighting function satisfies the three properties discussed in Section 2.2. 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 almost 1, results remain virtually unchanged (see Appendix F). Hence, we opt for the one-parameter version to keep the total number of parameters the same for CPT and ST. In ST, the decision weights depend on the degree of local thinking $0 < \delta \leq 1$ which we directly estimate using equation (1). In all 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 (Section 1 of the Online Appendix shows this for every binary choice we use). Hence, we do not need to specify a particular salience function.

5.2 Structural Model Results

We now present and interpret the result of the structural model. Table 2 exhibits the type-specific parameter estimates of the finite mixture model.[15] The results show that there is substantial heterogeneity in subjects' risk preferences. 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 the remaining 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.[16]

This classification confirms Result 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 (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 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 strongly inverse S-shaped probability weighting function with $\hat{\alpha} = 0.469$. This confirms that the CPT-types' choices are strongly influenced by probability weighting. With these parameter estimates, the average CPT-type displays the Common Consequence Allais Paradox discussed in the motivating example in Section 2.

The 95 ST-types display, on average, a strongly concave utility function with $\hat{\beta} = 0.870$ and a seemingly low 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

[15]Results of structural models neglecting heterogeneity and estimating at the aggregate level can be found in Appendix F. RT fits the data significantly worse than ST, justifying the choice of ST as the main representative of choice set dependent theories. Moreover, as indicated by the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), the finite mixture model fits the subjects' choices considerably better than any of the aggregate models.

[16]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 G shows histograms with the ex-post probabilities of type-membership.

Table 2: 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)).

example in Section 2. 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.[17]

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 with independent than with perfectly correlated lottery payoffs. 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 perfectly correlated – but also of the ST- and EUT-types.

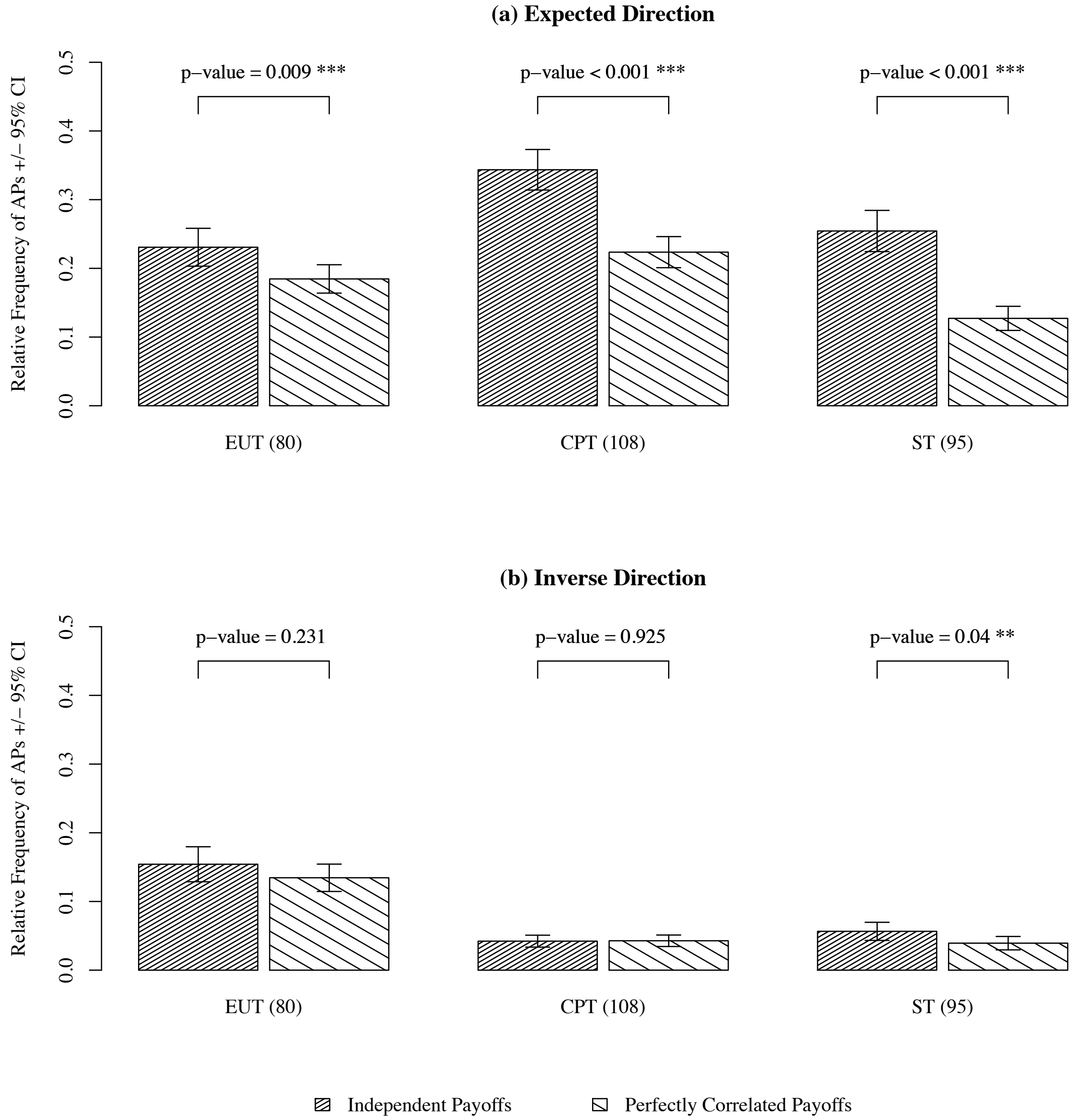
The relative frequency of Allais Paradoxes in the inverse direction, shown in Panel (b), reveals that EUT-types make noisier choices than CPT- and ST-types. This is consistent with the estimated choice sensitivity being much lower for the EUT-types. Moreover, it indicates that roughly two thirds of the 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' risk preferences 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 to a varying extent.*

[17]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.

5.3 Out-of-Sample Predictions

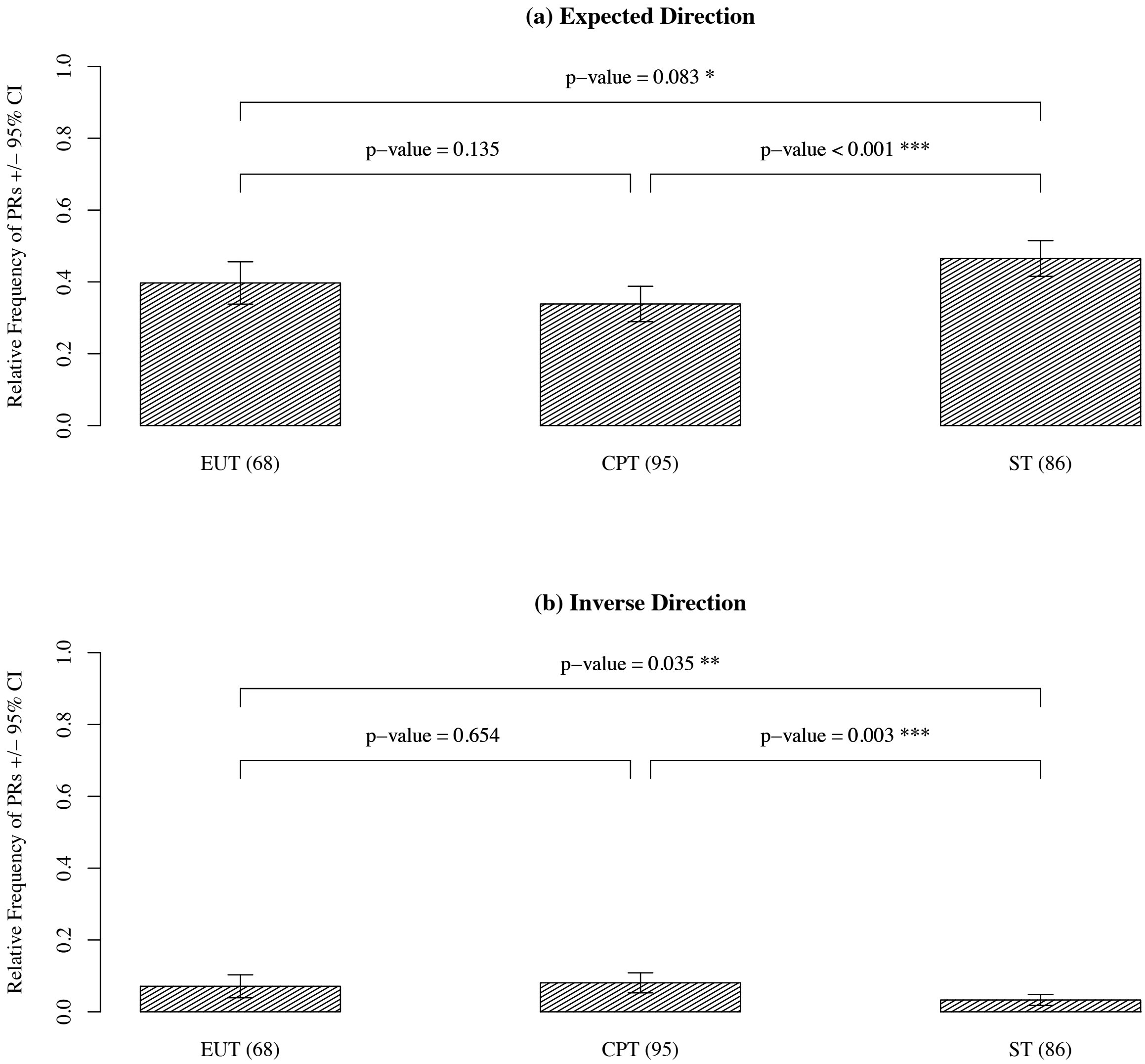
Next, we assess how well this parsimonious classification of subjects into types predicts the frequency of preference reversals out-of-sample, i.e., in the choices subjects made in additional part of the experiment described in Section 3.2.

We expect the ST-types to exhibit substantially more preference reversals than the EUT- 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, the frequency of preference reversals for these types should exceed the noise-level as well.

Figure 6 shows the relative frequency of preference reversals by type. 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 empirical patterns confirm our prediction. 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 relative frequency of preference reversals in the expected direction is similar for the EUT- and the CPT-types (t-tests: p-value = 0.135). In addition, across all types, the relative frequency of preference reversals is substantially lower in the inverse than in the expected direction, 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 in line with Result 2. That is, subjects classified as ST-types exhibit more preference reversals than those classified as EUT- and CPT-types. Moreover, 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.*

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 3.2). 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).

6 Conclusion

The paper assess the relative importance of probability weighting and choice set dependence both non-parametrically and with a structural model. This offers the first stringent test of the two main behavioral theories of choice under risk.

There are several main conclusions. First, for aggregate choices, both choice set dependence and probability weighting matter. This result neither relies on specific functional forms nor on the two presentation formats. Second, there is substantial individual heterogeneity which can be parsimoniously characterized by three types: 38% of subjects primarily exhibit probability weighting and are best described by CPT, 34% are influenced predominantly by choice set dependence and are best described by ST, and 28% are best described by EUT. 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.

Appendices

A 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.

A.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.

A.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.

A.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

$$V^{ST}(X) = \left[\pi_1^{ST}(6000, 0) + \pi_3^{ST}(6000, 3000)\right]\ v(6000) + \left[\pi_2^{ST}(0, 3000) + \pi_4^{ST}(0, 0)\right]\ v(0)\,.$$

and

$$V^{ST}(Y) = \left[\pi_2^{ST}(0, 3000) + \pi_3^{ST}(6000, 3000)\right]\ v(3000) + \left[\pi_1^{ST}(6000, 0) + \pi_4^{ST}(0, 0)\right]\ v(0)\,.$$

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 perfectly 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

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

and evaluates lottery Y as

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

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 perfectly correlated.

B 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 C).

C 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 B).

D Number of Choices

Table 3: Number of Binary Choices by Presentation Format and Type of Allais Paradox

	Canonical		
Allais Paradox	Independent Payoffs	Perfectly 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	Perfectly Correlated Payoffs	Preference Reversal
Common Consequence	18[b]	27	
Common Ratio[a]	15	18	
Total	33	45	6

[a] Three of the $3 \times 3 \times 2 = 18$ 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 B and C).

[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 perfectly 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.

E Comparison between the Two Presentation Formats

Since the states of the world presentation makes the common consequence more obvious (compare Figures 1 and 2), this could influence the number of Allais Paradoxes (Keller, 1985; Birnbaum, 2004; Leland, 2010; Birnbaum et al., 2017). When comparing the two presentation formats, we find two statistically significant but small differences. Table 4 exhibits these differences.

Table 4: Differences in the Frequency of Allais Paradoxes between the Canonical Presentation and the States of the World Presentation

Payoffs	independent	perfectly correlated
Expected direction	0.051	-0.036
Inverse direction	-0.029	0.011

With independent payoffs, the frequency of Allais Paradoxes in the expected direction is 5.1 percentage points higher in the canonical presentation than in the states of the world presentation (t-test: p-value = 0.005). However, this difference is much smaller than in Birnbaum et al. (2017) who argue that the states of the world presentation makes the common consequence more obvious and, thus, may lower the frequency of Allais Paradoxes. Perhaps the difference is much lower in our case because, when presenting the choices to the subjects, Birnbaum et al. (2017) place the common consequence always in the first column while we place it in a random column. This random placement may make the common consequence less obvious in our case. With perfectly correlated payoffs, the frequency of Allais Paradoxes in the expected direction is 3.6 percentage points lower in the canonical presentation than in the states of the world presentation (t-test: p-value = 0.006). The frequency of Allais Paradoxes in the inverse direction is 2.9 percentage points lower in the canonical presentation than in states of the world presentation with independent payoffs (t-test: p-value = 0.007), and 1.1 percentage points higher with perfectly correlated payoffs (t-test: p-value = 0.209).

F Structural Estimations at the Aggregate Level

Table 5: 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)		$0.681^{\circ\circ\circ}$	$0.692^{\circ\circ\circ}$	
		(0.027)	(0.030)	
Net index of convexity (γ)			0.962°	
			(0.020)	
Degree of local thinking (δ)				$0.931^{\circ\circ\circ}$
				(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: *** ($^{\circ\circ\circ}$); at the 5% level: ** ($^{\circ\circ}$); at the 10% level: * ($^{\circ}$)

[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 5 reveals that, at the aggregate level, all decision models fit the subjects' choices considerably worse than the finite mixture model (Table 2) 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 6: 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: * (°)

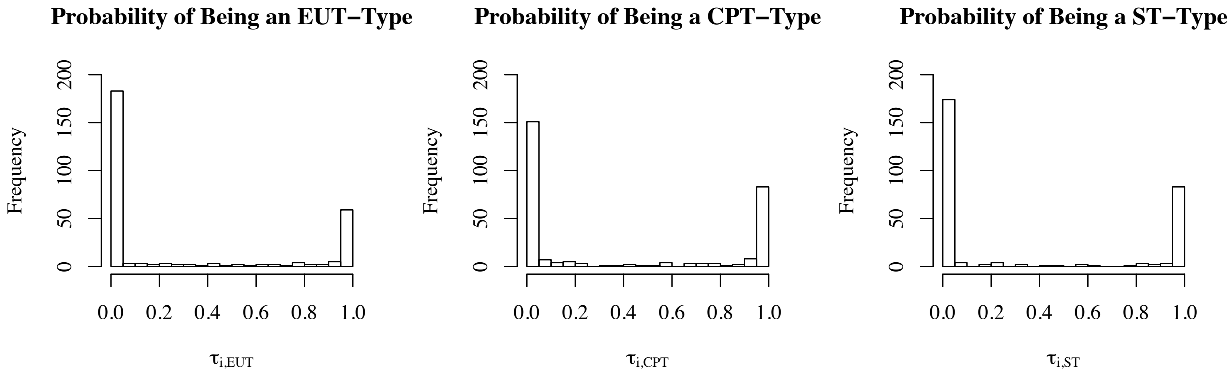
[b] RT denotes a specification of Regret Theory with a power regret function: $r(x) = x^{\zeta}$ if $x \geq 0$, otherwise $r(x) = -(-x)^{\zeta}$.

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

Table 6 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 5. Consequently, we opt for ST as our benchmark for choice set dependence.

G 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.

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.

BIRNBAUM, M. H. (2004): "Causes of Allais Common Consequence Paradoxes: An Experimental Dissection," *Journal of Mathematical Psychology*, 84, 87–106.

BIRNBAUM, M. H., U. SCHMIDT, AND M. D. SCHNEIDER (2017): "Testing independence conditions in the presence of errors and splitting effects," *Journal of Risk and Uncertainty*, 54, 61–85.

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., E. FEHR, AND D. SCHUNK (forthcoming): "The Many Faces of Human Sociality - Uncovering the Distribution and Stability of Social Preferences," *Journal of the European Economic Association*.

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.

CICCHETTI, C. J. AND J. A. DUBIN (1994): "A Microeconometric Analysis of Risk Aversion and the Decision to Self-Insure," *Journal of Political Economy*, 102, 169–186.

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.

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.

Forrest, D., R. Simmons, and N. Chesters (2002): "Buying a Dream: Alternative Models of the Demand for Lotto," *Economic Inquiry*, 485–496.

Frydman, C. and M. Mormann (2018): "The Role of Salience in Choice Under Risk: An Experimental Investigation," Working Paper.

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

Garrett, T. and R. Sobel (1999): "Gamblers Favor Skewness not Risk: Further Evidence from United States Lottery Games," *Economics Letters*, 63, 85–90.

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.

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 Macro- economics," *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. (2005): "Why We Should Not Be Silent About Noise," *Experimental Economics*, 8, 325–345.

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.

Huck, S. and W. Müller (2012): "Allais for all: Revisiting the paradox in a large representative sample," *Journal of Risk and Uncertainty*, 44, 261–293.

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.

Keller, R. L. (1985): "Testing of the "Reduction of Compound Alternatives" Principle," *Omega*, 13, 349–358.

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.

——— (2010): "The hunt for a descriptive theory of choice under risk – A view from the road not taken," *The Journal of Socio-Economics*, 39, 568–577.

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.

Santos-Pinto, L., A. Bruhin, J. Mata, and T. Astebro (2015): "Detecting heterogeneous risk attitudes with mixed gambles," *Theory and Decision*, 79, 573–600.

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

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 Descrip- tive 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*.

Sydnor, J. (2010): "(Over)Insuring Modest Risks," *American Economic Journal: Applied Economics*, 2, 177–199.

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.