Journal of Risk and Uncertainty
https://doi.org/10.1007/s11166-026-09514-9

RESEARCH

Less is more: Decision-making under increasingly complex ambiguity

Xueting Yang[1]

Received: 1 May 2026 / Accepted: 29 September 2026
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

Abstract

This paper examines whether increasingly complex decisions under ambiguity are better described by more complex behavioural models. Using an allocation experiment, we vary structural complexity through the number of possible outcomes and decision stages and compare Expected Utility (EU), αMaxMin (AMM), and Choquet Expected Utility (CEU). EU is the most frequently selected model according to AICc in all three ambiguous-lottery parts, although the strength of its relative advantage varies across tasks, with CEU remaining competitive in the most complex part. Estimated risk-attitude parameters exhibit a relatively stable central pattern across tasks, whereas ambiguity-attitude parameters vary more noticeably across task structures. Overall, the results show that explicitly incorporating ambiguity attitudes does not uniformly improve penalized model fit as structural complexity increases. Rather, the empirical value of more flexible ambiguity models depends on the structure of the choice environment. The findings highlight the importance of evaluating model complexity empirically rather than assuming that structurally more complex decisions necessarily require more flexible behavioural models.

Keywords Ambiguity · Complexity · Allocation elicitation method · Model fitting · Experiments

I am grateful to the Editor, Kip Viscusi, and two anonymous referees for their thoughtful and constructive comments and suggestions, which greatly improved the paper.

✉ Xueting Yang
xueting.yang@york.ac.uk

[1] Department of Economics and Related Studies, University of York, Heslington, York YO10 5DD, England

Published online: 21 October 2026

1 Introduction

Consider an investor allocating a retirement portfolio in a volatile market. They must weigh familiar risks, where outcomes have known probabilities, against ambiguity, where probabilities are not known with certainty. Is a new geopolitical conflict 10% likely or 50%? What are the possible consequences of a technological breakthrough, and how might uncertainty unfold over several stages? Such decisions can involve multiple outcomes and layers of uncertainty, making them structurally more complex than the simple lotteries commonly used to study decision-making under risk and ambiguity. This raises the central question of this paper: As ambiguous choice problems become more complex, do we need more complex behavioural models to describe individual behaviour, or can relatively simple models provide an equally good or better empirical description?

To investigate this question, we conduct a laboratory experiment using an allocation method. Participants allocate endowments across ambiguous lotteries of increasing structural complexity. For each ambiguous event, they are informed of two possible values of the objective probability: a lower probability and a higher probability. Structural complexity varies along two dimensions: the number of possible outcomes and the number of decision stages. We compare three models: Expected Utility (EU), which does not explicitly incorporate ambiguity attitudes, and the more flexible αMaxMin (AMM) and Choquet Expected Utility (CEU) models. By comparing their goodness-of-fit across the experimental parts, we examine whether the empirical value of models with explicit ambiguity parameters increases as the structure of the ambiguous choice problem becomes more complex.

The results reveal a "less is more" pattern. EU is the most frequently selected model according to AICc in all three ambiguous-lottery parts, although the strength of its relative advantage varies across task structures. The evidence in favour of EU is strongest in Part ALB, whereas CEU remains competitive with EU in Part ALC, the most structurally complex part. The parameter estimates provide complementary evidence. The estimated risk-aversion parameter under EU exhibits a relatively stable central pattern across the ambiguous-lottery parts, whereas the ambiguity-attitude parameters in AMM and CEU vary more noticeably across task structures.

These findings show that increasing structural complexity does not systematically increase the empirical advantage of more flexible ambiguity models. Instead, the relative performance of EU, AMM, and CEU depends on the structure of the choice environment. More complex ambiguous decisions therefore do not necessarily require more flexible behavioural models to provide a better empirical description. This result contributes to the literature on decision-making under ambiguity by shifting attention from whether richer models can accommodate ambiguity to whether their additional flexibility provides greater empirical value as ambiguous choice environments become more complex.

2 Related literature

Ambiguity refers to situations in which the probabilities of possible outcomes are not known with certainty. The classic Ellsberg paradox demonstrates that individuals often distinguish between risky prospects with known probabilities and ambiguous prospects with uncertain probabilities, in ways that cannot be accommodated by standard Expected Utility (EU) theory (Ellsberg, 1961; Savage, 1954). This evidence motivated the development of models that explicitly incorporate ambiguity attitudes.

Several influential models provide alternative representations of behaviour under ambiguity. Choquet Expected Utility (CEU) relaxes the additivity of probabilities and evaluates uncertain events using capacities (Schmeidler, 1989). MaxMin Expected Utility represents ambiguity through a set of possible probability distributions and evaluates an act according to its worst-case expected utility (Gilboa & Schmeidler, 1989). The αMaxMin (AMM) model generalizes this approach by allowing decision-makers to place different weights on the best- and worst-case expected utilities (Ghirardato et al., 2004). Other approaches, such as the Smooth Model of Ambiguity Attitudes, provide further ways to separate beliefs about uncertainty from attitudes towards ambiguity (Klibanoff et al., 2005). In the empirical specifications considered in this paper, these models provide richer descriptions of ambiguity attitudes than EU but require additional ambiguity parameters to be identified from observed choices.

A separate literature examines how the structure and complexity of decision problems affect behaviour. Complexity can arise from several features of a choice environment, including the number of possible outcomes, the number of stages through which uncertainty is resolved, and the cognitive demands associated with evaluating alternatives. More generally, bounded-rationality approaches suggest that individuals may simplify complex decision problems when fully processing all available information is costly (Gabaix, 2014). Experimental evidence also indicates that cognitive load can affect economic decision-making, including numerical performance and choices under risk (Deck & Jahedi, 2015). These findings suggest that increasing structural complexity may affect not only observed choices but also the precision with which richer behavioural models can be estimated.

The relationship between ambiguity and complexity has received comparatively less attention. Existing experimental studies have shown that ambiguity attitudes can depend on how ambiguous events and outcomes are structured. For example, Chew et al. (2017) show that behaviour under partial ambiguity depends on the position of an ambiguous event within the outcome distribution. Abdellaoui et al. (2011) develop methods for measuring source-dependent attitudes toward uncertainty across a rich domain of uncertain prospects, demonstrating substantial variation in attitudes across sources of uncertainty. These studies indicate that ambiguity attitudes need not be invariant to the structure or source of uncertainty. However, they do not directly address whether increasing the structural complexity of ambiguous choice problems changes the relative empirical performance of alternative models of ambiguity.

This paper contributes to these literatures by examining model performance across ambiguous choice environments that differ along two dimensions of structural complexity: the number of possible outcomes and the number of decision stages. The experiment compares EU with two models explicitly incorporating ambiguity

attitudes, AMM and CEU. Rather than assuming that structurally richer ambiguous choices necessarily require more flexible behavioural models, we ask whether the relative empirical performance of these models changes as the choice environment becomes more complex. This approach complements previous work on ambiguity attitudes by focusing on the relationship between task structure and model performance, while remaining agnostic about the specific cognitive mechanism generating any observed differences.

3 Experimental design

There were six sessions with 100 participants in total from the University of York. On average, each session lasted around 1 h, and participants received £16.83 pp including £2.50 show-up fee. The total payment of the experiment was £1661.01[1]. During the experiment, participants were asked to answer a series of allocation problems over four parts: Part RL, Part ALA, Part ALB and Part ALC. As this experiment mainly applied the allocation method, it is necessary to start with an explanation of how it is used to elicit preferences.

3.1 Allocation method

Participants are offered a certain number of tokens, F, as their endowment, and are asked to allocate it between two options with exchange rates between money and tokens. To demonstrate its application in preference elicitation, one of the allocation problems in Part RL will be explained in detail below. Options will be referred as colours.

For this problem, participants need to decide how to allocate $F = 100$ tokens between Red and Yellow. The probability of Red happening is p which is 70%, and that of Yellow happening is $1\text{-}p = 30\%$. The exchange rate of Red, E_1, is £0.06 so one token allocated to Red is worth 6 pence, and one token allocated to Yellow is worth 8 pence (i.e. $E_2 = £0.08$). If this problem is chosen for determining participants' payments for this part, the computer will choose either Red or Yellow. Suppose a participant allocates X tokens to Red: she/he will get £XE_1 if Red is played out by the computer. Otherwise, she/he will receive £$(F\text{-}X)E_2$. Figure 1 displays the software interface of this problem. In this case, $X = 50$, so the participant can get the amount of money allocated to Red, £3 = 50 × 0.06, with chance 70% and get £4 which is the amount of money allocated to Yellow with chance 30%. Participants can choose their allocation by moving the slider. It should be noted that changing the allocation only affects the amounts of money allocated to Red and Yellow but not probabilities and exchange rates. The right-hand side histogram in Fig. 1 visualizes the allocation. The horizontal axis shows the chances, and the vertical axis shows the amounts of money, so the chance of Red being chosen is fixed and shown by the horizontal distance of the Red rectangle – in this case 70%. Similarly, the chance of getting the money value of the tokens allocated to Yellow is 30%. Moving the slider will change heights

[1] I gratefully acknowledge Prof. John Hey's generous individual funding for this experiment.



Fig. 1 An example of a problem in the Risk part with an equal split allocation

of the rectangles change solely, but not their width. Probabilities of Red happening and Yellow happening, and exchange rates of them only change between problems. An allocation indicates a mixture of Red and Yellow. Participants can control how much risk they would expose to by changing the allocation. For example, an extreme risk-averse person may follow this half-allocation strategy to guarantee she/he can always get positive payments. If this problem were to be chosen by the computer at the end of Part RL, the lottery in Fig. 1 would be played out. The spinner would be spun by the computer, and where it stops determines the participant's payoff for this problem (Fig. 2).

3.2 The experiment

The experiment is in four parts: Part RL; Part ALA; Part ALB; Part ALC. In each part there are several allocation problems: 10 in Part RL; 10 in Part ALA; 15 in Part ALB; and 15 in Part ALC. After each part, the computer chooses a problem at random, and the payoff on that problem is the participant' payment for that part.

Overall, Part RL contains the simplest risk allocation problems, and the remaining three parts involve ambiguity which is reflected in the lack of knowledge of probabilities. Participants are informed the lowest probability and the highest probability of a colour happening. The complexity of problems increases over parts from ALA to ALC in terms of two dimensions: the number of possible outcomes and the number of stages. These four parts are discussed below.

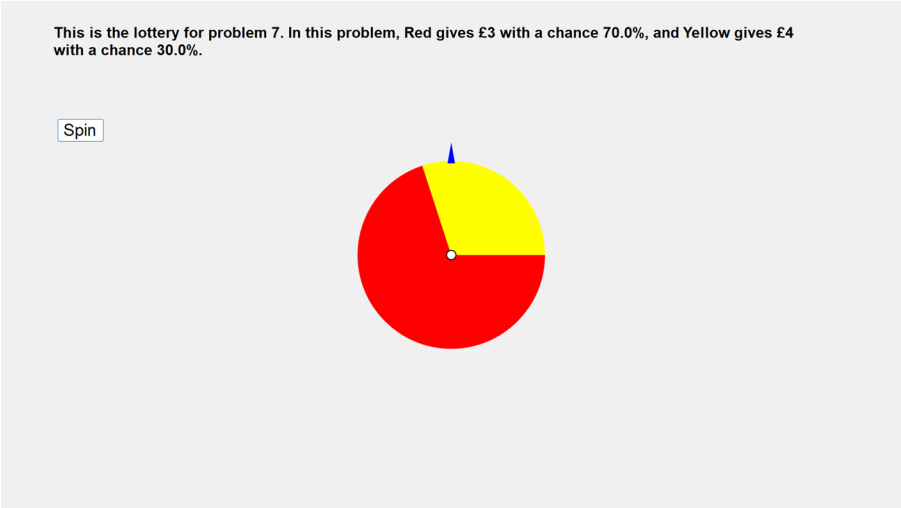


Fig. 2 The computer choosing the problem

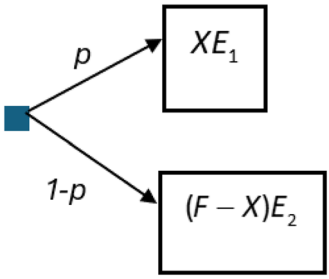


Fig. 3 Game tree of questions in Part RL

i. Part RL (Risky Lottery) (Fig. 3)

To summarise, this is a one-stage risky lottery. A decision-maker (DM) is offered the endowment F and asked to decide how many tokens allocated to Red X with the exchange rate E_1, and Yellow $(F\text{-}X)$ with the exchange rate E_2. The probability of Red happening is p and that of Yellow happening is $1\text{-}p$. If Red happens, the DM will get XE_1. If Yellow happens, the DM will get$(F\text{-}X)E_2$. Figure 1 has shown the software interface of this part.

ii. Part ALA (Ambiguous Lottery A) (Fig. 4)

This is a one-stage ambiguous lottery. A DM is offered the endowment F and asked to decide how many tokens allocated to Red X with the exchange rate E_1, and Yellow $(F\text{-}X)$ with the exchange rate E_2. The DM is only told two possible values to represent

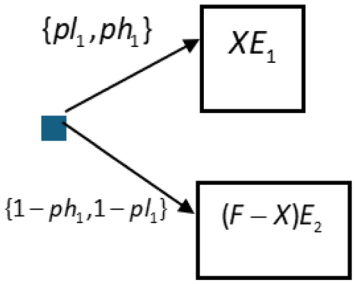


Fig. 4 Game tree of questions in Part ALA

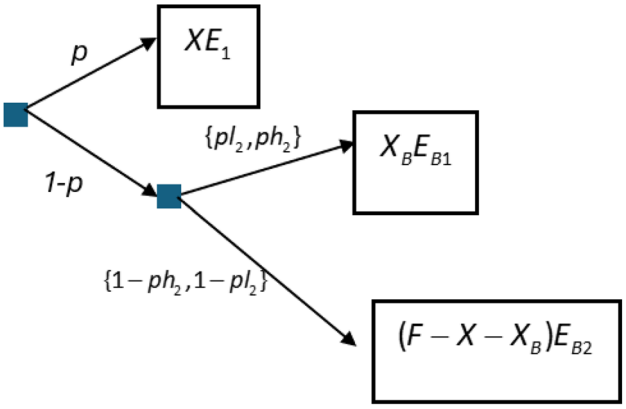


Fig. 5 Game tree of questions in Part ALB

the objective probability of Red happening: the highest probability, ph_1, and the lowest probability, pl_1. Therefore, the highest probability of Yellow happening is $1\text{-}pl_1$ and the lowest probability of Yellow happening is $1\text{-}ph_1$. If Red happens, the DM will get XE_1. If Yellow happens, the DM will get $(F\text{-}X)E_2$.

iii. Part ALB (Ambiguous Lottery B) (Fig. 5)

This is a two-stage ambiguous lottery. A DM is given an endowment of F tokens to allocate between Red, with probability of Red happening p, and another event which has the residual probability of $1\text{-}p$ of happening. Denote the allocation to Red by X. There is an exchange rate E_1 between tokens and money for Red. If Red happens, then the DM will get paid XE_1. If Red does not happen, the DM will have to take another allocation decision: out of the remainder of the endowment $(F\text{-}X)$, the DM must choose XB to allocate to Blue, which would occur with either the highest probability, ph_2, or the lowest probability, pl_2 - if it does occur the DM will get paid X_BE_{B1}; where E_{B1} is the exchange rate between tokens and money for Blue; if it does not occur, the DM will get paid $(F\text{-}X\text{-}X_B)E_{B2}$ where E_{B2} is the exchange rate between tokens and money for Purple.

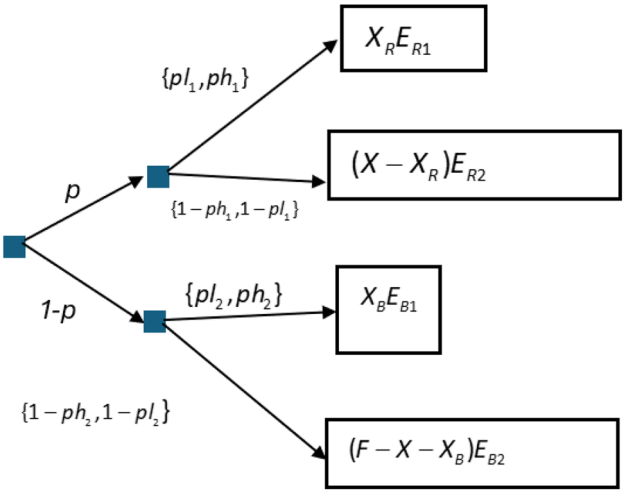


Fig. 6 Game tree of questions in Part ALC

iv. Part ALC (Ambiguous Lottery C) (Fig. 6)

This is a two-stage lottery with two ambiguous branches. A DM is given an endowment of F tokens to allocate between the event (Red or Orange) with probability p of happening, and another event (Blue or Purple) which has the residual probability of *1-p* of happening. Denote the allocation to Red or Orange by X. If Red or Orange does happen, the DM is asked to take another allocation decision: based on the endowment X to choose X_R to allocate to Red which would occur with either the highest probability, *ph 1*, or the lowest probability, *pl 1* , with the exchange rate E_{R1}. If Red happens, the DM will get $X_R E_{R1}$. If Red does not happen, the DM will get paid $(X\text{-}X_R)E_{R2}$, where E_{R2} is the exchange rate between tokens and money for Orange. Similarly, If Red or Orange does not happen that is, (Blue or Purple does happen), the DM is asked to take another allocation decision: out of the remainder of the endowment *(F-X)*, the DM must choose X_B to allocate to Blue which would occur with either highest probability, ph_2, or the lowest probability,pl_2, – if it does occur the DM gets paid $X_B E_{B1}$ where E_{B1} is the exchange rate between tokens and money for Blue; if it does not occur, the DM gets paid $(F\text{-}X\text{-}X_B)E_{B2}$ where E_{B2} is the exchange rate between tokens and money for Purple.

The software interfaces of these three ambiguous lottery parts are displayed in Figs. 7, 8 and 9 respectively.

4 Theoretical framework

The experiment is designed to examine decision-making under risk and ambiguity across choice environments with different levels of structural complexity. To compare the empirical performance of the alternative models, we derive the optimal allocation predicted by each model for each experimental problem and compare these

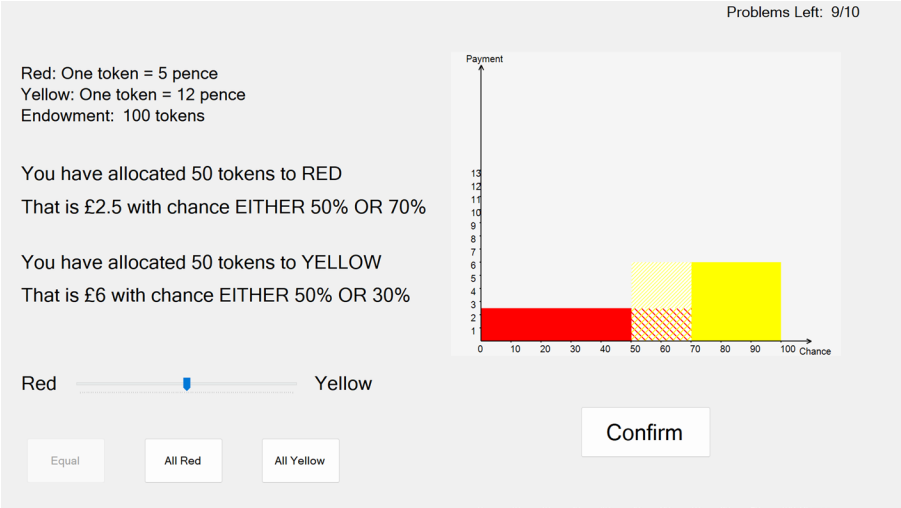


Fig. 7 Software interface of Part ALA

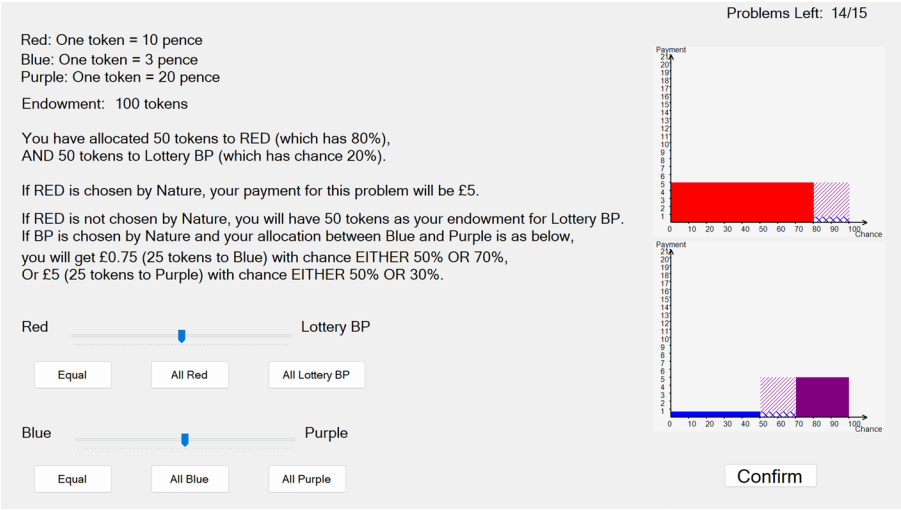


Fig. 8 Software interface of Part ALB

predictions with participants' observed allocations. The model-implied optimal allocations depend on the relevant preference parameters and on the probabilities and exchange rates specified in each problem. We are going to assume that these latter are: (Subjective) Expected Utility ((S)EU), αMaxMin (AMM) model and Choquet Expected Utility (CEU). The theoretical optimal allocation X for each problem in the experiment can be calculated based on these models.

To allow for stochastic variation in observed decisions, we adopt an additive error specification in the spirit of the Fechnerian stochastic choice approach commonly

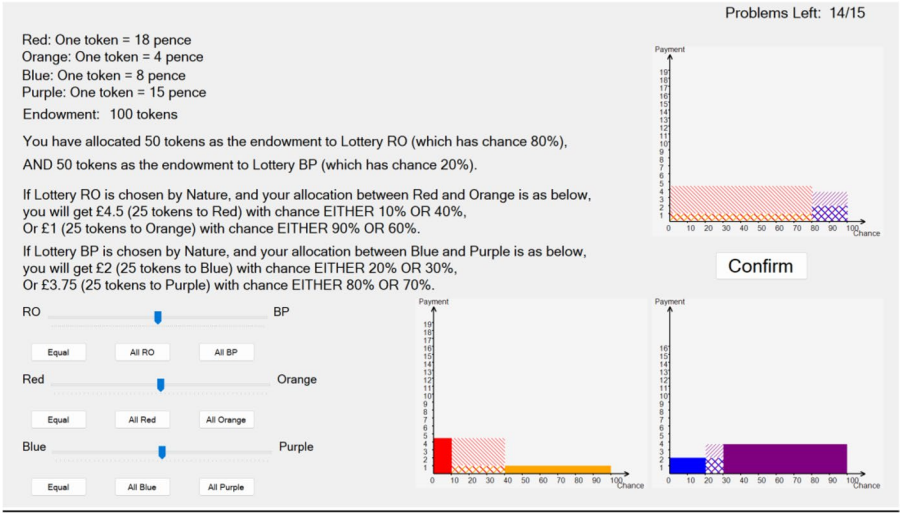


Fig. 9 Software interface of Part ALC

used in experimental decision theory (Hey & Orme, 1994). The actual allocation (AA) is assumed to equal the optimal allocation (X) plus noise with a normal distribution $(0, \sigma^2)$in each problem. Such noise would change over parts and be different among individuals: $AA_{ijk} = X_{ijk} + \varepsilon_{ik}$, where $\varepsilon_{ik} \sim (0, \sigma_{ik}^2)$, i indicates the subject number $i = \{1, 2, ..., 100\}$, j represents the problem number within each part and k here indicates the part $k = \{RL, ALA, ALB, ALC\}$. The optimal allocation is a function of key parameters from these models with given experiment parameters - probabilities and exchange rates – as we discuss below.

The methodology is as follows:

Estimation is conducted separately for each participant, experimental part, and model using maximum likelihood. To address identification issues, we adopt a sequential estimation procedure. We first use the observations from Part RL to estimate the CRRA risk-aversion parameter r. For the ambiguity models, the estimated r from Part RL is then held fixed when estimating the ambiguity parameters. Specifically, we estimate α for the AMM model and γ for the CEU model separately using the observations from Parts ALA, ALB, and ALC. The corresponding noise parameter is estimated separately for each participant, part, and model.

An inevitable by-product of the maximum likelihood estimation is, for each subject, each part and each model, the value of the maximised log-likelihood. This is an obvious measure of the goodness-of-fit of that model for that subject for that part and that model. To compare these goodness-of-fit across models and parts we need to work with a corrected Akaike Criterion considering differences in the number of parameters and the number of observations.

4.1 Model

The underlying utility function is assumed to be the CRRA model: $u(x,r) = \frac{x^{1-r}}{1-r}$ for $r \geqslant 0$ and $r \neq 1$;$u(x) = \ln(x)$ for $r = 1$. For the former, as a more general case, r is the CRRA index of risk aversion and u takes the form $u(x,r) = \frac{x^{1-r}}{1-r}$ and $u'(x,r) = x^{-r}$. For the latter, $u(x) = \ln(x)$ and $u'(x) = x^{-1}$. To find the optimal allocation, we took the first-order derivative of the expected utility with respect to X and put the answer equal to 0.

a. (S)EU model

It is the standard expected utility theory with the key parameter r from the CRRA model. If $r = 0$, individuals would be risk neutral. If $r > 0$, they would be risk averse, and the larger the r is, the more risk averse the DM. Otherwise, they are risk seeking In an ambiguous situation, it becomes the subjective EU model which assumes that individuals make decisions based on their *subjective* beliefs about the probabilities (Savage, 1954). As the target is to estimate r and check if r changes over three parts (Part ALA, Part ALB and Part ALC), the subjective probability in the ambiguous case was simply assumed to be the average of the highest probability, *ph*, and the lowest probability, *pl*.

b. AMM model

This model focuses on individual decision making under ambiguity. Developing from the MaxMin model (Gilboa & Schmeidler, 1989), it introduces a new parameter α to generalize this story (Ghirardato et al., 2004). The AMM model assumes that people consider both the best case and worst case in a weighted way. $\alpha \in [0,1]$ is not only the weighted average parameter but also the ambiguity aversion index. If $\alpha = 1$, people would be fully pessimistic, only looking at worst cases, and the AMM model reduces to the MaxMin model. In contrast, if $\alpha = 0$, only best cases matter, and the model reduces to the MaxMax model. When $\alpha = 0.5$, individuals weight these two extreme situations equally. Combining the setting of the experiment and the assumption in the EU model, under such situation, the AMM model reduces to the EU model mathematically. Individuals are treated as ambiguity neutral, as they behave similarly to a standard EU maximiser with a 'representative' prior $p = \frac{pl+ph}{2}$. Generally, if $0 \leqslant \alpha < 0.5$, people would be ambiguity seeking; if $0.5 < \alpha \leqslant 1$, they would be ambiguity averse.

c. CEU model

The standard EU model relies on additive probabilities: for disjoint events A and B, $P(A \cup B) = P(A) + P(B)$. CEU relaxes this additivity requirement by evaluating events using a capacity $v(.)$ rather than an additive probability measure. It states that under ambiguity, additivity is not necessary, and individuals are assumed to rely on capacities $v(.)$ to evaluate events instead of probabilities (Schmeidler, 1989). We applied the power capacity function to capture how people distort their subjective probabilities, which are same as those in the EU model, in ambiguous situations: $v(p) = p^{\gamma}$, where $\gamma > 0$. In addition to be treated as a probability distortion parameter, γ reflects the degree of ambiguity aversion. If $\gamma > 1$, individuals would be ambiguity averse. If $1 > \gamma > 0$, they would be ambiguity lovers. Otherwise, this model collapses to the (S)EU model, and individuals would be seen as ambiguity neutral. This function is commonly used. Given a two-positive-outcome lottery, the power capacity will be identical to the probability weighting function applied in Cumulative Prospect Theory (Tversky & Kahneman, 1992). Also, if probabilities become objective, the CEU model with the power capacity can be directly treated as Rank-Dependent Utility Theory.

4.2 Hypothesis

Compared with AMM and CEU, (S)EU provides a relatively parsimonious representation of behaviour because it does not explicitly incorporate a parameter capturing ambiguity attitudes. A natural intuition is that more complex choice problems may require more flexible behavioural models to describe individual behaviour. An alternative possibility, however, is that the empirical benefit of such flexibility decreases as ambiguous choice problems become more structurally complex. We therefore examine whether the relative goodness-of-fit of (S)EU, AMM, and CEU varies with the structural complexity of the ambiguous lotteries.

One motivation for this possibility comes from bounded rationality. Gabaix (2014) argues that individuals may simplify information processing when facing complex decision environments. Experimental evidence also shows that cognitive load can affect economic decision-making and the consistency of observed choices (Deck & Jahedi, 2015). In the present setting, increasing structural complexity may therefore reduce the extent to which additional flexibility in modelling ambiguity improves the empirical description of behaviour. Consequently, a relatively parsimonious model such as (S)EU may perform as well as, or better than, models that explicitly incorporate ambiguity attitudes in more complex choice environments.

A related consideration concerns the empirical value of additional model flexibility. AMM and CEU allow behaviour under ambiguity to be captured through additional ambiguity-attitude parameters. The usefulness of this flexibility depends on whether observed choices contain sufficient information for these parameters to improve model fit. As the structure of ambiguous lotteries becomes more complex, additional flexibility need not translate into a corresponding improvement in empiri-

cal performance. This provides a second motivation for examining whether the relative advantage of ambiguity models varies systematically with structural complexity.

Accordingly, we hypothesize that increasing structural complexity does not necessarily increase the relative empirical advantage of models with explicit ambiguity parameters. A relatively parsimonious model such as (S)EU may provide an equally good or better description of behaviour as ambiguous choice problems become more complex.

5 Econometric method

A key objective of the econometric analysis is to investigate whether estimated risk and ambiguity attitudes and the relative fit of alternative models differ across experimental parts as the structural complexity of the ambiguous lotteries increases. We estimate the relevant model parameters separately for each participant and experimental part using maximum likelihood estimation. Table 1 summarizes the parameters estimated under each model.

As an illustration, consider the simplest risk-averse case to show how model predictions and observed allocations can be used to infer the value of the parameter r. In Part RL, the expected utility is $EU(X) = pu(XE_1, r) + (1-p)u((F-X)E_2, r)$. After differentiating it w.r.t. X and setting the first-order condition equal to zero, $X = \frac{A}{A+B}F$, where $A = (1-p)^{-1/r}{E_2}^{1-1/r}$ and $B = p^{-1/r}{E_1}^{1-1/r}$. As values of p, E_1 and E_2 are known, the optimal allocation X can be expressed as a function of r, denoted *X(r)*. Based on each participant's observed allocations and the corresponding model-implied optimal allocations, maximum likelihood estimation is used to obtain the best-fitting value of r. The Technical Appendix provides further details on the

Table 1 Parameters estimated by model and experimental part

Model	EU	RN	AMM	CEU
Number of parameters	8	4	8	8
Part RL	(r_{RL}, σ_{rRL})	σ_{rRL}	(r_{rRL}, σ_{rRL})	(r_{RL}, σ_{rRL})
Part ALA	(r_{ALA}, σ_{rALA})	σ_{rALA}	$(\alpha_{ALA}, \sigma_{\alpha ALA})$	$(\gamma_{ALA}, \sigma_{\gamma ALA})$
Part ALB	(r_{ALB}, σ_{rALB})	σ_{rALB}	$(\alpha_{ALB}, \sigma_{\alpha ALB})$	$(\gamma_{ALB}, \sigma_{\gamma ALB})$
Part ALC	(r_{ALC}, σ_{rALC})	σ_{rALC}	$(\alpha_{ALC}, \sigma_{\alpha ALc})$	$(\gamma_{ALC}, \sigma_{\gamma ALC})$

RN indicates the risk neutral case based on the EU model (r = 0). It is considered as the nested benchmark of EU

For the EU model, r and the corresponding noise parameter σr are estimated separately in each experimental part. For the AMM model, r is first estimated using Part RL and then held fixed when estimating α and the corresponding noise parameter σα separately in Parts ALA, ALB, and ALC. The same sequential procedure is applied to the CEU model, with γ and the corresponding noise parameter σγ estimated separately in each ambiguous-lottery part conditional on the estimated r from Part RL. This sequential estimation procedure is adopted to avoid identification problems that arise when r, the ambiguity parameter α or γ, and the noise parameter are estimated simultaneously

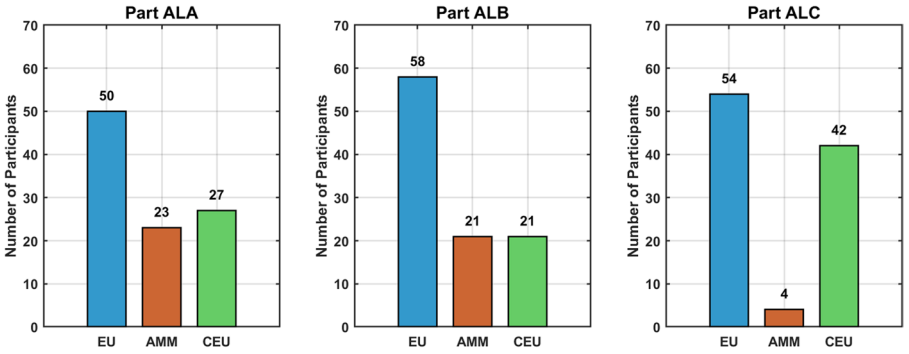


Fig. 10 Distribution of the best fitting model over parts

derivation of optimal allocations under the different models and experimental parts, including the risk-neutral and risk-seeking cases.

6 Results

The specific estimation results of parameters under different models are presented on the EXEC website. Because the number of observations available for each participant is relatively small and differs across experimental parts, we use the small-sample corrected Akaike Information Criterion (AICc) for model comparison. AICc is computed based on the regular AIC which is calculated by *2k – 2Loglikelihood*, where k is the number of parameters[2] of the model and *Loglikelihood* is the maximised log-likelihood. With AIC, AICc is given by $AICc = AIC + \frac{2k(k+1)}{n-k-1}$, where n indicates the sample size. The lower the value of AICc, the corresponding model fits better. Figure 10 shows the model-fit rankings across parts[3]. For example, the number '50' in Part ALA indicates that, among the 100 participants, EU is the best-fitting model according to AICc for 50 participants. Overall, EU is the most frequently selected model in all three parts, providing the lowest AICc for 50 participants in ALA, 58 in ALB, and 54 in ALC. CEU also performs relatively well in ALC, providing the lowest AICc for 42 participants. In contrast, AMM is selected as the best-fitting model for only 4 participants in ALC. Chi-square goodness-of-fit tests show that the distribution of best-fitting models differs significantly from an equal-frequency benchmark in all three parts (ALA, $\chi^2(2) = 12.74$, $p = 0.002$; ALB, $\chi^2(2) = 27.38$, $p < 0.001$; and ALC, $\chi^2(2) = 40.88$, $p < 0.001$).

To further assess whether EU provides a better fit than the alternative models, we compare participant-level AICc values. For each comparison, we calculate

[2] Unlike number of parameters showing in Table 1, for the part-specific AICc calculations, k denotes the number of parameters estimated from observations in the corresponding part; thus, $k=2$ for EU, AMM, and CEU and $k=1$ for RN.

[3] As complexity varies only across the ambiguous lotteries, the comparison of model fit focuses on Parts ALA, ALB, and ALC. The structural complexity of the ambiguous lotteries increases from ALA to ALC.

Table 2 Participant-level comparisons of AICc across models

Part	Comparison	Mean$\Delta AICc$	Median$\Delta AICc$	EU Lower AICc	t-test p	Wilcoxon p
ALA	EU-AMM	−10.214	−0.539	66	0.034	<0.001
ALA	EU-CEU	−9.085	−0.582	70	0.052	<0.001
ALA	EU-RN	−204.814	−250.144	96	<0.001	<0.001
ALB	EU-AMM	−12.394	−4.887	72	<0.001	<0.001
ALB	EU-CEU	−11.732	−6.043	74	<0.001	<0.001
ALB	EU-RN	−118.672	−60.691	97	<0.001	<0.001
ALC	EU-AMM	−33.159	−13.097	78	0.004	<0.001
ALC	EU-CEU	−13.328	−1.484	57	0.107	0.227
ALC	EU-RN	−124.024	−82.411	98	<0.001	<0.001

$\Delta AICc = AICc_{EU} - AICc_{comparison}$; negative values favour EU. "EU lower AICc" reports the number of participants, out of 100, for whom EU yields a lower AICc than the comparison model. p-values are one-sided, corresponding to the alternative that EU has lower AICc. RN is reported as a nested risk-neutral benchmark rather than as a separate substantive competing model

$\Delta AICc = AICc_{EU} - AICc_{comparison}$, such that a negative difference indicates a lower AICc for EU. We first conduct one-tailed paired t-tests against the alternative that the mean difference is negative. Because participant-level AICc differences may not be normally distributed and may be affected by large differences for a subset of participants, we additionally conduct one-tailed Wilcoxon signed-rank tests as a robustness check. Table 2 reports the mean and median AICc differences, the number of participants for whom EU yields a lower AICc, and the results of both tests.

The results provide the strongest evidence in favour of EU in Part ALB. EU yields a lower AICc than AMM for 72 participants and than CEU for 74 participants, with both the paired t-tests and Wilcoxon signed-rank tests rejecting the null hypotheses (all $p < 0.001$). In Part ALC, EU also performs significantly better than AMM according to both the paired t-test ($p = 0.004$) and the Wilcoxon signed-rank test ($p < 0.001$), with a lower AICc for 78 participants. In contrast, there is no significant difference between EU and CEU in ALC according to either the paired t-test ($p = 0.107$) or the Wilcoxon signed-rank test ($p = 0.227$), and EU has a lower AICc for 57 participants.

In Part ALA, EU performs significantly better than AMM according to both the paired t-test ($p = 0.034$) and the Wilcoxon signed-rank test ($p < 0.001$), with a lower AICc for 66 participants. For EU versus CEU, the evidence differs slightly across the two tests. The paired t-test does not reject the null at the 5% level ($p = 0.052$), whereas the Wilcoxon signed-rank test favours EU ($p < 0.001$). EU yields a lower AICc than CEU for 70 participants. Thus, the evidence in ALA generally favours EU, although the EU–CEU comparison is sensitive to the statistical test employed.

As discussed above, RN is treated as the nested risk-neutral benchmark of EU rather than as a separate competing behavioural model. We therefore report EU–RN comparisons separately as a supplementary benchmark. EU yields a lower AICc than RN for 96, 97, and 98 participants in ALA, ALB, and ALC, respectively, and both the paired t-tests and Wilcoxon signed-rank tests favour EU in all three parts (all $p < 0.001$).

Because AICc accounts for model complexity, the comparisons provide a measure of the relative penalized fit of the competing models. The results show that the empirical benefit of explicitly incorporating ambiguity attitudes varies across the

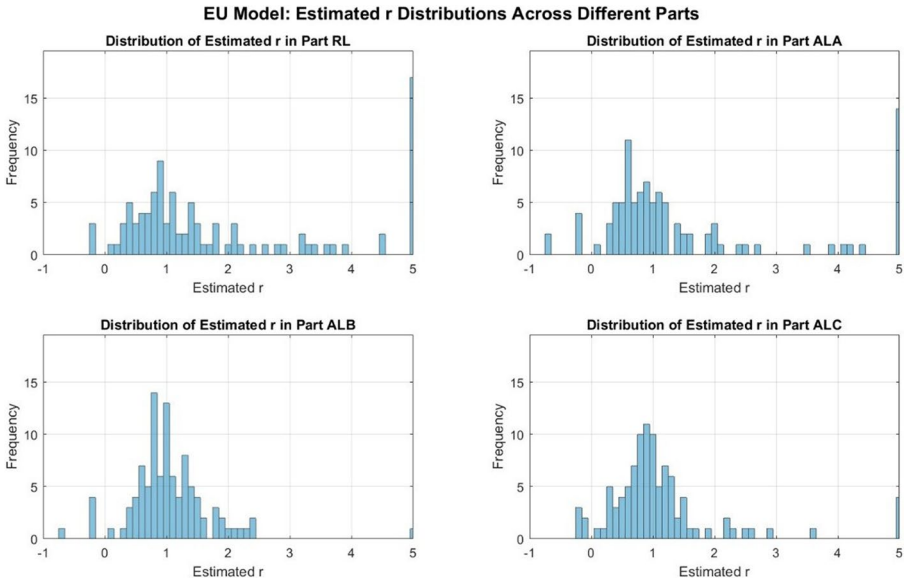


Fig. 11 Distribution of estimated r across parts for all participants ($N = 100$)

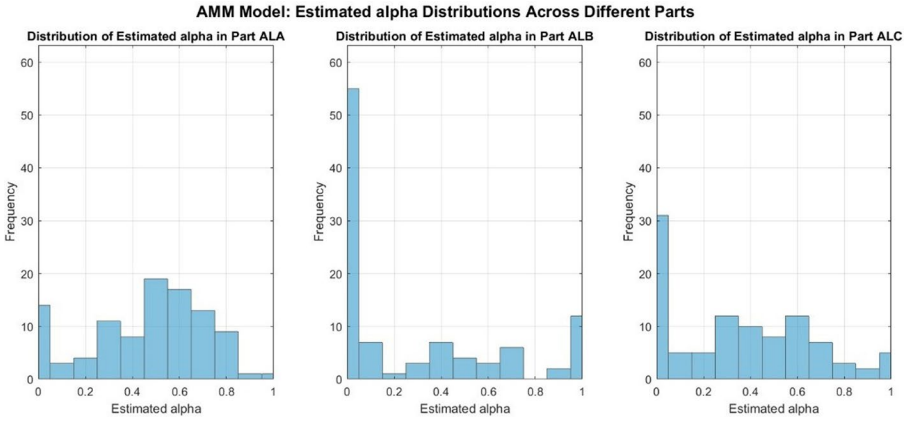


Fig. 12 Distribution of estimated α across parts for all participants ($N = 100$)

three ambiguous-lottery parts. EU performs particularly well in Part ALB, whereas CEU becomes more competitive in Part ALC. Thus, increasing structural complexity does not systematically increase the relative empirical advantage of models with explicit ambiguity parameters.

The observed parameter distributions provide additional descriptive evidence on how model estimates vary across parts. Figures 11, 12, and 13 report the relevant parameter estimates for all 100 participants, irrespective of which model provides the lowest AICc for a given participant. Thus, the distributions describe the estimated parameters under each model for the full sample rather than only for

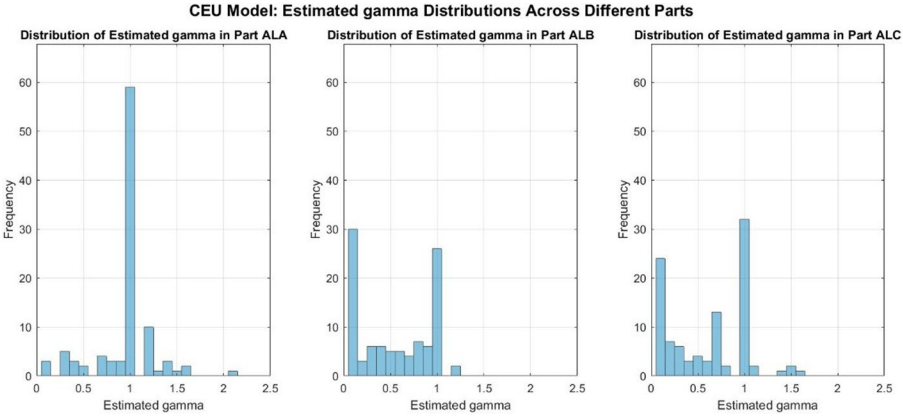


Fig. 13 Distribution of estimated γ across parts for all participants ($N = 100$)

Table 3 Hypothesis tests for the estimated AMM parameter α

Test	Part	Mean α	Std α	95% CI for mean	t-statistic	*p*-value	BH-FDR *p*
H_0: $\alpha=0$; H_1: $\alpha>0$	ALA	0.454	0.251	[0.405, 0.504]	18.131	<0.001	<0.001
H_0: $\alpha=0.5$; H_1: $\alpha\neq0.5$					−1.816	0.072	0.079
H_0: $\alpha=1$; H_1: $\alpha<1$					−21.763	<0.001	<0.001
H_0: $\alpha=0$; H_1: $\alpha>0$	ALB	0.256	0.357	[0.185, 0.327]	7.168	<0.001	<0.001
H_0: $\alpha=0.5$; H_1: $\alpha\neq0.5$					−6.828	<0.001	<0.001
H_0: $\alpha=1$; H_1: $\alpha<1$					−20.823	<0.001	<0.001
H_0: $\alpha=0$; H_1: $\alpha>0$	ALC	0.327	0.297	[0.269, 0.386]	11.037	<0.001	<0.001
H_0: $\alpha=0.5$; H_1: $\alpha\neq0.5$					−5.818	<0.001	<0.001
H_0: $\alpha=1$; H_1: $\alpha<1$					−22.672	<0.001	<0.001

Table 4 Hypothesis tests for the estimated CEU parameter γ

Test	Part	Mean γ	Std γ	95% CI for mean	t-statistic	*p*-value	BH-FDR *p*
H_0: $\gamma=1$; H_1: $\gamma\neq1$	ALA	0.964	0.320	[0.901, 1.028]	−1.119	0.266	0.266
H_0: $\gamma=1$; H_1: $\gamma\neq1$	ALB	0.540	0.375	[0.466, 0.615]	−12.257	<0.001	<0.001
H_0: $\gamma=1$; H_1: $\gamma\neq1$	ALC	0.571	0.392	[0.493, 0.649]	−10.945	<0.001	<0.001

The 95% confidence intervals refer to the mean parameter estimates across participants. Raw *p*-values are reported together with Benjamini–Hochberg false discovery rate adjusted *p*-values. The BH-FDR adjustment is applied jointly across the 12 hypothesis tests reported in Tables 3 and 4

participants best fitted by that model. The estimated risk-aversion parameter r under EU remains concentrated around broadly similar values across the three ambiguous-lottery parts, although the distributions also contain some estimates at or near the parameter bounds, particularly in Part ALA (Fig. 11). By contrast, the distributions of the ambiguity parameters α and γ change more noticeably across parts (Figs. 12 and 13). In Part ALA, both parameters are relatively concentrated around intermediate or ambiguity-neutral values, whereas Parts ALB and ALC exhibit substantially greater concentrations away from these values, including many estimates at or near parameter boundaries.

Table 5 Aggregate and individual-level McFadden pseudo-R^2 for the EU model

Part	N	Aggregate pseudo-R^2	Mean	Median	Q1	Q3	Min	Max
ALA	100	0.722	0.556	0.725	0.263	0.825	0.016	0.898
ALB	100	0.353	0.266	0.214	0.139	0.300	0.000	0.827
ALC	100	0.282	0.228	0.200	0.131	0.281	0.006	0.842

Aggregate pseudo-R^2 is calculated from log-likelihoods summed across participants, whereas the remaining statistics summarize pseudo-R^2 calculated separately for each participant. Q1 and Q3 denote the 25th and 75th percentiles, respectively

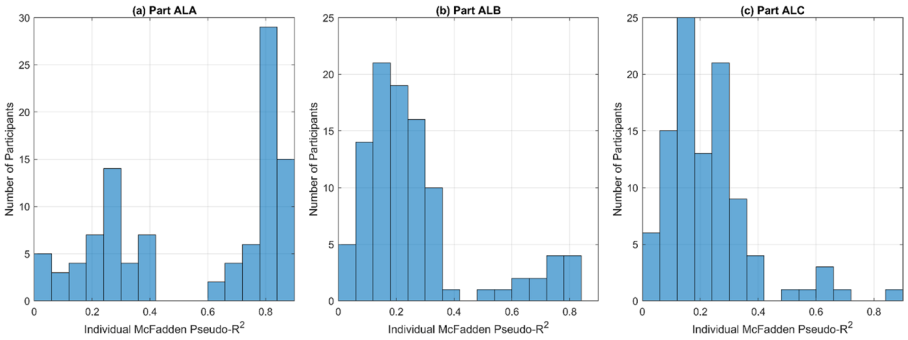


Fig. 14 Distribution of individual-level Pseudo-R^2 for the EU model across parts

These distributions suggest that the central pattern of estimated risk attitudes is relatively stable across the ambiguous-lottery parts, while the estimated ambiguity parameters are more sensitive to changes in task structure.

We further examine the estimated ambiguity parameters using a series of t-tests reported in Tables 3 and 4. To provide information on the precision of the estimated mean parameters, we also report 95% confidence intervals. Because these tables contain 12 hypothesis tests in total, we apply the Benjamini–Hochberg false discovery rate (BH-FDR) correction jointly across all 12 tests and report both the unadjusted and BH-FDR-adjusted p-values. The correction does not change any of the original significance decisions.

In Part ALA, the mean estimates of both ambiguity parameters are consistent with their respective ambiguity-neutral benchmarks. The mean AMM estimate is $\alpha = 0.454$ (95% CI [0.405, 0.504]) and is not significantly different from 0.5 after BH-FDR correction ($p = 0.079$). Similarly, the mean CEU estimate is $\gamma = 0.964$ (95% CI [0.901, 1.028]) and is not significantly different from 1 ($p = 0.266$). In Parts ALB and ALC, however, both ambiguity parameters differ significantly from their respective neutrality benchmarks (all BH-FDR-adjusted $p < 0.001$).

The preceding model comparisons assess the relative fit of EU against the alternative models. We next examine the goodness-of-fit of the EU model relative to the nested risk-neutral benchmark using McFadden's pseudo-R^2, defined as $R^2{}_{McFadden} = 1 - \frac{\ln L_{EU}}{\ln L_{RN}}$, where $\ln L_{EU}$ is the maximized log-likelihood under the EU

model and $\ln L_{RN}$ is the log-likelihood under the risk-neutral specification ($r=0$), which serves as the null model. Higher values indicate a greater improvement in fit relative to the risk-neutral benchmark.

Table 5 reports aggregate and individual-level McFadden pseudo-R^2 values for the EU model relative to the nested risk-neutral benchmark, while Fig. 14 shows the distributions of individual-level values. The aggregate pseudo-R^2 is 0.722 in ALA, 0.353 in ALB, and 0.282 in ALC. At the individual level, the mean pseudo-R^2 values are 0.556, 0.266, and 0.228, respectively, with corresponding medians of 0.725, 0.214, and 0.200. The distributions show substantial heterogeneity across participants, particularly in ALA, where individual-level values span from 0.016 to 0.898. Overall, EU provides the largest improvement in fit relative to the risk-neutral benchmark in ALA, while the magnitude of this improvement is smaller in ALB and ALC.

7 Conclusion

This study examines how the relative performance of alternative decision models varies across ambiguous choice problems of increasing structural complexity. Using an allocation experiment, structural complexity is varied along two dimensions: the number of possible outcomes and the number of decision stages. Behaviour is evaluated using EU, AMM, and CEU, allowing us to examine whether more structurally complex ambiguous decisions are better described by models that explicitly incorporate ambiguity attitudes.

The model-comparison results reveal a "less is more" pattern. EU is the most frequently selected model according to AICc in all three ambiguous-lottery parts, although the strength of its relative advantage varies across task structures. In Part ALB, EU performs significantly better than both AMM and CEU according to both paired t-tests and Wilcoxon signed-rank tests. In Part ALC, EU performs significantly better than AMM, while its performance is not significantly different from CEU. In Part ALA, EU performs significantly better than AMM according to both tests, while the comparison between EU and CEU differs across the two tests. These results show that the empirical benefit of explicitly incorporating ambiguity attitudes varies across choice environments and does not systematically increase with structural complexity.

The parameter estimates provide a complementary perspective. The central pattern of the estimated risk-aversion parameter r under EU is relatively stable across the ambiguous-lottery parts, whereas the distributions of the AMM parameter α and CEU parameter γ vary more noticeably across tasks. In Part ALA, the mean estimates are consistent with the respective ambiguity-neutral benchmarks, while these benchmarks are rejected in Parts ALB and ALC. Together with the model-comparison results, these patterns indicate that the empirical role of ambiguity attitudes differs across task structures rather than becoming uniformly more important as the ambiguous choice problem becomes more complex.

The individual-level goodness-of-fit analysis further shows substantial heterogeneity in model performance. EU provides the largest improvement in fit relative to the nested risk-neutral benchmark in Part ALA, while the magnitude of this improvement is smaller in Parts ALB and ALC. Individual-level pseudo-R^2 values also vary

considerably across participants within each part. This heterogeneity complements the aggregate model comparisons and emphasizes the value of examining both overall model performance and individual-level fit.

The findings also point to several directions for further research. The present experiment varies structural complexity through the number of outcomes and decision stages; future designs could isolate these dimensions more finely or independently manipulate cognitive demands while holding the underlying ambiguous lottery fixed. Counterbalanced task orders could further distinguish structural differences from learning or adaptation, while richer choice data could improve the precision with which ambiguity parameters are estimated. Extending the analysis to alternative ambiguity models and stochastic specifications would also provide a useful test of the generality of the model-ranking patterns documented here.

Overall, the results provide a simple but important message: more complex ambiguous decisions do not necessarily require more flexible behavioural models to achieve a better empirical description. The relative performance of EU, AMM, and CEU depends on the structure of the choice environment, and additional flexibility in modelling ambiguity does not systematically translate into better penalized model fit as structural complexity increases. In this sense, "less is more" describes an empirical relationship between decision complexity and model performance: richer decision environments do not automatically imply greater empirical value from richer behavioural specifications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11166-026-09514-9.

Acknowledgements The author is grateful to Professor John Hey for generous financial support of the experiment, as well as for valuable discussions on an earlier draft.

Author contributions I am the only one author for writing and reviewing this manuscript.

Data availability The data generated and analyzed during this experiment are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors declare no competing interests

References

Abdellaoui, M., Baillon, A., Placido, L., & Wakker, P. P. (2011). The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation. *American Economic Review*, *101*(2), 695–723. https://doi.org/10.1257/aer.101.2.695

Chew, S. H., Miao, B., & Zhong, S. (2017). Partial ambiguity. *Econometrica, 85*(4), 1239–1260.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review, 78*, 97–119. https://doi.org/10.1016/j.euroecorev.2015.05.004

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms*. *The Quarterly Journal of Economics, 75*(4), 643–669. https://doi.org/10.2307/1884324

Gabaix, X. (2014). A sparsity-based model of bounded rationality *. *The Quarterly Journal of Economics, 129*(4), 1661–1710. https://doi.org/10.1093/qje/qju024

Ghirardato, P., Maccheroni, F., & Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, *118*(2), 133–173. https://doi.org/10.1016/j.jet.2003.12.004

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics, 18*(2), 141–153. https://doi.org/10.1016/0304-4068(89)90018-9

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62*(6), 1291–1326. https://doi.org/10.2307/2951750

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, *73*(6), 1849–1892. https://doi.org/10.1111/j.1468-0262.2005.00640.x

Savage, L. (1954). *The Foundations of Statistics* (Vol. Vol. 11, pp. 86–89). Wiley Publications in Statistics.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica*, *57*(3), 571–587. https://doi.org/10.2307/1911053

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.