August 2024

iRisk WORKING PAPER SERIES

2024-iRisk-03

Pay all subjects or pay only some? An experiment on decision-making under risk and ambiguity

Ilke Aydogan
IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie´ Management, F-59000 Lille, France; and iRisk Research Center on Risk and Uncertainty (i.aydogan@ieseg.fr)

Loïc Berger
CNRS, Univ. Lille, IESEG School of Management, UMR 9221 - LEM - Lille Economie´ Management, F-59000 Lille, France; iRisk Research Center on Risk and Uncertainty; RFF-CMCC European Institute on Economics and the Environment (EIEE), and Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy (l.berger@ieseg.fr)

Vincent Theroude
Université de Lorraine, Université de Strasbourg, CNRS, BETA, 54000, Nancy, France, (vincent.theroude@univ-lorraine.fr).

IÉSEG School of Management Lille Catholic University 3, rue de la Digue F-59000 Lille Tel: 33(0)3 20 54 58 92 www.ieseg.fr

Pay all subjects or pay only some? An experiment on decision-making under risk and ambiguity*

Ilke Aydogan† Loïc Berger‡ Vincent Théroude§

Abstract

We investigate the validity of a double random incentive system where only a subset of subjects is paid for one of their choices. By focusing on individual decision-making under risk and ambiguity, we show that using either a standard random incentive system, where all subjects are paid, or a double random system, where only 10% of subjects are paid, yields similar preference elicitation results. These findings suggest that adopting a double random incentive system could significantly reduce experimental costs and logistic efforts, thereby facilitating the exploration of individual decision-making in larger-scale and higher-stakes experiments.

Keywords: Experimental methodology, Payment methods, Incentives, Ambiguity elicitation

JEL Classification: C90; D81

*We are grateful to Nahed Eddai, Emmanuel Kemel, and Thibault Richard for helpful comments and discussions. Logistic support from the Anthropo-Lab (ETHICS EA 7446) for hosting our experimental sessions is kindly acknowledged. This research is supported by a grant from the French Agence Nationale de la Recherche (ANR-21-CE03-0018 ENDURA), the Region Hauts-de-France (2021.00865 CLAM), the I-SITE UNLE (project IBEBACC), and the European Union's Horizon Europe research and innovation programme under grant agreement (No 101056891 CAPABLE). Declarations of interest: none. Material for replication is available at osf.io/r5ws9/.

†IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Économie Management, F-59000 Lille, France; and iRisk Research Center on Risk and Uncertainty, France (i.aydogan@ieseg.fr).

‡Corresponding author. Univ. Lille, CNRS, IESEG School of Management, UMR 9221 - LEM - Lille Économie Management, F-59000 Lille, France; iRisk Research Center on Risk and Uncertainty, France; RFF-CMCC European Institute on Economics and the Environment (EIEE), and Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy (loic.berger@cnrs.fr).

§Université de Lorraine, Université de Strasbourg, CNRS, BETA, 54000, Nancy, France (vincent.theroude@univ-lorraine.fr).

1 Introduction

Data collected in experiments are relevant only if subjects provide considered and truthful responses to the questions they are asked. Consequently, offering real monetary rewards based on the outcomes of subjects' decisions is often regarded as the gold standard in experimental economics. The rationale is that right financial incentives motivate subjects to exert cognitive effort to achieve desirable outcomes, mirroring actual decision-making processes in real-world situations (Edwards 1953; Slovic 1969, see Moffatt et al., 2009 for a review).

The most straightforward way to encourage truthful responses is to provide incentives for single-choice experiments. However, since these allow only for between-subject comparisons, it is common to require participants to perform a series of choices within the same experiment. Two main incentive mechanisms are typically implemented in such experiments: Subjects are either paid based on the outcome of each of the decisions they make ('pay all') or based on a subset of their decisions (typically one randomly selected decision). The latter mechanism is known as the within-subject random incentive system (wRIS, see Table 1). This mechanism has several attractive features: it allows for higher stakes in each task, avoids wealth and portfolio effects, and eliminates hedging opportunities. From a theoretical standpoint, Azrieli et al. (2018) demonstrated that wRIS is essentially the only incentive-compatible mechanism if the subject's preference respects dominance.

Several papers have investigated the validity of the wRIS, by comparing the data it generates with data from single-choice experiments (Beattie and Loomes, 1997; Cubitt et al., 1998b; Brokesova et al., 2017) or multiple-choice experiments with 'pay all' systems (Laury, 2006). Overall, the results commonly show that the wRIS does not distort preferences, suggesting that subjects treat each choice in isolation (see Charness et al., 2016, for a review). As a result, the wRIS has become the most commonly used incentive mechanism in individual-choice laboratory experiments.

Table 1: Incentive mechanisms in experimental economics

		Pay all subjects?	
		Yes	No
Pay all decisions?	Yes	Pay all	bRIS
	No	wRIS	wbRIS

Alternatively, incentive mechanisms can be distinguished based on who is actually paid. The between-subject random incentive system (bRIS) operates on the same principles as the wRIS, except that, the randomization determines which subset of participants is selected for payment. The two systems can be combined, creating the *within-between*-subject random incentive system (wbRIS, see Table 1). Paying only a subset of participants has the potential to significantly decrease logistic efforts without sacrificing the benefits of a proper incentive system. However, it has been argued that a fundamental psychological difference may exist between knowing you will surely be paid based on (one of) your choices and having only a chance of receiving payment (Charness et al., 2016). Thus, whether isolation also holds or not when randomization applies to who is paid remains an open question. Specifically, if subjects treat the entire experiment as a meta-lottery and consider only the *expected* payoffs, discounting the values at stake by the probability of being selected for payment, they may act differently when only a subset of subjects is paid compared to when every subject is paid. However, if subjects perceive each choice situation as isolated and consider the face values of the stakes, then the between-subject randomization should introduce no distortion.

This paper investigates the validity of the between-subject payment system by comparing decisions obtained under a standard wRIS with those obtained using a wbRIS, combining both *within-* and *between*-subject randomization. Because decision-making under uncertainty is central to modern decision theory, and ambiguity aversion is one of the most extensively studied empirical phenomena in behavioral economics (Wakker, 2010; Trautmann and van de Kuilen, 2015), our study focuses explicitly on eliciting risk and ambiguity preferences.

Our data show that paying all subjects or only a subset of them leads to the same elicited preferences. Specifically, we find no differences in attitudes towards risk and ambiguity across treatments using either a standard wRIS or a wbRIS, where only 10% of the subjects are paid. Hence, we provide additional evidence supporting the isolation effect in wbRIS and broaden the applicability of between-subject payments to experiments on risk and ambiguity. With the recent rise of large scale data collection using online platforms (e.g., Baillon et al., 2022; Fanghella et al., 2023; Burgstaller and Pfeil, 2024; Parra, 2024), these results have important implications for experimental designs in economics and psychology. In particular, they allow for studying decision-making in a wider variety of contexts (e.g., large-scale or high-stake experiments) than what is typically feasible when all participants must be paid.

2 Related literature

Previous experimental evidence on the impact of paying only a subset of participants is limited and mixed. In the context of social preferences, the early work of Bolle (1990) showed no behavioral differences in an ultimatum game whether all participants or only 10% of them received payment for one of their choices. Armantier (2006) later confirmed these findings. Conversely, Sefton (1992) found that participants in a dictator game tended to be more generous when a bRIS with 25% chance to be paid was employed, as opposed to a 'pay all' system. However, this result has been recently contested by Clot et al. (2018), who found that bRIS had no effect on behavior in a standard dictator game.

Several studies have examined the impact of not paying all subjects on individual risk preferences. Among studies that document no impact of between-subject incentive system, Harrison et al. (2007, footnote 16) reported no significant differences in risk preferences, elicited by Holt and Laury (2002) tasks, when comparing wRIS and wbRIS, where 1 out of 10 subjects was paid. Beaud and Willinger (2015) also found no difference in behavior when only 10% of subjects were randomly selected for payment (wbRIS) compared to when all subjects were paid (wRIS) in an elicitation of risk vulnerability (attitudes towards background risk) using the Gneezy and Potters (1997) method. Similarly, Brokesova et al. (2017) studied risk-taking behavior by using a binary choice between a certain amount and a lottery yielding different payoffs with different probabilities and found that subjects' choices did not differ when all participants were paid ('pay all') or when only one participant per session was paid (bRIS). More recently, Berlin et al. (2024) documented no differences between paying all subjects and paying only half or 10% of them while eliciting risk and time preferences with an adaptive bisection procedure entailing binary choice questions.

Studies documenting the impact of between-subject incentives on risk preferences have produced somewhat conflicting results. Cubitt et al. (1998a) observed lower levels of risk aversion in a single binary choice problem where all participants were compensated on a 'pay all' basis compared to a situation where only a subset of them was paid (bRIS). Baltussen et al. (2012) reported significant but opposite results in a dynamic risky choice problem modeled after the TV show *'Deal or No Deal,'* observing significantly more risk aversion under wRIS than under wbRIS. Interestingly, Baltussen et al. (2012) also found no difference between the 'pay all' and wRIS mechanisms, suggesting that risk preferences may fluctuate under different incentive schemes. Finally, the recent study of Anderson et al. (2023) reported that subjects tend to be less risk-averse in Holt and Laury (2002) tasks when a between-subject incentive randomization is introduced. However, these results should be interpreted with caution, as the experiment was conducted on a within-subject basis, with different parts of the experiment incentivized using different systems (wRIS or wbRIS). In this case, the entire experiment is effectively incentivized with a wRIS, which is not immune to wealth, portfolio, and hedging effects. Table 2 summarizes the existing literature by grouping studies according to treatment comparisons.

Table 2: Existing literature by treatment comparisons

Study	Domain	Treatment	Main result
Sefton (1992)	Social preferences	Pay all vs. bRIS	More generosity with bRIS
Clot et al. (2018)	Social preferences	Pay all vs. bRIS	No behavioral difference
Cubitt et al. (1998a)	Risk preferences	Pay all vs. bRIS	Less risk aversion with pay all
Brokesova et al. (2017)	Risk preferences	Pay all vs. bRIS	No behavioral difference
Baltussen et al. (2012)	Risk preferences	Pay all vs. wRIS	No behavioral difference
Bolle (1990)	Social preferences	wRIS vs. wbRIS	No behavioral difference
Armantier (2006)	Social preferences	wRIS vs. wbRIS	No behavioral difference
Harrison et al. (2007)	Risk preferences	wRIS vs. wbRIS	No behavioral difference
Baltussen et al. (2012)	Risk preferences	wRIS vs. wbRIS	More risk aversion with wRIS
Beaud and Willinger (2015)	Risk preferences	wRIS vs. wbRIS	No behavioral difference
Anderson et al. (2023)	Risk preferences	wRIS vs. wbRIS	More risk aversion with wRIS
Berlin et al. (2024)	Risk preferences	wRIS vs. wbRIS	No behavioral difference
Berlin et al. (2024)	Time preferences	wRIS vs. wbRIS	No behavioral difference

Our study complements and extends the existing literature in several ways. First, we compare the impact of incentive mechanisms on uncertainty preferences by using a between-subject design, while keeping absolute stake sizes constant across treatments. This approach addresses potential concerns about contamination across different incentive systems and confounding effects that arise from simultaneously altering the incentive system and the face-value rewards. Second, our design minimizes decision-making efforts by employing a simple and widely-used preference elicitation method. Specifically, we use straightforward choice lists to elicit certainty equivalents, presenting subjects with a series of choice between a simple lottery and various certain amounts. This design avoids unusual probabilities and complex or dynamic elements. Finally, our study goes beyond examining risk preferences by also considering preferences towards ambiguity and compound risk, which have been widely associated with each other (Halevy, 2007; Abdellaoui et al., 2015; Chew et al., 2017).

3 Experimental design

Our experiment compares individual choices under risk and ambiguity using two distinct incentive mechanisms: the within-subject random incentive system (wRIS) and the within-between-subject random incentive system (wbRIS). The comparison is made on a between-subject basis.

3.1 Stimuli and choice tasks

Subjects are asked to bet on the color of a ball drawn from an urn under different conditions. We examine five uncertain situations generated by urns containing 100 balls, each of which is either red or blue.[1] In each situation, subjects select themselves the color on which to bet. The situations are as follows:

1. *Risk (R):* The urn contains 50 red and 50 blue balls.
2. *Ambiguity-Uniform (A-U):* The proportion of red and blue balls in the urn is unknown.
3. *Ambiguity-Degenerate (A-D):* The urn is composed of either only red or only blue balls, with unknown probabilities. Thus, two potential degenerate compositions exist: 100 red and 0 blue balls, or 100 blue and 0 red balls.
4. *Compound Risk-Uniform (CR-U):* The proportion of red and blue balls in the urn is unknown, but all the possible compositions are equally likely. The urn is constructed as follows: A ticket is drawn from a bag containing 101 tickets numbered from 0 to 100. The number of that ticket determines the number of red balls.
5. *Compound Risk-Degenerate (CR-D)*: The urn is composed of either only red or only blue balls, with equal probabilities. The urn is constructed as follows: A ticket is drawn from a bag containing 2 tickets numbered 0 or 100. The number of that ticket determines the number of red balls.

The situations involving risk (*R, CR-U,* and *CR-D)* differ from those involving ambiguity (*A-U* and *A-D)* in terms of the uncertainty they entail. Specifically, whereas the probabilities in the risk and compound risk situations are *known*, they are *unknown* in the ambiguity situations. Based on a widely used (and empirically supported) symmetry assumption in Ellsberg experiments, a subjective uniform distribution can be considered over the potential urn compositions in *A-U* and *A-D* (see for example Abdellaoui et al., 2011; Chew et al., 2017; Aydogan et al., 2023). The situations *R* and *A-U* correspond to the ones originally proposed by Ellsberg (1961), whereas *A-D* was recently proposed by Chew et al. (2017). Situations *CR-U* and *CR-D* were previously considered by Halevy (2007), Abdellaoui et al. (2015), and Chew et al. (2017).

We measure uncertainty preferences using certainty equivalents (CEs) elicited through a choice list design. This method is easy to construct and to implement. It

[1] The experiment also included five additional situations with urns containing only two balls, which are used in the context of another study. Our conclusions about the validity of the wbRIS also hold for those situations (see Online Appendix).

consists in asking respondents to compare a fixed gamble with a series of increasing sure amounts. In our case, the gambles consisted in a €20 bet on each of the five situations described above. In each situation, the subjects were asked to make 22 binary choices between the prospect of receiving €20 and a sure amount ranging from €0 to €20. The sure amounts were incremented by €1 between €0.5 and €19.5. We implemented a fast filling system that automatically selects the dominating options once a choice in the list is made. We identify the CEs by looking at the switching points on the list between the bet and the sure amounts. The CE is determined as the midpoint between the highest sure amount not chosen and the lowest sure amount chosen over the gamble.

3.2 Incentive systems

All subjects received a €5 flat payment for their participation. In addition, they were paid a variable amount based on one of the two following incentive mechanisms.

Treatment '100%' In the baseline, we adopt the standard wRIS. In other words, all participants are paid based on one of their decisions in the experiment. To enhance isolation and minimize potential biases, the choice question implemented for payment was determined *prior* to the experiment (Johnson et al., 2021). In practice, at the beginning of each session, a volunteer from the subject pool randomly picked two sealed envelopes: one containing the description of one of the five potential uncertain situations and the other containing one of the 22 binary choices from the choice list. The two sealed envelopes were then attached to the wall and remained visible to all participants until the end of the experiment. Subjects were informed that the choice question that would ultimately be used to determine their payment would be the one corresponding to the combined contents of the two envelopes. The content of the two envelopes was revealed only when all subjects had completed all the tasks. At the end of the experiment, a ball was randomly drawn from the corresponding urn, and subjects were paid according to their decision in the selected choice question and the color of the ball drawn.

Treatment '10%' The procedure in the wbRIS treatment is similar to that in the wRIS baseline, except that only one-in-ten subjects is paid. In other words, only 10% of the participants in this treatment are randomly selected to be paid based on one of their decisions in the experiment. The between-subject randomization also took place *prior* to the experiment. Specifically, upon arriving at the lab, each subject was asked to draw an individual sealed envelope, among which 10% contained an image of a happy face, allowing them to implement one of their decisions for real.

These envelopes were kept sealed by the subjects until the end of the experiment. The content of all the envelopes was revealed at the end of the session. The choice question selected for payment was then implemented only for participants whose envelope contained a happy face.

3.3 Empirical strategy and power analysis

We aimed for a minimum of 86 subjects per treatment, which allows us to identify a medium effect size of 0.5 standard deviations (i.e., a Cohen's d of 0.5) with power of 0.9 at the 5% significance level in a standard t-test for two independent samples. This means that the t-tests are able to correctly reject the null hypothesis with 90% probability when the effect size is at least 0.5. A non-rejection of the null hypothesis in these standard tests thus suggests no significant difference between treatments. Nevertheless, these tests fall short in further informing us about the precise likelihood of the null hypothesis being true in case of its non-rejection. Therefore, we also provide a Bayesian estimation analysis, as proposed in Kruschke (2013). Specifically, we estimate the Bayesian posterior probability of observing an effect size smaller than 0.5 in absolute value to evaluate the evidence in favor of the null hypothesis. The details of the procedure of our Bayesian estimations are provided in Online Appendix.

3.4 Procedure and further experimental details

The experiment was run on computers at the Anthropo-Lab (Lille, France). In total, 182 university students participated in the experiment (91 subjects per treatment). Upon arrival at the lab, each subject was assigned an individual cubicle. Subjects could not communicate with each other during the experiment. Each session started with the signature of a consent form, the reading of the experimental instructions, some examples of the stimuli, and comprehension questions. At the end of the experiment, subjects answered a short survey with a few socio-economic questions.

The urns representing the uncertain situations were constructed in advance by an assistant, who was not present in the room during the experiment. Thus, no one in the room (including the experimenters) had any additional information about the content of the urns than what was described in the instructions. At the end of the experiment, subjects had the opportunity to open the urns to verify the truthfulness of the instructions. Subjects were paid in cash at the end of the experiment. Average earnings were €15 in the '100%' treatment, and €6.8 in the '10%' treatment. Each session lasted approximately 40 minutes, including instructions and payment. Complete instructions are available in the Online Appendix.

4 Results

4.1 Data

The data consists of five choice lists per subject. The analysis focuses on the lists indicating a precise indifference interval. We exclude the lists where a sure €0 was chosen over the uncertain bet or where the uncertain bet was chosen over a sure €20, as they do not correspond to any meaningful indifference measure. Such lists amount to 3.3% (15/455) and 3.5% (16/455) of all observations in the '100%' and '10%' treatments, respectively.

4.2 Attitudes towards uncertainty

We present the results of the mean elicited CEs in Table 3 and the cumulative distribution functions in Figure 1.

Table 3: Results on Elicited Certainty Equivalences

	Mean CE		t-test	Bayesian Estimation		
	'100%'	'10%'	p-value	Estimated effect size	95% CI on posterior effect sizes	% of posterior effect sizes within $[-0.5; 0.5]$
R	8.48	8.85	0.43	0.15	$[-0.30; 0.58]$	94%
$A-U$	7.19	7.82	0.21	0.27	$[-0.17; 0.71]$	85%
$A-D$	7.67	8.04	0.49	0.13	$[-0.30; 0.57]$	95%
$CR-U$	8.40	8.45	0.94	0.04	$[-0.41; 0.48]$	97%
$CR-D$	8.54	8.28	0.61	-0.16	$[-0.61; 0.33]$	92%

Notes: p-values are based on two-sided t-tests for two independent samples. Estimated effect sizes in Bayesian Estimations are based on the mean of the posterior distributions and computed as the ratio of mean differences to the combined standard deviations of the two samples (i.e., $(\mu_1 - \mu_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)2}$). They indicate differences in terms of standard deviations.

In both treatments '100%' and '10%', the CEs tend to be lower than the expected payoff €10 (two-sided binomial test, $p<0.002$ for all), indicating general uncertainty aversion. Testing for the impact of the incentive system on the experimental results, we observe that the CE distributions do not differ across the two treatments (Kolmogorov-Smirnov test, $p>0.353$ for all), and that the t-tests do not indicate any significant difference between the mean CEs ($p>0.211$ for all).

The absence of differences between treatments is confirmed by the Bayesian estimation analysis, where the 95% highest density credibility intervals (CIs) on effect sizes always include zero (right panel of Table 3). To evaluate the strength of evidence in favor of the null hypothesis, we look at the percentages of the posterior

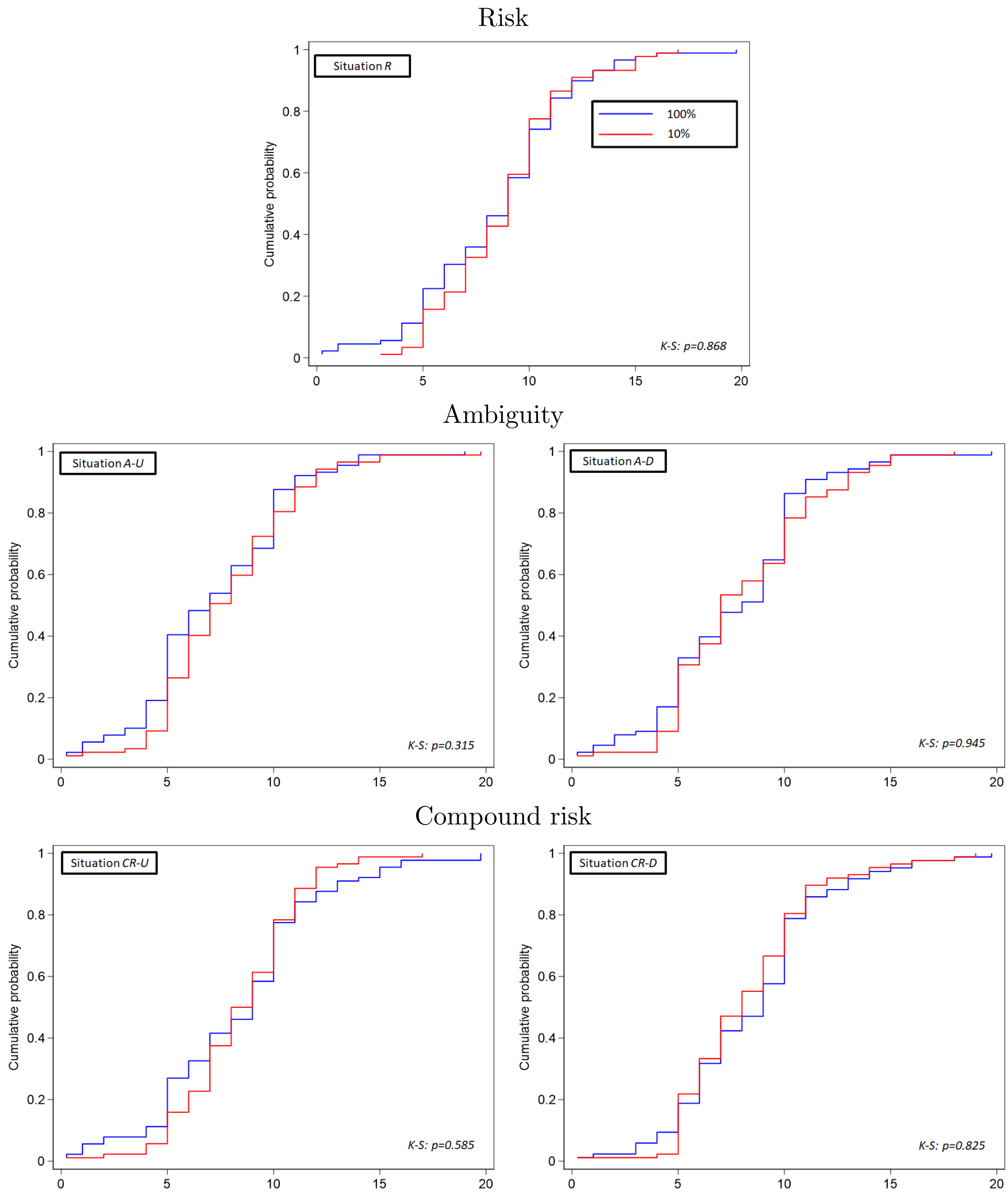

Figure 1: CDFs of the elicited CEs of the five uncertain situations. *Note*: K-S refers to the exact p-value of the Kolmogorov-Smirnov test

effect sizes that are smaller than an effect size of 0.5 in absolute value. As can be observed, the estimated effect sizes indicate, at most, a 0.27 standard deviation difference between the two treatments. A large proportion of the posterior effect sizes also fall within the interval [-0.5; 0.5], indicating convincing evidence in favor of the equivalence of the two treatments.

To test the overall effect of the incentive system, we moreover run a regression analysis based on the pooled data with random effects at the individual level. The

CEs are regressed on a dummy variable for the '10%' treatment, a dummy for ambiguous situations, and a dummy for compound risk situations. The baseline thus corresponds to the behavior under simple risk in the '100%' treatment. Table 4 presents the results. The first column pools all the CEs without differentiating between the uncertain situations. This model indicates overall no difference between the CEs in the two treatments (p=0.429). The treatment effect furthermore remains insignificant when including the dummies for ambiguity and compound risk in the regression (second column, p=0.432). Finally, we find no interactions between the treatment effect and the different uncertain situations (third column, p=0.730 and p=0.190 for ambiguity and compound risk, respectively).

Table 4: Random effect regressions

	(1)	(2)	(3)
'10%'	0.328	0.326	0.456
	(0.415)	(0.414)	(0.484)
Ambiguity		-0.914***	-0.977***
		(0.186)	(0.285)
Compound Risk		-0.241	-0.010
		(0.177)	(0.258)
'10%' × Ambiguity			0.129
			(0.373)
'10%' × Compound Risk			-0.461
			(0.352)
Constant	8.039***	8.503***	8.437***
	(0.312)	(0.341)	(0.373)
Observations	879	879	879

Notes: Cluster-robust standard errors are in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3 Specific attitudes towards ambiguity and compound risk

We next examine the effect of the incentive system on specific attitudes towards ambiguity and compound risk. To measure these attitudes, we approximate the ambiguity and compound risk premia by the difference between the CE of the bets on risk and on ambiguity or compound risk, respectively, i.e., $\pi_i = CE_R - CE_i$ for $i = \{A\text{-}U, A\text{-}D, CR\text{-}U, CR\text{-}D\}$. In words, each premium indicates what an individual is ready to pay to be confronted with the risk situation R rather than ambiguity or compound risk situation.

Table 5 presents the results on the premium measures. In line with what has been documented in the literature (e.g., Trautmann and van de Kuilen, 2015), the ambiguity premia are overall positive, revealing that our subjects usually exhibit ambiguity aversion (two-sided binomial tests, p<0.001 for $\pi_{A\text{-}U}$ and $p < 0.048$ for $\pi_{A\text{-}D}$ in both treatments). On the contrary, we mostly observe compound risk neutrality (i.e., zero compound risk premium, $p > 0.111$ for all, except for $CR - U$ in the '10%' treatment, p=0.048). Comparing the premia across the two treatments, we observe no differences (two-sided t-tests, $p > 0.155$ for all). These results are confirmed by examining the effect sizes obtained in the Bayesian estimations, where the 95% CIs always include the zero effect and a large proportion of the posterior effect sizes do not exceed 0.5 in absolute value.

Table 5: Results on Estimated Premia

	Mean CE		t-test	Bayesian Estimation		
	'100%'	'10%'	p-value	Estimated effect size	95% CI on posterior effect sizes	% of posterior effect sizes within $[-0.5; 0.5]$
π_{A-U}	1.22	0.94	0.51	0.05	$[-0.49; 0.58]$	93%
π_{A-D}	0.71	0.79	0.88	0.18	$[-0.35; 0.69]$	88%
π_{CR-U}	0.04	0.38	0.42	0.12	$[-0.41; 0.66]$	90%
π_{CR-D}	-0.07	0.54	0.16	0.24	$[-0.30; 0.78]$	83%

Notes: p-values are based on two-sided t-tests for two independent samples. Estimated effect sizes in Bayesian Estimations are based on the mean of the posterior distributions and computed as the ratio of mean differences to the combined standard deviations of the two samples (i.e., $(\mu_1 - \mu_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)2}$). They indicate differences in terms of standard deviations.

Table 6 reports the proportions of aversion ($\pi_i > 0$), neutrality ($\pi_i = 0$), and seeking ($\pi_i < 0$) attitudes towards ambiguity and compound risk. In line with what precedes, our data do not indicate any difference in the distribution of attitudes across the two treatments (last column). These results are also confirmed in pooled multinomial logistic regressions of a categorical attitude variable with random effects at individual level (p=0.689, see Online Appendix). Overall, our data thus suggest that the incentive system does not distort preferences towards ambiguity and compound risk.

4.4 The relationship between ambiguity and compound risk attitudes

As the relationship between attitudes towards ambiguity and compound risk has received ample attention in the literature (Halevy, 2007; Abdellaoui et al., 2015; Chew et al., 2017; Aydogan et al., 2023), we here investigate the robustness of this relationship to the incentive system. Table 7 summarizes the association between

Table 6: PROPORTIONS OF ATTITUDES

	'100%'			'10%'			Chi-square test
	Aversion $\pi_i > 0$	Neutrality $\pi_i = 0$	Seeking $\pi_i < 0$	Aversion $\pi_i > 0$	Neutrality $\pi_i = 0$	Seeking $\pi_i < 0$	p-value
π_{A-U}	52.3%	34.1%	13.6%	55.8%	31.4%	12.8%	0.895
π_{A-D}	42.5%	33.3%	24.1%	48.8%	25.6%	25.6%	0.523
π_{CR-U}	37.5%	38.6%	23.9%	42%	34.1%	23.9%	0.787
π_{CR-D}	31.8%	41.2%	27.1%	40.7%	33.7%	25.6%	0.447

ambiguity neutrality and compound risk reduction. A subject is here considered

Table 7: ASSOCIATION BETWEEN AMBIGUITY NEUTRALITY AND REDUCTION OF COMPOUND RISK

	Reduction of compound lotteries					
	'100%'			'10%'		
Ambiguity neutrality	No	Yes	Total	No	Yes	Total
No	60	8	68	66	8	74
	[48.9]	[19.1]		[58.7]	[15.3]	
Yes	4	17	21	3	10	13
	[15.1]	[5.9]		[10.3]	[2.7]	
Total	64	25	89	69	18	87
Independence test:	Fisher's exact test (2-tailed): p<0.001			Fisher's exact test (2-tailed): p<0.001		

Notes: In brackets, the expected frequency given the population.

ambiguity neutral if she exhibits ambiguity neutrality in the two ambiguous situations (i.e., $\pi_{A\text{-}U} = \pi_{A\text{-}D} = 0$). Compound risk reduction is defined analogously. The data from both groups show a strong relationship between ambiguity and compound risk attitudes. In particular, among subjects who do not reduce compound risk, respectively 94% (60 out of 64) and 96% (66 out of 69) exhibit ambiguity nonneutrality (two-sample Z-test of proportions, p=0.624). Similarly, among subjects who exhibit ambiguity nonneutrality, the proportions of nonreduction of CR are respectively 88% (60 out of 68) and 89% (66 out of 74). Those proportions do not differ from each other (two-sample Z-test of proportions, p=0.858). Thus, the relationship between ambiguity neutrality and compound risk reduction typically obtained in the literature is preserved under the between-subject random incentive system.

5 Concluding remarks

This paper contributes to the limited research on the within-between-subject random incentive system and presents new evidence supporting its validity. We demonstrate that paying all subjects or just a subset of them (e.g., one in ten) based on a standard within-subject random incentive system does not affect (i) the elicited certainty equivalents of risky and ambiguous bets, (ii) attitudes towards ambiguity and compound risk, or (iii) the association between ambiguity neutrality and compound risk reduction. These findings provide a strong foundation for using the between-subject incentive mechanism in future experiments exploring individual decision-making.

References

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review 101*(2), 695–723.

Abdellaoui, M., P. Klibanoff, and L. Placido (2015). Experiments on compound risk in relation to simple risk and to ambiguity. *Management Science 61*(6), 1306–1322.

Anderson, L. R., B. A. Freeborn, P. McAlvanah, and A. Turscak (2023). Pay every subject or pay only some? *Journal of Risk and Uncertainty 66*(2), 161–188.

Armantier, O. (2006). Do wealth differences affect fairness considerations? *International Economic Review 47*(2), 391–429.

Aydogan, I., L. Berger, and V. Bosetti (2023). Unraveling Ambiguity Aversion. *The Review of Economics and Statistics*, 1–32.

Aydogan, I., L. Berger, V. Bosetti, and N. Liu (2023). Three layers of uncertainty. *Journal of the European Economic Association 21*(5), 2209–2236.

Azrieli, Y., C. P. Chambers, and P. J. Healy (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy 126*(4), 1472–1503.

Baillon, A., H. Bleichrodt, and G. D. Granic (2022). Incentives in surveys. *Journal of Economic Psychology 93*, 102552.

Baltussen, G., G. T. Post, M. J. Van Den Assem, and P. P. Wakker (2012). Random incentive systems in a dynamic choice experiment. *Experimental Economics 15*(3), 418–443.

Beattie, J. and G. Loomes (1997). The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty 14*, 155–168.

Beaud, M. and M. Willinger (2015). Are people risk vulnerable? *Management Science 61*(3), 624–636.

Berlin, N., E. Kemel, V. Lenglin, and A. Nebout (2024). Paying none, some or all? between-subject random incentives and preferences towards risk and time. *SSRN Working Paper*.

Bolle, F. (1990). High reward experiments without high expenditure for the experimenter? *Journal of Economic Psychology 11*(2), 157–167.

Brokesova, Z., C. Deck, and J. Peliova (2017). Comparing a risky choice in the field and across lab procedures. *Journal of Economic Psychology 61*, 203–212.

Burgstaller, L. and K. Pfeil (2024). You don't need an invoice, do you? an online experiment on collaborative tax evasion. *Journal of Economic Psychology 101*, 102708.

Charness, G., U. Gneezy, and B. Halladay (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization 131*, 141–150.

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

Clot, S., G. Grolleau, and L. Ibanez (2018). Shall we pay all? An experimental test of random incentivized systems. *Journal of Behavioral and Experimental Economics 73*, 93–98.

Cubitt, R., C. Starmer, and R. Sugden (1998a). Dynamic choice and the common ratio effect: An experimental investigation. *The Economic Journal 108*(450), 1362–1380.

Cubitt, R. P., C. Starmer, and R. Sugden (1998b). On the validity of the random lottery incentive system. *Experimental Economics 1*(2), 115–131.

Edwards, W. (1953). Probability-preferences in gambling. *American Journal of Psychology 66*(3), 349–364.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics 75*, 643–669.

Fanghella, V., C. Faure, M.-C. Guetlein, and J. Schleich (2023). Locus of control and other-regarding behavior: Experimental evidence from a large heterogeneous sample. *Journal of Economic Psychology 95*, 102605.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics 112*(2), 631–645.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica 75*(2), 503–536.

Harrison, G. W., M. I. Lau, and E. E. Rutström (2007). Estimating risk attitudes in Denmark: A field experiment. *The Scandinavian Journal of Economics 109*(2), 341–368.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Johnson, C., A. Baillon, H. Bleichrodt, Z. Li, D. Van Dolder, and P. P. Wakker (2021). Prince: An improved method for measuring incentivized preferences. *Journal of Risk and Uncertainty 62*(1), 1–28.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General 142*(2), 573.

Laury, S. (2006). Pay one or pay all: Random selection of one choice for payment. *Andrew Young School of Policy Studies Research Paper Series* (06-13).

Moffatt, P., C. Starmer, R. Sugden, N. Bardsley, R. Cubitt, and G. Loomes (2009). *Experimental economics: Rethinking the rules.* Princeton University Press.

Parra, D. (2024). Eliciting dishonesty in online experiments: The observed vs. mind cheating game. *Journal of Economic Psychology 102*, 102715.

Sefton, M. (1992). Incentives in simple bargaining games. *Journal of Economic Psychology 13*(2), 263–276.

Slovic, P. (1969). Differential effects of real versus hypothetical payoffs on choices among gambles. *Journal of Experimental Psychology 80*(3p1), 434.

Trautmann, S. T. and G. van de Kuilen (2015). *Ambiguity Attitudes*, Chapter 3, pp. 89–116. The Wiley Blackwell Handbook of Judgment and Decision Making, John Wiley & Sons, Ltd.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity.* Cambridge university press.