Finance Research Letters 50 (2022) 103270

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

On the consistency of the individual behavior when facing higher-order risk attitudes

Annarita Colasante [a,1], Jaume García-Segarra [b,*,1], Luca Riccetti [c,1], Alberto Russo [b,d,1]

[a] Department of Law and Economics, Università degli studi di Roma "Unitelma Sapienza", Piazza Sassari 4, 00161 Roma, Italy
[b] LEE & Department of Economics, Universitat Jaume I. Av. Vicent Sos Baynat, s/n 12071, Castelló de la Plana, Spain
[c] Department of Economics and Law, Università degli Studi di Macerata, Via Crescimbeni 14, 62100 Macerata, Italy
[d] Department of Management, Universitá Politecnica delle Marche, Piazzale Martelli 8, 60121 Ancona, Italy

ARTICLE INFO

Keywords:
Indifferent choice
Risk aversion
Prudence
Temperance

ABSTRACT

We propose and analyze three procedures to elicit participants' behavior regarding their higher-order risk attitudes. Our procedures relax the rigidity of having to choose among only two alternatives (as usual in the literature), and allow for choosing a third or even a fourth alternative. This way, the remaining choices for the two original alternatives carry lower levels of noise, i.e., lower randomness in behavior from the subjects without a strong preference in the binary setting, that are forced to choose among only two alternatives. Although the proposed procedures only filter out a small proportion of the choices, the effect on the consistency of the individual behavior across moments is highly significant.

1. Introduction

We propose an experiment that sheds light on a relevant question approached by the behavioral econ literature on decision-making under uncertainty. Starting from the well-known mean–variance paradigm developed in Markowitz (1952), the theory usually assumes that investors try to increase the expected return (first moment of the return distribution) and to decrease the risk computed as return variance (second moment of the return distribution). However, many pieces of research show that higher-order risk preferences are important determinants of choices under uncertainty in a large domain of economic and financial issues, e.g. portfolio management models (Ñíguez et al., 2016), professional path decisions (Schaap, 2021), cryptocurrencies market oscillations (Nagy and Benedek, 2021), etc.

Preference for a higher third moment is usually called prudence (Kimball, 1990), while preference for a lower fourth moment is usually called temperance (Kimball, 1992). Relying on the model-independent concept of risk apportionment proposed by Eeckhoudt and Schlesinger (2006), many experiments in the literature elicit the preferences of individuals on risk aversion, prudence, and temperance by employing binary lotteries (Noussair et al., 2014; Deck and Schlesinger, 2014; Ebert and Wiesen, 2014; Deck and Schlesinger, 2018). However, these experiments usually elicit behavior forcing participants to choose between one of two proposed options (binary comparisons), ignoring the possibility that sometimes subjects might not be able to express a clean, unequivocal preference. This happens if the subjects are either close to be indifferent between the two offered alternatives, or they have some trouble in understanding complex lotteries. Hence, the binary comparison procedure naturally incorporates high levels of noise

* Corresponding author.
E-mail addresses: annarita.colasante@unitelmasapienza.it (A. Colasante), jagarcia@uji.es (J. García-Segarra), luca.riccetti@unimc.it (L. Riccetti), russo@uji.es (A. Russo).

[1] All the authors contributed equally.

https://doi.org/10.1016/j.frl.2022.103270
Received 8 June 2022; Received in revised form 8 August 2022; Accepted 19 August 2022
Available online 27 August 2022
1544-6123/

whenever subjects happen not to have a strong preference between *Left* and *Right*, but are still forced to choose among only these two alternatives.

Such noise could have a critical effect on the individual consistency of the higher-order risk preferences. We refer to individually consistent behavior (as in Noussair et al., 2014) when a subject prefers low variance (risk aversion), high skewness (prudence), and low kurtosis (temperance) of the return distribution.[2] The current work presents an innovative experimental design that, by relaxing the constraint of having only two alternatives, aims at checking the robustness of results collected in previous experiments. Hence, we provide the possibility to choose *Left*, *Right*, *I am indifferent* (Treatment 2), *I don't know* (Treatment 3), or both (Treatment 4), in order to check whether the consistency of the individual behavior improves when facing these elicitation tasks.

There is a large strand of literature focusing on the elicitation of higher moments: even though results show correlation between risk aversion and prudence (i.e., consistent behavior), results about temperance are not so explicit (see, for instance, Deck and Schlesinger, 2014; Haering et al., 2017; Baillon et al., 2018). Noussair et al. (2014) run an experiment with a large and heterogeneous sample where students (the usual target population of lab experiments) are only a small part. Unlike results from other experiments, their results show that individuals' decisions are consistent with risk aversion, prudence, and temperance. However, Deck and Schlesinger (2014) contested the conclusions in Noussair et al. (2014) suggesting that the individual consistency they found may be due to their large online sample. Deck and Schlesinger (2014) pointed out that this individual consistency partially vanished when solely the sample from the lab experiment was considered. Indeed, when focusing on the laboratory data, only remains a significant strong positive correlation between risk aversion and temperance. Moreover, Deck and Schlesinger (2014) stated that there is no significant correlation between risk aversion and prudence in the laboratory data of Noussair et al. (2014), and although the correlation between prudence and temperance is positive, it is low and marginally significant.

This paper sheds light on this specific question. Following the approach proposed in Noussair et al. (2014) we used the very same lotteries in compounded form, and we filtered some choices over alternatives that are not unequivocal preferences for *Left* or *Right*. We rely on the findings of Alós-Ferrer and Garagnani (2021) who show that the strength of the preference over binary alternatives is monotonically related to choice consistency, i.e., choices are more inconsistent (consistent) when the preferences over the alternatives are closer to (further from) indifference. Hence, when a subject chooses *I am indifferent* or *I don't know*, she is actually revealing that her preference over *Left* and *Right* is closer to indifference than when a subject directly chooses *Left* or *Right*. Therefore, we hypothesized that if we allow non-binary comparisons (to be filtered), then the choices that remain for *Left* or *Right* will display a significant improvement in the consistency of individual behavior on higher-order tasks.

The rationale for our prediction is that by relaxing the rigidity of the binary comparisons, we can distinguish the choices with a strong and unequivocal preference for *Left* or *Right*. Therefore, according to Alós-Ferrer and Garagnani (2021), a higher degree of consistency in the individual behavior should emerge.

The reader might wonder whether our manipulation also produces significant treatment effects in the prevalence of the different higher moments. We did not have a preexisting hypothesis about this. However, Deck and Schlesinger (2018) would suggest that there are no grounds to expect a systematic shift in behavior across the different higher moments. Finally, we conjectured that the higher the proportion of non-unequivocal choices elicited by one of these three competing methods, the higher the degree of consistency in the individual behavior. However, we obtained that the method filtering fewer choices was the one eliciting higher individual consistency.

2. Design and procedures

The experiment was conducted in the Laboratori d'Economia Experimental (LEE) at Universitat Jaume I and programmed in z-Tree (Fischbacher, 2007). The experiment consisted of four treatments. Treatment 1, with $N = 80$, consisted of the exact setting in Noussair et al. (2014).[3] The other three treatments, with $N = 120$ each, were three different competing elicitation methods. We included these different manipulations to test the main hypothesis. Hence, this experiment had in total $N = 440$ participants. The experiment had a between-subjects design (allowing for checking treatment effects), jointly with the individual decision-making approach required to check the main hypothesis. Sessions lasted around 90 min (including instruction reading), and the average payoff, including the show-up fee,[4] was EUR 9 (USD 10.5 at the time of the experiment) ranging from EUR 5 to EUR 175.

Fig. 1 illustrates the different manipulations in the treatments. Treatment 1 (T1) forces subjects to choose among the canonical dichotomized alternatives (as in Noussair et al., 2014). Treatment 2 (T2) allows showing *indifference* among the two alternatives by clicking both boxes. Treatment 3 (T3) does not allow showing indifference, but includes a third box, titled *I don't know*, between the options *Left* and *Right*. Finally, Treatment 4 (T4) allows for both showing *indifference* by clicking *Left* and *Right*, and also displays the third box in between to select *I don't know*.[5]

[2] Previous literature refers to this behavior as *all moments risk aversion* (Colasante and Riccetti, 2020) or *mixed risk aversion* (Caballe and Pomansky, 1996).

[3] We run an ex-ante power test to determine the sample size. We used the G*Power software. By setting $\alpha = 0.05$ (error probability), $(1 - \beta) = 0.9$ (the power of the test), and expecting an alternative hypothesis of $\rho = 0.33$, the recommended size is around 75 observations. We rounded up to 80 for Treatment 1 (where subjects can only choose one out of two alternatives) and up to 120 for the other three treatments because in these treatments subjects can pick among 3 or 4 alternatives and potential issues of lack of consistency might arise.

[4] We deploy the same mechanism payment of Noussair et al. (2014): besides the show up fee, each subject has a 1 in 10 chance to win an amount between 5 and 175 euros.

[5] Of course, in Treatment 4 the participants had the freedom to click all three boxes and report inconsistent preferences. For instance, they could choose to be simultaneously in favor of *Left*, *Right*, or *I'm indifferent*, but on top of that, at the same time without knowing whether they actually preferred *Left*, *Right*, or indifference. We only had 4 out of 120 participants in the whole experiment who showed this inconsistent behavior in some of their choices. More concretely, only one subject clicked all the three boxes, and the other three clicked *I don't know* plus *Right* or *Left*. If we remove data from these four participants, the results do not change.

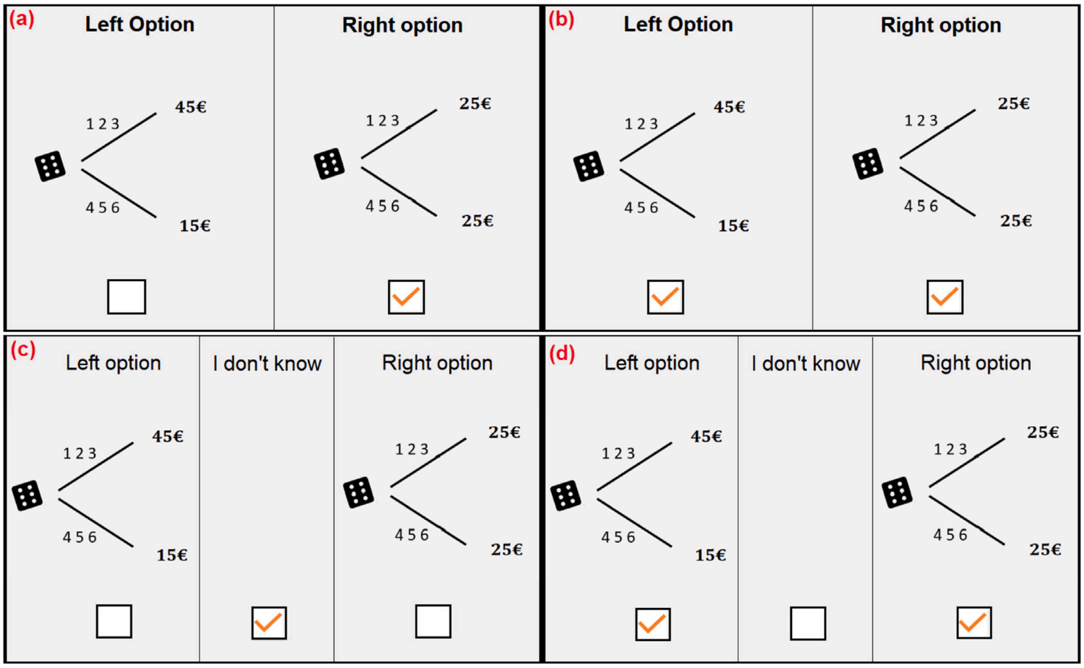

Fig. 1. Sketch of the decision screen in each treatment. Panel (a) shows the classical binary choice between two lotteries corresponding to Treatment 1. Panel (b) refers to Treatment 2 in which subjects may manifest *indifference*. Panel (c) refers to Treatment 3 in which we include the additional option *I don't know*. Panel (d) shows the screen in Treatment 4 in which both *I am indifferent* and *I don't know* have been introduced.

Table 1
Spearman rank correlation coefficients across treatments. The Table 1 shows the consistency of the individual behavior by comparing the Spearman rank correlation coefficients between the higher-order risk attitudes in each treatment. The table includes the p-values and the adjusted p-values using the Bonferroni method to correct for multiple-hypotheses testing.

	Treat. 1		Treat. 2		Treat. 3		Treat. 4	
	Risk	Prudence	Risk	Prudence	Risk	Prudence	Risk	Prudence
Prudence	−0.044		0.175		**0.217**		0.204	
(*p*-value)	(0.699)		(0.056)		(0.017)		(0.025)	
(adj. *p*-value)	(1.000)		(0.168)		(0.051)		(0.075)	
Temperance	0.208	0.198	0.307	0.261	**0.348**	**0.465**	0.154	0.358
(*p*-value)	(0.064)	(0.077)	(0.000)	(0.004)	(0.000)	(0.000)	(0.099)	(0.000)
(adj. *p*-value)	(0.191)	(0.233)	(0.002)	(0.012)	(0.000)	(0.000)	(0.297)	(0.000)

3. Results

Table 1 illustrates our main result. In T1, when only binary comparison among alternatives is allowed, the relative Spearman rank correlation coefficients deliver similar conclusions to the previous work, i.e., very mild individual consistent behavior that is not significant at all once controlled for multiple-hypotheses testing (Noussair et al., 2014; Deck and Schlesinger, 2014, 2018). In contrast, once we allow for non-binary comparisons (as in T2, T3, and T4), by allowing a third or fourth alternative to filter the non-unequivocal choices, the coefficients reveal a stronger consistency of the individual behavior across the moments. This phenomenon can be observed in all three treatments and supports our main hypothesis. Therefore, according to the data, the consistency of the individual behavior when facing higher-order elicitation tasks improves when allowing for intermediate alternatives.[6]

At first glance, T1 of our experiment succeeded in replicating the existing results in Noussair et al. (2014), as intended. The number of safe, prudent, and temperate choices in each treatment (jointly with other descriptive statistics) are included in the online Appendix A1 (Table 4).

Regarding the existence of significant treatment effects in the prevalence of the different higher moments, Table 2 displays the pairwise comparisons across treatments. None of these comparisons found significant differences in the prevalence of the higher moments. Hence, data show no significant differences in the prevalence of risk aversion, prudence, and temperance when allowing

[6] The drop in the correlation's coefficients (and significance) when varying the setting from T3 to T4 suggests that an optimal degree of intervention might exist to raise individual consistency, which consists of allowing only a third option under the more general label of "I do not know".

Table 2
Prevalence of risk aversion, prudence, and temperance across treatments. The first column shows the pairwise comparison between Treatment 1 and the three proposed manipulations. We run nine Mann Whitney Wilcoxon tests. The second column shows the statistical z of every test. The third column displays the p-value. Finally, we adjusted the p-values in the fourth column using the Bonferroni method to correct for multiple-hypotheses testing.

	Compared treatments	z	p-value	adj. p-value
Risk	T1 vs. T2	1.247	0.212	0.637
	T1 vs. T3	1.105	0.269	0.808
	T1 vs. T4	1.852	0.064	0.192
Prudence	T1 vs. T2	0.589	0.556	1.000
	T1 vs. T3	0.744	0.457	1.000
	T1 vs. T4	1.467	0.142	0.427
Temperance	T1 vs. T2	−0.769	0.442	1.000
	T1 vs. T3	0.239	0.811	1.000
	T1 vs. T4	0.705	0.481	1.000

Table 3
Proportion of non-unequivocal choices across treatments. The table displays the proportion of non-unequivocal choices for each lottery among treatments.

	T2	T3	T4		
	Indiff.	Don't know	Indiff.	Don't know	T4 Total
Risk:					
Lottery 1	5/120	6/120	8/120	8/120	16/120
Lottery 2	6/120	3/120	6/120	6/120	12/120
Lottery 3	2/120	4/120	3/120	1/120	4/120
Lottery 4	4/120	2/120	2/120	3/120	5/120
Lottery 5	3/120	2/120	5/120	2/120	7/120
Prudence:					
Lottery 6	3/120	3/120	9/120	3/120	12/120
Lottery 7	10/120	7/120	13/120	3/120	16/120
Lottery 8	7/120	2/120	11/120	1/120	12/120
Lottery 9	10/120	1/120	13/120	3/120	16/120
Lottery 10	11/120	2/120	8/120	6/120	14/120
Temperance:					
Lottery 11	7/120	3/120	8/120	7/120	15/120
Lottery 12	8/120	6/120	12/120	4/120	16/120
Lottery 13	6/120	1/120	6/120	4/120	10/120
Lottery 14	7/120	2/120	5/120	6/120	11/120
Lottery 15	3/120	2/120	10/120	3/120	13/120
Total	**92/1800**	**78/1800**	119/1800	60/1800	**179/1800**
% non-uneq.	**5.11%**	**4.33%**			**9.94%**

for binary comparisons compared to the new procedures allowing for three or four alternatives. Although we did not have a preexisting hypothesis on this issue, we can conclude that the results of Noussair et al. (2014) and Deck and Schlesinger (2018) are robust, and they still hold when relaxing the rigidity of having to pick among two alternatives.

Finally, Table 3 shows the proportion of non-unequivocal choices in each lottery across the three competing methods to elicit behavior. First, as one might expect, we observe that the number of unequivocal choices is greater for lotteries eliciting risk than in lotteries eliciting prudence and temperance. This could be due both to the fact that the structure of the lottery is more complex, and to the fact that the preferences for higher-order moments could be less intense. Interestingly, contrary to our expectations, the method that elicited greater consistency of the individual behavior across higher-order risk attitudes is the one that filtered the lower proportion of non-unequivocal choices (4,33%).[7] More concretely, it is in T3 (i.e., when allowing a third alternative as a box in between *Left* and *Right* under the label *I don't know*) that the correlation coefficients are higher (and significant) across all the moments (risk vs. prudence, $\rho = 0.021$, adj p-value $= 0.051$; risk vs. temperance, $\rho = 0.348$, adj p-value < 0.001; and prudence vs. temperance, $\rho = 0.465$, adj p-value < 0.001). This result is remarkable since the manipulation included in T3 only collected 78 non-unequivocal choices out of 1800 (this represents only 4.33% of the choices). In contrast, T2 collected 92 non-unequivocal choices out of 1800 (5.11%) and T4, 179 non-unequivocal choices out of 1800 (9.94%), which is more than twice the number of

[7] Filtered means that our procedures allow for distinguishing the strength of the preference for the alternatives included in the binary paradigm. That is, when an intermediate alternative, e.g., *I don't know* is included in the choice set, the collected *Left* choices capture a higher strength of the preference for *Left* in comparison with a binary setting where some *Left* choices can just be the result of random behavior of the subjects without a clear preference of *Left* over *Right*. This filtering is what produces an improvement in the degree of individual consistency across the high-order moments.

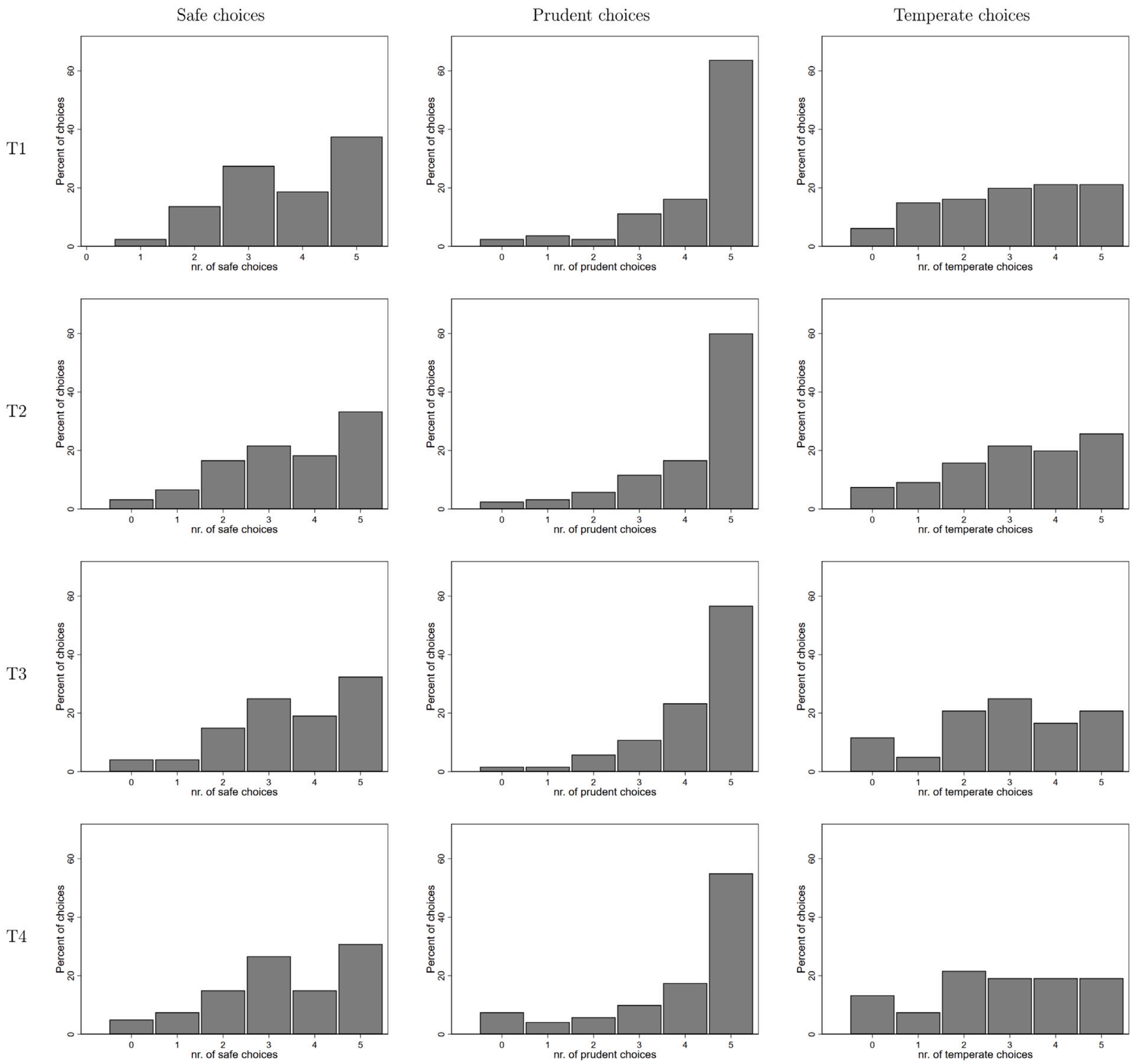

Fig. 2. Histograms of choices in the four Treatments.

non-unequivocal choices in T3. Therefore, data suggest that relaxing the rigidity of having to choose among only two alternatives only affects a small proportion of the choices. However, it remarkably affects the consistency of individual behavior across moments.

As a robustness check, Fig. 2 shows the distribution of the safe, prudent, and temperate choices, for each of the five lotteries eliciting each high-order risk attitude, and for each of the four treatments. Data show that the distribution of choices is very similar across the treatments, suggesting that our manipulation (rather than sample assumptions) is what drives our results.[8]

The Appendix A.2 includes some analyses considering the demographic characteristics of the subjects, mainly gender. As in Noussair et al. (2014), and Ebert and Wiesen (2014), we obtain that in general, women are significantly more temperate than men ($N = 430$, 214 males and 216 females, $Z = -2.254$, p-value $= 0.024$).

4. Conclusion

Our experiment contributes to the debate on whether lab experiments display a statistically significant pairwise correlation between risk aversion, prudence, and temperance. This fact would imply consistency of the individual behavior when facing

[8] We thank an anonymous referee for suggesting this picture.

elicitation tasks dealing with higher-order risk preferences. The results of Noussair et al. (2014) that claimed such effect in their study have been contested because the evidence was only robust for their larger (and over-powered) online set of responders, and it failed to find robust evidence in the experimental data collected in the lab (Deck and Schlesinger, 2014). In contrast, we propose three alternative procedures that obtain a higher and statistically significant pairwise correlation between the moments. Counterintuitively, the procedure that filters less than 5% of the choices displays the higher correlation coefficients and the most solid evidence about the existence of consistent individual behavior when facing higher-order risk preferences in the lab.

In a nutshell, we find that, contrary to the lack of individual consistency reported by existing literature (i.e., lack of individual correlation between risk aversion, prudence, and temperance), in general, people who display preferences for a low variance of the return distribution also prefer high skewness (prudence) and low kurtosis (temperance) of such distribution. As an application of our findings, professionals in the finance sector could use our manipulation by means of questionnaires to elicit the degree of individual consistency of their clients, and then offer them products that fit their preferences better, taking into account not only the variance of the return distribution but also the skewness and kurtosis.

Regarding future avenues of research, beyond the seminal contributions of Kahneman and Tversky (1979), recent literature still reports differences in behavior depending on whether the payoffs are framed as gains or losses in a wide variety of environments, from voting decisions (Alós-Ferrer et al., 2021) to higher-order risk preferences (Bleichrodt and van Bruggen, 2022). It remains an open question whether the procedures introduced here have the same effect on improving individual consistency in lotteries framed as losses from a reference point.

CRediT authorship contribution statement

Annarita Colasante: Conceptualization, Stated the hypotheses, Supervised the final version, Programmed the experiment, Data curation. **Jaume García-Segarra:** Conceptualization, Stated the hypotheses, Supervised the final version, Drafted the paper. **Luca Riccetti:** Conceptualization, Stated the hypotheses, Supervised the final version. **Alberto Russo:** Conceptualization, Stated the hypotheses, Supervised the final version.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We thank Carlos Alós-Ferrer, Alexander Ritschel, the editor, and two anonymous referees for their insightful comments. Jaume García-Segarra and Alberto Russo acknowledge financial support from Universitat Jaume I, Spain and Generalitat Valenciana, Spain under the grants UJI-B2020-16 and AICO/2021/005.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.frl.2022.103270.

References

Alós-Ferrer, C., Garagnani, M., 2021. Choice consistency and strength of preference. Econom. Lett. 198, 109672.
Alós-Ferrer, C., Garagnani, M., García-Segarra, J., 2021. The role of payoff valence on voting: Egalitarian for gains and selfish for losses. Front. Psychol. 12, 737225.
Baillon, A., Schlesinger, H., van de Kuilen, G., 2018. Measuring higher order ambiguity preferences. Exp. Econ. 21 (2), 233–256.
Bleichrodt, H., van Bruggen, P., 2022. The reflection effect for higher-order risk preferences. Rev. Econ. Stat. 104 (4), 705–717.
Caballe, J., Pomansky, A., 1996. Mixed risk aversion. J. Econom. Theory 71 (2), 485–513.
Colasante, A., Riccetti, L., 2020. Risk aversion, prudence and temperance: It is a matter of gap between moments. J. Behav. Exp. Financ. 25 (100262), 1–22.
Deck, C., Schlesinger, H., 2014. Consistency of higher order risk preferences. Econometrica 82 (5), 1913–1943.
Deck, C., Schlesinger, H., 2018. On the robustness of higher order risk preferences. J. Risk Insurance 85 (2), 313–333.
Ebert, S., Wiesen, D., 2014. Joint measurement of risk aversion, prudence, and temperance. J. Risk Uncertaint. 48 (3), 231–252.
Eeckhoudt, L., Schlesinger, H., 2006. Putting risk in its proper place. Amer. Econ. Rev. 96 (1), 280–289.
Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Haering, A., Heinrich, T., Mayrhofer, T., 2017. Exploring the consistency of higher-order risk preferences. Ruhr Economic Papers 688.
Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–291.
Kimball, M., 1990. Precautionary saving in the small and in the large. Econometrica 58 (1), 53–73.
Kimball, M., 1992. Precautionary motives for holding assets. In: Newman, P., Milgate, M., John Eatwell (Eds.), The New Palgrave Dictionary of Money and Finance. Stockton Press, pp. 158–161.
Markowitz, H., 1952. Portfolio selection. J. Finance 7 (1), 77–91.
Nagy, B.Z., Benedek, B., 2021. Higher co-moments and adjusted Sharpe ratios for cryptocurrencies. Finance Res. Lett. 39, 101543.
Ñíguez, T.-M., Paya, I., Peel, D., 2016. Pure higher-order effects in the portfolio choice model. Finance Res. Lett. 19, 255–260.
Noussair, C., Trautmann, S., Van de Kuilen, G., 2014. Higher order risk attitudes, demographics, and financial decisions. Rev. Econom. Stud. 81 (1), 325–355.
Schaap, R.-J., 2021. The prevalence of prudence in a risky occupation. Econom. Lett. 207, 110019.