# Online Appendix: Correction and Re-examination of Stationary Concepts for Experimental 2x2 Games: A Reply

*By* Reinhard Selten, Thorsten Chmura and Sebastian J. Goerg

*JEL: C70, C91*
*Keywords: stationary concepts, action-sampling, payoff-sampling, impulse balance, quantal response, Nash*

### The Fisher-Pitman Permutation test for paired replicates

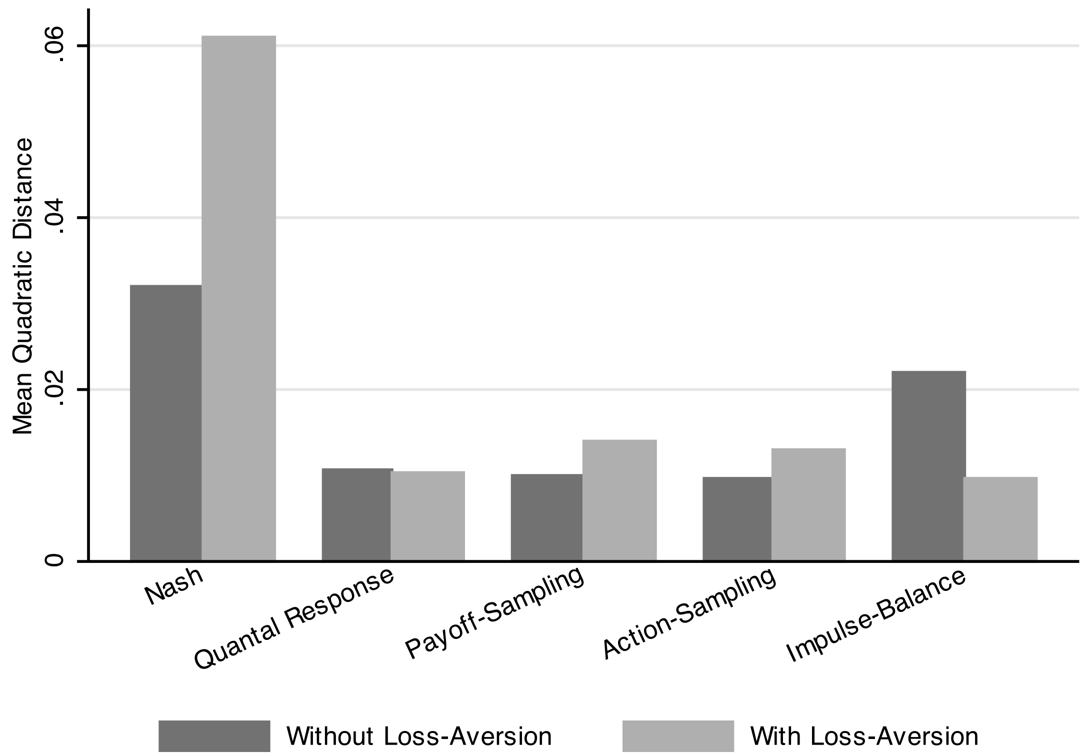
Non-parametric statistical tests are free of assumptions about the probability distributions of the investigated variables. But this does not mean that those tests are in general free of assumptions. Most non-parametric tests have assumptions about the population, from which the variables are drawn. Such assumptions are, e.g., that subjects are randomly drawn from the population and that the distribution of the population is symmetric around the median. Especially the assumption of symmetry around the median might be difficult in the context of quadratic distances, and therefore one might want to avoid the application of the Wilcoxon signed-rank test.

The Fisher-Pitman permutation test for paired replicates (FP)[1] does not draw any conclusions about the underlying population. Therefore it does not depend on assumptions about the population (e.g., subjects being randomly drawn and symmetry around the median). The results of the FP test are only valid for the investigated samples. For the pairwise comparisons in Table 1, the null hypothesis of the FP test is that the paired quadratic distances for one matching group are randomly assigned to the two stationary concepts.

The rationale of the test is as follows: the FP test keeps the paired distances constant, but randomly assigns the quadratic distance to the two labels (our labels are “Quadratic distance to theory A” and ‘Quadratic distance to theory B”) and calculates the difference between the two quadratic distances. This procedure is repeated with a Monte-Carlo algorithm (for Table 1, it was repeated 20.000 times). The p-value is the proportion of permutations, which provide a test statistic at least as large as the one for the correct assignment of quadratic distance and stationary concept.

[1] Refer to Kaiser (2007), The Stata Journal, 7(3) for a Stata implementation of the test.

**Appendix: Original versus Transformed Games**




Note that parameters are adjusted for the following concepts:
Quantal response equilibrium without loss-aversion ($\lambda = 1.05$) and with loss-aversion ($\lambda = 0.845$)
Payoff-sampling equilibrium without loss-aversion ($n = 6$) and with loss-aversion ($n = 3$)
Action-sampling equilibrium without loss-aversion ($n = 12$) and with loss-aversion ($n = 5$)

Figure A.1. Advantages and Disadvantages of Applying a Concept to the Transformed Game Rather Than the Original One (Figure 11 in SC)

Table A.1—Two-sided significances for the comparison of stationary concepts with and without loss-aversion, Monte-Carlo approximation of the Fisher-Pitman permutation test for paired replicates (Rounded to the next higher level among 0.1 percent, 0.2 percent, 1 percent, 2 percent, 5 percent, and 10 percent)

| Loss-aversion | Impulse-balance equilibrium | Action-sampling equilibrium | Payoff-sampling equilibrium | Quantal response equilibrium | Nash equilibrium |
|---|---|---|---|---|---|
| in favor | **0.1 percent**<br>*0.1 percent*<br>0.1 percent | | | **n.s.**<br>-<br>2 percent | |
| in disfavor | | **0.1 percent**<br>*0.1 percent*<br>n.s. | **0.1 percent**<br>*0.1 percent*<br>n.s. | **n.s.**<br>*0.1 percent*<br>- | **0.1 percent**<br>*0.1 percent*<br>0.1 percent |

Notes: Above: all 108 Experiments; Middle: 72 constant-sum game experiments;
Below: 36 non-constant sum game experiments.
Comparison for action-sampling, payoff-sampling and Nash equilibrium in favor of the concepts without loss-aversion, and for impulse-balance equilibrium in favor of loss-aversion. For quantal response equilibrium, the comparison in the constant-sum games is in favor of the concept without loss-aversion and in the non-constant sum games it is in favor of the concept with loss-aversion.