Experimental Economics (2024) 27:582–603
https://doi.org/10.1007/s10683-024-09837-x

Measuring decision confidence

Sara Arts[1] · **Qiyan Ong**[1] · **Jianying Qiu**[1]

Received: 5 March 2022 / Revised: 25 June 2024 / Accepted: 3 July 2024 /
Published online: 19 July 2024

Abstract

We examine whether the way individuals randomize between options captures their decision confidence. In two experiments in which subjects faced pairs of options (a lottery and a varying sure payment), we allowed subjects to choose randomization probabilities according to which they would receive each option. Separately, we obtained two measures of self-reported confidence - confidence statements and probabilistic confidence - for choosing between the two options. Consistent with the predictions of two theoretical frameworks incorporating preference uncertainty, the randomization probabilities correlated strongly with both self-reported measures (median Spearman correlations between 0.86 to 0.89) and corresponded in absolute levels to probabilistic confidence. This relationship is robust to two exogenous manipulations of decision confidence, where we varied the complexity of the lottery and subjects' experience with the lottery.

Keywords Decision confidence · Randomization · Incentivized approach · Preference uncertainty

JEL Classification B40 · C91 · D81

We sincerely thank the editor and three anonymous reviewers for their extensive and constructive comments. Their feedback has substantially improved our paper. The replication and supplementary material for the study is available at DOI 10.17605 (OSF.IO/F3CHX).

✉ Jianying Qiu
jianying.qiu@ru.nl

Sara Arts
sara.arts@ru.nl

Qiyan Ong
qiyan.ong@ru.nl

[1] Department of Economics, IMR, Radboud University, EOS 02.577, Heyendaalseweg 141, 6525 AJ Nijmegen, The Netherlands

1 Introduction

There are many decisions in life that people may not be able to make with full confidence. These decisions often involve difficult trade-offs among conflicting objectives, such as price vs. quality when buying goods, risk vs. return when investing, and efficiency vs. equality when making policy decisions. As more studies suggest that decision confidence has the potential to explain a wide range of behavioral anomalies, there is growing interest in eliciting and accounting for decision confidence when studying people's choices.[1]

Past studies so far have mostly relied on non-incentivized self-reports to elicit decision confidence. For example, some studies (Cohen et al., 1987; Dubourg et al., 1994, 1997; Cubitt et al., 2015) allowed subjects to indicate whether they were unsure of their choices. Butler and Loomes (2007, 2011) had subjects indicate their decision confidence using the ordinal terms "definitely" and "probably." More quantitatively, Enke and Graeber (2021, 2023) had subjects rate how certain (from 0% to 100% in increments of 5%) they were that their actual valuation of an option was within the interval obtained from the choices they made earlier in a price list.

While asking people explicitly about how confident they are about their decisions directly elicits decision confidence, incentivized behavioral measures that elicit decision confidence without referring to confidence may encourage people to contemplate their decisions carefully and thereby reduce behavioral biases such as priming effects and experimenter demand effects (Camerer and Hogarth, 1999). Yet, finding behavioral measures of decision confidence "in a form simple and transparent enough to work without creating additional uncertainty" is not trivial (Butler and Loomes, 2011, p. 516).

Building on earlier studies, we propose to use the randomization probability assigned to an option in a choice pair as a behavioral measure of decision confidence, and test its validity in two experiments. For each pair of options (a lottery x and a sure payment y), we obtained subjects' binary choice, self-reported decision confidence, and randomization probabilities. Self-reported decision confidence was elicited by having subjects select a confidence statement from "Surely x," "Probably x," "Unsure," "Probably y," "Surely y," after they chose an option and by having them report how confident they were in choosing either options as a probabilistic confidence of p% x and 100-p% y, where p ranged from 0 to 40 and 60 to 100 in increments of 10. Separately, subjects chose a randomization probability $0 \leq \lambda \leq 1$ with which they would receive x (and with probability $1 - \lambda$ receive y) for each pair of options (Miao and Zhong, 2018; Agranov and Ortoleva, 2023; Feldman and Rehbeck, 2022; Ong and Qiu, 2023). Unlike the two self-reported confidence

[1] These anomalies include the willingness to accept (WTA) - willingness to pay (WTP) gap (Dubourg et al., 1994) preference reversals (Butler and Loomes, 2007), stochastic choices (Agranov and Ortoleva, 2017), insensitivity to variation in probabilities (Enke and Graeber, 2023), anomalies in intertemporal choices (Enke and Graeber, 2021), small-stakes risk aversion (Khaw et al., 2021), and many other violations of standard decision theory (Butler and Loomes, 2011).

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

measures, the elicitation of randomization probabilities made no reference to decision confidence.

We exogenously manipulated decision confidence by a) having a simple lottery with two outcomes and a complex lottery with the same expected value but with more payoff outcomes over a wider range of possible values (Fudenberg and Puri, 2021), and by b) increasing subjects' experience with the lottery by allowing them to either observe the outcome draws of the lottery or to make hypothetical choices and observe the payoffs of their choices and the counterfactual (Myagkov and Plott, 1997; Plott and Zeiler, 2005; van de Kuilen and Wakker, 2006; van de Kuilen, 2009).

To ensure that experimenter demand effects and order effects were not driving our results, in one of our experiments, we elicited the three measures separately over three sessions (at least seven days apart). Subjects were randomly assigned to one of the three decision orders, which differed by whether the confidence statements, the probabilistic confidence, or the randomization probabilities were elicited in the first session.

We structure the analyses of our experiments through two theoretical frameworks, one based on Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015) and the other based on Fudenberg et al. (2015). We illustrate how randomization emerges from the optimization behavior of an individual who faces uncertainty regarding her preference between the two options. Based on our expectations that randomization probabilities serve as a good proxy for decision confidence, and that complexity increases preference uncertainty while experience decreases it, we have four hypotheses. First, randomization probabilities correlate positively with both self-reported measures; second, subjects choose randomization probabilities around 0.5 for choice pairs in which the sure payment has a similar decision utility as the lottery (and one does not dominate the other); third, subjects randomize over a wider range of sure payments and choose randomization probabilities closer to 0.5 for the complex lottery than for the simple lottery; finally, subjects randomize over a smaller range of sure payments and choose randomization probabilities further away from 0.5 in the experience treatments.

Our experimental results support all four hypotheses. Subjects' randomization probabilities were strongly and positively correlated with both confidence statements and probabilistic confidence (median Spearman correlation between 0.86 to 0.89). In line with our expectations, the two exogenous manipulations affected self-reported decision confidence. Increasing the complexity of the lottery led to a decrease in self-reported decision confidence, while increasing experience with lotteries led to an increase in self-reported decision confidence. These exogenous changes in self-reported decision confidence were met with corresponding changes in randomization probabilities: subjects randomized over a larger range of sure payments and the randomization probabilities were closer to 0.5 for the complex lottery than for the simple lottery, while the opposite occurred when subjects had more experience. As a result, the correlations between randomization probabilities and self-reported decision confidence measures were robust to the exogenous manipulations of decision confidence.

Our study builds on the growing literature on preferences for randomization, implying preference functionals that are convex with respect to probabilistic mixing,

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

which is a violation of the betweenness axiom (Chew, 1983; Dekel, 1986; Chew, 1989). Preferences for randomization have been documented over wide ranges, across different domains, in experimental settings as well as in real life decisions. In a multiple-decision setting, Rubinstein (2002) suggested that randomization (diversification in his term) by choosing differently across five independent and identical decisions is "an expression of a more general phenomenon in which people tend to diversify their choices when they face a sequence of similar decision problems and are uncertain about the right action" (Rubinstein, 2002, p. 1370). Dwenger et al. (2018) found that their experimental subjects preferred to randomize via an external randomization device rather than making choices themselves, and the authors reported similar behavior among German university applicants. Miao and Zhong (2018) showed that randomization could be used to balance ex-ante and ex-post social preferences. Feldman and Rehbeck (2022) elicited individuals' attitudes toward reduced mixtures over two lotteries in the space of three-outcome lotteries (the Marschak-Machina triangle) and found pervasive evidence of a preference for non-degenerate mixing over lotteries. The studies closest to ours are Agranov and Ortoleva (2023) and Ong and Qiu (2023), who also allowed subjects to choose randomization probabilities when deciding between two options. Both studies found that subjects often randomized and did so over large ranges. Ong and Qiu (2023) further found that subjects were willing to pay to randomize, suggesting that randomization was deliberate and not merely a result of indifference. Popular explanations for convex preferences include hedging in the face of preference uncertainty (Cerreia-Vioglio et al., 2015; Fudenberg et al., 2015; Cerreia-Vioglio et al., 2019), non-linear probability weighting (Kahneman and Tversky, 1979; Quiggin, 1982; Tversky and Kahneman, 1992), and responsibility aversion (Dwenger et al., 2014). Our study is the first to provide experimental evidence linking preference uncertainty and randomization behavior.[2]

We also contribute to the literature on stochastic choice, which examines why individuals change their decisions when they face the same decision situation repeatedly. The relationship we found between randomization probabilities and sure payments bears a remarkable resemblance to results reported in studies on stochastic choice, for example, Mosteller and Nogee (1951, Fig. 2) and Loomes and Pogrebna (2017, Table 1).[3] The similarity between the choice proportion in repeated choices and the randomization probability in a one-shot decision suggests that decision confidence may have the potential to explain stochastic choices. Consistent with this interpretation, we find that, across subjects and decisions, higher decision confidence in an option corresponded to choosing that option more frequently (but

[2] Agranov and Ortoleva (2023) also explicitly discuss this link. Based on reports from the end-of-experiment questionnaire, they found that many of their subjects randomized because they were unsure of their preferences (Agranov and Ortoleva, 2023, Appendix A.8).

[3] Note that these results come from entirely different designs. In Mosteller and Nogee (1951) and Loomes and Pogrebna (2017) individuals repeatedly faced a lottery and a sure payment, with the sure payment varying from one question to another, and the results are about the proportion of accepting the lottery across decisions, whereas in our experiment subjects faced the lottery and a sure payment once and chose the randomization probability of receiving the lottery.

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

not always) in binary choices. Meanwhile, random (expected) utility models (see, e.g., Eliashberg and Hauser, 1985; Loomes and Sugden, 1995; Gul and Pesendorfer, 2006; Apesteguia and Ballester, 2018), which are the standard explanations for stochastic choices, do not predict randomization in a one-shot decision as observed in our experiment. This is because, while individuals may be considered to have a set of utility functions in this literature, at the moment of decision-making, they rely on one utility function randomly realized from the set.

Overall, our study provides direct evidence to the connections between some important concepts in the literature, such as decision confidence, cognitive uncertainty, preference uncertainty/imprecision, incomplete preference, preference for randomization, and stochastic choice. While there have been notable theoretical advancements and accumulating empirical evidence in this field, the precise interplay and relationships between these concepts as well as how they are connected to choices remain ambiguous. Our finding of a systematic relationship between the randomization, alternative measures of decision confidence, and stochastic choice suggests that there may exist a common psychological underpinning for these various concepts.

The rest of the paper proceeds as follows. Section 2 describes the experimental procedure. Section 3 provides the theoretical basis for how randomization probabilities may be linked to decision confidence. The results are reported in Sect. 4. Finally, Sect. 5 concludes the paper.

2 Experimental design

We had two experiments. We first describe the general structure of the experiments before detailing the differences.

2.1 General structure of the experiments

In each decision, subjects faced a pair of options: a lottery x and a sure payment y. The lottery was paired with 13 values of sure payments (0, 2, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, and 10 euros). For each type of decisions that we will describe below, subjects faced these pairs in a random sequence. Each decision was made on a separate screen, and subjects were not allowed to review or change their decisions once they were made. Each subject made three types of decisions: binary choices, self-reported decision confidence, and randomized choices.

2.1.1 Binary choices

The binary choices required subjects to choose either x or y. If x was chosen, the computer would draw a random number to determine x's outcome. For example, for x that has a 50% chance of paying 9 euro and a 50% chance of paying 1 euro, if the randomly drawn number falls between 1 and 50, the subject would receive 1 euro, and

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

if the randomly drawn number falls between 51 and 100, the subject would receive 9 euros.

2.1.2 Two measures of self-reported decision confidence

After making the binary choices, we asked subjects how confident they felt about their choices. The confidence statements they could choose were "Surely x," "Probably x," "Unsure," "Probably y," or "Surely y." Similar statements were used in Dubourg et al. (1994), Butler and Loomes (2007), and Butler and Loomes (2011). Confidence statements were not incentivized and could not affect payoffs.

In addition to the confidence statements, subjects in Experiment 2 also had to report their probabilistic confidence in a separate experimental decision. Instead of making a direct binary choice, subjects had to choose how confident they felt about choosing x versus y. They had to choose between ten levels of probabilistic confidence: "100% x, 0% y," "90% x, 10% y," ... "60% x, 40% y," "40% x, 60% y," ..., "0% x, 100% y." Subjects were considered to have chosen the option for which they indicated more than 50% probabilistic confidence. For example, if a subject chose "60% x, 40% y," she was considered to have chosen x over y in that decision. To use the probabilistic confidence as a measure of decision confidence as well as an indicator of a subject's choice between x and y, we omitted the option "50% x, 50% y".

2.1.3 Randomized choices

The randomized choices required subjects to choose a randomization probability λ, based on which they would receive x (and hence with a probability $1 - \lambda$ of receiving y). For example, choosing $\lambda = 0.40$ means the subject would receive x with a chance of 40% and y with a chance of 60%. The subjects used a slider from 0% to 100% with increments of 1% to choose the randomization probability in each choice. In experiment 1, the slider was set in the middle at the start. To reduce anchoring bias, in experiment 2, the slider had no initial position, and subjects needed to click on the slider and move the bar to determine the randomization probability. If the randomized choice was chosen for payment, the computer would draw a random number between 1 and 100. If the drawn number was between 1 and 100 λ, x would be chosen over y in that decision. If x was chosen for payment, a second random draw determined the outcome of the lottery. Figure 1 shows the decision screen for the randomized choice. To ensure that subjects understand the payoff structure of randomized choices, we provided two examples as well as reminders in the lower part of the decision screens on how randomized choices affect payment (see Figs. C.4 and C.5 in Appendix C).

2.2 Manipulating decision confidence

In the baseline treatment, the subjects faced a simple lottery with two outcomes (a 50% chance of 9 euro and a 50% chance of 1 euro). They received a complete description of the lottery before making their decisions.

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

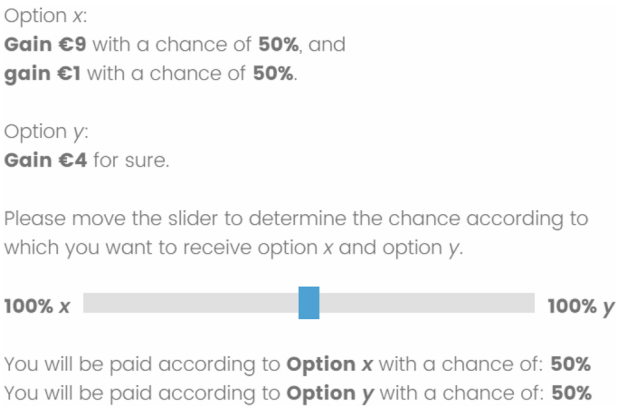


Fig. 1 An example of the randomized choice decision screen, where option x is a lottery with a 50% chance of gaining 9 euro and a 50% chance of gaining 1 euro. Option y is a sure payment and varies across choices

2.2.1 Varying the complexity of the lottery

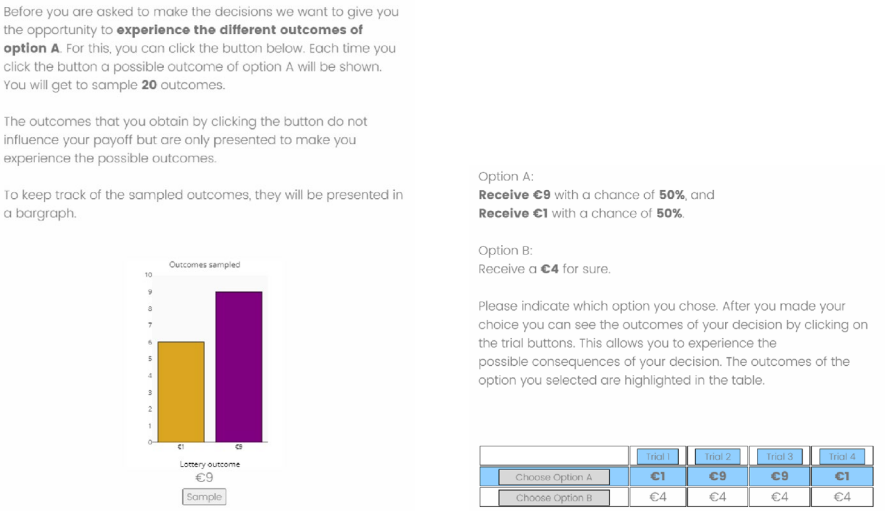
We manipulated, within subject, the complexity of the lottery by asking all subjects to make decisions involving a complex lottery with four outcomes in addition to decisions involving the simple lottery. The complex lottery has the same expected value as the simple lottery. It offers 9.75 euros with a chance of 20%, 7.50 euros with a chance of 30%, 2.50 euros with a chance of 30%, and 0.25 euros with a chance of 20%. The order of the lotteries was randomized at individual subject level: some subjects proceeded from the simple lottery to the complex lottery, while others completed the decisions in the reverse order.[4]

2.2.2 Varying subjects' experience with the lottery

For this manipulation, we adopted a between-subject design, where subjects were randomly assigned to either the baseline treatment or the experience treatment. In Experiment 1, after learning the probability distribution of the lottery and prior to making actual decisions, subjects in the partial-experience treatment had to click and view 20 draws of the lottery. As the subjects viewed each lottery draw, an accompanying bar chart which recorded each lottery outcome was updated. By the 20th draw, the bar chart reflected the probability distribution of the lottery. Figure 2a shows an example of the partial-experience treatment.

[4] The subjects in Experiment 1 also made decisions involving a loss lottery and a mixed lottery. We included these lotteries because they might lead to larger preference uncertainty due to the additional uncertainty in attitudes toward losses. We omitted these two lotteries in Experiment 2 because, as pointed out by one reviewer, the theoretical analysis of these two lotteries requires a more general approach than what we had relied on. The main results involving these lotteries are provided in Appendix B.4. Further details can be found in the working paper version (Arts et al., 2020).

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press



(a) The partial-experience treatment (b) The full-experience treatment

Fig. 2 Panel **a** illustrates what subjects in the partial-experience treatment saw when they generated the outcomes of the lottery. Panel **b** illustrates what subjects saw in the full-experience treatment. The numbers highlighted in blue in the table show a subject's hypothetical decision and her four potential payoffs, and the not highlighted numbers show the counterfactuals

In Experiment 2, subjects in the full-experience treatment had to make five hypothetical decisions per lottery, with each decision involving a different sure payment (3, 4, 5, 6 or 7 euros in a random sequence). For each decision, they observed four potential realizations of the payoff of their choice as well as the counterfactual in a payoff table. Figure 2b shows the decision screen and the payoff table viewed by a subject who chose the lottery over the sure payment of 4 euros in the hypothetical decision.

2.3 Design considerations

A few design features deserve some discussion. First, our within-subject design of eliciting self-reported confidence measures and randomization probabilities for each subject raises the concern that experimenter demand effects or order effects may unintentionally influence subjects to give similar responses, resulting in a systematic relationship between them. We took several measures in Experiment 2 to make it more obscure and costly for subjects to connect self-reported confidence measures and randomization probabilities (Zizzo, 2010), such as spreading the decisions over three sessions (at least seven days apart), including a cost for randomizing, and randomly assigning subjects to one of the three different orders. The key features of the two experiments are summarized in Figure 3.

Second, varying the complexity of the lottery could induce changes in behavior through channels other than decision confidence. For example, subjects may

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

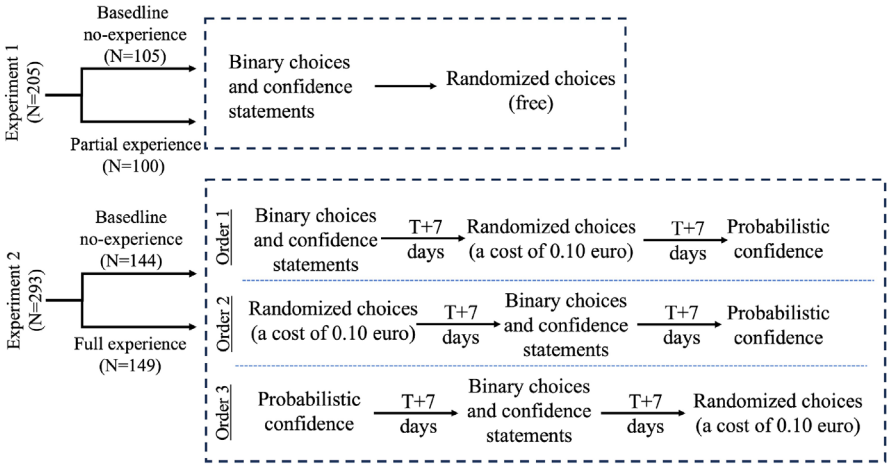


Fig. 3 Summary of the treatments and experimental procedure in Experiment 1 and 2. The dotted rectangle highlights the types of decisions that subjects in each experiment made for both lotteries. The sequence of the simple lottery and the complex lottery in each type of decision was randomized at individual subject level

valuate the two lotteries differently, as found in studies documenting complexity seeking and complexity averse behaviors (see e.g. Abdellaoui et al., 2020, and the references therein). This difference is less relevant for our purpose because our focus is on decision confidence (e.g., the range of sure payments that subjects do not have full confidence) rather than the average valuation of the lotteries. Another concern is that varying the complexity of the lottery induces different randomization behavior. While this may occur in some non-EUT models (Machina, 1985; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Quiggin, 1982), they do not predict a close relationship between randomization probability and self-reported decision confidence, or how this relationship changes with the complexity of the lottery. We will return to this point in subsection 4.4 to discuss other interpretations of randomization probabilities that could predict different randomization behavior between the two lotteries.

A final concern is the choice of the experience treatments. The description-experience gap literature suggests that description and experience may induce different risk preferences (see e.g., Hertwig et al., 2004; Wulff et al., 2018), which could reduce decision confidence. Our partial experience design is unlikely to have this issue because subjects' experience of the lottery realizations were equivalent to its description. Recent studies by Aydogan and Gao (2020) and Cubitt et al. (2022) show that the description-experience gap should be small in this case. Examining more closely, our experience treatments differ from the standard design in the Description-Experience gap literature as we provided subjects with the full description of the lottery's probability distribution in addition to the opportunity to experience the realizations of the lottery. In this sense, our experience treatments are closer to studies showing that experience in addition to a full description of the lottery could help subjects develop a better

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

understanding about their preference (e.g., van de Kuilen and Wakker, 2006; van de Kuilen, 2009).

2.4 Sample and procedure

The data were collected from a sample of 498 subjects of the ID lab at Radboud University. A total of 205 subjects participated in Experiment 1 and 293 in Experiment 2. Invitations were sent in batches via ORSEE (Greiner, 2015). The experiment was conducted using Qualtrics and lasted approximately 20 min for Experiment 1 and about 30 min in total for Experiment 2. In Experiment 2, subjects were also asked to answer a post-experiment questionnaire at the end of each of the three sessions based on the type of decision confidence they reported in that session. Appendix C contains the experimental instructions and decision screens. Each subject received a participation fee of 1 euro and an additional payment based on one of the decisions they made in the experiment. In Experiment 1, the additional monetary compensation was based on a decision randomly selected from their binary choices or randomized choices. In Experiment 2, it was based on a decision randomly selected from their binary choices, randomized choices, or probabilistic confidence decisions. The average additional payment was 6.28 euros. We made the payment via bank transfers.

3 Theoretical analysis

Under expected utility theory (EUT) which ignores decision confidence and assumes that a unique utility function (subject to positive affine transformation) captures an individual's preference, it is straightforward to show that the individual chooses $\lambda^* \in (0, 1)$ for at most one value of the sure payment in the 13 choice pairs. Thus, under EUT, strict randomization ($\lambda^* \in (0, 1)$) rarely occurs, and randomization probabilities do not contain additional information beyond indifference.

We present two theoretical analyses of our experiments that provide an explicit link between randomization probabilities and decision confidence. Both analyses assume that the individual is uncertain about her preference. Appendix A.1 presents a theoretical framework based on Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015). In this framework, the individual has multiple utility functions that we call multiple selves. She is not fully confident about her choice when some selves prefer one option while others prefer the other. In such instances, the individual prefers randomization over selecting a particular option because it offers a "fair" way to resolve internal conflicts among her different selves. This approach of capturing the lack of decision confidence from unsureness about preferences is closely related but is different from models of ambiguity (e.g., Gilboa and Schmeidler, 1989; Klibanoff et al., 2005), which focus on unsureness about beliefs (e.g., Halevy, 2007; Chew et al., 2017; Cubitt et al., 2020, and the references therein).

Appendix A.2 presents the extension of Fudenberg et al.'s (2015) model. Fudenberg et al. (2015) axiomatized a choice rule of deliberate randomization called additive

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

perturbed utility (APU). Their representation corresponds to a form of ambiguity-averse preferences for an individual who is uncertain about her true utility function. The individual randomizes to balance the probability of errors due to preference uncertainty against the cost of avoiding them (Fudenberg et al., 2015, p. 2373).

Both analyses suggest that the preference for randomization is motivated by the hedging of preference uncertainty. In particular, randomization probabilities are affected by the perceived preference uncertainty of the options, attitudes towards preference uncertainty, as well as the utility difference between the options. The two theoretical analyses suggest that our proposed link between randomization probabilities and decision confidence could hold under a broad class of decision models that incorporate preference uncertainty.

We expect that subjects perceive more preference uncertainty with the complex lottery than with the simple lottery, and experience with the lottery reduces preference uncertainty regarding the lottery. With these expectations, our theoretical analyses show that randomization probabilities share three important properties of decision confidence: a) subjects choose randomization probabilities close to 0.5 when the sure payment has a similar decision utility to the lottery; b) they randomize over a wider range of sure payments, with randomization probabilities closer to 0.5, when they face the complex lottery compared to the simple lottery; and c) with experience and less preference uncertainty about the lottery, subjects' randomization probabilities may be stretched away from 0.5 as they randomize over a smaller range of sure payments. If randomization probabilities and the two self-report measures both capture decision confidence, we expect the following:

Hypotheses

1. Randomization probabilities are positively correlated with the self-reported confidence measures.
2. When two choice options are more similar, for example, around the switching choices where subjects switch between the lottery and the sure payment, the subjects have lower decision confidence, a higher likelihood of randomizing, and randomization probabilities closer to 0.5.
3. Compared to the decisions about the simple lottery, the decisions about the complex lottery exhibit lower decision confidence, as measured by the self-reported confidence measures, and randomization probabilities are affected in the same direction, maintaining a strong association between them.
4. Compared to the decisions in the no-experience treatment, the decisions in the experience treatments exhibit higher decision confidence, as measured by the self-reported confidence measures, and randomization probabilities are affected in the same direction, maintaining a strong association between them.

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

4 Experimental results

We report the results in two steps. We begin by showing the systematic link between randomization probabilities and the two measures of self-reported confidence in the baseline no-experience treatment for decisions about the simple lottery (Hypotheses 1 and 2). We then show that decision confidence responded to our treatment manipulations in the expected direction by comparing the two measures of self-reported decision confidence across treatments. We demonstrate that exogenous shifts in self-reported decision confidence are paired with corresponding shifts in randomization probabilities, maintaining their systematic link (hypotheses 3 and 4). We pool subjects in different orders in our main analyses and discuss order effects in subsection 4.3.

4.1 Randomization probabilities and self-reported confidence

Below, we report two empirical observations that are consistent with Hypothesis 1 and 2.

Result 1 In the baseline no-experience-simple-lottery treatment, randomization probabilities were significantly and positively correlated with confidence statements and probabilistic confidence among the large majority of subjects. Further, on average, randomization probabilities corresponded to probabilistic confidence in absolute levels.

To obtain the correlation between randomization probabilities and confidence statements, we transformed the confidence statements to a scale of 1 to 5, with "Surely y" taking the value of 1 and "Surely x" taking the value of 5 to represent one's decision confidence in choosing x. We computed for each subject the non-parametric Spearman correlation between confidence statements and randomization probabilities in Experiment 1 and 2, and between probabilistic confidence and randomization probabilities in Experiment 2. The results for the baseline treatment are illustrated in Figure 4, Panel (a). Table B.1 in Appendix B summarizes the cross-measure correlations across treatments. Consistent with Hypothesis 1, confidence statements and randomization probabilities have a high and positive correlation. Moderate to strong correlations of 0.60 in Experiment 1 and 0.71 in Experiment 2 were found at the 10th percentile level, which increased to 0.91 in Experiment 1 and 0.89 in Experiment 2 at the median level. Since the subjects in Experiment 2 reported confidence statements and randomization probabilities in different sessions separated by at least seven days and in different orders, the similarities between the correlations found in Experiments 1 and 2 suggest that confidence statements and randomization probabilities are associated in ways beyond experimenter demand effects and order effects. In Experiment 2, we also found high correlation between self-reported probabilistic confidence and

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

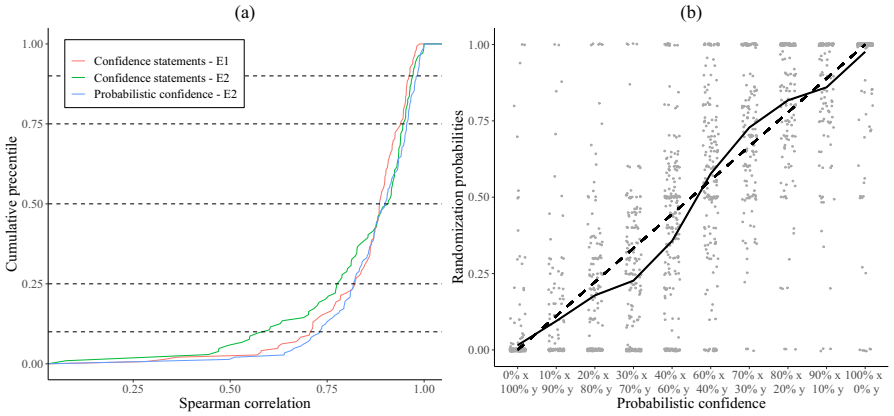


Fig. 4 **a** Cumulative distributions of subjects' Spearman correlations between randomization probabilities and confidence statements or probabilistic confidence in the baseline no-experience treatment for decisions about the simple lottery. E1 and E2 refer to data from Experiment 1 and 2 respectively. **b** Scatter plot of randomization probabilities, with the mean randomization probability (in solid line) at each probabilistic confidence level in Experiment 2 in the baseline no-experience treatment for decisions about the simple lottery. The dashed line is a 45-degree line

randomization probabilities: the correlation is 0.73 at the 10th percentile and 0.90 at the median.

As correlations do not describe the correspondence between randomization probabilities and self-reported decision confidence in absolute levels, we also computed the mean randomization probability at each level of probabilistic confidence for each subject and took the mean across subjects. This is shown in Panel (b) of Fig. 4. Overall, the mean randomization probability for x is close to the probabilistic confidence of choosing x: subjects who chose a randomization probability of, for example, 0.7 would report probabilistic confidence of 70% on average. Examining the absolute correspondence between randomization probabilities and confidence statements gives a similar result, as summarized in Fig. B.1 and Table B.2 in Appendix B. These results suggest randomization probabilities can be used as a direct proxy for probabilistic confidence.

Next, we turn to Hypothesis 2 and examine the randomization probabilities around the switching choices. Intuitively, x and y are harder to compare around the switching choices. Reflecting this, subjects reported lower decision confidence and chose randomization probabilities close to 0.5 around the switching choices, as indicated in Result 2.

Result 2 On average, the subjects reported lower decision confidence around the switching choices based on the self-reported confidence measures and were more likely to randomize and chose randomization probabilities close to 0.5 around the switching choices.

We study the switching choice of each subject by considering two levels of sure payments: $\underline{y_b}$ and $\bar{y_b}$. We let $\underline{y_b}$ denote the highest sure payment at and below which

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

Table 1 Median behavior around the switching choices in the baseline no-experience treatment for decisions about the simple lottery

Experiment	Median behavior around the switching range					
	Confidence statements		Probabilistic confidence		Randomization probabilities	
	$\underline{y_b}$	$\bar{y_b}$	$\underline{y_b}$	$\bar{y_b}$	$\underline{y_b}$	$\bar{y_b}$
Experiment 1 (N = 105)	Probably x	Probably y	–	–	0.67	0.46
Experiment 2 (N = 145)	Probably x	Probably y	60% x	40% x	0.65	0.45

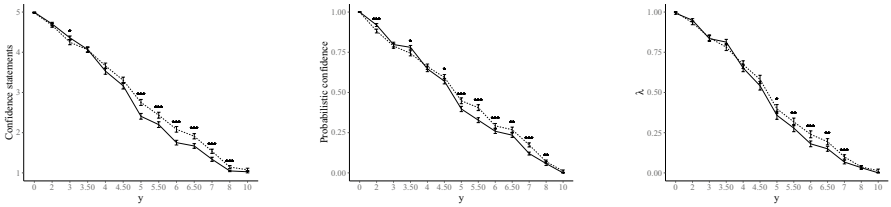
the subject always preferred x over y, and $\bar{y_b}$ denote the lowest sure payment amount at and above which the subject always chose y over x in the binary choices. We henceforth refer to the values of y between $\underline{y_b}$ and $\bar{y_b}$ as the subject's switching range. This approach allows us to accommodate subjects who switched once as well as those who switched multiple times between lottery x and the sure payments (for the simple and complex lotteries, respectively, 19% and 25% in Experiment 1 and 14% and 23% in Experiment 2).[5]

As expected, decision confidence was lower within the switching range than outside it. In Experiment 2 (Experiment 1), 88% (85%) of the confidence statements within the switching range were "Probably x," "Unsure," or "Probably y," compared to 41% (40%) outside the switching range. In Experiment 2, "60% x, 40% y" and "40% x, 60% y" were selected for 53% of the values within the switching range, compared to 13% outside the switching range. Table 1 shows the median randomization probabilities, probabilistic confidence, and confidence statements around the switching range. The median responses to the self-reported confidence measures indicate a lack of confidence around the switching range.

The randomization probabilities within the switching range resemble the two self-reported confidence measures. In Experiment 2 (Experiment 1), 67% (85%) of randomization probabilities reported for values of y within the switching range fell between 0.1 and 0.9, whereas this only holds for 33% (47%) outside the switching range. Further, in Experiment 2 (Experiment 1), the subjects assigned a median randomization probability of 0.65 (0.67) to x at $\underline{y_b}$, and a median randomization probability of 0.45 (0.46) to x at $\bar{y_b}$. The median randomization probability for all the choices that fell within the switching range is 0.5. These results are consistent with Hypothesis 2: subjects were more likely to choose randomization probabilities close to 0.5 for choices that they found difficult to compare.

[5] It is important to include these subjects, because when subjects are not fully confident about their choices, they may switch between x and y multiple times. For the subjects who switched from the lottery to the sure payments once, the switching range simply includes the two sure payments around the switching choice (e.g., if a subject chooses the lottery at $y = 4$ and switches to the sure payment at $y = 4.5$ euros, this means that $\underline{y_b} = 4$, $\bar{y_b} = 4.5$, and the switching range is [4, 4.5]).

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press



(a) Confidence statements (b) Probabilistic confidence (c) Randomization probabilities

Fig. 5 The mean self-reported decision confidence and randomization probabilities for each value of y for the simple lottery (solid line) and complex lottery (dashed line). Wilcoxon signed-rank tests were performed to test the treatment difference for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2 Manipulating decision confidence

In this section, we examine whether our exogenous manipulations of the decision situation affected self-reported decision confidence in the expected direction and whether randomization probabilities were affected in similar ways to maintain a systematic relationship with the self-reported confidence measures.

4.2.1 The complex lottery versus the simple lottery

Consistent with Hypothesis 3, we find that subjects had lower decision confidence when making decisions about the complex lottery compared to the simple lottery. Result 3 summarizes our finding.

Result 3 Compared to decisions about the simple lottery, the subjects revealed less than full decision confidence over a wider range of sure payments for decisions about the complex lottery, and their reported decision confidence were more compressed toward "Unsure" or (50% x, 50% y). Likewise, the subjects randomized over a wider range of sure payments and chose randomization probabilities closer to 0.5 for decisions about the complex lottery.

We find that the range of sure payments over which the subjects chose confidence statements "Unsure" or "Probably" is larger for decisions about the complex lottery than for decisions about the simple lottery in both Experiment 1 and 2, and it is statistically significant in Experiment 2 (Experiment 1: 3.62 vs 3.36, Wilcoxon signed-rank test $p = 0.150$; Experiment 2: 3.58 vs 3.15, Wilcoxon signed-rank test $p < 0.01$). Comparing the range of sure payments for which subjects did not indicate probabilistic confidence of (100% x, 0% y) or (0% x, 100% y) gives similar results: the subjects were not fully confident over a wider range of sure payments for decisions about the complex lottery than decisions about the simple lottery (4.63 vs. 4.37, Wilcoxon signed-rank test $p < 0.01$).

Panel (a) and Panel (b) of Fig. 5 illustrate how confidence statements and probabilistic confidence varied with different sure payment amounts across the two

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

lotteries in Experiment 2. Compared to the simple lottery, self-reported decision confidence measures were more compressed towards "Unsure" or (50% x, 50% y) when the subjects faced the complex lottery. The difference in decision confidence across the two lotteries is statistically significant between sure payments of 5 and 8 euros, and is less often statistically significant for lower sure payment amounts. The results for Experiment 1 are similar, albeit weaker, and can be found in Fig. B.2 of Appendix B.

We proceed to examine the randomization probabilities chosen for each lottery. As we can see from Table B.3 in Appendix B, in both experiments, the range of sure payments over which subjects chose a randomization probability between 0.1 and 0.9 was significantly larger for decisions about the complex lottery than for decisions about the simple lottery (Experiment 1: 4.06 vs. 3.63, Wilcoxon signed-rank test $p < 0.01$; Experiment 2: 3.19 vs. 3.03, Wilcoxon signed-rank test $p < 0.10$). This is consistent with the findings from self-reported decision confidence reported above.

Further, the randomization probabilities were also more compressed towards 0.5 when the subjects faced the complex lottery compared to the simple lottery, as shown in Panel (c) of Fig. 5. The difference in randomization probabilities across the two lotteries is statistically significant between sure payments of 5 and 7 euros, coinciding with the range obtained from probabilistic confidence. Panel (c) of Figure 5 also shows asymmetric treatment effects on randomization probabilities for sure payments above 5 euros and sure payments below 5 euros. We show in Appendix A.3 that this asymmetric treatment effect can be consistent with the theoretical analysis when the treatment manipulation affects both preference uncertainty and the average valuation of the lotteries.

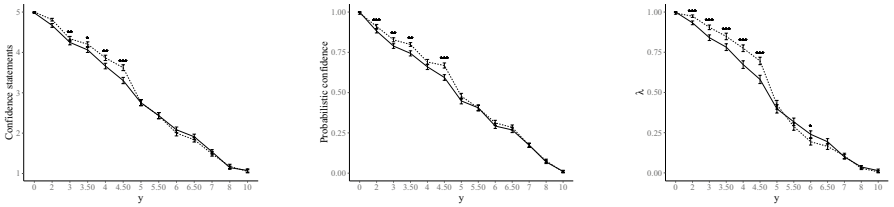
Despite our manipulation, the correlations between the two decision confidence measures and randomization probabilities remain similar. Comparing decisions about the simple lottery with those about the complex lottery, the median correlations between randomization probabilities and confidence statements are 0.86 versus 0.82 in Experiment 1 and 0.89 versus 0.88 in Experiment 2. The median correlations between randomization probabilities and probabilistic confidence are 0.90 versus 0.89 in Experiment 2. More results can be found in Table B.1, B.2, and B.4 in Appendix B.

4.2.2 Experience and no experience

Hypothesis 4 states that, compared to the baseline no-experience treatment, gaining experience with the lotteries increases decision confidence. The results about Hypothesis 4 are summarized in Result 4.

Result 4 Decisions in the partial-experience treatment and decisions about the simple lottery in the full-experience treatment did not exhibit significant treatment effects. Comparing decisions about the complex lottery in the full-experience treatment and the no-experience treatment, the subjects (1) revealed less than full decision confidence over a narrower range of sure payments, and their self-reported decision confidence were stretched further away from "Unsure" or (50% x, 50% y);

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press



(a) Confidence statements (b) Probabilistic confidence (c) Randomization probabilities

Fig. 6 The mean self-reported decision confidence and randomization probabilities for each value of y for the complex lottery in Experiment 2. The graphs show the baseline no-experience treatment (solid line) compared to the full-experience treatment (dashed line). Wilcoxon rank-sum tests were performed to test the treatment difference for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and (2) they randomized over a narrower range of sure payments and chose randomization probabilities further away from 0.5.

We report the results about the partial-experience treatment and the simple lottery in Figs. B.3, B.4, and Table B.5 in Appendix B, and report the comparison between the full-experience treatment and the no-experience treatment about the complex lottery here. We find that the range of sure payments over which subjects reported confidence statements of "Probably" or "Unsure" is significantly narrower (3.16 vs. 3.58, Wilcoxon rank-sum test $p < 0.05$). The range of sure payments over which subjects chose probabilistic confidence between 0.1 and 0.9 did not differ significantly between the full-experience treatment and the no-experience treatment (4.58 vs. 4.63, Wilcoxon rank-sum test $p = 0.590$).

Panels (a) and (b) of Fig. 6 show how self-reported decision confidence differs between the full-experience and the no-experience treatment. Compared to the no-experience treatment, self-reported decision confidence was stretched further away from "Unsure" or (50% x, 50% y) for the subjects in the full-experience treatment and these differences in decision confidence were significantly different for sure payments between 2 and 4.5 euros. This implies that the subjects in the full-experience treatment were more confident about which option they preferred than subjects in the no-experience treatment.

Next, we examine these treatment effects on randomization probabilities. Like decision confidence, we find that the range of sure payments over which subjects chose randomization probabilities between 0.1 and 0.9 was significantly narrower in the full-experience treatment than in the no-experience treatment for decisions about the complex lottery (2.67 vs. 3.19, Wilcoxon rank-sum test $p < 0.05$). Panel (c) of Fig. 6 shows that the difference in mean randomization probabilities across sure payments between the full-experience treatment and the no-experience treatment resembles that shown in Panel (a) and Panel (b). Compared to the no-experience treatment, randomization probabilities were also stretched further away from 0.5 among the subjects in the full-experience treatment. Significant differences in the randomization probabilities between subjects in the full-experience treatment and the no-experience treatment were also observed between 2 euros to 4.5 euros.

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

Asymmetric treatment effects on randomization probabilities could also be observed here for sure payments above 5 compared to those below 5, which we discuss further in the theoretical models in Appendix A.3.

The increase in decision confidence from subjects' experience with the lotteries did not affect the high correlation between self-reported decision confidence and randomization probabilities. The median correlation between self-reported decision confidence and randomization probabilities was broadly similar in the two treatments. More details can be found in Table B.1, B.2, and B.4 in Appendix B.

4.3 Order effects

An important concern of the within-subject design is order effects: subjects' earlier decisions may affect their subsequent decisions. We are especially concerned with the order effects arising from priming: when randomization probabilities were elicited after self-reported confidence measures, subjects could be primed to link randomization probabilities to decision confidence and consequently reported randomization probabilities that cohered with self-reported decision confidence measures. Below we highlight the key findings about order effects and leave the details to Appendix B.1.

We find some order effects, suggesting that priming effects on randomized choices could be present. For example, subjects randomized strictly ($0 <$ randomization probability< 1) in fewer choices when randomized choices were presented first (Order 2) compared to later (Order 1 and 3) in all treatments: averaging across treatments, 34% in Order 2 compared to 41% and 45% in Order 1 and 3 respectively (see Table B.6 and Fig. B.5 in Appendix B.1 for details). Also, the cumulative distributions of the correlations in Order 2 tend to be lower (on the left of the other two orders), with a larger difference at the lower percentiles and in the complex lottery treatments.

Despite these differences, we find support for our hypotheses among the subjects in Order 2 where the aforementioned priming effects were absent, although the support is sometimes weaker than in the full sample, possibly due to the reduction of sample size. Figure B.6 in Appendix B.1 shows that across treatments, the median cross-measure correlations in Order 2 are high, consistent with H1. Subjects reported low decision around the switching choices and chose randomization probabilities around 0.5 in these choices, supporting H2 (see Table B.7). Table B.8 suggests that subjects in Order 2 had lower decision confidence in decisions involving the complex lottery compared to the simple lottery, consistent with H3. In Order 2, similar to the other two orders, decisions involving the complex lottery showed higher confidence in the full experience treatment than in the non-experience treatment, consistent with H4, although the treatment effect is not statistically significant in any individual order. This high level of consistency between randomization probabilities and self-reported confidence measures as well as their similar reactions to exogenous change of decision confidence suggest that these measures likely share common psychological foundations, even if not identical.

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

Finally, it is worth noting that the order effects discussed above do not necessarily suggest that randomization probabilities are a poorer proxy for decision confidence than the self-reported confidence measures. Confidence statements and probabilistic confidence are also noisy proxies of decision confidence, and it is not obvious what the "right" amount of strict randomization is. When we assess the value of decision confidence based on its correspondence with actual choices (when subjects report lower confidence in choosing an option, they should be less likely to choose that option), we find both randomization probabilities and probabilistic confidence corresponded to actual choices. Importantly, randomization probabilities exhibited a closer correspondence to binary choices than probabilistic confidence, especially in Order 2. See Appendix B.1 for more details.

4.4 Alternative interpretations of randomization probabilities

We have interpreted randomization behavior as a lack of decision confidence in the face of preference uncertainty. This is consistent with the findings from our post-experiment questionnaire as well as findings in Agranov and Ortoleva (2023) where many subjects explicitly mentioned unsureness, complexity, difficulty, and hedging as reasons for randomization (see Appendix B.3 for details). However, subjects may have other reasons for randomization. While it is not possible to eliminate all alternative interpretations, we show that indifference, random errors, or utility difference alone cannot be the driving force behind subjects' randomization behavior and the treatment effects.

First, indifference is not the driving reason for randomization because the majority of the subjects randomized at least twice (see Table B.9 in Appendix B.2), while randomization from indifference should occur for at most one value of sure payments. Second, randomization was unlikely a result of random errors because subjects' randomization probabilities of choosing x decreased monotonically with the value of y, even though they faced a random sequence of y (see e.g., Panel (c) in Fig. 5). Further, the treatment effects on randomization probabilities in the expected directions suggest that subjects' randomization was likely a deliberate choice. Third, if randomization probabilities were due to utility differences alone, randomization probabilities should increase for each value of sure payments when the utility over the lottery increases (e.g., moving from the simple lottery to the complex lottery or the no-experience treatment to the full-experience treatment), which increases the lower bound and the upper bound of randomization as well. Our results clearly reject these predictions. Randomization probabilities were compressed toward 0.5 rather than increased monotonically, and the two bounds often moved in opposite directions. We elaborate on this in Appendix B.2.

5 Conclusion

We propose that letting individuals assign randomization probabilities according to which they receive options can be an incentivized way to elicit decision confidence. In two experiments, we elicited randomization probabilities as well as

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press

two self-reported confidence measures and further manipulated decision confidence exogenously.

We find that most subjects randomized frequently, and their randomization probabilities and self-reported confidence measures were linked in ways that are consistent with the hypotheses derived from two theoretical analyses. While there were some order effects depending on whether randomization probabilities were elicited before or after the self-reported confidence, cross-measure correlations were high, and randomization probabilities corresponded closely to probabilistic confidence in absolute levels, with a high randomization probability assigned to an option associated with high self-reported probabilistic confidence. Our further examination suggests that alternative interpretations of randomization such as indifference, random errors, or differences in utility alone are unlikely to be the driving factors. Overall, our results suggest that decision confidence can be meaningfully and accurately inferred from randomization probabilities.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-024-09837-x.

References

Abdellaoui, M., Li, C., Wakker, P., & Wu, G. (2020). A defense of prospect theory in Bernheim and Sprenger's experiment. Working paper, Erasmus University Rotterdam.

Agranov, M., & Ortoleva, P. (2023). Ranges of preferences and randomization. *Review of Economics and Statistics*, Forthcoming.

Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy, 125*(1), 40–68.

Apesteguia, J., & Ballester, M. A. (2018). Monotone stochastic choice models: The case of risk and time preferences. *Journal of Political Economy, 126*(1), 74–106.

Arts, S., Ong, Q., & Qiu, J. (2020). Measuring subjective decision confidence. MPRA Paper 106811, University Library of Munich, Germany.

Aydogan, I., & Gao, Y. (2020). Experience and rationality under risk: Re-examining the impact of sampling experience. *Experimental Economics, 23*(4), 1100–1128.

Butler, D. J., & Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review, 97*(1), 277–297.

Butler, D., & Loomes, G. (2011). Imprecision as an account of violations of independence and betweenness. *Journal of Economic Behavior & Organization, 80*(3), 511–522.

Camerer, C. F., & Hogarth, R. (1999). The effects of financial incentives in experiments: A review and capital-labor-roduction framework. *Journal of Risk and Uncertainty, 19*, 7–42.

Cerreia-Vioglio, S., Dillenberger, D., & Ortoleva, P. (2015). Cautious expected utility and the certainty effect. *Econometrica, 83*(2), 693–728.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., & Riella, G. (2019). Deliberately stochastic. *American Economic Review, 109*(7), 2425–2445.
Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica, 51*(4), 1065–1092.
Chew, S. H. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research, 19*(1), 273–298.
Chew, S. H., Miao, B., & Zhong, S. (2017). Partial ambiguity. *Econometrica, 85*(4), 1239–1260.
Cohen, M., Jaffray, J.-Y., & Said, T. (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational Behavior and Human Decision Processes, 39*(1), 1–22.
Cubitt, R., Kopsacheilis, O., & Starmer, C. (2022). An inquiry into the nature and causes of the description—experience gap. *Journal of Risk and Uncertainty, 65*(2), 105–137.
Cubitt, R., van de Kuilen, G., & Mukerji, S. (2020). Discriminating between models of ambiguity attitude: A qualitative test. *Journal of the European Economic Association, 18*(2), 708–749.
Cubitt, R., Navarro-Martinez, D., & Starmer, C. (2015). On preference imprecision. *Journal of Risk and Uncertainty, 50*, 1–34.
Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory, 40*(2), 304–318.
Dubourg, W. R., Jones-Lee, M. W., & Loomes, G. (1994). Imprecise preferences and the WTP-WTA disparity. *Journal of Risk and Uncertainty, 9*, 115–133.
Dubourg, W. R., Jones-Lee, M. W., & Loomes, G. (1997). Imprecise preferences and survey design in contingent valuation. *Economica, 64*(256), 681–702.
Dwenger, N., Kübler, D., & Weizsäcker, G. (2014). Flipping a coin: Theory and evidence. CESifo Working Paper 4740, Center for Economic Studies and ifo Institute (CESifo), Munich.
Dwenger, N., Kübler, D., & Weizsäcker, G. (2018). Flipping a coin: Evidence from university applications. *Journal of Public Economics, 167*, 240–250.
Eliashberg, J., & Hauser, J. (1985). A measurement error approach for modeling consumer risk preference. *Management Science, 31*(1), 1–25.
Enke, B., & Graeber, T. (2021). Cognitive uncertainty in intertemporal choice. Working paper, Harvard University.
Enke, B., & Graeber, T. (2023). Cognitive uncertainty. *Quarterly Journal of Economics, Forthcoming.*
Feldman, P., & Rehbeck, J. (2022). Revealing a preference for mixtures: An experimental study of risk. *Quantitative Economics, 13*(2), 761–786.
Fudenberg, D., & Puri, I. (2021). Evaluating and extending theories of choice under risk. Working paper.
Fudenberg, D., Iijima, R., & Strzalecki, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica, 83*(6), 2371–2409.
Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics, 18*(2), 141–153.
Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association, 1*(1), 114–125.
Gul, F., & Pesendorfer, W. (2006). Random expected utility. *Econometrica, 74*(1), 121–146.
Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica, 75*(2), 503–536.
Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15*(8), 534–539.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.
Khaw, M. W., Li, Z., & Woodford, M. (2021). Cognitive imprecision and small stakes risk aversion. *The Review of Economic Studies, 88*(4), 1979–2013.
Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica, 73*(6), 1849–1892.
Loomes, G., & Pogrebna, G. (2017). Do preference reversals disappear when we allow for probabilistic choice? *Management Science, 63*(1), 166–184.
Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review, 39*(3–4), 641–648.
Machina, M. J. (1985). Stochastic choice functions generated from deterministic preferences over lotteries. *Economic Journal, 95*(379), 575–94.

Miao, B., & Zhong, S. (2018). Probabilistic social preference: How Machina's mom randomizes her choice. *Economic Theory, 65*(1), 1–24.
Mosteller, F., & Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy, 59*(5), 371–404.
Myagkov, M., & Plott, C. R. (1997). Exchange economies and loss exposure: Experiments exploring prospect theory and competitive equilibria in market environments. *American Economic Review, 87*(5), 801–828.
Ong, Q., & Qiu, J. (2023). Paying for randomization and indecisiveness. *Journal of Risk and Uncertainty, 67*, 45–72.
Plott, C. R., & Zeiler, K. (2005). The willingness to pay - willingness to accept gap, the "endowment effect,'' subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review, 95*(3), 530–545.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization, 3*(4), 323–343.
Rubinstein, A. (2002). Irrational diversification in multiple decision problems. *European Economic Review, 46*(8), 1369–1378.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.
van de Kuilen, G. (2009). Subjective probability weighting and the discovered preference hypothesis. *Theory and Decision, 67*, 1–22.
van de Kuilen, G., & Wakker, P. (2006). Learning in the Allais paradox. *Journal of Risk and Uncertainty, 33*, 155–164.
Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin, 144*(2), 140–176.
Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*(1), 75–98.

https://doi.org/10.1007/s10683-024-09837-x Published online by Cambridge University Press