FRAMING AND REPETITION EFFECTS ON RISKY CHOICES: A BEHAVIOURAL APPROACH

Noemí Herranz-Zarzoso[a], Gerardo Sabater-Grande[b] and Ainhoa Jaramillo-Gutiérrez[c]

Abstract

Framing effects play an important role in individual decision-making under risk. This investigation revisits framing effects caused by two versions of the choice list procedure, lottery vs. lottery (LL) and lottery vs. certainty (LC). In the first, subjects face pairwise choices between lotteries within a choice list. In the second, subjects face pairwise choices between a safe amount and a lottery. In order to measure the sensitivity of subjects' choices to the structure of the tasks, we implement an incentive-compatible experiment using repetition in order to have a robust measure of the subjects' propensity to make a choice. Particularly, it is tested whether variations in the number of options offered in a choice list with and without variations in the range of options affect subjects' choices. Our results suggest that changes in framework disturb subjects' risk preferences only in the LC version when the range of options presented has been varied.

Keywords: risk aversion, framing effects, risk task repetition, choice list procedure.

JEL classification: C91, D03, D81

Noemí Herranz-Zarzoso expresses her gratitude to the Ministerio de Educación Cultura y Deporte for the financial support (FPU14/02966).

[a] LEE & Economics Dept., Universitat Jaume I, Castellon, Spain. e-mail: nherranz@uji.es, phone: +34 964387636

[b] LEE & Economics Dept., Universitat Jaume I, Castellon, Spain. e-mail: sabater@uji.es, phone: +34 964387632

[c] LEE & Economics Dept., Universitat Jaume I, Castellon, Spain. e-mail: jaramill@uji.es, phone: +34 964387635

1. Introduction

Risk attitude is known to be a key determinant of various economic and financial choices. Behavioural studies that aim to evaluate the role of risk attitude in contexts of this type require tools for measuring risk aversion at both the individual and the aggregate levels. The most frequent procedure to elicit individual risk attitudes is referred to as the ***choice list procedure***. The choice list method presents a table of binary choices designed so that, as a respondent works through the table, he/she can be expected to switch at some point from "one side" to the other. Two alternative versions of this procedure are analysed in this study: the lottery vs. lottery (LL) and lottery vs. certainty (LC) methods. In the LL method, subjects face pairwise choices between lotteries within a choice list. A famous example of this method was proposed by Holt and Laury (2002, HL hereafter), in which subjects are given a list of 10 choices between paired lotteries where payoffs are constant and probabilities vary systematically across the successive decisions in steps of 10%. Our LC method is based on a Becker-DeGroot-Marschak (1964, BDM) auction in which subjects face pairwise choices between a safe amount and a lottery. In order to determine the subject's payoff in a choice list, the Random Lottery Incentive (RLI) payment mechanism is used to pick one decision at random from the list. If the isolation hypothesis from prospect theory[1] (which implies that subjects evaluate each risk task independently of the other tasks) is maintained, each pairwise choice a subject makes in the list can be interpreted as if he/she had faced only that binary choice.

This paper investigates whether subjects' choices are influenced by framing effects originated by the choice list procedure. A framing effect is a decision-making bias in which subjects choose differently when the same problem is presented in diverse ways.[2] This effect can refer to multiple issues that may affect the presentation of the risk task implemented.[3] The best known framing effects relating to choice lists flown from variations in the order, the number and the range of the options presented in the risk choice task. We aim to distinguish between framing effects and the effects of underlying stochastic variability of choice (which may be present even if the choice list is held constant). In order to disentangle these effects, in our experiment subjects face repetitions of the different choice lists, in which the number, the range of options or both have changed. This allows us to obtain a measure of subjects' propensity to make a choice that is richer than if they have only faced the choice list once.

[1] See Kahneman and Tversky (1979).

[2] Tversky and Kahneman (1986) argued that framing effects violate the normative condition of description invariance, which stipulates that the same problem should be evaluated in like manner regardless of its description.

[3] Although Cason and Plott (2014) argued that subjects' game form misconception can be misinterpreted as a framing effect using a BDM value elicitation experiment, Bartling *et al.* (2015) demonstrated that subject misconception is not necessary to produce such an effect.

Results on framing effects are a mixed bag depending on the method used to elicit subjects' risk aversion and the type of framing effect analysed (ordering effects, changes in the number of options presented with or without affecting their range, simultaneous versus sequential presentation of lotteries, etc.). Hey and Orme (1994) found that when the same 100 pairs of ternary lotteries were repeated twice on separate days[4] (with a possibility to declare indifference) in a different order, subjects chose identical options for each pair in around 75% of all cases. On increasing the number of repetitions[5] with respect to his previous paper, Hey (2001) found that some individuals maintained a constant variability in their responses to 100 pairwise risky-choice questions in spite of their being repeated. Using the HL method, Andersen *et al.* (2006)[6] found that choices were affected by ordering effects[7] and the range of a given lottery. Specifically, they found that the deletion of the worst pairs (with the lowest expected value) of lotteries increased risk aversion. Additionally, the authors showed that by enforcing only one switching point (strict monotonicity and transitivity) had no systematic effect. Lévy-Garboua *et al.* (2012) presented experimental evidence of how framing[8] affected decisions in the context of the HL procedure. They found that presenting lotteries simultaneously induced significantly less inconsistency than showing lotteries sequentially. Furthermore, both repetition of identical choices and high payoffs also reduced inconsistency. Bosch-Domènech and Silvestre (2013) found what they called "an embedding bias". This bias implies that when some specific pairs of alternative lotteries are removed, risk aversion becomes less frequent and the ranking of individuals by risk aversion is not preserved. However, the aforementioned bias was not found when they analysed the certainty equivalent (CE) elicitation method. Contrary to these results, Freeman *et al*. (2019) found that embedding a pairwise choice in a choice list increased the fraction of subjects choosing the riskier lottery when the safer alternative was certain, but it did not significantly affect choices when the safer alternative is risky. Erev *et al.* (2008) and Blavatskyy and Köhler (2009) analysed the robustness of the CE mechanism to elicit risk preferences and found that elicited payoffs were systematically affected by the range of certain payoffs to which the lottery was compared. Beauchamp *et al.* (2012) studied how risk aversion parameters were affected by manipulating the intermediate pairs of options without affecting the range of options. They found that

[4] However, their design included confounding the wealth effects from paying all experiments after subjects had performed the final one.
[5] The same set of ternary lotteries was presented to subjects in five sessions separated by at least two days and the authors did not give participants the opportunity to indicate indifference.
[6] In this paper, a budget constraint precluded paying all subjects, so each subject is given only a 10% chance of actually receiving the payment associated with his/her decision.
[7] These order effects are consistent with findings reported in Harrison *et al.* (2005).
[8] In this experiment lottery choices were presented either simultaneously or sequentially and the probabilities of winning are ranked either in increasing, decreasing or in random order.

when the endpoints of the multiple price list were fixed and intermediate outcomes were decreased, participants' choices became significantly more risk averse. Finally, Loomes and Pogrebna (2014) used three elicitation methods[9] and found considerable variability within (and even more, between) the results they produced. This finding suggested that not only different elicitation instruments but also framing-specific issues could interact with imprecise underlying preferences.[10] Lastly, Brown and Healy (2018) found significant violations of isolation when all decisions were displayed in a standard list format, but not when the rows of the list were randomized and shown on separate screens.

However, analyses of framing effects in the literature rely on two crucial assumptions: (1) the supposition that subjects always choose the same answer to exactly the same question, and (2) the assumption that subjects consider each choice list in isolation from other ones.

Regarding the first assumption, Lévy-Garboua *et al.* (2012) pointed out that "even in decision experiments where subjects make repeated independent and identically distributed (*i.i.d.*) decisions among pairs of lotteries without any alteration" (p. 129) an estimable number of subjects reported different options upon repetition. The results from experiments conducted by Ballinger and Wilcox (1997) and Loomes and Sugden (1998) supported this evidence and sustained that repetition drove subjects towards increasingly safer choices. Later, Loomes and Pogrebna (2014) found that most subjects showed variability when they answered some questions aimed at eliciting their risk attitude.

The second crucial assumption is related to the fact that a possible contamination effect across choice lists can be observed when subjects do not consider each choice list in isolation from the others. The extreme case of such a contamination effect is considered in Holt (1986) by means of the reduction hypothesis, by which the subject views the whole experiment as a single compound lottery. Starmer and Sugden (1991) were the first to test subjects' behaviour in random-lottery experiments, discarding Holt's conjecture. However and following Cubitt *et al.* (1998), even if the reduction hypothesis is rejected, Starmer and Sugden's (1991) results still allow for some degree of across-choice-list contamination. Although in general this weaker hypothesis is not supported by Loomes (1997) and Cubitt *et al*. (1998), more recent experimental papers by Harrison and Swarthout (2014) and Cox *et al.* (2014, 2015) showed consistent evidence of across-choice-list contamination.

[9] The choice list procedure, the ranking procedure (presenting a set of options and asking the respondent to identify which option he/she ranks top) and the allocation procedure (providing the respondent with a budget and allowing him/her to distribute it between different state-contingent claims).
[10] In this online experiment only 1/8 of the randomly selected subjects were invited to the laboratory to play out their decisions for real money.

Our study argues that it cannot be concluded that changes in decisions are necessarily due to changes in framing if subjects make different decisions in identical sequentially repeated risk tasks. In this sense, we depart from previous literature because we analyse framing effects taking into account across-choice-list contamination effects. Isolating subjects who reported in the questionnaire that they did not consider each choice list separately from the others, we implement a second step of the analysis in order to perform a refined check of our results. From these findings it is seen that the LL elicitation method is robust to manipulations in the number and/or the range of options offered in the list. Nevertheless, the LC method is not so robust because changes in the task structure modify subjects' revealed preferences.

2. Experimental Design

In order to study framing effects in the multiple choice list procedure, we tested for shifts in risk preferences due to: (1) a (a)symmetric variation in the number of pairs offered while keeping the range of options constant (CR), and (2) a (a)symmetric variation in the number of pairs while varying the range (VR) of options offered.

Changes in CR and VR are analysed using both the LL and the LC elicitation methods. Treatment 1 (T1) and treatment 2 (T2) correspond to the LL elicitation method for CR and VR changes, respectively. Treatment 3 (T3) and treatment 4 (T4) are related to the LC method for CR and VR modifications, respectively. Following Gonzalez and Wu (1999), we asked subjects to choose which row they wanted to switch to in order to fill in the remaining choices for the subject.[11]

A total of 137 subjects (31 in T1, 36 in T2 and in T3, and 34 in T4) were recruited among undergraduate students from different economics or business-related courses at the Universitat Jaume I (UJI), using standard recruitment procedures with an open call for subjects through the LEE (Laboratorio de Economía Experimental) website. Before the beginning of each session subjects were given written instructions, which were also read aloud by the organisers. Any remaining questions were answered privately.

At the end of each session, subjects answered a questionnaire, which asked them to report whether they had varied their choices across the different repetitions. If so, they were asked to choose between

[11] These authors found that enforcing strict monotonicity and transitivity had no systematic effect on responses. However, Lévy-Garboua *et al.* (2012) showed that a non-negligible part of players exhibited inconsistent behavior when monotonicity was not imposed. Andersson *et al.* (2016) reported evidence that lower cognitive ability was significantly correlated with subjects having multiple switching points.

three possible causes of this variability. The first potential cause was that they did not have a clear choice between some pairs of options. The second one was that their decision in one choice list was affected by decisions in other choice lists. The last cause was the category "other", which allowed for free-text responses.[12] After that, subjects were paid privately in cash. All sessions were computerized and carried out in a specialized computer laboratory, using software based on the Z-Tree toolbox by Fischbacher (2007).

In the case of lottery vs. lottery, LL_{17} is shown in Table 1 below. The name refers to the fact that in this task subjects face a list of seventeen pairs of lotteries, which we number from one to seventeen, each pair involving a "safe" lottery (S) and a "risky" lottery (R). These labels are provided because if we compare lottery R with S, R offers both the best and the worst (null) payoffs.

Lottery Pair	Safe lottery (S)				Risky lottery (R)				EV_S	EV_R	EV_S- EV_R
	Prob.	Payoff	Prob.	Payoff	Prob.	Payoff	Prob.	Payoff			
1	10%	€17.50	90%	€26.70	10%	€0.00	90%	€100.0	*€25.78*	*€90.00*	*- €64.22*
2	15%	€17.50	85%	€26.70	15%	€0.00	85%	€100.0	*€25.32*	*€85.00*	*- €59.68*
3	20%	€17.50	80%	€26.70	20%	€0.00	80%	€100.0	*€24.86*	*€80.00*	*- €55.14*
4	25%	€17.50	75%	€26.70	25%	€0.00	75%	€100.0	*€24.40*	*€75.00*	*- €50.60*
5	30%	€17.50	70%	€26.70	30%	€0.00	70%	€100.0	*€23.94*	*€70.00*	*- €46.06*
6	35%	€17.50	65%	€26.70	35%	€0.00	65%	€100.0	*€23.48*	*€65.00*	*- €41.52*
7	40%	€17.50	60%	€26.70	40%	€0.00	60%	€100.0	*€23.02*	*€60.00*	*- €36.98*
8	45%	€17.50	55%	€26.70	45%	€0.00	55%	€100.0	*€23.56*	*€55.00*	*- €31.44*
9	50%	€17.50	50%	€26.70	50%	€0.00	50%	€100.0	*€22.10*	*€50.00*	*- €27.90*
10	55%	€17.50	45%	€26.70	55%	€0.00	45%	€100.0	*€21.64*	*€45.00*	*- €23.36*
11	60%	€17.50	40%	€26.70	60%	€0.00	40%	€100.0	*€21.18*	*€40.00*	*- €18.82*
12	65%	€17.50	35%	€26.70	65%	€0.00	35%	€100.0	*€20.72*	*€35.00*	*- €14.28*
13	70%	€17.50	30%	€26.70	70%	€0.00	30%	€100.0	*€20.26*	*€30.00*	*- €9.74*
14	75%	€17.50	25%	€26.70	75%	€0.00	25%	€100.0	*€19.70*	*€25.00*	*- €5.30*
15	80%	€17.50	20%	€26.70	80%	€0.00	20%	€100.0	*€19.34*	*€20.00*	*- €0.66*
16	85%	€17.50	15%	€26.70	85%	€0.00	15%	€100.0	*€18.88*	*€15.00*	*€3.88*
17	90%	€17.50	10%	€26.70	90%	€0.00	10%	€100.0	*€18.42*	*€10.00*	*€8.42*

Table 1. Pairs of lotteries offered in LL_{17}.

[12] This category was not chosen by any subject.

The last three columns in Table 1 (not shown to the experimental subjects) indicate the expected euro values of the safe lottery in the pair (denoted EV_S) and that of the risky lottery (denoted EV_R), as well as the difference between the two. For the first fifteen rows, the risky option offers the higher expected value (EV) while for the last two rows the safe option offers the higher EV, with the difference between EVs decreasing as we go down the list. Thus, a risk-neutral individual would select the R lottery in all pairs with the exception of the last two. Subjects' payoffs are selected in order to offer: (1) a sufficient reward to subjects in an experiment with multiple risk task repetitions and random lottery incentive (RLI) as the payment mechanism, and (2) a wide number of pairs of lotteries where EV_R exceeds EV_s.

Using LL_{17}, we construct four additional risk tasks: LL_{13}, LL_9, LL_7 and LL_5. LL_{13} contains the odd pairs of LL_{17} plus all pairs from ten to seventeen; LL_9 is composed of all odd pairs from one to seventeen, LL_7 comprises odd pairs from five to seventeen and LL_5 contains odd pairs from five to thirteen. Treatment 1 (2) is formed by tasks LL_9, LL_{17} (LL_5) and LL_{13} (LL_7). As can be seen by the design, the baseline task is LL_9 since it is used in both treatments. Thus, task LL_{17} (LL_{13}) increases the number of pairs of lotteries symmetrically (asymmetrically) keeping the range of options from 10-90% constant relative to the benchmark. In the same way, task LL_5 (LL_7) decreases the number of options symmetrically (asymmetrically) reducing the range of options to 30-90% and 30-70%, respectively, with respect to the baseline task.

Subjects face repeated *i.i.d.* decisions among lists of pairs of options without any alteration. Specifically, in treatment 1 (2), each subject deals with tasks LL_9, LL_{17} (LL_5) and LL_{13} (LL_7) six times for each of them in a totally random order. Repetition of identical choice lists allows us to study the variability of subjects' choices across identical tasks. Subjects were classified in two categories: "*constant*" (C) subjects, who are individuals who always choose the same in all identical tasks, and "*inconstant*" (IC) subjects, who are individuals showing variability in their responses across identical tasks. In the latter category, there are subjects who self-report in the questionnaire that they do not think about each choice list separately, i.e. who show across-choice-list contamination. Since the concern of the paper is to analyse contamination effects within choice lists, we exclude those subjects in part of the analysis in order to have a refined check of the framing effects.

In both treatments, we informed subjects that three draws would be implemented to determine their payment. A first draw was carried out to choose which of their 18 choice list tasks will be selected; a second draw was used to randomly choose one from all the pairs of lotteries contained in the selected task; a third draw, given the odds of the lottery preferred by the subject in the pair, was applied to determine individual payoffs. This design rules out possible wealth effects due to subjects' (expected) earnings from previous periods.

As regards the LC elicitation method, we used two lotteries: a safe lottery (S) and a risky one (R). In Table 2 (Table 3), we present LC^S_{21} (LC^R_{31}). In these tasks, subjects must choose a lottery (S or R) and safe amounts of money that increase by €0.50 in LC^S_{21} and €1.50 in LC^R_{31} for each additional pair. The last columns of these tables (not shown to the experimental subjects) indicate the difference between the expected euro value of the lottery in the pair (denoted $EV_{S/R}$) and the safe amount of money. A risk-neutral individual would select the safe lottery in all the pairs in Table 2 except the last two. However, a risk-neutral subject would choose the R lottery only in the three first pairs of Table 3.

Pair	Safe lottery				Safe amount (SA)	EV_S-SA
1	81%	€10.00	19%	€0.00	€0.00	*€8.10*
2	81%	€10.00	19%	€0.00	€0.50	*€7.60*
…	…				…	…
20	81%	€10.00	19%	€0.00	€9.50	*-€1.40*
21	81%	€10.00	19%	€0.00	€10.00	*-€1.90*

Table 2 Pairs of options offered in LC^S_{21}

Pair	Risky lottery				Safe amount (SA)	EV_R-SA
1	19%	€45.00	81%	€0.00	€0.00	€8.55
2	19%	€45.00	81%	€0.00	€1.50	€7.05
…	…				…	…
30	19%	€45.00	81%	€0.00	€43.50	-€35.05
31	19%	€45.00	81%	€0.00	€45.00	-€36.55

Table 3. Pairs of options offered in LC^R_{31}

Using LC^S_{21} to construct the safe lotteries vs. certainty, we create four additional risk tasks: LC^S_{11}, LC^S_{16}, LC^S_5 and LC^S_8. Task LC^S_{11} contains all the odd pairs from LC^S_{21}; LC^S_{16} comprises all the odd pairs of LC^S_{21} and all the pairs from one to eleven; LC^S_5 comprises all the odd pairs from seven to fifteen; and LC^S_8 contains all the odd pairs from one to fifteen. Treatment 3 (4) for safe lotteries contains LC^S_{11}, LC^S_{21} (LC^S_5) and LC^S_{16} (LC^S_8). As can be observed, the baseline task for safe lotteries is LC^S_{11}. Compared with this, task LC^S_{21} (LC^S_{16}) increases the number of certainty payoffs symmetrically (asymmetrically) without varying their range with respect to LC^S_{11} (the range goes from €0.00–10.00). Additionally, task LC^S_5 (LC^S_8) decreases the number of options symmetrically (asymmetrically) by diminishing the range of options offered to subjects with respect to LC^S_{11}. Indeed, the latter goes from €3.00–7.00 in the case of LC^S_5 and from €0.00–7.00 in the case of LC^S_8.

In a similar manner, we use LC^R_{31} to build four additional risk tasks: LC^R_{16}, LC^R_{23}, LC^R_{10} and LC^R_9. LC^R_{16} comprises all the odd pairs from LC^R_{31}; LC^R_{23} contains all the odd pairs from LC^R_{31} and all the pairs from one to fifteen, LC^R_{10} contains all the odd pairs from seven to twenty-five, and lastly LC^R_9 comprises all the odd pairs from one to seventeen. Treatment 3 (4) for risky lotteries contains LC^R_{16}, LC^R_{31} (LC^R_{10}) and LC^R_{23} (LC^R_9). The task repeated in two treatments and the one that is considered our benchmark is LC^R_{16}. Therefore, task LC^R_{31} (LC^R_{23}) increases the number of safe amounts symmetrically (asymmetrically) without varying their range with respect to LC^R_{16}, the latter going from €0.00–€45.00. Additionally, task LC^R_{10} (LC^R_9) decreases the range of options offered to subjects with respect to LC^R_{16} by symmetrically (asymmetrically) diminishing the number of certainty payoffs featured.

In treatment 3 (treatment 4), all subjects complete tasks LC^S_{11}, LC^S_{21} (LC^S_5), LC^S_{16} (LC^S_8), LC^S_{16}, LC^R_{31} (LC^R_{10}) and LC^R_{23} (LC^R_9) in random order. All tasks are repeated six times in both treatments. In this case, subjects are informed that up to three draws could be necessary to calculate their payment, thus avoiding the aforementioned wealth effects. A first draw was used to choose which of their 36 tasks will be selected; a second draw was conducted to choose one from all the pairs of options contained in the selected task; if the chosen option is the lottery, a third draw was implemented to obtain subjects' payoffs.

To sum up the experimental design, a summary of the treatments is presented in Table 4.

Treatment	Subjects	Tasks	Type of framing effect
T1	31	LL_9, LL_{17}, LL_{13}	Constant range (CR)
T2	36	LL_9, LL_5, LL_7	Varying range (VR)
T3	36	LC^S_{11}, LC^S_{21}, LC^S_{16} LC^R_{16}, LC^R_{31}, LC^R_{23}	Constant range (CR)
T4	34	LC^S_{11}, LC^S_5, LC^S_8 LC^R_{16}, LC^R_{10}, LC^R_9	Varying range (VR)

Table 4: Summary treatments

3. Data analysis

3.1. Statistical tests

In order to analyse framing effects, we use a Wilcoxon test. Specifically, it is used to compare the percentage of safe choices (in the LL method) or the certainty choices (in the LC method) taking place under two different frameworks for the same sample of subjects. With the aim of treating observations independently, we calculate the percentage of safe choices made by each subject for the six repetitions in a common pair. Those percentages are then compared between different lotteries for each subject. We apply a Bonferroni correction[13] to take into account the problem of false positives in multiple pair comparisons.

By repeating the same risk task six times, we are able to analyse in a within-subject framework the variability of subjects' choices within each *i.i.d.* decision. This allows us to classify subjects according to their variability within *i.i.d.* decisions. We use the term *"constant"* (C) subjects to refer to individuals who always choose the same in *i.i.d*. decisions and *"inconstant"* (IC) subjects for participants who do not choose the same in *i.i.d.* decisions. In the IC category, there are subjects that create contamination across choice lists because they report that their decisions in one choice list were influenced by the decisions in other choice lists. Then, in a second step of the analysis, we will exclude them to eliminate this potential problem and to have a robustness check of contamination within choice lists.

3.1.1. Lottery vs. lottery method

Figures 1 and 2 present the average rate of safe choices per pair of options in each LL task included in T1 and T2. In both treatments, the benchmark choice list is LL_9. In T1 we symmetrically (asymmetrically) vary the number of pairs offered while keeping the range of options constant by means of LL_{17} (LL_{13}). In T2, we symmetrically (asymmetrically) vary the number of pairs while diminishing the range of options offered by means of LL_5 (LL_7).

[13] The Bonferroni correction consists in multiplying the p-value by the number of pair comparisons, resulting in a rather demanding threshold for rejection. We apply Bonferroni corrections to all the tests performed.

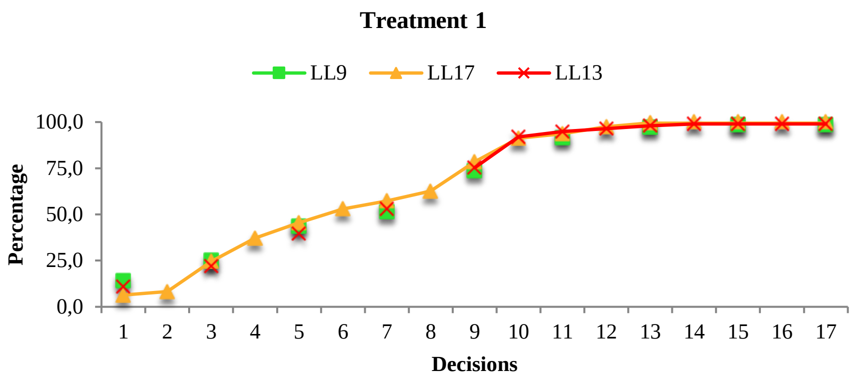

Figure 1. Average rate of safe choices per pair in the LL task in T1.

In T1, where the range is constant, the differences across the choice lists in the average rate of safe choices in the individual decisions are, in general, unnoticeable. Specifically, when we expand the number of pairs symmetrically (from LL_9 to LL_{17}) we do not find any significant differences between the percentages of safe lotteries chosen by subjects in the same pair.[14] An identical result is obtained when the number of options increases asymmetrically (from LL_9 to LL_{13} or from LL_{13} to LL_{17}). Therefore, we can conclude that:

Result 1: *A (symmetric or asymmetric) variation in the number of pairs offered in the lottery vs. lottery method, keeping the range of options constant, does not produce any framing effects.*

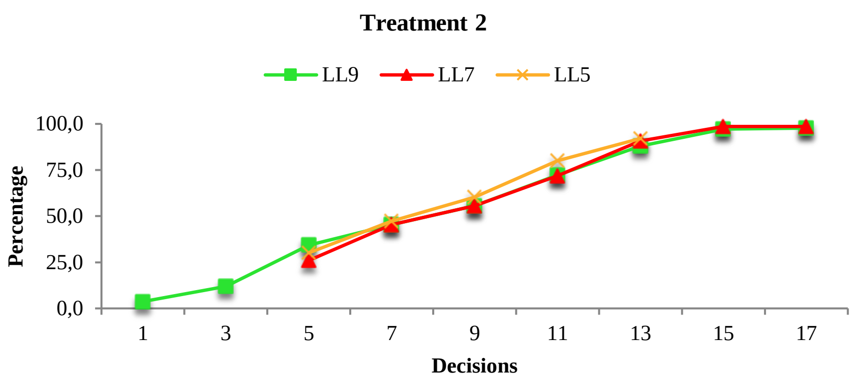

Figure 2. Average rate of safe choices per pair in the LL task in T2.

In T2, we present a symmetric or an asymmetric variation in the number of pairs offered varying the range of options. Specifically, in task LL_5 (LL_7) the range of options offered decreases with respect to the baseline task diminishing the number of pairs symmetrically (asymmetrically). Comparing LL_9

[14] All Bonferroni-corrected Wilcoxon test p-values corresponding to each pair are above 0.05.

and LL_7, no significant differences[15] between the percentages of safe lotteries chosen by subjects in the common pairs are found. The same results occur when we compare LL_9 with LL_5. In consequence, we can state that:

Result 2: *A (symmetric or asymmetric) variation in the number of pairs offered in the lottery vs. lottery method, varying the range of options, does not generate any framing effects.*

The previous analysis is based on the original sample. Nevertheless, IC subjects who self-report that they do not consider each choice list in isolation from other choice lists are removed from our original sample to implement a check for the robustness of our previous results. This refined sample (RS) ensures that our results account only for within-choice-list contamination, which is the actual concern of our paper.

Figures 3 and 4 reformulate the empirical evidence of Figures 1 and 2, presenting the average rate of safe choices per pair of options for the RS.

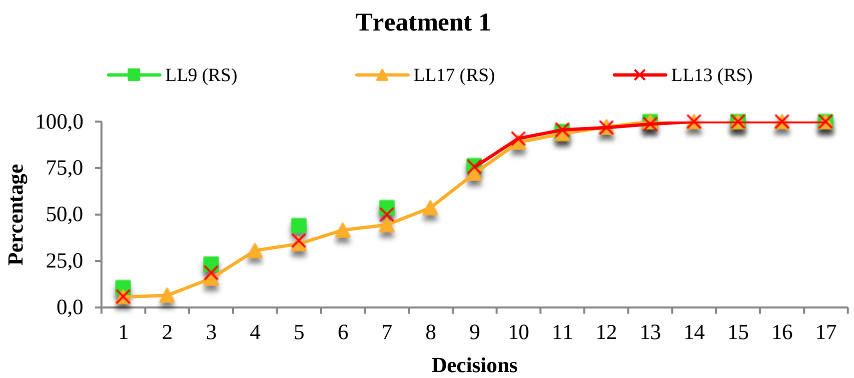

Figure 3. Average rate of safe choices per pair in the LL task for the RS in T1.

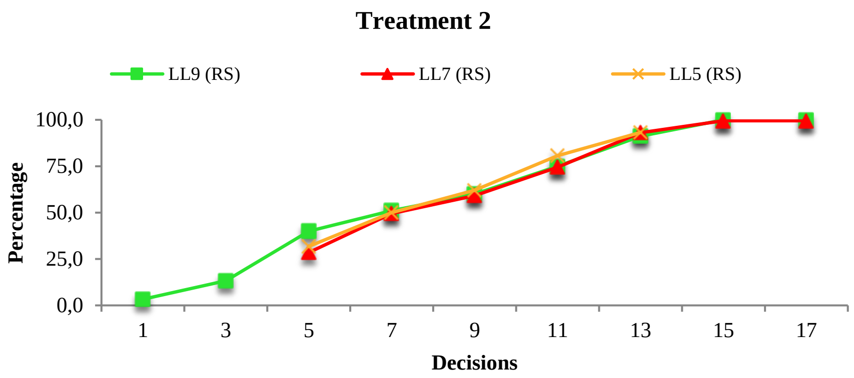

Figure 4. Average rate of safe choices per pair in the LL task for the RS in T2.

[15] All Bonferroni-corrected Wilcoxon test p-values corresponding to each pair are above 0.05.

In all the aforementioned comparisons, we obtain identical results to those of the original sample case, i.e. no framing effects are found.

Result 3: *Considering and disregarding subjects who self-report across-choice-list contamination, no framing effects are found in the lottery vs. lottery method.*

These results contrast with those obtained by some authors who have analysed the same method searching for framing effects. Andersen *et al*. (2006) found that choices were affected by order and lottery range when they deleted the two worst pairs. More recently, Bosch-Domènech and Silvestre (2013) found that when some pairs were removed, subjects' choices change, which they called embedding bias.

Lastly, we check whether our non-significant differences are due to a lack of statistical power by running an ex-post power analysis. This allows us to find the minimum sample size required to detect the smallest effect that, if true, has a 90% chance of producing an impact estimate which is statistically significant at 5%. For the case of lottery vs. lottery, sample sizes would have to increase up to at least N=318 in order to find framing effects.

3.1.2. Lottery vs. certainty method

Framing effects in the LC method are analysed by means of T3 and T4. In T3, we vary the number of certainty payoffs symmetrically/asymmetrically keeping their range constant with respect to the baseline tasks (LC^{S}_{11} or LC^{R}_{16}). In T4, we vary the range of options offered to subjects with respect to the baseline tasks symmetrically/asymmetrically by varying the number of certainty payoffs.

Figures 5 and 6 display the average percentage of certain choices in both the safe and the risky lottery, respectively, per pair in the LC tasks presented in T3.

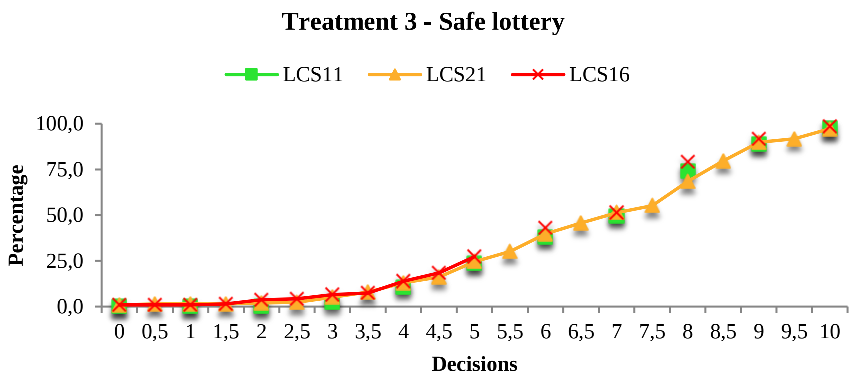

Figure 5. Average percentage of certain choices per pair preferred to the safe lottery in T3.

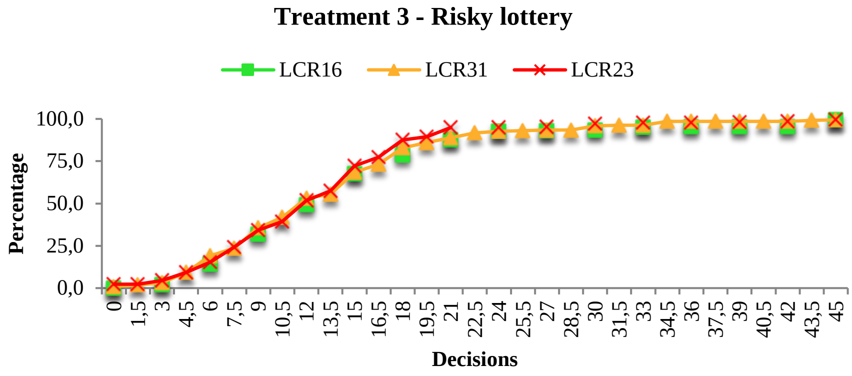

Figure 6. Average percentage of certain choices per pair preferred to the risky lottery in T3.

In general, when we increase the number of certainty payoffs symmetrically (from LC^{S}_{11} to LC^{S}_{21} and from LC^{R}_{16} to LC^{R}_{31}) or asymmetrically (from LC^{S}_{11} to LC^{S}_{16} or from LC^{S}_{16} to LC^{S}_{21} and from LC^{R}_{16} to LC^{R}_{23} or from LC^{R}_{23} to LC^{R}_{31}), without changing the range of options, we do not find any significant differences between the percentage of safe amounts chosen by subjects in the same pair. However, a framing effect is found when we compare LC^{R}_{16} and LC^{R}_{23} for a safe amount of €18.[16]

Result 4: *In general, no framing effects are found in the lottery vs. certainty method when the number of pairs varies symmetrically keeping the range of options constant. A marginally significant effect is found in one case when there is an asymmetric variation in the number of pairs offered keeping the range of options constant.*

Figures 7 and 8 present the average percentage of certain choices in both the safe and the risky lotteries, respectively, per pair in the LC task in T4, in which the range of options has been reduced with respect to the baseline task.

[16] There is a framing effect after computing the Bonferroni-corrected Wilcoxon at a 10% level of significance.

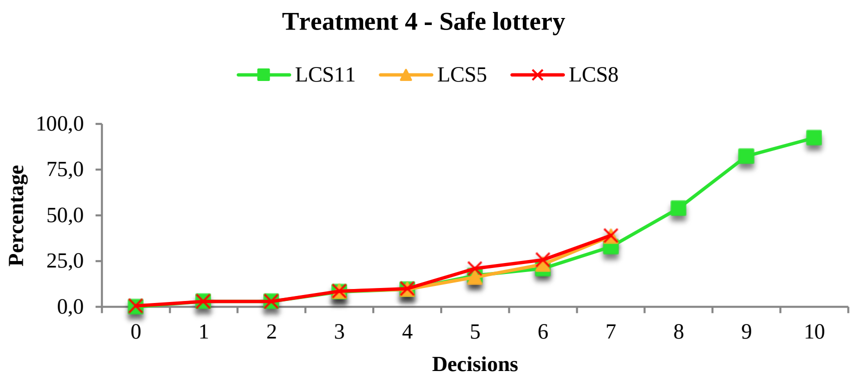

Figure 7. Average percentage of certain choices per pair preferred to the safe lottery in T4.

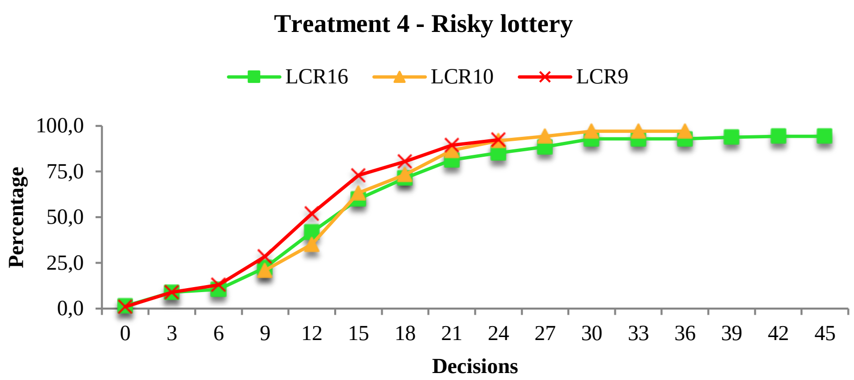

Figure 8. Average percentage of certain choices per pair preferred to the risky lottery in T4.

It is important to note that in the safe (risky) lottery, we reduce the number of pairs symmetrically between LC^S_{11} and LC^S_5 (LC^R_{16} and LC^R_{10}), whereas the number of certainty payoffs is decreased asymmetrically between LC^S_{11} and LC^S_8 (LC^R_{16} and LC^R_9) and between LC^S_5 and LC^S_8 (LC^R_{10} and LC^R_9).

We find significant differences between the percentage of certain choices selected only for the risky lottery in the following cases: (a) comparing LC^R_{16} and LC^R_{10} for safe amounts €24, €27, €30, €33 and €36;[17] (b) comparing LC^R_{10} and LC^R_9 for safe amounts of €12 and €18;[18] and (c) in the comparison between LC^R_{16} and LC^R_9 for safe amounts of €9, €12, €15 and €18.[19]

Result 5: *A (symmetric or asymmetric) variation in the number of safe amounts varying the range of*

[17] Bonferroni-corrected Wilcoxon test p-values are 0.07, 0.06, 0.07, 0.07 and 0.07 respectively after multiplying the original p-values by 10.
[18] Bonferroni-corrected Wilcoxon test p-values are 0.012 and 0.090 after multiplying the original p-values by 6.
[19] Bonferroni-corrected Wilcoxon test p-values are 0.063, 0.012, 0.090 and 0.072 after multiplying the original p-values by 9.

options produces marginal framing effects for a large number of pairs *in the lottery vs. certainty method.*

As in the lottery vs. lottery method, we remove from our sample those subjects who would be prone to contamination effects across choice lists.

In Figures 9 and 10, we present the average percentage of certain choices per pair preferred to the safe and the risky lottery, respectively, for the refined sample in T3.

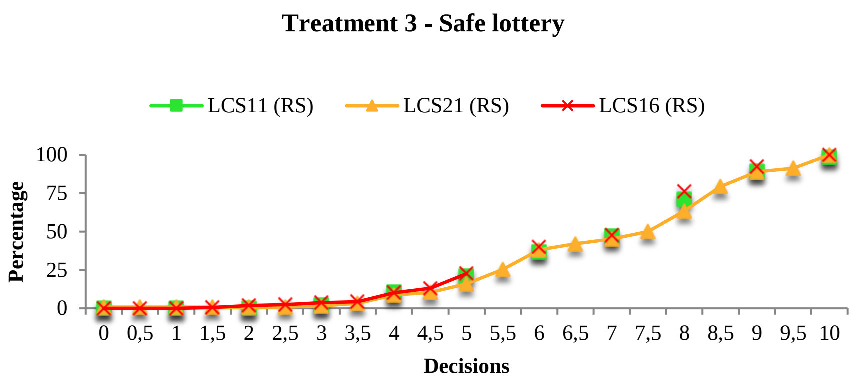

Figure 9. Average percentage of certain choices per pair preferred to the safe lottery for the RS in T3.

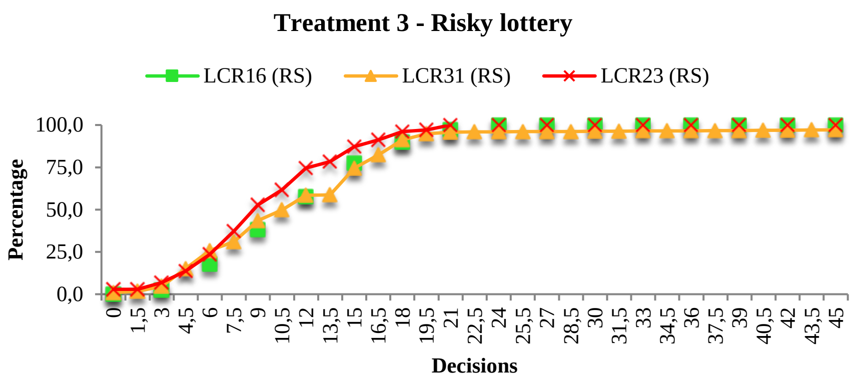

Figure 10. Average percentage of certain choices per pair preferred to the risky lottery for the RS in T3.

A (a)symmetric variation in the number of certainty payoffs without changing the range of options and removing inconsistent subjects does not produce any framing effects.[20]

Result 6: *A (symmetric or asymmetric) variation in the number of safe amounts keeping the range of options constant does not produce any framing effects in the lottery vs. certainty method*[21] *when subjects self-reporting across-choice-list contamination are removed.*

In Figures 11 and 12, we present the average percentage of certain choices per pair preferred to the safe and the risky lottery, respectively, for the RS in T4.

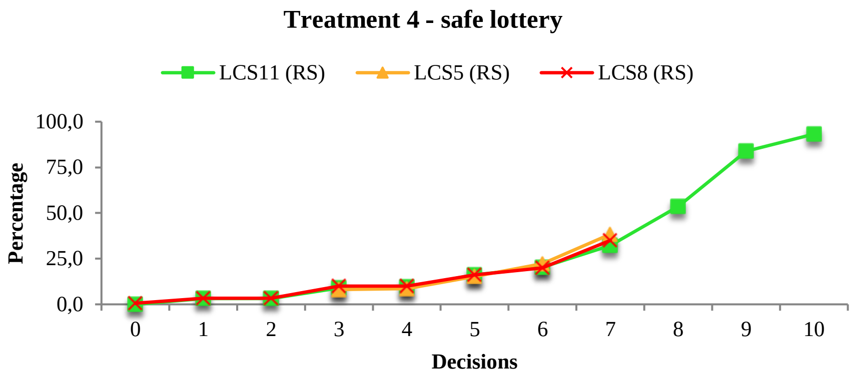

Figure 11. Average percentage of certain choices per pair preferred to the safe lottery for the RS in T4.

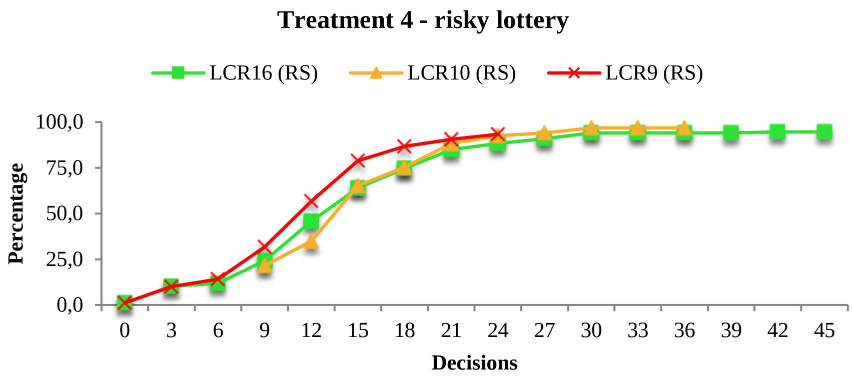

Figure 12. Average percentage of certain choices per pair preferred to the risky lottery for the RS in T4.

[20] All Bonferroni-corrected Wilcoxon test p-values are above 0.05.

[21] Although in Figure 10, there seem to be framing effects, these disappear once we apply the Bonferroni correction to avoid the existence of false positives.

Unlike the original sample case, when we exclude subjects self-reporting across-choice-list contamination and vary the number of certainty payoffs symmetrically by varying the range of options offered, no significant differences[22] are found in the average rate of adoption of the safe amount. However, the removal of these subjects cannot completely eliminate all the framing effects generated by varying the number of safe payoffs offered asymmetrically by decreasing the range: we find that the previous differences obtained in the comparison of LC^{R}_{16} and LC^{R}_{9} disappear, but the ones between LC^{R}_{10} and LC^{R}_{9} still remain.[23]

Result 7: *On removing subjects self-reporting across-choice-list contamination, a symmetric variation in the number of safe amounts varying the range of options does not generate any framing effects in the lottery vs. certainty method. Nevertheless, the elimination of these subjects reduces, but does not completely eliminate, the framing effects if the number of* safe amounts *is varied asymmetrically, by varying the range of options.*

Our results are consistent with those of Blavatskyy and Köhler (2009) by inferring that the range of feasible minimum safe amounts systematically affected the elicited prices and those of Bosch-Domènech and Silvestre (2013), concluding that the CE method was robust.

As in the previous method, in the lottery vs. certainty method we test whether our previous non-significant differences are due to a lack of statistical power. In this regard, we run an ex-post power analysis using power set at 90% and probability at 5%. For the lottery vs. certainty method, sample sizes would have to increase up to at least N=353 in order to find framing effects in the ones in which there are none.

3.2. Logit models.

In this section, we estimate different logit models to shed light on the determinants of framing effects and to corroborate our previous results based on statistical tests.

3.2.1. Lottery vs. Lottery (LL)

Table 5 includes, as explanatory variables, the tasks subjects undertake in random order in each treatment and period, which means all the tasks that modify our baseline lottery (LL_9) and the

[22] All Bonferroni-corrected Wilcoxon test p-values are above 0.05.
[23] Bonferroni-corrected Wilcoxon test p-values corresponding to €9, €12, €15 and €18 are 0.066, 0.018, 0.054 and 0.042 respectively after multiplying the original p-values by 6.

dependent variable is the number of safe choices. The modifications are based on changes in the number of options offered with and without varying the range. Additionally, we have two different models for each treatment: one includes the original sample and the other includes the refined one.

LL decision	T1	T1 (RS)	T2	T2 (RS)
LL_{13}	0.0560	0.0215		
	(0.0789)	(0.104)		
LL_{17}	-0.0473	-0.0270		
	(0.0794)	(0.104)		
LL_5			-0.163	0.175
			(0.114)	(0.140)
LL_7			0.0851	0.263
			(0.114)	(0.140)
Period	-0.0222	-0.000924	0.00823	0.0117
	(0.0189)	(0.0248)	(0.0274)	(0.0333)
Constant	-0.846	-0.762	-1.015	-1.426
	(0.206)***	(0.292)***	(0.469)**	(0.603)**
N	31	17	36	26

Standard errors in parentheses adjusted within-subjects; *** p<0.01, ** p<0.05, * p<0.1

Table 5. LL models for the original and for the refined sample.

These models corroborate our *Results 1*, *2* and *3*. Any modifications in the number of pairs offered with and without changing the range of options do not produce any framing effects in the lottery vs. lottery method. These results hold not only for the refined sample, but also for the original one. Thus, this version of the multiple choice list procedure is robust to framing effects.

3.2.2. Lottery vs. Certainty (LC)

The aim of this subsection is the same as in the previous one, but now for the lottery vs. certainty version.

Table 6 includes, as explanatory variables, the different tasks undertaken by the subjects in random order in each different treatment (for safe and risky lotteries) and the period. The dependent variable is the number of safe choices, that is, all the modifications made to our baseline lotteries (LC^S_{11} and LC^R_{16}) in the number of options offered with and without varying the range. Furthermore, we have two different models for each treatment: one includes the original sample and the other includes the refined one.

LC decision	T3 Safe	T3 Safe (RS)	T3 Risky	T3 Risky (RS)	T4 Safe	T4 Safe (RS)	T4 Risky	T4 Risky (RS)
LC^{S16}	-0.123 (0.0644)*	-0.0559 (0.0927)						
LC^{S}_{21}	-0.0258 (0.0651)	0.0790 (0.0937)						
LC^{R}_{23}			-0.134 (0.0554)**	0.0143 (0.111)				
LC^{R}_{31}			-0.110 (0.0551)**	0.00703 (0.109)				
LC^{S}_{5}					-0.180 (0.172)	-0.366 (0.230)		
LC^{S}_{8}					-0.378 (0.168)**	-0.665 (0.222)***		
LC^{R}_{10}							-0.740 (0.0897)***	-0.815 (0.123)***
LC^{R}_{9}							-0.246 (0.0752)***	-0.266 (0.106)**
Period	-0.00800 (0.0156)	0.00571 (0.0223)	-9.81e-05 (0.0132)	-4.78e-05 (0.0262)	0.0974 (0.0401)**	0.0764 (0.0521)	0.00378 (0.0195)	0.000547 (0.0274)
Constant	0.706 (0.144)***	0.856 (0.203)***	-0.835 (0.127)***	-1.386 (0.191)***	3.429 (0.638)***	4.177 (0.687)***	0.620 (0.236)***	0.646* (0.351)*
N	36	18	36	11	34	26	32	16

Standard errors in parentheses adjusted within-subjects; *** p<0.01, ** p<0.05, * p<0.1

Table 6. LC models for the original and for the refined sample.

From the previous models our *Results 4, 5* and 6 are corroborated with this additional analysis. The only difference found is with respect to our *Result 7*. The deletion of subjects self-reporting across-choice-list contamination does not eliminate all the framing effects in the regression analysis. This is because, in the statistical tests, the Bonferroni correction was applied and it is quite restrictive. Nevertheless, in general terms, the same conclusion is found: this version of the multiple choice list procedure is not robust to framing effects. Modifications to the structure (number of options and range) of the LC used produce modifications in the risk attitude of subjects.

4. Conclusions

In this study, the robustness of two different choice list methods has been analysed. The first is the lottery vs. lottery method, where subjects are faced with pairwise choices between gambles within a choice list. The second is the lottery vs. certainty method, where subjects are faced with pairwise choices between a safe amount and a lottery. In order to analyse the framing effects, we have implemented a within-subjects experiment allowing variability in responses across repeated identical tasks. The framing effects analysed include shifts in risk preferences due to a (a)symmetric variation in the number of pairs offered keeping the range of options constant, and a (a)symmetric variation in the number of pairs varying the range of options offered.

By repeating each identical risk task six times, we classify subjects according to their variability within *i.i.d*. decisions in two categories: constant and inconstant subjects. The latter category includes some subjects self-reporting across-choice-list contamination in the questionnaire. Thus, they were discarded from the sample in part of the analysis. By so doing, we provide a refined check on the contamination effects within choice lists (framing effects), thereby eliminating a potential source of across-choice-list contamination.

In the LL elicitation method, we do not find any framing effects when this type of subject is removed from the sample. However, the LC method does not seem as robust as the LL method. Particularly, if we account for across-choice-list contamination, all framing effects found in the full sample analysis disappear with the exception of those that appear when the range varies.

Summing up, some changes in the revealed risk preferences attributed to framing effects in the literature may really correspond to a confounding across-choice-list contamination effect.

References

Andersen, S., Harrison, G. W., Lau, M. I., & Elisabet, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics, 9(4),* 383–405.

Andersson, O., Holm, H. J., Tyran, J. R., & Wengström, E. (2016). Risk aversion relates to cognitive ability: Preferences or Noise? *Journal of the European Economic Association, 14(5),* 1129-1154

Ballinger, T. P., & Wilcox, N. T. (1997). Decisions, error and heterogeneity. *Economic Journal, 107*, 1090–1105.

Bartling, B., Engl, F., & Weber, R. A. (2015). Game form misconceptions are not necessary for a willingness-to-pay vs. willingness-to-accept gap. *Journal of the Economic Science Association, 1*(1), 72–85.

Beauchamp, J. P., Benjamin, D. J., Chabris, C. F., & Laibson, D. I. (2012). *How malleable are risk preferences and loss aversion?* Harvard University Mimeo.

Blavatskyy, P. R. & Köhler, W. R. (2009). Range effects and lottery pricing. *Experimental Economics, 12(3),* 332–349.

Bosch-Domènech, A. & Silvestre, J. (2013). Measuring risk aversion with lists: a new bias. *Theory and Decision, 75(4),* 465-496.

Brown, A. L., & Healy, P. J. (2018). Separated decisions. *European Economic Review*, 101, 20-34.

Cason, T. N., & Plott, C. R. (2014). Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing. *Journal of Political Economy*, *122*(6), 1235–1270.

Cox, J. C., Sadiraj, K., & Schmidt, U. (2014). Asymmetrically dominated choice problems, the isolation hypothesis and random incentive mechanisms. *PLOS ONE, 9(3)*, e90742.

Cox, J.C., Sadiraj V. & Schmidt U. (2015). Paradoxes and mechanisms for choice under risk, *Experimental Economics, 18(2)*, 215-250.

Cubitt, R. P., Starmer, C. & Sugden, R (1998). On the validity of the Random Lottery Incentive System. *Experimental Economics, 1(2)*, 115-131.

Erev, I., Glozman, I., & Hertwig, R. (2008). What impacts the impact of rare events. *Journal of Risk and Uncertainty, 36(2),* 153–177.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10(2)*, 171-178.

Freeman, D. J., Halevy, Y., & Kneeland, T. (2019). Eliciting risk preferences using choice lists. *Quantitative Economics*, *10*(1), 217–237.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology, 38(1),* 129–166.

Harrison, G.W., Lau, M. I., Rutström, E. E., & Sullivan, M. B. (2005). Eliciting risk and time preferences using field experiments: Some methodological issues. In Carpenter, J., Harrison, G.W., & List, J.A., (Eds.), Field Experiments in Economics (Greenwich, CT: JAI Press, Research in Experimental Economics, Volume 10).

Harrison, G. W., & Swarthout, J. T. (2014). Experimental payment protocols and the Bipolar Behaviorist. *Theory and Decision*, *77*(3), 423–438.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62(6),* 1291–1326.

Hey, J. D. (2001). Does repetition improve consistency? *Experimental economics, 4(1),* 5-54.

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

Holt, C. A. (1986). Preference reversals and the independence axiom. *The American Economic Review, 76(3),* 508-515.

Kahneman D & Tversky A (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*, 263–291.

Lévy-Garboua, L., Maafi, H., Masclet, D., & Terracol, A. (2012). Risk aversion and framing effects. *Experimental Economics, 15(1),* 128–144.

Loomes, G. & Pogrebna G. (2014). Measuring individual risk attitudes when preferences are imprecise. *The Economic Journal, 124(576),* 569-593.

Loomes, G., & Sugden, R. (1998). Testing different stochastic specifications of risky choice. *Economica, 65(260),* 581–598.

Starmer, C. & Sugden, R., (1991). Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation. *American Economic Review, 81(4),* pp. 971-78.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59,* 251–278.