Safe models for risky decisions

Steingröver, H.M.

Appendix C

Appendix to Chapter 4: "Validating the PVL-Delta Model for the Iowa Gambling Task"

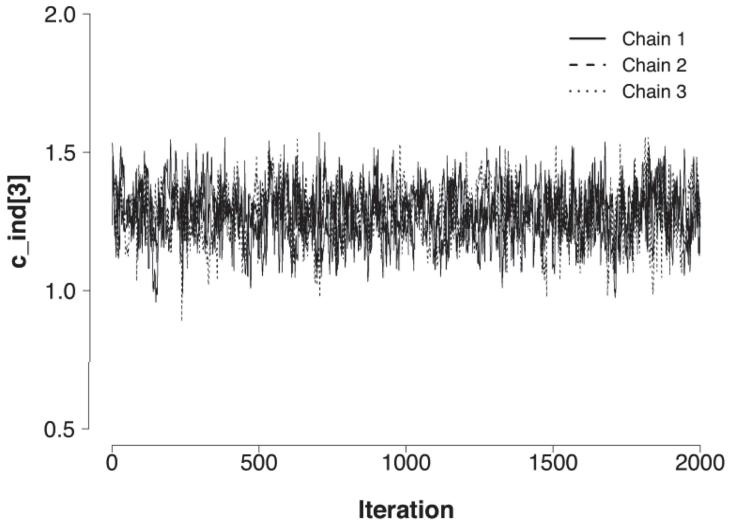
In this appendix, we present how we visually assessed convergence, additional absolute model performance checks and the results of two parameter-recovery studies that confirm that we correctly implemented the PVL-Delta model. For the parameter-recovery studies, we used two synthetic data sets that were generated with the PVL-Delta model. The data-generating parameters correspond to the medians of the individual-level joint posteriors that were obtained by fitting the PVL-Delta model to two real data sets.

Figure C.1 shows the HMC chains of one individual-level parameter. From the figure it is evident that the chains have converged successfully from their starting values to their stationary distribution, looking like "hairy caterpillars" that are randomly intermixed. We inspected this type of plot for every parameter to visually assess convergence in addition to the formal diagnostic measure of convergence $\hat{R}$.

Figure C.2 presents the fit performance and simulation performance of the PVL-Delta model that was obtained with random draws from the joint posterior distributions over the group-level parameters (hereafter group-level joint posteriors). Note that Figure 4.9 presents the fit performance and simulation performance based on the individual-level joint posteriors. A comparison of both figures reveals that the fit performance based on the group-level joint posteriors (Figure C.2) closely matches the fit performance based on the individual-level joint posteriors (Figure 4.9). However, there are a few discrepancies in the case of the simulation performance: From Figures 4.9 and C.2 it is evident that the simulation performance based on the group-level joint posteriors is more extreme, that is, the most preferred deck is preferred even stronger, whereas the least preferred deck is avoided even stronger. However, it is evident that in general Figure C.2 mirrors the conclusion drawn from Figure 4.9.

Figure C.3 presents the results of the first recovery study. This data set contains 18 synthetic participants. The figure contains four panels; each panel illustrates the recovery of one of the four model parameters. In each panel, the mode of the group-level posterior is represented by the dotted line, whereas the solid line represents the true group-level parameter. In addition, the panels can also be used to assess the individual-level recovery: The unfilled dots represent the modes of the

Figure C.1: HCM chains of the individual-level consistency parameter c of the third participant in the consistency condition. In addition to the formal diagnostic measure of convergence $\hat{R}$, we inspected this type of plot for every parameter to visually assess convergence.



individual-level posteriors, whereas the filled dots represent the true individual-level parameters.

Note that the individual-level posterior distributions are not sorted by the subject ID; in order to visualize the degree of individual differences in each model parameter, we sorted the individual-level posterior distributions by the true individual-level parameters.

From Figure C.3 it is evident that the group-level updating parameter is slightly underestimated, but the remaining group-level parameters are recovered very accurately. However, the recovery of the individual-level parameters is less accurate. Especially in the case of the shape parameter, most of the individual-level modes differ from the true individual-level parameters by regressing to the mode of the group-level parameter (i.e., shrinkage); small deviations are noticeable in the case of the individual-level loss aversion parameters and the individual-level updating parameter. Yet, in the case of the consistency parameter, most individual-level parameters are recovered very accurately.

Figure C.4 presents the results of the second recovery study. This data set contains 30 synthetic participants. It is evident that all group-level parameters are recovered very accurately. However, the recovery of the individual-level parameters is less accurate. Especially in the case of the individual-level shape parameters and the individual-level loss aversion parameters, it is evident that the individual-level modes differ from the true individual-level parameters. Yet, the recovery of the individual-level updating parameters and the individual-level consistency parameters is adequate.

Figure C.2: Observed choice behavior and assessment of absolute model performance. The first column shows the mean proportion of choices from each deck within 15 blocks as observed in the four experimental conditions reported by Wetzels, Vandekerckhove, et al. (2010). Each block contains 10 trials. The second and third column show the fit performance and simulation performance, respectively, for each of the four conditions. Fit performance and simulation performance are based on random draws from the Fit performance and simulation performance are based -level joint posteriors.

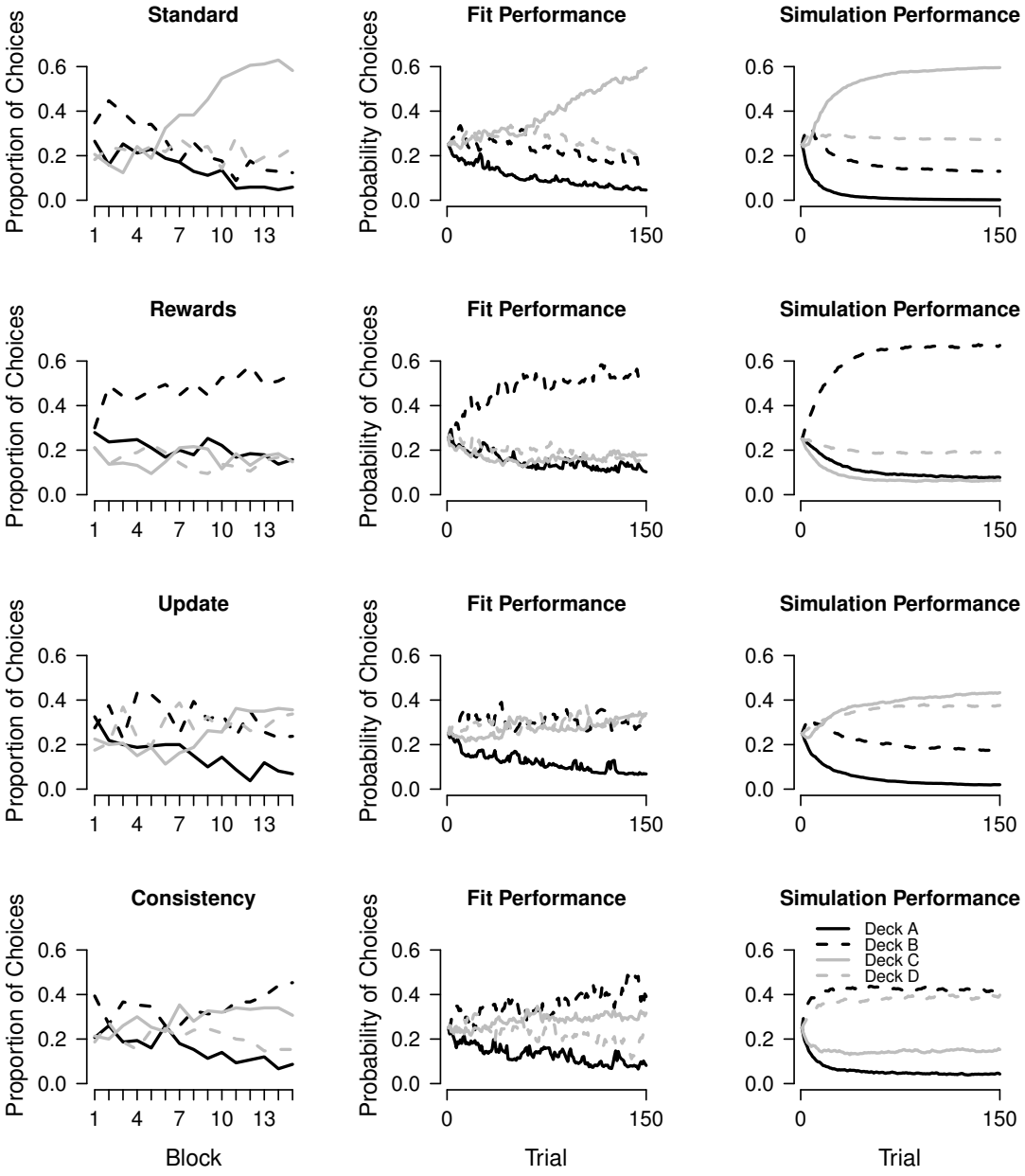


Figure C.3: Recovery of individual and group-level parameters of the PVL-Delta model. Data of 18 participants completing a 100-trial IGT. The dotted lines and the unfilled dots represent the modes of the group-level posteriors and of the individual-level posteriors, respectively. The black lines and black dots represent the true group-level and true individual-level parameters, respectively.

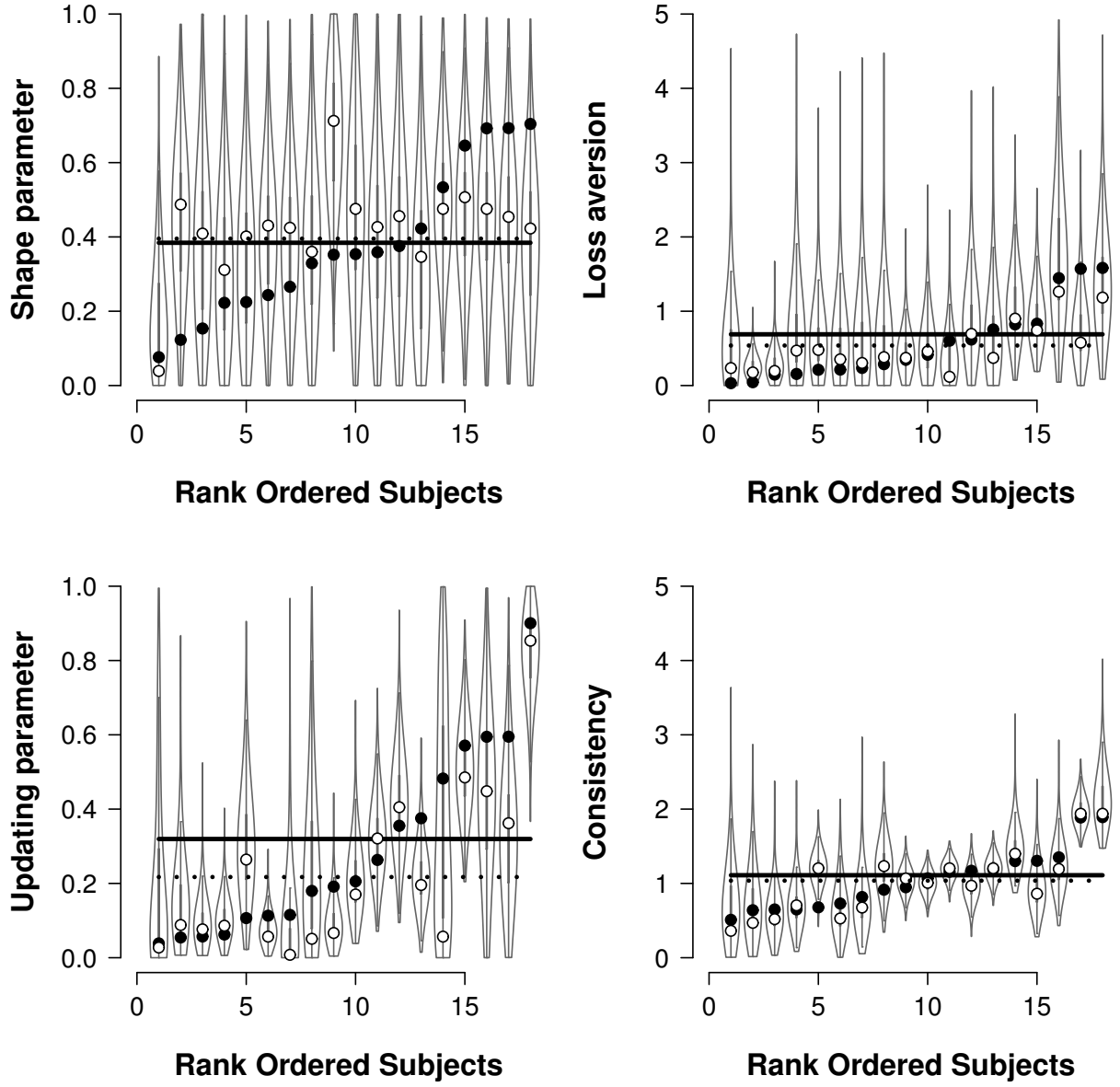


Figure C.4: Recovery of individual and group-level parameters of the PVL-Delta model. Data of 30 participants completing a 100-trial IGT. The dotted lines and the unfilled dots represent the modes of the group-level posteriors and of the individual-level posteriors, respectively. The black lines and black dots represent the true group-level and true individual-level parameters, respectively.