Measuring subjective decision confidence

Arts, Sara and Ong, Qiyan and Qiu, Jianying

Radboud University Nijmegen, National University of Singapore,
Radboud University Nijmegen

August 2020

Measuring subjective decision confidence

Sara Arts[*] Qiyan Ong[†] Jianying Qiu[‡]

August, 2020

Abstract

To understand decision-making, it is important to determine the degree to which individuals are confident in making choices. In place of self-reported confidence statements, on which most studies have relied, this study examines an incentivized measure to elicit quantitative decision confidence theoretically and experimentally. We demonstrate the feasibility of this measure in a setting where individuals are allowed to choose randomization probabilities for two options according to which they may receive either option. Our theoretical analysis demonstrates that individuals randomize when they are not 100% confident about their choices, and the randomization probability reveals their level of decision confidence. We tested this inference in an experiment that elicited both confidence statements and randomization probability from our subjects. Our experimental results provide strong evidence that one could interpret the randomization probability for an option as the probabilistic confidence of choosing that option.

Keywords: decision confidence, randomization, incentivized mechanism

JEL Classification B40, C91, D81

[*]Department of Economics, IMR, Radboud University, Heyendaalseweg 141, 6525 AJ Nijmegen, the Netherlands.

[†]Social Service Research Centre, Faculty of Arts and Social Sciences, National University of Singapore, VENTUS, #03-01, 8 Kent Ridge Drive, Singapore 119246 Email: qyong@nus.edu.sg.

[‡]Jianying Qiu, email: j.qiu@fm.ru.nl, Department of Economics, IMR, Radboud University, EOS 02.577, Heyendaalseweg 141, 6525 AJ Nijmegen, the Netherlands.

1 Introduction

Although they are ignored in standard economics, there are many important life decisions that people are unable to make with full confidence. This is because many choices involve trade-offs between conflicting objectives, and resolving these trade-offs is difficult. Studies have increasingly shown that decision confidence – or the lack of it – has the potential to explain a wide range of anomalies, such as the WTA-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, 2019), and many other violations of standard decision theory (Butler and Loomes, 2011). The widespread influence of confidence on decision-making and behaviour suggests that policy design should account for decision confidence in order to improve decision-making and social welfare.[1]

Given the important role of decision confidence in decision-making, there is increasing interest in its elicitation. Most existing studies have elicited decision confidence using non-incentivized self-reported confidence statements. For example, Dubourg et al. (1994, 1997) allowed subjects to indicate whether they were unsure of their choices. Butler and Loomes (2007) and Butler and Loomes (2011) asked subjects to indicate their decision confidence about their choice between two options in the ordinal terms of surely, probably, and unsure. Instead of asking subjects to state how confident they were, Cohen et al. (1987), Cubitt et al. (2015), and Enke and Graeber (2019) obtained confidence intervals from their subjects. Cohen et al. (1987) and Cubitt et al. (2015) had subjects report the range of choices over which they were unsure of their preferences, whereas Enke and Graeber (2019) had subjects report the range of values over which they were certain of their preferences.

We build on these studies to propose an incentivized quantitative measure to elicit deci-

[1]For example, insofar as individuals feel less confident about their choices when excessive options are available, it may be better to limit the number of investment options in retirement plans (Iyengar and Lepper, 2000; Iyengar et al., 2004; Iyengar and Kamenica, 2010)

sion confidence. Instead of requiring subjects to commit to one option out of two available options, this measure allows the individual to choose the randomization probabilities according to which she receives each option. We show theoretically and experimentally that the individual may prefer to randomize when she is not 100% confident in choosing one option over the other and that the randomization probabilities vary according to how confident she is about this decision.

Theoretically, we capture the lack of confidence about choices by assuming that an individual has multiple selves. Each self represents one particular way to trade off between conflicting objectives, and different trade-offs imply different optimal choices. The more strongly the multiple selves disagree with each other, the less confident the individual feels about choosing one option over the other. Furthermore, the individual dislikes disagreement among the multiple selves, because individual decision-making in the presence of multiple selves is similar to group decision-making, which requires members with different opinions to reach a consensus. In both situations, stronger disagreement requires more time and cognitive effort for decision-making. Our theoretical analysis demonstrates that randomization reveals the individual's lack of confidence in choosing one option over the other because randomizing reveals her preference to "hedge" across multiple selves when they disagree with each other. Moreover, the randomization probability assigned to each option measures the level of decision confidence about that option, with a smaller randomization probability for an option indicating weaker confidence in choosing that option over the other.

We tested the link between randomization probability and decision confidence in an experiment. The experiment requires subjects to make a choice between pairs of options: a lottery x and a sure payment y. The choice between a lottery and a sure payment involves a trade-off between risk and return, as in many investment decisions. For each lottery x, we kept the lottery the same in each pair and varied the sure payment y with 13 possible values in a random sequence. Subjects first made a standard binary choice in which they chose either the lottery x or the sure payment y. To elicit subjects' decision confidence, we followed Dubourg et al. (1994) and Butler and Loomes (2007) by allowing subjects to state how confident they were about each choice (surely, probably, or unsure). After the binary choices and confidence statements were made, subjects proceeded to the randomized choices, in which they 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, again in random sequence. The binary choices, confidence statements, and randomization probabilities allowed us to test for the presence of a systematic relationship between the randomization probabilities and confidence statements. We further checked the robustness of this relationship by varying across four different lotteries of different cognitive demands and two experimental conditions in which half of the subjects were provided with the simulated experience of a lottery and the other half did not.

Our experimental results suggest a systematic relationship between randomization probabilities and confidence statements. We find an economically important and statistically significant correlation between subjects' randomization probabilities and confidence statements (median Spearman correlation of 0.85). The randomization probability associated with an option increases as the stated confidence for that option increases from "Unsure" to "Probably" to "Surely." The results suggest empirical validity in interpreting randomization probability as probabilistic confidence, in the sense that assigning a randomization probability of 0.5 to an option corresponds to being 50% confident in choosing that option over the other. Additionally, the confidence intervals – the range of sure payments y for which subjects do not feel fully confident about their choices – identified from randomization probabilities are close to those identified from confidence statements. These findings are largely robust to lottery types and the provision of experience sampling. Further analysis also suggests that subjects' randomization pattern is consistent with our theoretical analysis, but not with indifference, errors, or nonlinear probability weighting (Kahneman and Tversky, 1979; Quiggin, 1982; Tversky and Kahneman, 1992). Overall, our results suggest that randomization probability serves well as an incentivized quantitative measure of decision confidence about choices.

Our study contributes to the literature on decision confidence in several ways. First, our measure substantiates and complements the earlier non-incentivised measures. If different measures aiming to capture decision confidence are systematically related, we are more confident that they indeed capture what we intend to capture. Our measure further complements earlier measures because proper material incentives improve data quality by motivating subjects to devote more time and cognitive effort to making the choices. Second, our measure is quantitative and continuous, offering flexibility for data analysis. For example, depending on the needs of the research, we can focus on variant confidence intervals, such as 90% or 95%. Achieving the same flexibility with confidence intervals revealed from self-reported statements might be difficult, because such confidence intervals might not be sensitive to manipulation. Indeed, Enke and Graeber (2019) elicited confidence intervals of 70%, 90%, 95%, 99%, and 100% via confidence statements and found that subjects are unresponsive to the manipulation and always report similar confidence intervals. Furthermore, the quantitative randomization probability and its associated lottery outcomes are objective. They can be compared across individuals when needed.[2] Self-reported confidence statements are, however, subjective and could have a different meaning for different individuals.[3] For example, some consider a probabilistic confidence level of 68% sufficient for stating "sure," whereas others may require 85%. Finally, decision confidence is traditionally an intuitive idea with many different interpretations, and our theoretical analysis provides a concrete conceptual definition.[4] We show that decision confidence revealed through randomization probability is the (expected) utility difference weighted by a measure of disagreement among multiple selves.

Although our paper builds directly on studies of decision confidence (Dubourg et al., 1994, 1997; Butler and Loomes, 2007, 2011; Cubitt et al., 2015), it is also closely related to studies that have investigated preference uncertainty, convex preferences, and/or conflicts in choices (Bewley, 2002; Eliaz and Ok, 2006; Ok et al., 2012; Cerreia-Vioglio et al., 2015; Qiu and Ong, 2017; Cerreia-Vioglio et al., 2019; Agranov and Ortoleva, 2020). In particular, both Qiu and Ong (2017) and Agranov and Ortoleva (2020) allowed subjects to assign randomization probabilities to the two options that they face in a decision. Qiu and Ong

[2]Needless to say, comparing any measures across individuals can be problematic and requires a strong motivation.

[3]Indeed, our experimental result presented in Appendix D suggests that different individuals could associate the same confidence statement with different probabilistic confidence levels.

[4]For example, decision confidence could be related to the utility difference between options (as the strength of preferences in Butler et al., 2014). Although the two are clearly related, the utility difference does not directly translate into decision confidence. The utility difference between two sure payments of 10 euro and 10.01 euro is small, but subjects are likely to be 100% confident about their preferences.

(2017) discussed how randomization reveals a difficult trade-off between conflicting values in choices, whereas Agranov and Ortoleva (2020) investigated the ranges of values for which subjects randomize between two options and related these ranges to certainty bias and nonmonotonic choices. Our study contributes to this line of research by showing how randomization may also be related to decision confidence and how eliciting randomization probabilities may be an alternative measure of decision confidence.

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

2 Experimental design

The experiment consisted of eight treatments of four lotteries (with different cognitive demands) $\times$ two conditions (with/out experience sampling). We describe the general structure of the experiment in all treatments before detailing the conditions of each treatment.

2.1 General structure of the experiment

Subjects faced a pair of options in each decision: option x (a lottery) and option y (a sure payment). Within each treatment, option x was kept the same, whereas option y went through a random sequence of 13 possible values, which were lottery dependent. For each pair of options, subjects needed to make three decisions. First, subjects made a binary choice in which they chose either x or y. Next, subjects reported their level of confidence about their binary choice. They could report "Surely x," "Probably x," "Unsure," "Probably y" and "Surely y", statements that were also used in, e.g., Dubourg et al. (1994) and Butler and Loomes (2007). Finally, they made randomized choices in which they chose a randomization probability λ according to which they receive x (and hence with chance $1 - \lambda$ receive y). For example, a value of $\lambda = 0.75$ means subjects may

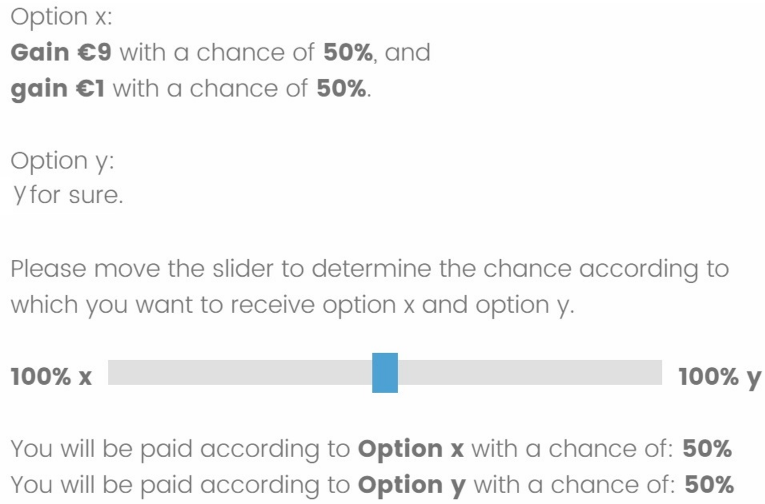


Figure 1: An example of the decision screen, where option x is a lottery to gain 9 euro with a chance of 50% and 1 euro with a chance of 50%. Option y is a sure payment and varies across choices. Subjects had to move the slider to determine the randomization probability. The randomization probability changed at an increment of 1%. Changes in the randomization probability were reflected in the descriptions below the slider.

receive x with a chance of 75% and receive y with a chance of 25%.[5] The randomization probability changes at an increment of 1%. Figure 1 shows an example of the decision screen for a randomized choice.

The three sets of decisions were made in two stages. In each treatment, subjects stated the binary choice and then provided their confidence statements for each pair of options in the first stage. When all the binary choices and confidence statements were made, they went on to the randomization stage, in which they assigned a randomization probability for each pair of options x and y.

2.2 Treatments

There were two types of treatment variations in our experiment: four lotteries with different cognitive demands and the condition of allowing for or not allowing for sampling potential outcomes of the lottery.

[5] In the experiment, we explain the chance as a computer drawing a random number between 1 and 100.

We vary cognitive demand for decision-making by changing the number (two or four) and the payoff (loss or gain) domain of the lottery outcomes. This allows us to examine whether the association between confidence and randomization, if present, is robust to the different levels of cognitive demand required for different lotteries. There were four lotteries, and all subjects went through them in a sequence randomized across subjects. Subjects completed all the decisions as described in the previous subsection associated with each lottery before moving on to the next lottery.

The baseline lottery is a simple lottery with two outcomes (gain 9 euro with a chance of 50%, and gain 1 euro with a chance of 50%). Because some studies have shown that people find it hard to evaluate lotteries with more outcomes and dislike them (Huck and Weizsäcker, 1999; Sonsino et al., 2002; Moffatt et al., 2015), we introduce a second lottery with four outcomes, which we refer to as the complex lottery (gain 9.75 euro with a chance of 20%, gain 7.50 euro with a chance of 30%, gain 2.50 euro with a chance of 30%, and gain 0.25 euro with a chance of 20%).[6] The third and fourth lottery are similar to the simple lottery in the number of outcomes, but they involve potential losses. Establishing the association between confidence and randomization in lotteries involving losses is important in view of evidence showing that people behave differently when faced with losses compared to gains (Gonzalez et al., 2005; Pabst et al., 2013) and that they face emotional trade-off difficulty with losses (Luce et al., 1999). The third lottery, which we refer to as the loss lottery, has all its outcomes in the domain of losses: lose 9 euro with a chance of 50% and lose 1 euro with a chance of 50%. This lottery is constructed by subtracting 10 euro from payoffs of the simple lottery. The fourth lottery – the mixed lottery – involves both payoff domains (gain 4 euro with a chance of 50%, and lose 4 euro with a chance of 50%). This lottery is constructed by subtracting 5 euro from the payoffs of the simple lottery. Because there may be losses in these two lotteries, subjects received an initial endowment for these two lotteries to prevent them from incurring out-of-pocket losses (an initial endowment of 10 euro for the loss lottery and of 5 euro for the mixed lottery). To make the initial endowment salient, a 10-euro bill and 5-euro bill was displayed on the decision screens for the loss lottery and mixed lottery, respectively. Table 1 summarizes the different lotteries

[6]The simple and the complex lotteries have the same expected value. The standard deviation of the simple lottery is 4.0, slightly larger than that of the complex lottery (3.6).

Lotteries	Option x	Option y
The simple lottery	(50%, 9 ; 50%, 1)	Receiving y for sure (y is between 0 and 10)
The complex lottery	(20%, 9,75 ; 30%, 7,50 ; 30%, 2,50 ; 20%, 0,25)	Receiving y for sure (y is between 0 and 10)
The loss lottery	(50%, -9 ; 50%, -1)	Receiving y for sure (y is between -10 and 0)
The mixed lottery	(50%, 4 ; 50%, -4)	Receiving y for sure (y is between -5 and 5)

Table 1: Summary of options x and y in the four lotteries. In the simple lottery and the complex lottery, y takes the value of $Y \in \{0, 2, 3, 3.5, 4, 4.5, 5.0, 5.5, 6, 6.5, 7, 8, 10\}$, in the loss lottery y is each value in Y minus 10, and in the mixed lottery y is each value in Y minus 5.

and the set of sure payment amounts for y in the experiment.

The second type of treatment we introduce in the experiment examines whether allowing subjects to experience the potential outcomes of the lotteries improves their confidence in decision-making. Previous experiments have found that trading experience reduces the WTA-WTP gap (List, 2003, 2011) and that simulated experience makes subjects feel more informed about their decisions and encourages them to take more risks (Bradbury et al., 2014). We adopt the experience sampling approach of Hertwig et al. (2004). Experience sampling is consistent with a notion in accumulator models that contemplatesthat the utility distribution of the alternative is built on past experiences (Busemeyer and Townsend, 1993).

Half of the subjects were assigned to the experience sampling treatment. After a lottery was explained to these subjects and before they had to make choices between the lottery and the sure payment, they viewed a screen on which they had to click a button to generate the potential outcomes of the lottery. A bar chart on the same page recorded each lottery outcome that was simulated and illustrated the frequency of each lottery outcome. Subjects had to click 20 times to complete the experience sampling exercise, at which point they could proceed to making their choices. At the 20th click, the final bar chart displayed the distribution of the lottery outcomes the same as the probability distribution of the lottery.For example, the simulated 20 outcomes of the complex lottery were always 4 times 9.75 euro, 6 times 7.50 euro, 6 times 2.50 euros, and 4 times 0.75 euro. Subjects went through experience sampling for each lottery and made decisions for that lottery before

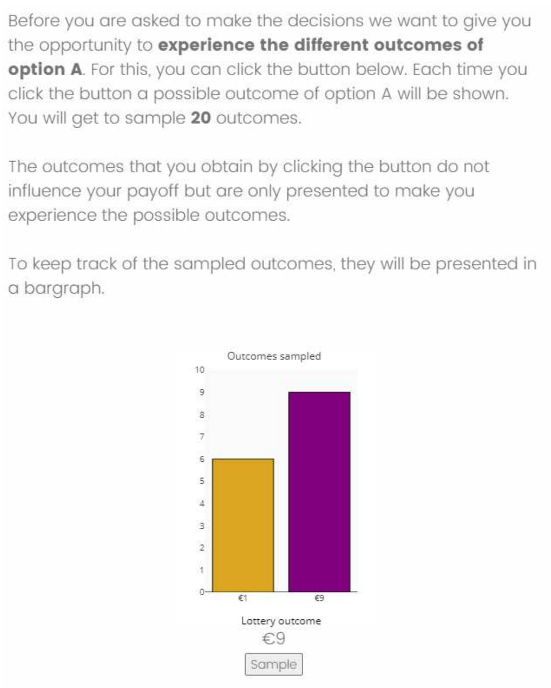


Figure 2: An example of the screen on which subjects are allowed to sample outcomes of the simple lottery. The final bargraph of the sampled outcomes is displayed on the subsequent decision screens.

going through experience sampling for the next lottery. To remind the subjects about the simulated experience of a lottery, the bar chart showing the distribution of outcomes for that lottery was made available at the side of the decision screens when subjects made their binary choices and the randomization probabilities between that lottery and the sure payment.

2.3 Sample and procedure

The experiment was conducted with a sample of 205 subjects of the DISCON lab at Radboud University. Invitations were sent in batches via ORSEE (Greiner, 2015). About half of the subjects were male and about half were female. The experiment was conducted using Qualtrics, and lasted approximately 20 minutes. Subjects made binary choices, confidence statements, and randomization probabilities for 13 pairs of options for each lottery. Examples from the experimental materials can be found in Appendix D.

Each student received a participation fee of 1 euro and monetary compensation based on their binary and randomization decisions in the experiment. Specifically, for each subject, one decision out of all the binary and randomization decisions was randomly selected for payment. We made the payment via bank transfers.

3 Theoretical analysis

In this section, we demonstrate the concrete link between decision confidence and the randomization probability in the randomized choices. Building on Cerreia-Vioglio et al. (2015) and Klibanoff et al. (2005), we propose an approach to capturing the decision-making of an individual who might have limited confidence for some choices. We then show the link between randomization probability and decision confidence. We discuss alternative interpretations of randomization, such as indifference, errors, and nonlinear probability weighting, in subsection 4.2.

3.1 Decision-making with limited confidence

Standard economics assumes that a unique utility function (subject to positive affine transformation) captures the individual's preference. The individual always makes choices with full confidence. It can be shown that under standard economic models, the individual chooses $\lambda^* \in (0,1)$ at most once in the 13 choice pairs in most treatments.[7]

When the individual might feel unsure about her choices, the assumption of a unique utility function is no longer appropriate. To accommodate the possibility that a decision-maker might not be fully confident about her choices, we assume the individual has multiple utility functions that we call multiple selves, with each self representing one particular way to trade off between conflicting objectives in choices. Such a modelling technique is popular in models of incomplete preferences (see e.g., Bewley, 2002; Dubra et al., 2004;

[7]Such models include the expected utility theory, cumulative prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), and rank dependent utility theory (Quiggin, 1982). Appendix B discusses this claim in more detail.

Cerreia-Vioglio et al., 2015). The individual is fully confident about her choices when all selves choose in the same way. The individual is not fully confident about her choices when some selves choose one option but others choose other options.

Specifically, let u_τ denote the utility function of the self τ and Γ denote the set of selves. Let π denote the probability distribution over Γ. Given a utility function u_τ, we follow the standard assumption that the self behaves according to the expected utility theory (EUT). Let $EU_\tau(l)$ denote the expected utility of an option $l \in L$. We further assume that the individual dislikes disagreement among selves, because arriving at a choice in the presence of multiple selves with different preferences is, in essence, similar to situations in which a group of people with different opinions tries to reach a consensus. The more strongly members disagree with each other, the more difficult it is for the group to make compromises and agree on a single opinion. Aversion to disagreement among selves can then be interpreted as the cost of forcing different selves to make compromises and agree on a single choice. With the above assumptions, we can write the individual's preference over an option l as

$$V(l) = \int_\Gamma \phi\left[EU_\tau(l)\right] d\pi, \tag{1}$$

, where concave $\phi(\cdot)$ implies an aversion to disagreement – deviations from the mean expected utility – among different selves. Similar to the connection between the concavity of utility function and risk aversion, the concavity of $\phi(\cdot)$ implies that the individual places more weight on the selves who give a lower value to l.Such a cautious attitude is consistent with Levitt (2020), who showed that subjects who experience difficulty in making a decision are often excessively cautious with respect to maintaining the status quo.

Equation 1 extends directly from Cerreia-Vioglio et al. (2015) and Klibanoff et al. (2005). It can be seen as a smooth version of the cautious expected utility model (Cerreia-Vioglio et al., 2015). It is also a parallel of the smooth ambiguity model of Klibanoff et al. (2005), in which the state space is the set of multiple selves. Indeed, in ambiguity models of multiple priors, an individual is unsure about the probability distribution of the states of nature and has multiple priors. In the current approach, an individual is unsure about her utility function and has multiple selves.

3.2 Linking the randomization probability to decision confidence

We are now ready to establish the link between decision confidence and the randomization probability in the randomized choices. In particular, we will show that a smaller randomization probability for an option is associated with lower confidence in choosing that option.

Specifically, recall that in our mechanism the individual chooses a randomization probability $\lambda \in [0,1]$ and builds a lottery $(\lambda, x; (1-\lambda), y)$: she receives x with probability λ and y with probability $1-\lambda$. Because for any given self τ the individual's preference over the lottery $(\lambda, x; (1-\lambda), y)$ satisfies EUT, we have $EU_t\left[\lambda x + (1-\lambda)y\right] = \lambda EU_t(x) + (1-\lambda)EU_t(y)$.[8] The individual's decision is then to maximize her utility by choosing the optimal randomization probability $0 \leq \lambda \leq 1$:[9]

$$Max_{\lambda}\ V\left[\lambda x + (1-\lambda)y\right] = \int_{\Gamma} \phi\left[\lambda EU_{\tau}(x) + (1-\lambda)EU_{\tau}(y)\right] d\pi.$$

In the experiment, y is a sure payment. Sure monetary payments are probably the easiest options to evaluate, and we thus assume the individual is always confident about her evaluation of a sure payment: $EU_{\tau}(y) = u(y),\ \forall \tau \in \Gamma$ when y is a sure payment. In this case, the optimal λ is simply:[10]

$$\lambda^* \approx \frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}} \times \frac{\Delta_u}{\sigma_x^2} \tag{2}$$

, where $\Delta_u = E_t\left[EU_t(x)\right] - u(y)$ captures the expected utility difference of x and y, $\sigma_x^2 = E_t\left[EU_t(x) - E_t(EU_t(x))\right]^2$ is the standard deviation of $EU_t(x)$ and approximates

[8]When x is a lottery, $(\lambda, x; (1-\lambda), y)$ is a compound lottery. We follow EUT and assume reduction of compound lotteries for these choices.

[9]Taking Cerreia-Vioglio et al. (2015) literally, the negative certainty independence axiom implies no preference for randomization when a lottery is matched with a sure payment.

[10]More precisely, because $0 \leq \lambda \leq 1$, $\lambda^* \approx min\left\{max\left\{0, \frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}} \times \frac{\Delta_u}{\sigma_x^2}\right\}, 1\right\}$. The detailed derivation is presented in Appendix A.

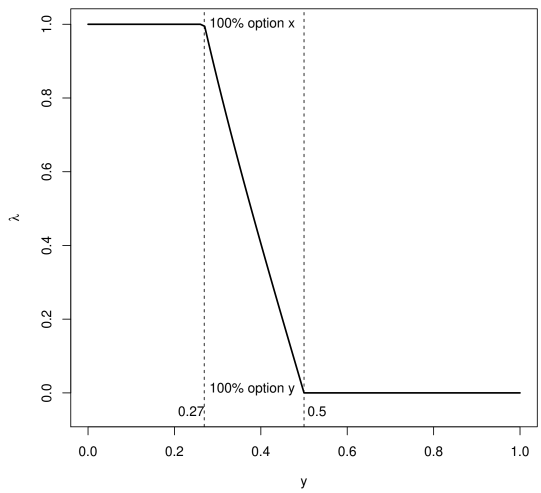


Figure 3: The optimal λ^* depending on the value of y. The figure is produced by assuming $\phi(EU_t) = 1 - e^{-EU_t}$, $Prob(u_1) = 0.5$, $Prob(u_2) = 0.5$, $EU_1(x) = 1$,$EU_2(x) = 0$, and $EU_1(y) = EU_2(y) = y$. The optimal randomization probability $\lambda^* = -ln(\frac{y}{1-y})$.

how strongly different selves disagree with each other. Similarly to decision-making under risk, $-\frac{\phi''(u(y))}{\phi'(u(y))}$ can be interpreted as a metric of attitudes toward disagreement among selves.

To see the link between the randomization probability λ and decision confidence, note that intuitively the individual should be less confident about choosing x when x becomes less attractive relative to y and should be more uncertain about her evaluation of x. Equation 2 reflects this intuition exactly: Weaker confidence in choosing x is associated with smaller randomization probability for x (λ is small when Δ_u is small and σ_x^2 is large). It is in this sense that we state λ^* reveals and measures confidence in choices.

As a concrete illustration, consider the following numerical example: The individual has two selves $t = 1, 2$, and considers them equally likely. Option x is a lottery, and $EU_1(x) = 1$ and $EU_2(x) = 0$. Option y is a sure payment, and $u_1(y) = u_2(y) = y$. The function $\phi(EU_t) = 1 - e^{-EU_t}$, where the concavity of $\phi(\cdot)$ captures her attitude toward the disagreement among multiple selves. The decision utility of choosing option y is $V(y) = 1 - e^{-y}$, and the decision utility of choosing option x is $V(x) = 0.5(1 - e^{-1}) + 0.5 \times 0 = 0.316$. When y is sufficiently similar to x ($0.27 < y < \frac{1}{2}$), the individual will have incentive to randomize over x and y. The decision utility of such a lottery is $V(\lambda x + (1-\lambda)y) = 0.5\left[1 - e^{-[\lambda \times 1 + (1-\lambda)\times y]}\right] + 0.5 \times \left[1 - e^{-[\lambda \times 0 + (1-\lambda)\times y]}\right]$. A simple calculation shows that the optimal $\lambda^* = -ln(\frac{y}{1-y})$, subject to $0 \leq \lambda^* \leq 1$. Figure 3 shows the relationship between the optimal λ^* and y. As one can see, when y becomes more attractive, the value of λ^* decreases. Moreover, λ^* approaches 0.5 when the two options become similar in terms of their decision utilities ($V(y) = 0.314$ versus $V(x) = 0.316$).

4 Experimental results

We report our experimental results in two steps. First, we link randomization probabilities directly to self-reported confidence statements and show that there exists a systematic relationship between the two. Second, we discuss alternative interpretations of randomization and show why randomization is unlikely to reflect indifference, errors, or nonlinear probability weighting.

4.1 Randomization probabilities and confidence statements

We provide four empirical observations that are consistent with a systematic relationship between randomization probabilities and confidence statements. First, we show that there is an economically important and statistically significant correlation between randomization probabilities and confidence statements. Second, we demonstrate that it is empirically valid to interpret randomization probability directly as probabilistic confidence. Third, around the sure payments where subjects switched between the lottery and the sure payments (switching choices),we find that subjects tended to report the least confidence about their choice and choose randomization probabilities around 0.5. Finally, we compare the confidence intervals defined by randomization probabilities and confidence statements and show that they are closely related to each other. Because these results are robust to the four lotteries and the experience sampling condition, most of the results discussed henceforth are based on the aggregated data from all lotteries and experimental conditions. Where relevant, we briefly discuss the presence of treatment differences. The detailed results for

		Simple	Complex	Loss	Mixed
No experience sampling	10th percentile	0.60	0.69	0.67	0.35
	median	0.91	0.90	0.90	0.85
	90th percentile	0.97	0.97	0.97	0.96
Experience sampling	10th percentile	0.60	0.63	0.65	0.47
	median	0.93	0.88	0.87	0.81
	90th percentile	0.97	0.96	0.97	0.95

Table 2: Nonparametric Spearman correlation at the 10th percentile, median, and 90th percentile by lottery and experience sampling conditions

each lottery and the experience sampling conditions that are not presented in this section are provided in the Appendix.

Our first result summarizes the correlation between randomization probabilities and confidence statements.

Result 1. *Across all lotteries and experience sampling conditions, there exists an economically important and statistically significant correlation between randomization probabilities and confidence statements.*

Support: To compute the correlation between the randomization probabilities and the confidence statements, we assigned values to the confidence statements of "Surely x," "Probably x," "Unsure," "Probably y," and "Surely y" on a scale of 5 to 1, with "Surely x"' taking the value of 5 and "Surely y" taking the value of 1. For each subject, we computed the nonparametric Spearman correlation for each lottery. A positive correlation is consistent with our hypothesis that the more confident the subject is about choosing the lottery over the sure payment, the higher the randomization probability she assigns to x. The pooled median Spearman correlation across lotteries and subjects is 0.85 ($p < 0.01$), with a correlation of 0.93 at the 90th percentile and a correlation of 0.67 at the 10th percentile.[11] Table 2 reports the 10th percentile, medians, and 90th percentile of the correlations for the four lotteries in the two experimental conditions. Across lotteries and conditions, median correlations range from 0.81 to 0.93. When we examine the correlation at the 10th percentile, correlation is above 0.60 for most lotteries.[12] Overall, the high

[11]Three subjects did not vary their confidence statements, and thus no correlation can be estimated.

[12]This is true except for the mixed lottery, which showed a markedly weaker correlation between confi-

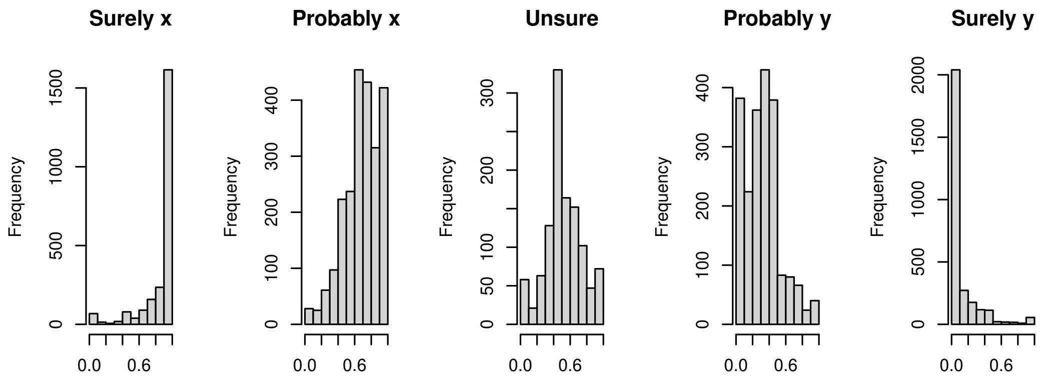


Figure 4: Histograms of randomization probability of each confidence statement across lotteries and conditions. The x-axis is the randomization probability ($0 \leq \lambda \leq 1$).

correlation found between confidence statements and randomization probabilities is consistent with a systematic relationship between randomization probabilities and confidence statements.

Our second result demonstrates how randomization probabilities vary with confidence statements.

Result 2. *Randomization probabilities relate systematically to confidence statements, with a high randomization probability for an option corresponding to high confidence in choosing that option. Furthermore, it is empirically valid to interpret randomization probability for an option directly as the probabilistic confidence of choosing that option.*

Support: We look at the mean randomization probability that corresponds to each confidence statement. On aggregate, consistent with our hypothesis, the randomization probability for x decreases with the confidence about choosing x: The randomization probability is 0.89 for "Surely x," 0.70 for "Probably x," 0.54 for "Unsure," 0.33 for "Probably y," and 0.11 for "Surely y." Figure 4 shows the histograms of the randomization probabilities across the five confidence statements. The randomization probabilities show a clear shift from the right to the left as we move from "Surely x" to "Surely y."

dence statements and randomization probabilities.

Figure 4 shows that the majority of the decisions (64%) for which subjects chose "Surely x" were given the randomization probability of 1 for x and that the majority of the decisions (63%) for which subjects chose "Surely y" were given the randomization probability of 0 for x. The reverse is also true. Among the decisions in which subjects assigned a randomization probability of 1 for x, 78% had "Surely x" as the corresponding confidence statement. 83% of the decisions in which subjects assigned a randomization probability of 0 for x were rated "Surely y." These results show a strong association between randomization probability and confidence statements at the extreme ends.

Compared to "Surely x," the distribution of the randomization probability corresponding to the confidence statement "Probably x" is also skewed to the left. However, the randomization probability for "Probably x" does not have a clear peak, as "Surely x" does.Likewise, the randomization probability distribution corresponding to the confidence statement "Probably x" is skewed to the right but without a clear peak. The results suggest that subjects may have less consensus over the randomization probabilities associated with "Probably x" and "Probably y" than over those associated with "Surely x" and "Surely y," even though there is general agreement that the randomization probabilities associated with "Probably x" are smaller than those of "Surely x" and that the randomization probabilities associated with "Probably y" are larger than those of "Surely y."

Finally, the randomization probability distribution corresponding to "Unsure" is a bell-shaped distribution, with a clear peak between 0.4 to 0.5. The results suggest that the consensus over the randomization probability associated with "Unsure" is weaker than the consensus over that associated with "Surely x" and "Surely y" but stronger than the consensus over that associated with "Probably x" and "Probably y". The positive relationship between the randomization probability for x and the confidence in choosing x is observed for all lotteries and experience sampling conditions. Table 3 reports the mean and standard deviation of the randomization probability for each confidence statement in aggregate and for each lottery and experience sampling condition. The mean randomization probability for each confidence statement does not differ significantly across lotteries. Experience sampling also does not appear to have any effect on the relationship between the randomization probability and the confidence statements.

Conditions	Lotteries	Surely x	Probably x	Unsure	Probably y	Surely y
Aggregate		0.89 (0.22)	0.70 (0.21)	0.54 (0.22)	0.33 (0.22)	0.11 (0.20)
No sampling	Simple	0.92 (0.20)	0.72 (0.20)	0.50 (0.23)	0.35 (0.22)	0.10 (0.18)
	Complex	0.90 (0.19)	0.72 (0.19)	0.56 (0.20)	0.34 (0.19)	0.11 (0.20)
	Loss	0.92 (0.17)	0.72 (0.20)	0.52 (0.23)	0.31 (0.23)	0.08 (0.18)
	Mixed	0.88 (0.26)	0.71 (0.22)	0.53 (0.23)	0.35 (0.27)	0.12 (0.24)
Sampling	Simple	0.87 (0.24)	0.68 (0.19)	0.55 (0.19)	0.31 (0.19)	0.10 (0.19)
	Complex	0.86 (0.23)	0.67 (0.21)	0.51 (0.20)	0.32 (0.19)	0.12 (0.19)
	Loss	0.89 (0.19)	0.70 (0.22)	0.57 (0.23)	0.33 (0.23)	0.08 (0.16)
	Mixed	0.88 (0.24)	0.71 (0.22)	0.55 (0.23)	0.34 (0.24)	0.12 (0.25)

Table 3: Mean randomization probability at each confidence level for the four lotteries in the two experience sampling conditions. The values in parentheses are the standard deviations of randomization probabilities.

In addition, we found that the association between the randomization probabilities and the confidence statements is broadly consistent with the way probabilistic confidence levels were implemented empirically in other studies. For example, Vanberg (2008) used the probabilistic confidence of 0.85 as the cutoff level between sure and probably, 0.68 as the cutoff level between probably and unsure, and 0.50 as unsure (Vanberg, 2008, Footnote 10). These values correspond to the mean randomization probabilities reported above for each confidence statement. For example, the mean randomization probability on aggregate is 0.89 for "Surely x," 0.70 for "Probably x," and 0.54 for "Unsure." Because we did not intentionally prime subjects to relate the randomization probability of choosing x to the probabilistic confidence of choosing x, the similarities in the empirical findings of the two measures in two separate studies suggest that a common cognitive pathway may have been responsible for both types of decision, linking randomization probability to confidence. In particular, it is empirically valid to interpret randomization probability directly as probabilistic confidence (e.g., a randomization probability 0.8 of choosing an option corresponds to being 80% confident about choosing that option).

We proceed to check the connection between subjects' behaviour around the switching choices in the binary choices and their corresponding randomization probabilities. Recall that subjects faced 13 binary choices. Subjects may prefer the lottery x over y in some choices and prefer y over x in others. If subjects are confident about their choices, they may switch from preferring x to preferring y at one value of y (one switching point). However, when subjects are not fully confident about their choices, they may switch between x to y multiple times (multiple switching points). Among the 205 subjects who participated in the experiment, 140 switched multiple times for at least one of the four lotteries.[13]

Due to the prevalence of multiple switching points, we study the switching decisions of each subject at two levels of sure payments: $\underline{y}$ is the highest sure payment amount at and below which subjects consistently preferred x over y; $\bar{y}$ is the lowest sure payment amount at and above which subjects consistently chose y over x. We henceforth refer values of y between $\underline{y}$ and $\bar{y}$ as the switching range. Because utility of x and utility of y are the closest within the switching range, we postulate that subjects were less likely to be fully confident about their choices when y was between $\underline{y}$ and $\bar{y}$ and were hence more likely to randomize between x and y facing these values compared to other values of y. Result 3 summarizes this result.

Result 3. *Subjects were less confident about their choices within the switching range, and they chose a randomization probability around 0.5 when they were least confident.*

Support: Across lotteries and experience sampling conditions, the median confidence statements were "Probably x" at $\underline{y}$ and "Probably y" at $\bar{y}$, with "Unsure" selected for 20% of values of y within the switching range. 73% of all confidence statements within the switching range were "Probably x," "Unsure," or "Probably y," compared to 45% of confidence statements outside the switching range. These confidence reports support our hypothesis that subjects were less confident about their choices within the switching range.

Examining the randomization probabilities at $\underline{y}$ and $\bar{y}$, we find that across treatments subjects assigned a median randomization probability of 0.62 to x (0.38 to y) at $\underline{y}$, and a

[13] Specifically, there were 36 subjects who switched multiple times for the simple lottery, 50 for the complex lottery, 56 for the loss lottery, and 98 for the mixed lottery. In the no sampling condition and sampling condition, 71 and 69 subjects switched multiple times, respectively.

median randomization probability of 0.39 to x (0.61 to y) at $\bar{y}$. In other words, subjects were more likely to choose a randomization probability which lies between 0.39 to 0.62 for x for a choice they were less confident about. The median randomization probability for all the choices that fall within the switching range, $\underline{y}$ to $\bar{y}$, is 0.5.

This result is consistent with the example in Section 2.2, in which individuals choose $\lambda^* = 0.5$ when the decision utilities from x and y were close. This result holds for all lotteries and experimental conditions. Table 5 in the Appendix reports subjects' confidence statements and randomization probabilities corresponding to the switching ranges for the four lotteries and the experience sampling conditions.

Consistent with the above result, Figure 5 shows that higher proportions of subjects randomize and choose "Probably x," "Unsure," and "Probably y" around the switching choices than for y values further away from the switching range. In addition, the proportion of subjects choosing a randomization probability between 0.15 and 0.85 closely matches the proportion of subjects choosing "Probably x," "Unsure," or "Probably y" (see Table 6 in the Appendix for the exact proportion at each value of y). This result provides further support for Result 2 to interpret randomization probability as the probabilistic confidence of choices.

There is growing interest in finding ways to elicit confidence intervals, that is, the range of values between which subjects do not feel fully confident about their choices (Cohen et al., 1987; Butler and Loomes, 2011; Cubitt et al., 2015; Enke and Graeber, 2019). These studies have relied mostly on self-reported confidence statements to determine these confidence intervals. We show that randomization probabilities offer an alternative approach to determining the confidence intervals. Result 4 summarizes the findings.

Result 4. *On average, the confidence intervals defined by confidence statements are closely related to the confidence intervals defined by randomization probabilities of* $0.15 < \lambda < 0.85$.

Support: There are two ways to define the confidence intervals: using the self-reported confidence statements and defining them as the interval for which subjects do not consistently state "Surely x or y", or using the randomization probability and defining it as the

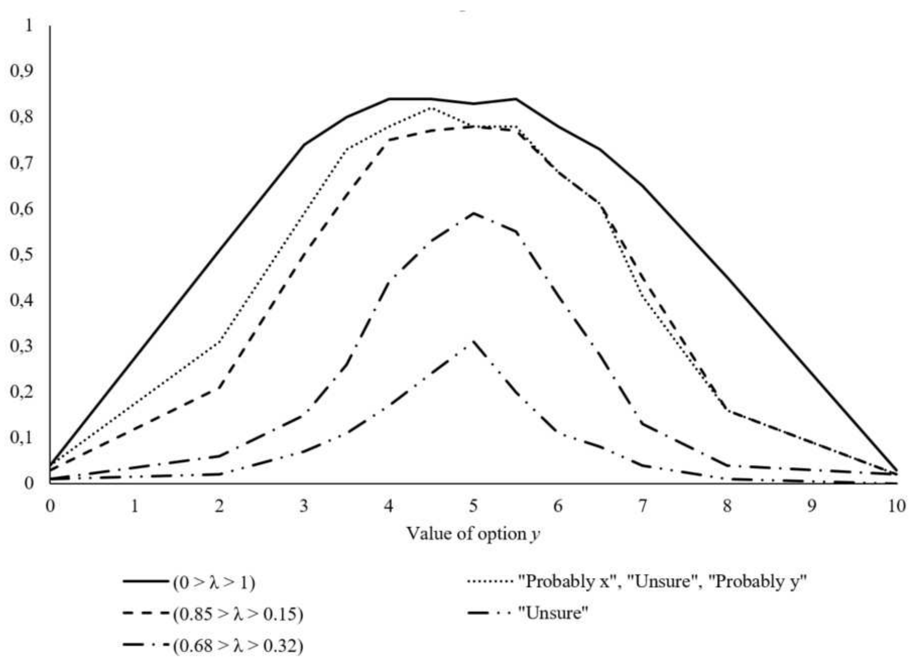


Figure 5: Proportion of subjects' decisions corresponding to different levels of randomization and confidence statements at the values of y.

interval for which subjects give randomization probabilities not consistently outside the given threshold ($0 < \lambda < 1$ or $0.15 < \lambda < 0.85$). Before we report the results, recall that the y values of the loss lottery and mixed lottery are simply those of the simple lottery and complex lottery subtracted by 10 euro and 5 euro, respectively. To make the confidence intervals across lotteries comparable, we add 10 and 5 to the confidence intervals of the loss lottery and the mixed lottery, respectively. Taking $0 < \lambda < 1$ to define the confidence interval based on randomization probabilities, we find a larger interval than the confidence interval defined by the confidence statements. On aggregate, the confidence interval defined by randomization probabilities ($0 < \lambda < 1$) ranges from 2.0 to 7.0. By contrast, the confidence interval corresponding to the confidence statements "Probably x," "Unsure," and "Probably y" ranges from 3.0 to 6.5. This is not surprising, because as we reported above, the confidence statements of "Probably x" and "Probably y" correspond to less extreme randomization probabilities (See result 2). Once we restrict the randomization probabilities to the same values defining the confidence thresholds in Vanberg (2008), that is, $0.15 < \lambda < 0.85$, we obtain a confidence interval that ranges from 3.0 to 6.5, similar to the confidence interval defined by the confidence statements. Looking at the confidence intervals for each subject, we see that the median interval defined by the confidence statements is 4.5. For the confidence intervals defined by the randomization probabilities ($0 < \lambda < 1$ and $0.15 < \lambda < 0.85$), we find a median confidence interval of 6 and 4.5, respectively. A two-sided Wilcoxon signed rank sum test of the confidence intervals derived from confidence statements and randomization probabilities shows a significant difference between the confidence statements and randomization probabilities with $0 < \lambda < 1$. Even though the median ranges are the same, we find an economically weak (mean values of 4.73 for confidence statements and 4.27 for randomization probabilities) but statistically significant difference between the confidence intervals derived from the confidence statements and the randomization probabilities with $0.15 < \lambda < 0.85$ ($p < 0.01$). This is probably due to the high power from the large sample and the paired test. The similarity between the two confidence intervals derived from confidence statements and randomization probabilities with $0.15 < \lambda < 0.85$ is robust to the type of lotteries and experience sampling conditions. Table 7 in the Appendix reports the confidence intervals defined by randomization probabilities and confidence statements for the four lotteries and the experience sampling conditions separately.

4.2 Alternative interpretations of randomization

We have interpreted randomization as a lack of decision confidence in the face of preference uncertainty. Our theoretical analysis provides an explicit link between the randomization probability and decision confidence. Our experimental results show a systematic relationship between randomization probabilities and confidence statements. However, subjects might randomize for reasons other than decision confidence, and the above relationship could be merely a coincidence. In this subsection, we consider interpretations of randomization other than decision confidence. Overall, these alternative interpretations suggest either randomization at most once or no systematic relationship between randomization probabilities and confidence statements.

The first alternative interpretation of randomization is indifference. In expected utility theory, individuals can choose any randomization probability when they consider two op-

Randomization Interval	The number of subjects who chose randomization			
	0 times	1 time	2 times or more	3 times or more
$0 < \lambda < 1$	2	1	202	196
$0.15 < \lambda < 0.85$	3	1	201	196
$0.32 < \lambda < 0.68$	9	3	193	188

Table 4: The distribution of subjects who chose $0 < \lambda < 1$, $0.15 < \lambda < 0.85$, and $0.32 < \lambda < 0.68$ zero times, one time, two times or more, and three times or more across the four lotteries and two experimental conditions. The results for the two conditions and four lotteries separately are in Table 8 in the Appendix.

tions indifferent, and indifference arises only in one pair of x and y. Our next result shows that most of our subjects randomize more than once.

Result 5. *Inconsistent with expected utility theory, the majority of subjects randomized over x and y at least over two values of y.*

Support: Table 4 shows that the majority (over 99% on aggregate) of subjects assigned a randomization probability strictly within 0 and 1.0 to y two times or more in at least one lottery or experimental condition. Table 8 in the Appendix reports the results for the two conditions and four lotteries separately. As we can see, the vast majority (more than 86%) of subjects assigned a randomization probability strictly within 0 and 1 to y two times or more in all lotteries and conditions. This is inconsistent with the prediction of standard economic models, according to which subjects randomize at most once.

The second possibility is that those randomization choices were random errors. If so, there should be no relationship between randomization probabilities and confidence statements, which is inconsistent with our results in the previous section. Furthermore, Result 6 below shows that despite the random sequence in which the different values of y were presented, the randomization probabilities of choosing x decreased monotonically with the value of y. In addition, there is evidence that randomization probabilities and confidence statements responded to experimental treatments.

Result 6. *Consistent with our theoretical analysis, the randomization probability of choosing x decreased with the value of y; and there is evidence that the confidence intervals in the complex lottery, the loss lottery, and the mixed lottery were wider than those in the simple lottery.*

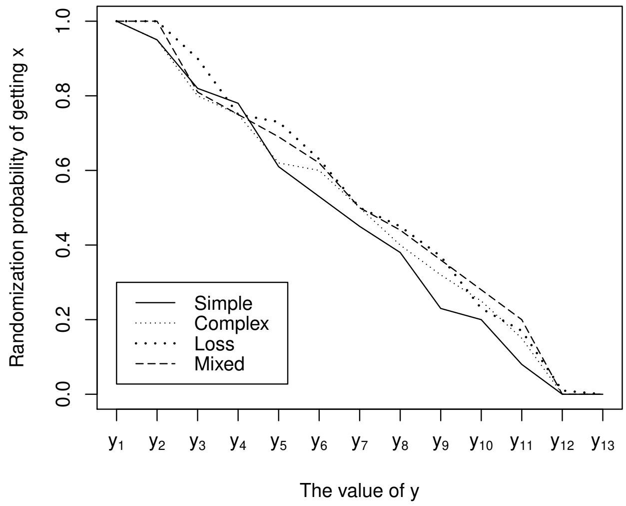


Figure 6: Median randomization probabilities that subjects assign to x, as a function of the value of y, for the four lotteries across the two conditions. For the simple and complex lottery, y takes the value of $Y \in \{0, 2, 3, 3.5, 4, 4.5, 5.0, 5.5, 6, 6.5, 7, 8, 10\}$, for the loss lottery y is each value in Y minus 10, and for the mixed lottery y is each value in Y minus 5. The figure for each condition separately can be found in the Appendix.

Support: Figure 6 reports the median randomization probability of choosing x in relation to y in the four lottery treatments. As we can see, despite the random sequence in the randomized choices, subjects' λ^* decreased monotonically with the value of y in all treatments. This result is consistent with Equation 2 and Figure 3.

To investigate the treatment differences across the experience sampling conditions and lotteries, we compare the confidence intervals obtained from the self-reported confidence statements and randomization probabilities ($0 < \lambda < 1$ or $0.15 < \lambda < 0.85$) across treatments. We hypothesized that subjects experience less decision confidence for the complex lottery, the loss lottery, and the mixed lottery compared to the simple lottery. Using the confidence intervals defined by confidence statements, we find some evidence that the confidence intervals of the complex lottery, the loss lottery, and the mixed lottery were wider than those of the simple lottery (one-sided Wilcoxon signed rank sum tests, $p = 0.06$ for the comparison of the complex versus the simple lottery, $p < 0.01$ for the comparison of the loss lottery and the mixed lottery versus the simple lottery). Using the confidence intervals defined by randomization probabilities, we find that confidence intervals of the complex lottery are wider than those of the simple lottery for both $0 < \lambda < 1$ and $0.15 < \lambda < 0.85$ (one-sided Wilcoxon signed rank sum test, $p < 0.05$ and $p < 0.01$, respectively). However, we find no significant differences for the loss lottery and the mixed lottery compared to the simple lottery (one-sided Wilcoxon signed rank sum test, $p > 0.10$). In light of the numerical simulations in Appendix B, we think risk attitude in the loss domain may distort confidence-driven randomization. We further hypothesized that subjects who were allowed to gain experience with the lottery outcomes perceive more decision confidence than those who did not get to sample outcomes. However, we do not find any significant differences between the sampling and no sampling conditions for any of the confidence intervals (one-sided Wilcoxon rank sum test, $p > 0.10$ for all tests). In Table 9 in the Appendix, the comparisons for all treatments and conditions are summarized.

The third interpretation of deliberate randomization can be found in nonexpected utility theories, in particular nonlinear probability weighting. Allowing for sufficient flexibility, subjects could randomize multiple times when they face a sequence of randomization decisions between x and y. To examine whether nonlinear probability could account for the randomization pattern in our experiment, we use the most popular parametric forms in the literature. If nonlinear probability weighting is responsible for randomization, as positive theories that aim to capture individuals' actual choices, those forms should be able to account for the randomization pattern of a large proportion of subjects. However, our numerical calculation in the Appendix B suggests that probability weighting is unlikely to be the driving factor for subjects' randomization. In fact, individuals behaving according to the popular parametric probability weighting functions should not randomize in the simple lottery and in the complex lottery. They may randomize when losses are involved, due to the convex value function in the loss domain. In that case, they are more likely to randomize in the loss lottery than in the mixed lottery. These predictions are inconsistent with our experimental results.

Result 7. *Nonlinear probability weighting is unlikely the main reason for randomization.*

Support: see the numerical calculations in the Appendix B.

5 Conclusion

We have shown in this study that letting individuals assign randomization probabilities according to which they receive each option is an incentive-compatible way to elicit decision confidence. We show the link between randomization probability and decision confidence theoretically in a framework extended from Cerreia-Vioglio et al. (2015) and Klibanoff et al. (2005) and demonstrate this relationship empirically through an experiment.

Our experimental results provide strong evidence that one could interpret the randomization probability for an option as the probabilistic confidence of choosing that option. We find that the majority of subjects randomize frequently, and the randomization pattern is consistent with our theoretical analysis. We further find that randomization probabilities are highly correlated and vary systematically with self-reported qualitative confidence statements in our experiment, with high randomization probabilities for options associated with statements indicating higher confidence. Our further examination of alternative interpretations of randomization suggests that indifference, errors, and nonlinear probability weighting are unlikely to be driving factors. Overall, our results suggest that decision confidence can be meaningfully and accurately inferred from randomization probability.

References

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

Agranov, M. and Ortoleva, P. (2020). Ranges of preferences and randomization. Working paper, California Institute of Technology.

Bewley, T. F. (2002). Knightian decision theory: Part i. *Decisions in Economics and Finance*, 25:79–110.

Bradbury, M. A. S., Hens, T., and Zeisberger, S. (2014). Improving Investment Decisions with Simulated Experience*. *Review of Finance*, 19(3):1019–1052.

Busemeyer, J. and Townsend, J. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100(3):432–459.

Butler, D., Isoni, A., Loomes, G., and Tsutsui, K. (2014). Beyond choice: investigating the sensitivity and validity of measures of strength of preference. *Experimental Economics*, 17(4):537–563.

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

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

Cerreia-Vioglio, S., Dillenberger, D., and ortoleva, P. (2015). Cautious Expected Utility and the Certainty Effect. *Econometrica*, 83:693–728.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., and Riella, G. (2019). Deliberately stochastic. *American Economic Review*, 109(7):2425–45.

Cohen, M., Jaffray, J.-Y., and 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., Navarro-Martinez, D., and Starmer, C. (2015). On preference imprecision. *Journal of Risk and Uncertainty*, 50(1):1–34.

Dubourg, W. R., Jones-Lee, M. W., and Loomes, G. (1994). Imprecise preferences and the wtp-wta disparity. *Journal of Risk and Uncertainty*, 9(2):115–33.

Dubourg, W. R., Jones-Lee, M. W., and Loomes, G. (1997). Imprecise Preferences and Survey Design in Contingent Valuation. *Economica*, 64(256):681–702.

Dubra, J., Maccheroni, F., and Ok, E. A. (2004). Expected utility theory without the completeness axiom. *Journal of Economic Theory*, 115(1):118–133.

Eliaz, K. and Ok, E. A. (2006). Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences. *Games and Economic Behavior*, 56(1):61–86.

Enke, B. and Graeber, T. (2019). Cognitive uncertainty. Working paper, Harvard University.

Gonzalez, C., Dana, J., Koshino, H., and Just, M. (2005). The framing effect and risky decisions: Examining cognitive functions with fmri. *Journal of Economic Psychology*, 26(1):1 – 20.

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

Hertwig, R., Barron, G., Weber, E. U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8):534–539. PMID: 15270998.

Huck, S. and Weizsäcker, G. (1999). Risk, complexity, and deviations from expected-value maximization: Results of a lottery choice experiment. *Journal of Economic Psychology*, 20(6):699 – 715.

Iyengar, S., Huberman, G., and Jiang, W. (2004). *How much choice is too much: determinants of individual contributions in 401K retirement plans.* in O.S. Mitchell and S. Utkus (Ed.), Pension Design and Structure: New Lessons from Behavioral Finance. Oxford University Press, Oxford (2004), pp. 83-95.

Iyengar, S. S. and Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, 94(7-8):530–539.

Iyengar, S. S. and Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6):995–1006.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Levitt, S. D. (2020). Heads or Tails: The Impact of a Coin Toss on Major Life Decisions and Subsequent Happiness. *The Review of Economic Studies*. rdaa016.

List, J. A. (2003). Does Market Experience Eliminate Market Anomalies?*. *The Quarterly Journal of Economics*, 118(1):41–71.

List, J. A. (2011). Does market experience eliminate market anomalies? the case of exogenous market experience. *American Economic Review*, 101(3):313–17.

Luce, M. F., Payne, J. W., and Bettman, J. R. (1999). Emotional trade-off difficulty and choice. *Journal of Marketing Research*, 36(2):143–159.

Moffatt, P., Sitzia, S., and Zizzo, D. (2015). Heterogeneity in preferences towards complexity. *Journal of Risk and Uncertainty*, 51:147–170.

Ok, E. A., Ortoleva, P., and Riella, G. (2012). Incomplete Preferences Under Uncertainty: Indecisiveness in Beliefs versus Tastes. *Econometrica*, 80(4):1791–1808.

Pabst, S., Brand, M., and Wolf, O. (2013). Stress effects on framed decisions: there are differences for gains and losses. *Front Behav Neurosci*, 7(142).

Qiu, J. and Ong, Q. (2017). Indifference or indecisiveness: a strict discrimination. MPRA Paper 81790, University Library of Munich, Germany.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343.

Segal, U. (1990). Two-stage lotteries without the reduction axiom. *Econometrica*, 58(2):349–377.

Sonsino, D., Benzion, U., and Mador, G. (2002). The Complexity Effects on Choice with Uncertainty — Experimental Evidence. *Economic Journal*, 112(482):936–965.

Tversky, A. and Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Vanberg, C. (2008). Why Do People Keep Their Promises? An Experimental Test of Two Explanations -super-1. *Econometrica*, 76(6):1467–1480.

Appendices

A Derivation of the optimal λ^*

Taking first order derivative of the optimisation equation gives:[14]

$$\frac{dV\left[\lambda x+(1-\lambda)y\right]}{d\lambda}=\sum_T \phi'\left[\lambda EU_t(x)+(1-\lambda)u(y)\right]\times\left[EU_t(x)-u(y)\right]d\pi(t)=0.$$

In some cases, preferences of x over y can be straightforward, e.g., when options can be ordered by some dominance rules. For example, when options x and y are risky lotteries and option x first degree stochastically dominates option y, it seems natural that individuals have $EU_t(x)>u(y)$, for $\forall t\in T$. Since $\phi'\left[\lambda EU_t(x)+(1-\lambda)u(y)\right]>0$, this leads to a positive first order condition and, hence, $\lambda=1$. Unfortunately, two options cannot in general be ordered via simple dominance rules. In such situations the choice of λ would give insights on decision confidence between x and y.

Note that $EU_t(x)$ is a random variable governed by the probability distribution π. Let $X=EU_t(x)$, and $\Delta_t=X-u(y)$. With these notations, we have

$$\phi'\left[\lambda EU_t(x)+(1-\lambda)u(y)\right]=\phi'\left[u(y)+\lambda\Delta_t\right].$$

We are mostly interested in the scenario where the individual finds choices between x and y difficult, i.e., when the two options are close. Specifically, we are interested in those situations where Δ_t is small relative to X and $u(y)$. When this is the case, we can use the

[14]The second-order derivative is

$$\frac{d^2V\left[\lambda x+(1-\lambda y)\right]}{d\lambda^2}=\sum_T \phi''\left[\lambda EU_t(x)+(1-\lambda)u(y)\right]\times\left[EU_t(x)-u(y)\right]^2d\pi(t).$$

Since $\phi(\cdot)$ is concave, $\phi''(\cdot)$ is negative. We are interested in situations where options x and y are not the same, i.e., $EU_t(x)\neq u(y)$ for some $t\in T$. Together we have $\phi''\left[\lambda EU_t(x)+(1-\lambda)u(y)\right]\times\left[EU_t(x)-u(y)\right]^2\leq 0$, and the inequality is strict for some $t\in T$. Consequently, $\frac{d^2V[\lambda x+(1-\lambda y)]}{d\lambda^2}=\sum_T\phi''\left[\lambda EU_t(x)+(1-\lambda)u(y)\right]\times\left[EU_t(x)-u(y)\right]^2d\pi(t)<0$. This ensures we are indeed seeking for the maximum.

Taylor expansion and obtain

$$\phi'\left[u(y)+\lambda\Delta_t\right]=\phi'(u(y))+\phi''(u(y))\lambda\Delta_t+O\left(\lambda\Delta_t\right)\approx\phi'(u(y))+\phi''(u(y))\lambda\Delta_t,$$

where $O\left(\lambda\Delta_t\right)$ is the sum of the terms that have $\lambda\Delta_t$ with a power of two or higher. The above first order condition can then be written as

$$\begin{aligned}\tfrac{dV[\lambda x+(1-\lambda)y]}{d\lambda} &= \textstyle\sum_T \phi'\left[u(y)+\lambda\Delta_t\right]\Delta_t d\pi(t), \\ &\approx \textstyle\sum_T \left[\phi'(u(y))+\phi''(u(y))\lambda\Delta_t\right]\Delta_t d\pi(t) \\ &= E_t\left[\phi'(u(y))\Delta_t\right]+\lambda E_t\left[\phi''(u(y))\Delta_t^2\right]=0,\end{aligned}$$

where $E_t(\cdot)$ is the expectation operator with respect to the distribution π. Solving for λ, and we have:

$$\lambda^*\approx min\left\{max\left\{0,\frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}}\times\frac{\Delta_u}{\sigma_x^2-\Delta_u^2}\right\},1\right\}\approx min\left\{max\left\{0,\frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}}\times\frac{\Delta_u}{\sigma_x^2}\right\},1\right\},$$

where $\Delta_u = E_t\left[EU_t(x)\right]-u(y)$ is the (expected) utility difference of x and y, $\sigma_x^2 = E_t\left[EU_t(x)-E_t(EU_t(x))\right]^2$ is the standard deviation of $EU_t(x)$.

B Predictions under CPT and RDU.

Predicted randomization pattern under CPT and RDU: under popular parametric forms and parameters, subjects do not randomize in the simple lottery and the complex lottery. They may randomize when losses are involved, due to the convex value function in the loss domain, and they are more likely to randomize in the loss lottery than in the mixed lottery.

Support: When x is a lottery and y is a sure payment, randomization over x and y creates a compound lottery. Let $v(\cdot)$ denote the value function, $V(\cdot)$ denote the prospect value of a lottery, $w(\cdot)$ denote the probability weighting function. If we assume compound independence (Segal, 1990), an axiom weaker than reduction of compound lotteries, it can be easily shown that individuals have no strict incentive to randomize. To see this, note that under compound independence subjects evaluate the compound lottery by first evaluating the simple lottery x and obtain the prospect value $V(x)$. Then, subjects evaluate the simple lottery $(\lambda, V(x); 1-\lambda, y)$ as:

$$V(\lambda, V(x); 1-\lambda, y) = w(\lambda)V(x) + [1-w(\lambda)]V(y) = V(y) + w(\lambda)\left[V(x) - V(y)\right],$$

where $V(x)$ $(V(y))$ is the prospect value of x (y, respectively).

If we assume reduction of compound lotteries, in principle it is possible for individuals to have a preference to randomize. For example, when the probability weighting function is concave. To examine whether non-linear probability could account for the randomization pattern in our experiment, we use the most popular parametric forms in the literature and adopt the median estimated parameters to calculate the optimal randomization probability between the lottery and the sure payment. If non-linear probability weighting is responsible for randomization, median estimations which intend to capture average group behavior should account for the randomization pattern of a large proportion of subjects. Our numerical calculations below show that, with most empirical parameters of the probability weighting function, subjects do not randomize in the simple lottery and the complex lottery. They may randomize when losses are involved, due to the convex value function in the loss domain, and they are more likely to randomize in the loss lottery than in the mixed lottery. However, these predictions are clearly inconsistent with our experimental results. Table 8 suggests that subjects did not randomize more frequently in the loss lottery and mixed lottery than in the simple lottery and the complex lottery. Furthermore, Figure 8 suggests that the randomization probabilities that subjects chosen were comparable across lotteries.

In the calculation below we assume the value function:

$$V_(x) = \begin{cases} x^{0.88}, & x \geqslant 0 \\ -2(-x)^{0.88}, & x < 0 \end{cases}$$

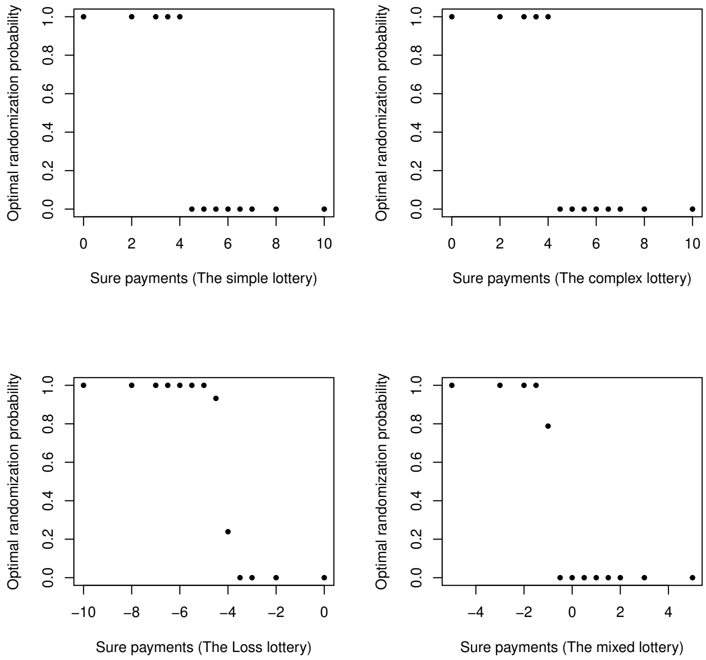


Figure 7: Calculations of optimal randomization probabilities. In the calculation we assume the value function: $v(x) = x^{0.88}, x \geq 0$, and $= -2(-x)^{0.88}, x < 0$. The probability weighting function: $w(p)^{+} = w(p)^{-} = \frac{0.88p^{0.65}}{[0.88p^{0.65}+(1-p)^{0.65}]^{1/0.65}}$. We tried a number of forms and different parameter values. The results are qualitatively the same under all popular specifications.

The probability weighting function:

$$\omega_(p)^{+} = \omega_(p)^{-} = \frac{0.88p^{0.65}}{[0.88p^{0.65} + (1-p)^{0.65}]^{\frac{1}{0.65}}}.$$

We tried a number of forms and different parameter values. The results are qualitatively the same under all popular specifications. Figure 7 summarizes our calculations.

Consider first the simple lottery $x = (10, 0.5; 0, 0.5)$. When a subject assign randomization probability λ to x (and $1 - \lambda$ to y), the reduced compound lottery is: $L =$

$(0.5\lambda, 10; 1-\lambda, y; 0.5\lambda, 0)$. The decision utility of the lottery is:

$$\begin{aligned} V(L) &= w(0.5\lambda)v(10) + [w(1-0.5\lambda) - w(0.5\lambda)]\, v(y) + [1 - w(1-0.5\lambda)]\, v(0) \\ &= w(0.5\lambda)v(10) + [w(1-0.5\lambda) - w(0.5\lambda)]\, v(y), \end{aligned}$$

where the last equation obtains by making the standard assumption of $v(0) = 0$. For the complex lottery the compound lottery is $L = (0.2\lambda, 10; 0.3\lambda, 7.50; 1-\lambda, y; 0.3\lambda, 2.5; 0.2\lambda, 0.25)$ when $2.5 < y < 7.5$, and the decision utility is:

$$\begin{aligned} V(L) &= w(0.2\lambda)v(10) + [w(0.5\lambda) - w(0.2\lambda)]\, v(7.5) + [w(1-0.5\lambda) - w(0.5\lambda)]\, v(y) \\ &\quad + [w(1-0.2\lambda) - w(1-0.5\lambda)]\, v(2.5) + [1 - w(1-0.2\lambda)]\, v(0.25). \end{aligned}$$

For the loss lottery the reduced compound lottery is $L = (0.5\lambda, -10; 1-\lambda, y; 0.5\lambda, 0)$. The decision utility of the lottery is:

$$V(L) = w(0.5\lambda)v(-10) + [w(1-0.5\lambda) - w(0.5\lambda)]\, v(y).$$

For the mixed lottery the reduced compound lottery is $L = (0.5\lambda, 5; 1-\lambda, y; 0.5\lambda, -5)$, the decision utility of the lottery is:

$$V(L) = w(0.5\lambda)v(-5) + w(0.5\lambda)v(5) + [w(1-0.5\lambda) - w(0.5\lambda)]\, v(y).$$

C Tables and figures

Condition	Treatment	Behavior around the switching choice			
		Randomization probability		Confidence statements	
		$\underline{y}$	$\bar{y}$	$\underline{y}$	$\bar{y}$
Aggregate		0.62	0.39	Probably x	Probably y
No sampling	Simple	0.67	0.46	Probably x	Probably y
	Complex	0.63	0.43	Probably x	Probably y
	Loss	0.55	0.38	Probably x	Probably y
	Mixed	0.67	0.37	Probably x	Probably y
Sampling	Simple	0.60	0.40	Probably x	Probably y
	Complex	0.60	0.35	Probably x	Probably y
	Loss	0.58	0.40	Probably x	Probably y
	Mixed	0.69	0.30	Probably x	Probably y

Table 5: Behavior around the switching choice. The value $\underline{y}$ is the sure payment at and below which subjects consistently choose x and switch to y for the first time when y is just above $\underline{y}$; and the value $\bar{y}$ is the sure payment at and above which subjects consistently choose y and choose x for the last time when y is just below $\bar{y}$. The reported valued values are medians across subjects.

Value y	Randomization intervals			Confidence statements	
	(1>λ>0)	(0.85>λ>0.15)	(0.68>λ>0.32)	Probably or unsure	Unsure
0	0.04	0.03	0.01	0.04	0.01
2	0.51	0.21	0.06	0.31	0.02
3	0.74	0.50	0.15	0.59	0.07
3.50	0.80	0.63	0.26	0.73	0.11
4	0.84	0.75	0.44	0.78	0.17
4.50	0.84	0.77	0.53	0.82	0.24
5	0.83	0.78	0.59	0.78	0.31
5.50	0.84	0.77	0.55	0.78	0.20
6	0.78	0.68	0.41	0.68	0.11
6.50	0.73	0.61	0.28	0.61	0.08
7	0.65	0.45	0.13	0.41	0.04
8	0.45	0.16	0.04	0.16	0.01
10	0.03	0.02	0.02	0.02	0.00

Table 6: Proportion of decisions made that correspond to the different levels of randomization and confidence statements for each value of y. The values of y of the loss and mixed lottery can be obtained by subtracting 10 euro and 5 euro respectively.

Treatments		Mean switching value	Confidence intervals		
			Randomization		Confidence statements
			($0 < \lambda < 1.0$)	($0.15 < \lambda < 0.85$)	
No sampling	Simple	4.5	[2.0, 7.0]	[3.5, 6.5]	[3.0, 6.5]
	Complex	4.8	[2.0, 7.0]	[3.0, 6.5]	[3.0, 6.5]
	Loss +10	5.2	[3.0, 7.0]	[3.5, 7.0]	[3.0, 7.0]
	Mixed +5	5.0	[3.0, 7.0]	[3.0, 7.0]	[3.0, 7.0]
Sampling	Simple	4.6	[2.0, 7.0]	[3.0, 6.5]	[3.0, 6.5]
	Complex	4.8	[2.0, 7.0]	[3.0, 6.5]	[3.0, 6.5]
	Loss +10	5.5	[3.0, 8.0]	[3.5, 7.0]	[3.5, 7.0]
	Mixed +5	5.2	[3.0, 8.0]	[3.0, 7.0]	[3.0, 7.0]

Table 7: Mean value in the switching range and confidence intervals across treatments. The value in the switching range is the mean of the ys in the switching range. The confidence intervals are the median values of $\underline{y}$ and $\bar{y}$ in which subjects choose $0 < \lambda < 1$, $0.15 < \lambda < 0.85$, and in which subjects reports "Probably x", "unsure", and "Probably y". The values of y for the loss and mixed lottery have been adjusted by +10 and +5 respectively to reflect the same range as the simple and complex lottery.

Condition	Treatments	The number of subjects who chose $\lambda^* \in (0,1)$ for			
		0 time	1 time	2 times or more	3 times or more
No sampling	Simple	2	6	97	95
	Complex	4	1	100	95
	Loss	3	3	99	95
	Mixed	6	6	93	89
Sampling	Simple	6	1	93	89
	Complex	3	5	92	89
	Loss	6	4	90	89
	Mixed	9	5	86	83

Table 8: The distribution of subjects who chose $0 < \lambda < 1$ for zero time, for one time, for two times and more, and for three times and more for different lotteries and conditions.

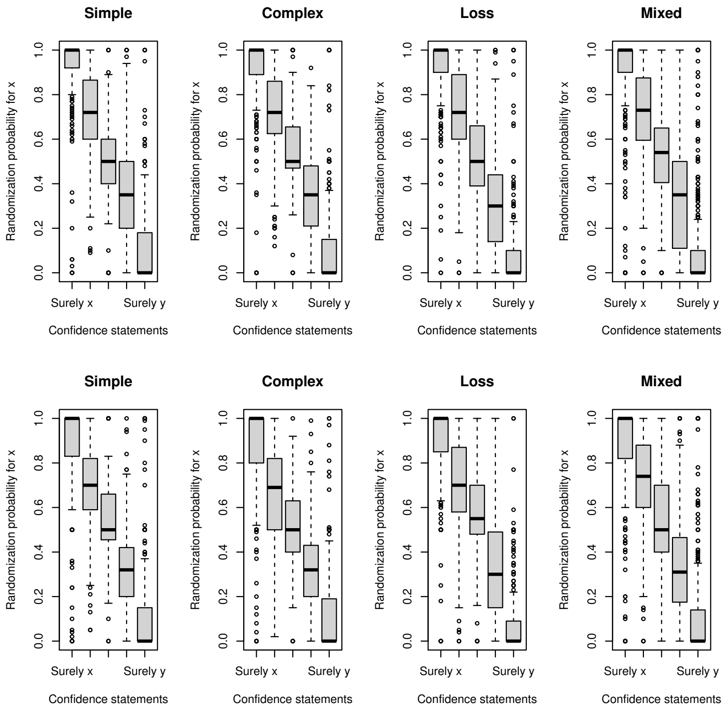


Figure 8: Boxplots of the randomization probability of choosing x, given each confidence statement. The thick line is median, the upper and lower bars are 2nd and 3rd quantiles, respectively. The upper panel is for the no sampling condition, and the lower panel is for the sampling condition.

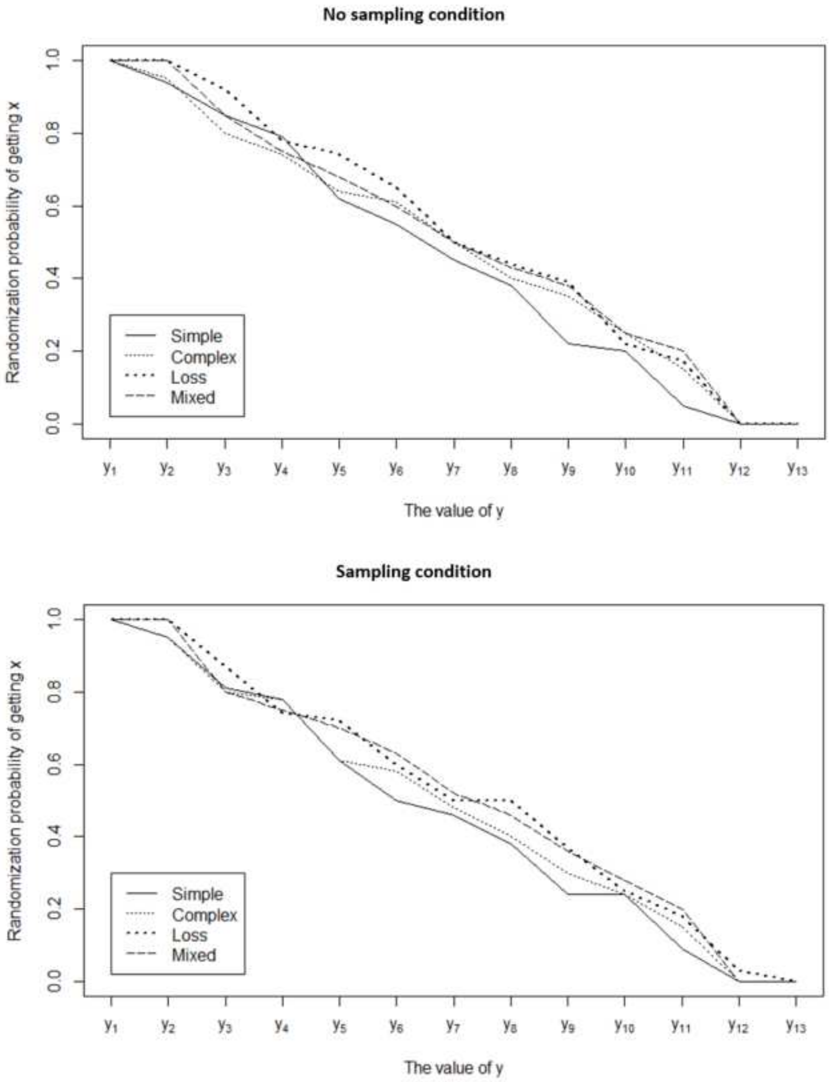


Figure 9: Median randomization probabilities that subjects assign to x, as a function of the value of y, for the four lotteries in the two conditions. For the simple and complex lottery, y takes the value of $Y \in \{0, 2, 3, 3.5, 4, 4.5, 5.0, 5.5, 6, 6.5, 7, 8, 10\}$, for the loss lottery y is each value in Y minus 10, and for the mixed lottery y is each value in Y minus 5.

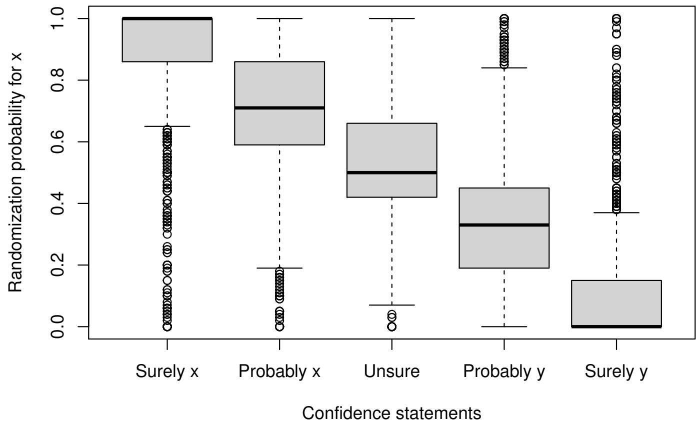


Figure 10: A boxplot of the randomization probability of choosing x across lotteries and conditions, given each confidence statement. The thick line is median, the upper and lower bars are 2nd and 3rd quantiles, respectively. More detailed figures for each lottery and condition are presented in Figure 8 in Appendix.

D Additional results: the subjectivity of confidence statements

In the introduction we point out that confidence statements are self-reports, and there is no universal understanding of confidence statements such as surely, probably, or unsure. Therefore, these statements could have different meaning for different individuals. To examine this hypothesis we look at the randomization probabilities at each level of confidence. Our following result summarizes this finding.

Result 8. *Given any confidence statement, there is substantial heterogeneity in the randomization probability among subjects.*

Support: Figure 10 gives the boxplot of the randomization probability of choosing x in treatment 1. Figure 8 in Appendix gives the boxplots of the other treatments. As we can clearly see, there is substantial variation of randomization probabilities across subjects at each confidence level. As we can also see from Table 3, the standard deviation of the randomization probabilities at each confidence level is substantial in all treatments, ranging from 0.11 to 0.20. The above results suggest that subjects indicating the same confidence level chose substantially different randomization probabilities.

Result 8 does not necessarily mean confidence statements are not informative. Results 2 to 4 suggest that confidence statements are surprisingly consistent. The problem is rather that each subject has her/his own subjective interpretation of confidence levels such as surely or probably, which might obscure the relationship between confidence statements and subjects' actual behavior. Moreover, the subjective interpretation of confidence levels might differ across tasks. For example, some subjects need to be 90% sure or higher to make the statement of surely, others might state surely with 80% confidence. Additionally, subjects might want to be highly certain for easy tasks to make the statement of surely, while they can be much more tolerant for difficult tasks. These issues make between-individual comparisons of confidence statements difficult,

E Experimental materials

Welcome

You are invited to participate in an experiment in which we examine how individuals make decisions. Your decisions in the experiment are about choices between different options. There are no right or wrong answers. The whole experiment will take approximately 20 minutes.

You will receive a participation fee of €1 for completing the survey. In addition you will receive monetary compensation up to €10 based on the decisions you make in the experiment. Specifically, one of the questions will be randomly selected. The decision you made in this question will determine your additional compensation.
You will receive the payment (the participation fee of €1 and the additional compensation) via bank transfer. For this we will ask your IBAN number. This information will only be used for payment, and will be permanently deleted afterwards.

Thank you for your participation!

Sincerely yours,

Associate Professor Dr. Jianying Qiu and PhD student Sara Arts
The Institute of Management Research
Radboud University Nijmegen.

Figure 11: Welcome screen of the experiment.

Voluntary participation

Your participation in this research is voluntary. This means that you can withdraw your participation and consent at any time during the survey, without giving a reason. All data we have collected from you will be deleted permanently. If you desire to withdraw, please simply close your internet browser. After completion of the survey it will not be possible to withdraw your data form the research.

What will happen to the data?

The research data we collect during this study will be used by scientists as part of data sets, articles and presentations. The anonymized research data is accessible to other scientists for a period of at least 10 years. When we share data with other researchers, these data cannot be traced back to you.

More information?

Should you want more information on this research study, please contact Sara Arts (email: s.arts@fm.ru.nl)

CONSENT:

Please select your choice below.

Checking "Agree" below indicates that:
you voluntarily agree to participate.

- ○ I agree with the provided information, and I would like to proceed to the survey
- ○ I do not agree with the above.

Figure 12: Informed consent.

The following questions are about the options below:

Option A:
Gain €4 with a chance of **50%**, and
lose €4 with a chance of **50%**.

Option B:
Lose or gain a **sure** amount (which varies across questions).

You will be asked to choose between the two options, and to describe how confident you feel about your choice.

Figure 13: Introduction screen of the mixed lottery in the no sampling condition.

The following questions are about the options below:

Option A:
Lose €9 with a chance of **50%,** and
lose €1 with a chance of **50%**.

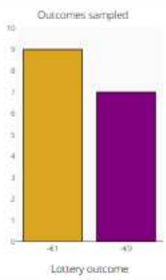
Option B:
Lose a **sure** amount (which is varied).

You will be asked to choose between the two options, and to describe how confident you feel about your choice.

Before you are asked to make the decisions we want to give you the opportunity to **experience the different outcomes of option A**. For this, you can click the button below. Each time you click the button a possible outcome of option A will be shown. You will get to sample **20** outcomes.

The outcomes that you obtain by clicking the button do not influence your payoff but are only presented to make you experience the possible outcomes.

To keep track of the sampled outcomes, they will be presented in a bargraph.



−€1

Sample

Figure 14: Introduction screen of the loss lottery in the sampling condition.

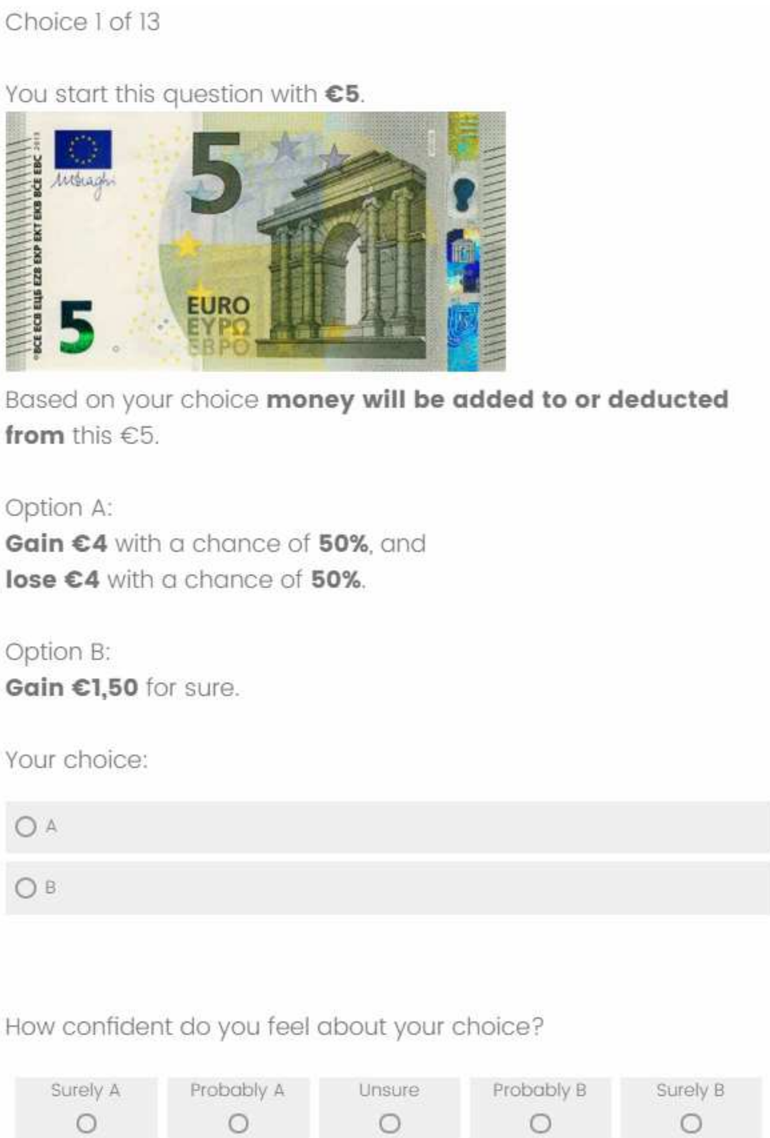


Figure 15: Example of the decision screen for the binary choices and confidence statements for the mixed lottery in the no sampling condition.

Figure 16: Example of the decision screen for the binary choices and confidence statements for the loss lottery in the sampling condition.

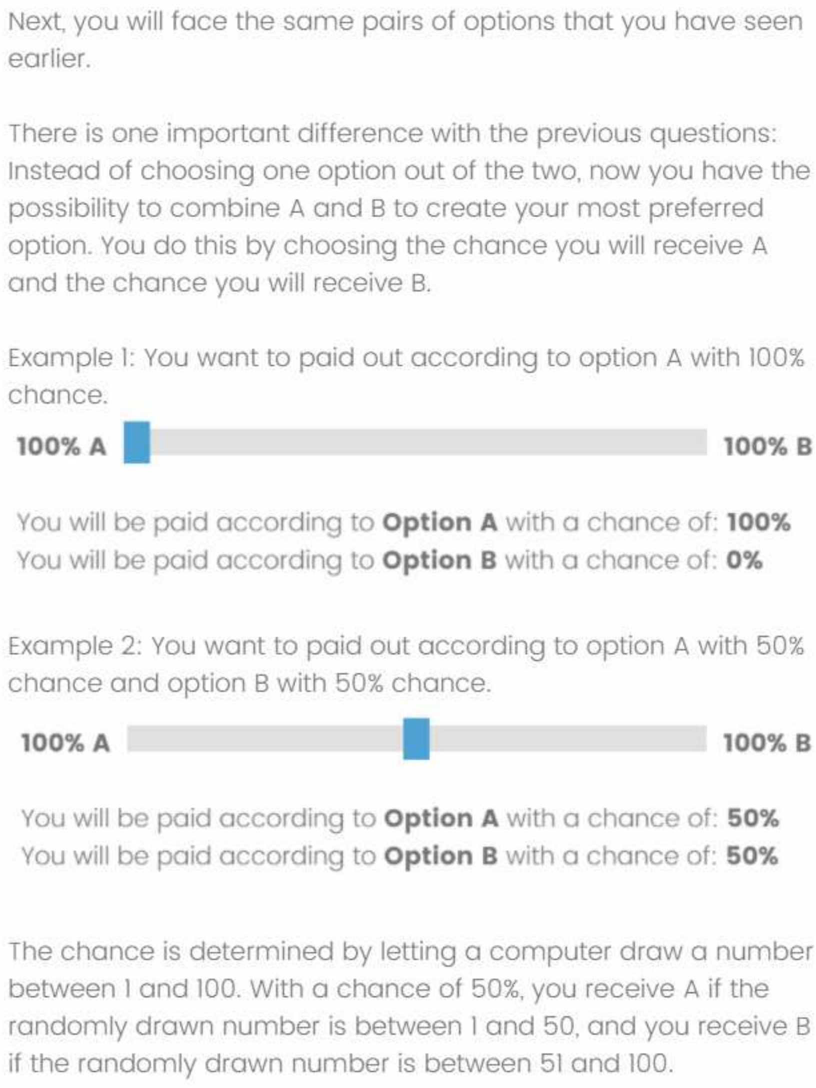


Figure 17: Explanation of randomization decisions.

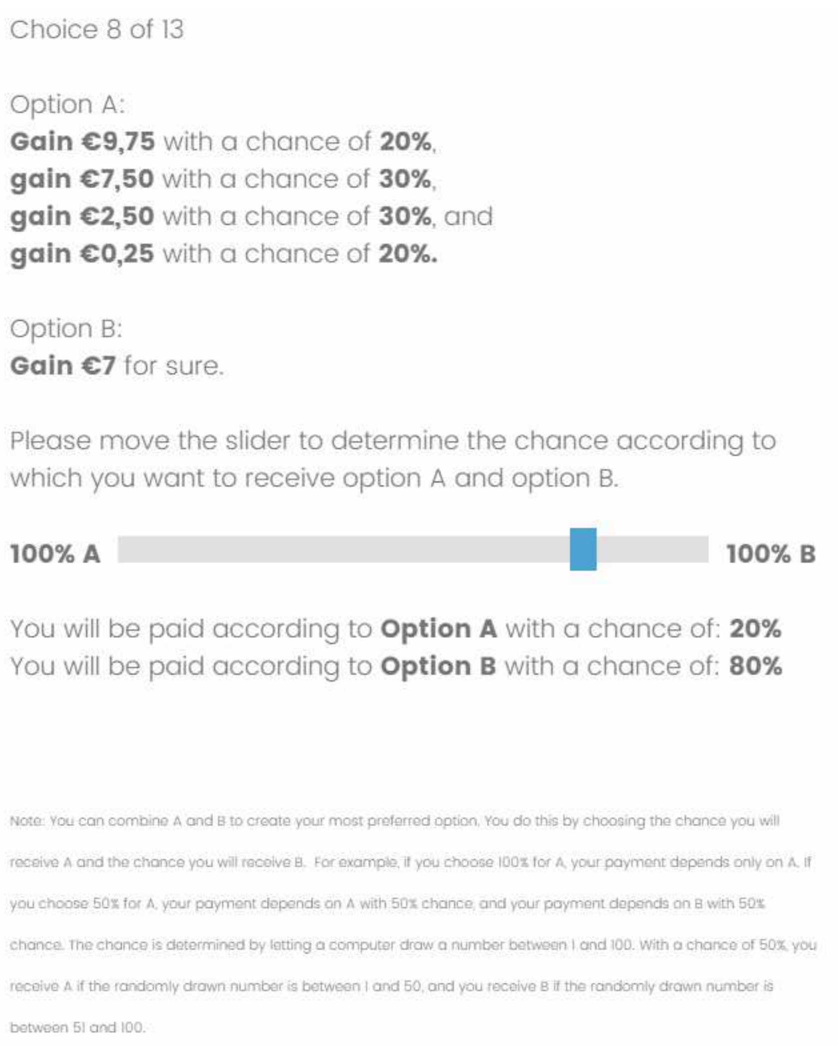


Figure 18: Example of the decision screen for the randomization choices for the complex lottery in the no sampling condition.

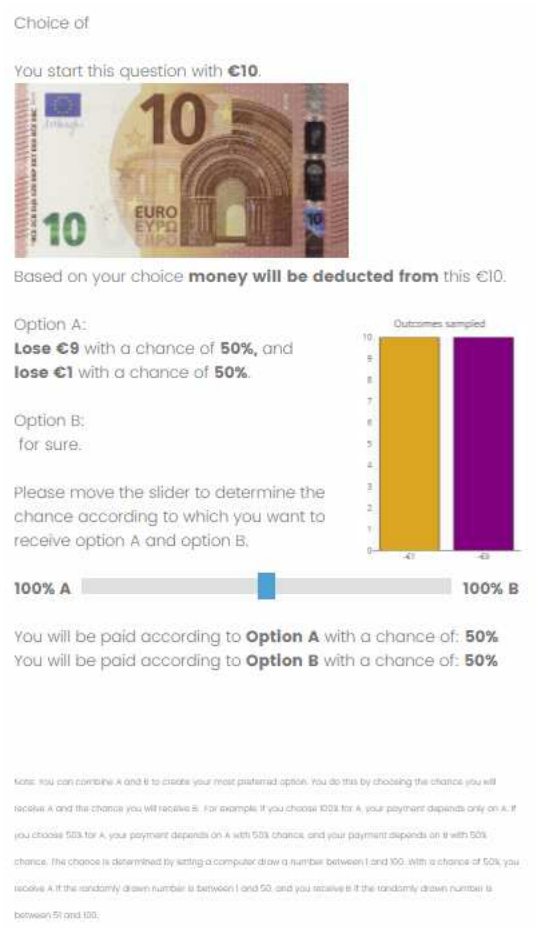
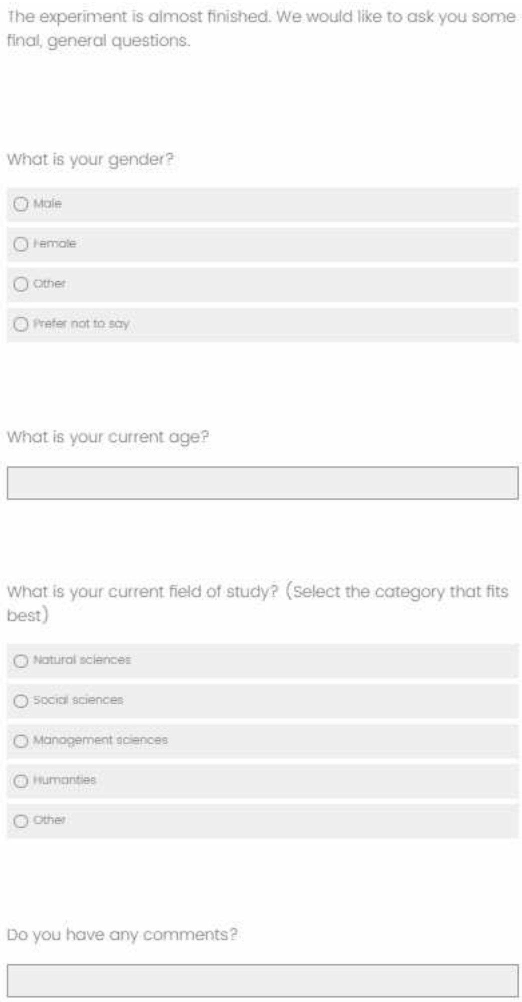


Figure 19: Example of the decision screen for the randomization choices for the loss lottery in the sampling condition.



The experiment is almost finished. We would like to ask you some final, general questions.

What is your gender?

- Male
- Female
- Other
- Prefer not to say

What is your current age?

What is your current field of study? (Select the category that fits best)

- Natural sciences
- Social sciences
- Management sciences
- Humanties
- Other

Do you have any comments?

Figure 20: Demographic questions asked at the end of the experiment.

Confidence statements				Wilcoxon tests	
Condition	Treatment	Median	Mean (sd)	Comparison to simple lottery	Comparison to sampling condition
No sampling	Simple	4.5	4.35 (1.69)	-	$p > 0.10$
	Complex	4.5	4.62 (1.91)	$p = 0.08$	$p > 0.10$
	Loss	4.5	4.82 (1.76)	$p < 0.01$	$p > 0.10$
	Mixed	5.0	5.18 (1.95)	$p < 0.01$	$p > 0.10$
Sampling	Simple	4.5	4.44 (2.10)	-	-
	Complex	4.5	4.49 (1.83)	$p > 0.10$	-
	Loss	5.0	4.85 (2.00)	$p < 0.05$	-
	Mixed	5.0	5.07 (2.21)	$p < 0.01$	-

Randomization probability: $0 < \lambda < 1$				Wilcoxon tests	
Condition	Treatment	Median	Mean (sd)	Comparison to simple lottery	Comparison to sampling condition
No sampling	Simple	5.5	5.20 (1.98)	-	$p > 0.10$
	Complex	5.5	5.41 (1.96)	$p = 0.07$	$p > 0.10$
	Loss	5.5	5.16 (2.00)	$p > 0.10$	$p > 0.10$
	Mixed	5.0	4.94 (2.32)	$p > 0.10$	$p > 0.10$
Sampling	Simple	6.0	5.39 (2.20)	-	-
	Complex	6.0	5.67 (2.19)	$p < 0.05$	-
	Loss	6.0	5.32 (2.23)	$p > 0.10$	-
	Mixed	6.0	5.47 (2.46)	$p > 0.10$	-

Randomization probability: $0.15 < \lambda < 0.85$				Wilcoxon tests	
Condition	Treatment	Median	Mean (sd)	Comparison to simple lottery	Comparison to sampling condition
No sampling	Simple	4.0	4.03 (1.93)	-	$p > 0.10$
	Complex	4.5	4.45 (1.77)	$p < 0.01$	$p > 0.10$
	Loss	4.0	4.05 (1.76)	$p > 0.10$	$p > 0.10$
	Mixed	4.5	4.08 (2.17)	$p > 0.10$	$p > 0.10$
Sampling	Simple	4.5	4.32 (1.89)	-	-
	Complex	4.5	4.60 (2.01)	$p = 0.06$	-
	Loss	4.5	4.14 (1.97)	$p > 0.10$	-
	Mixed	5.0	4.53 (2.24)	$p > 0.10$	-

Table 9: Comparison of confidence intervals defined by confidence statements and randomization probabilities across treatments and conditions. The median an mean values are based on the confidence intervals computed for each subject. The comparison of the lotteries is based on a one-sided Wilcoxon signed rank sum test. The comparison of the (no) sampling conditions is based on a one-sided Wilcoxon rank sum test.