# A Criticism of Bernheim, Royer, & Sprenger (2022 AEA P&P)

Peter P. Wakker
Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands; Wakker@ese.eur.nl
June, 2022

## 1 Introduction

Bernheim & Sprenger (2020), BS henceforth, claimed to falsify rank dependence. Abdellaoui, Li, Wakker, & Wu (2020), AL henceforth, criticized BS. Bernheim, Royer, & Sprenger (2022), BRS henceforth, redid part of BS's experiment. I first criticize BRS's experiment, and then criticize them for ignoring AL's other criticisms of BS.

## 2 Experimental problems

BRS improve stimulus explanations and avoid a cancellation heuristic and fatigue, as common in preceding studies (Diecidue, Wakker, & Zeelenberg 2007, DWZ henceforth; Weber & Kirsner 1997). However, contrary to their claims, BRS worsen rather than improve incentives. AL (p. 2 l. 8, p.8 l. 9) emphasized that *differences* in outcomes ($m, k$ in BRS's notation), rather than outcomes themselves, provide incentives. Unfortunately, BRS mostly increase outcomes but not their differences. In only one incentivized choice pair ("Condition 5") difference is increased ($m = 20$ iso of $m = 5$), though still below the minimal difference in DWZ. Incentives are worsened because only few subjects were paid, 1/20$^{th}$. Remarkably, BRS's Figure 1 does suggest rank dependence for this choice pair, but BRS give no statistical analysis of it.

Apart from a trivial test of stochastic dominance ($k^+$ vs. $k^-$), all statistical conclusions in BRS are based on accepted null hypotheses. Every statistics textbook warns against this (Greenland et al., 2016, Misinterpretation 4). Power analyses would have been warranted.

## 3 Ignoring preceding criticisms

Besides concerning the stimuli used, BRS ignore, and give no counterarguments to, all AL's warnings. Thus, AL (end of §2) pointed out that BS (§3.2) erroneously claimed to identify utility and probability weighting whereas those are unidentifiable from their stimuli. But BRS's Finding 2 still continues to use these invalid measurements.

AL (p. 11 last para) pointed out that BS's equalizing reductions are not new, but were used before by DWZ, who also tested quantitative measurements of rank dependence. BRS (2$^{nd}$ para) nevertheless do not cite DWZ and continue to falsely claim novelty. DWZ considered the more interesting domain of uncertainty rather than risk, used bigger outcome *differences*, and based their findings (both cases of rank dependence and violations thereof) on significant statistics. They also explained that, contrary to BS's claims, the method is not valid for general utility but only if utility is linear for the incentives used (AL para on pp. 11-12). This explains why the equalizing reductions method never became popular. Modern quantititive measurements allow for nonlinear utility.

So as to disregard novelty of many other preceding violations of prospect theory, BS claimed that counting statistics, used in such studies, are generally invalid, even though they are used in all empirical scientific disciplines and are explained in all textbooks. AL (§6.1) pointed out that BS's

claim is invalid, so that priority of the many preceding falsifications of prospect theory should be acknowledged. BRS (end of §1) nevertheless maintain their claim.

Several other problems of BS, pointed out by AL, are not reconsidered by BRS and remain in need of fixing. The alternatives to rank dependence suggested by BS, rank-independent probability weighting and complexity aversion[1] are invalid (AL §2 2nd problem; §6.3). BS erroneously equated separable prospect theory with 1979 prospect theory (AL §2 1st problem), and BRS continue to do so. BRS's null hypotheses and Finding 1 do not falsify rank dependence but only show neutrality, in the same way as expected value does not falsify expected utility but only shows neutrality (AL §6.4). For real falsification, BRS should not only have solved the aforementioned problems of their Finding 1, but also the many problems of their Finding 2 (which they do not reconsider).

Even if BRS had avoided all experimental problems, then their study still would have contributed nothing because of the remaining theoretical problems and the more informative preceding findings in the literature.

# 4 Conclusion

Bernheim, Royer, & Sprenger (2022) improve some aspects of BS's first experiment to make them meet the standards of the field, but unfortunately worsen incentives and only report null hypotheses. They repeat many claims by BS proved wrong by Abdellaoui et al. (2000). In particular, their novelty claims ignore many preceding negative, and many more positive, findings on rank dependence and prospect theory that, contrary to BRS's claims, were based on valid statistical significance.

It would have been surprising if BS and BRS had been the first to "properly" test rank dependence, 40 years after its introduction by Quiggin (1982) for risk[2], 30 years after its inclusion in prospect theory (Tversky & Kahneman 1992), and 20 years after its shared prize in memory of Nobel (2002). BS's suggested alternatives of rank-independent weighting and "complexity aversion" (their misnomer) have long been known to be inviable.

[1] The online appendix of AL, at the end of their file, explains that BS's term is a misnomer. The literature uses the term in a broader and more interesting way than BS do. Their dependence on number of outcomes (also their nonlinear weighting) relates more to what is known as event splitting effects.

[2] Schmeidler (1982, 1989) independently introduced rank dependence for the more general and more interesting context of uncertainty, studied by DWZ.

# References

Abdellaoui, Mohammed, Chen Li, Peter P. Wakker, & George Wu (2020) "A Defense of Prospect Theory in Bernheim & Sprenger's Experiment," working paper: https://personal.eur.nl/wakker/pdf/abd.li.wak.wu_bernh.sp.pdf

Bernheim, B. Douglas, Rebecca Royer, & Charles Sprenger (2022) "Robustness of Rank Independence in Risky Choice," *AEA Papers and Proceedings* 112: 415-420.

Bernheim, B. Douglas & Charles Sprenger (2020) "On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting," *Econometrica* 88, 1363–1409.

Diecidue, Enrico, Peter P. Wakker, & Marcel Zeelenberg (2007) "Eliciting Decision Weights by Adapting de Finetti's Betting-Odds Method to Prospect Theory," *Journal of Risk and Uncertainty* 34, 179–199.

Greenland, Sander, Stephen J. Senn, Kenneth J. Rothman, John B. Carlin, Charles Poole, Steven N. Goodman, & Douglas G. Altman (2016) "Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations," *European Journal of Epidemiology* 31, 337–350.

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

Quiggin, John (1982) "A Theory of Anticipated Utility," *Journal of Economic Behaviour and Organization* 3, 323–343.

Schmeidler, David (1982) "Subjective Probability without Additivity," Foerder Institut of Economic Research, Tel Aviv University, Tel Aviv, Israel.

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587.

Tversky, Amos & Daniel Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.

Weber, Elke U. & Britt Kirsner (1997) "Reasons for Rank-Dependent Utility Evaluation," *Journal of Risk and Uncertainty* 14, 41–61.