ONLINE SUPPLEMENTARY DATA TO: "THE MANY FACES OF HUMAN SOCIALITY: UNCOVERING THE DISTRIBUTION AND STABILITY OF SOCIAL PREFERENCES"

Adrian Bruhin
(HEC Lausanne), University of Lausanne

Ernst Fehr
University of Zurich

Daniel Schunk
University of Mainz

E-mail: adrian.bruhin@unil.ch (Bruhin); ernst.fehr@econ.uzh.ch (Fehr); daniel.schunk@uni-mainz.de (Schunk)

Journal of the European Economic Association
Preprint prepared on 4 April 2019 using jeea.cls *v1.0.*

Contents

B.1 Monte Carlo simulations: Parameter recovery and impact of serially correlated errors in the aggregate model

This section presents the set-up and the results of a series of four Monte Carlo simulations in the context of the aggregate model. These Monte Carlo simulations serve two main goals: (i) to check whether the experimental design and the econometric strategy allow us to reliably recover the parameters of the behavioral model; and (ii) to assess the extent to which a potential misspecification due to serially correlated errors across games can bias our results.

B.1.1 Set-up

All four Monte Carlo simulations have an analogous set-up. First, we define a vector $\psi = (\alpha, \beta, \gamma, \delta, \sigma)$ containing the true parameters of the behavioral model and the choice sensitivity. The below table shows the true parameters for each of the four Monte Carlo simulations.

	True parameters				
Simulation	α	β	γ	δ	σ
1	0.00	0.00	0.00	0.00	0.01
2	0.20	0.50	0.15	-0.10	0.01
3	0.05	0.10	0.00	-0.05	0.01
4	-0.35	0.00	0.10	-0.05	0.01

The true parameters of Simulation 1 reflect completely selfish behavior. In contrast, the true parameters of Simulations 2, 3, and 4 correspond roughly to the parameter estimates of the SA-, MA-, and BA-type, respectively. The choice sensitivity σ is equal to 0.01 in all four simulations, implying a somewhat higher decision noise than estimated in the paper.

After defining the true parameters, we execute $R = 1{,}000$ simulation runs for each of the four simulations. Each simulation run r consists of two steps:

a) In the first step, we simulate the choices of $N \in \{1, 20, 160\}$ subjects in the 117 dictator and reciprocity games based on the true parameters and the random utility model presented in section 3.1 of the paper.

   Varying the number of subjects allows us to investigate how N affects the potential bias and the overall accuracy of our estimations: $N = 1$ corresponds to the amount of data available in the individual estimations; $N = 20$ roughly corresponds to the number of subjects classified as BA-types – the smallest preference type that the finite mixture model identifies; and $N = 160$ corresponds to the total number of subjects that we use in the estimations in the paper.

   We also intend to assess the extent to which a potential misspecification due to serially correlated errors across games can bias our results. Therefore, the simulated random errors in a subject's utility follow an AR(1) process across games with serial correlation $\rho \in \{0, 0.2, 0.4, 0.6, 0.8\}$ and type-1-extreme-value-distributed innovations. If $\rho = 0$, there is no serial correlation in the errors and the random utility model in section 3.1 is specified correctly. However, if $\rho > 0$, the errors are serially correlated across games and the model is misspecified, potentially leading to a biased estimator. Moreover, the order of games within the blocks of dictator and reciprocity games is randomly reshuffled in each simulation run. This random reshuffling takes into account that in the experiment, the games were presented in random order within the blocks of dictator and reciprocity games (see section 2.3 in the paper).

b) In the second step, we try to recover the true parameters by estimating the random utility model on the simulated choices. This yields a vector of estimated parameters $\hat{\psi}^{(r)}$ for each simulation run r.

After finishing all R simulation runs, we compare the true and the estimated parameters to assess the potential bias and the overall accuracy of our estimators. The bias in the estimator of parameter j,

$$Bias(\hat{\psi}_j) = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\psi}_j^{(r)} - \psi_j\right), \tag{B.1}$$

indicates whether, on average, we can recover the true parameter, or whether the estimated parameter systematically deviates from the true value. The overall accuracy of the estimator is given by its Mean Squared Error (MSE),

$$MSE(\hat{\psi}_j) = Bias(\hat{\psi}_j)^2 + Var(\hat{\psi}_j) = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\psi}_j^{(r)} - \psi_j\right)^2, \tag{B.2}$$

and corresponds to the sum of the estimator's squared bias plus its variance – an inverse measure for its precision. A MSE close to zero indicates great overall accuracy, while a large MSE indicates a low overall accuracy due to a bias in the estimator and/or a high variance. Considering the overall accuracy and not just bias is particularly important when performing horse races between different estimators, as a biased estimator with a relatively low variance may be overall more accurate than an unbiased estimator with a relatively high variance.

B.1.2 Results

Tables B.1 to B.4 summarize the results in terms of bias and overall accuracy, while Figures B.1 to B.4 visualize the bias in the estimators.

The results are qualitatively identical across all four Monte Carlo simulations. For the estimators of the behavioral model's parameters, α, β, γ, and δ, the results can be summarized as follows:

- If there is no serial correlation in the errors, i.e. $\rho = 0$, the bias is negligible in absolute value for estimations with $N = 1$ subject and approaches zero for estimations with $N = 20$ and $N = 160$ subjects. Hence, even estimations at the individual level allow us to on average recover the true parameters of the behavioral model. However, the MSEs reveal that estimations at the individual level are overall at least an order of magnitude less accurate than estimations at the level of $N = 20$ or $N = 160$ subjects. This highlights the benefit of using more parsimonious models instead of relying on estimations at the individual level.

- If the random errors are serially correlated, i.e. $\rho \in \{0.2,0.4,0.6,0.8\}$, estimations with $N = 1$ subject are biased. The bias is particularly severe for high serial correlations, i.e. $\rho = 0.6$ or

$\rho = 0.8$. However, the bias quickly vanishes if the number of subjects increases: with $N = 20$ subjects, it is negligible for all but the highest values of ρ; with $N = 160$ subjects it is negligible for all values of ρ. Similarly, the MSEs indicate that with $N = 20$ subjects, estimations are overall accurate for all but the highest values of ρ, while with $N = 160$ subjects estimations are overall highly accurate for all values of ρ.

In sum, the above results on the estimators for the behavioral model's parameters indicate that, as long as we do not rely on estimations at the individual level, the experimental design and the empirical strategy allow us to estimate the true parameters with high overall accuracy. This holds regardless of the considered preference type and even if the random errors in the subjects' utility are serially correlated across games.

The results on the estimator of the choice sensitivity σ are different, however. Serial correlation in the random error seems to cause a downward-bias in the estimator of σ that (i) becomes more pronounced when ρ gets bigger and (ii) persists if the number of subjects increases. Hence, we may underestimate how sensitive the subjects' choices react to utility differences. This would lead to overly conservative prediction intervals when using our model to make behavioral predictions.

True Parameter	ρ	Bias $N = 1$	Bias $N = 20$	Bias $N = 160$	MSE $N = 1$	MSE $N = 20$	MSE $N = 160$
$\alpha = 0.00$	0.0	-0.011	-0.002	0.000	0.025	0.001	0.000
	0.2	-0.012	0.000	0.000	0.024	0.001	0.000
	0.4	-0.010	-0.004	-0.002	0.030	0.002	0.000
	0.6	0.186	-0.004	-0.004	6.357	0.002	0.000
	0.8	1.979	0.095	-0.004	46.515	1.948	0.000
$\beta = 0.00$	0.0	-0.007	-0.002	0.000	0.022	0.001	0.000
	0.2	-0.014	-0.003	-0.001	0.025	0.001	0.000
	0.4	-0.026	-0.002	-0.002	0.029	0.001	0.000
	0.6	0.126	-0.003	-0.004	2.896	0.002	0.000
	0.8	2.065	0.157	-0.004	47.311	5.089	0.000
$\gamma = 0.00$	0.0	0.001	0.001	0.000	0.032	0.001	0.000
	0.2	0.002	0.000	0.000	0.033	0.002	0.000
	0.4	0.002	0.000	-0.001	0.038	0.002	0.000
	0.6	0.095	0.002	0.001	4.172	0.002	0.000
	0.8	1.448	0.061	0.001	34.555	0.695	0.000
$\delta = 0.00$	0.0	0.003	0.002	0.000	0.031	0.002	0.000
	0.2	-0.001	0.000	0.000	0.034	0.002	0.000
	0.4	0.004	0.001	0.000	0.041	0.002	0.000
	0.6	0.155	0.002	0.000	3.371	0.002	0.000
	0.8	1.349	0.133	0.001	32.445	3.659	0.000
$\sigma = 0.01$	0.0	0.001	0.000	0.000	0.000	0.000	0.000
	0.2	0.001	0.000	0.000	0.000	0.000	0.000
	0.4	0.000	-0.001	-0.001	0.000	0.000	0.000
	0.6	-0.001	-0.002	-0.002	0.000	0.000	0.000
	0.8	-0.004	-0.004	-0.004	0.000	0.000	0.000

Biases and Mean Squared Errors (MSE) are calculated based on 1,000 simulation runs, each with $N \in \{1,20,160\}$ subjects. The subjects' simulated choices are based on the random utility model presented in section 3.1 of the paper, using the true parameters shown in the first column. To simulate the effect of serially correlated errors across games, the errors of each subject follow an AR(1) process with serial correlation ρ and type-1-extreme-value-distributed innovations. The simulation considers that games were presented in random order within the blocks of dictator and reciprocity games.

Table B.1: Results of Simulation 1 with selfish subjects.

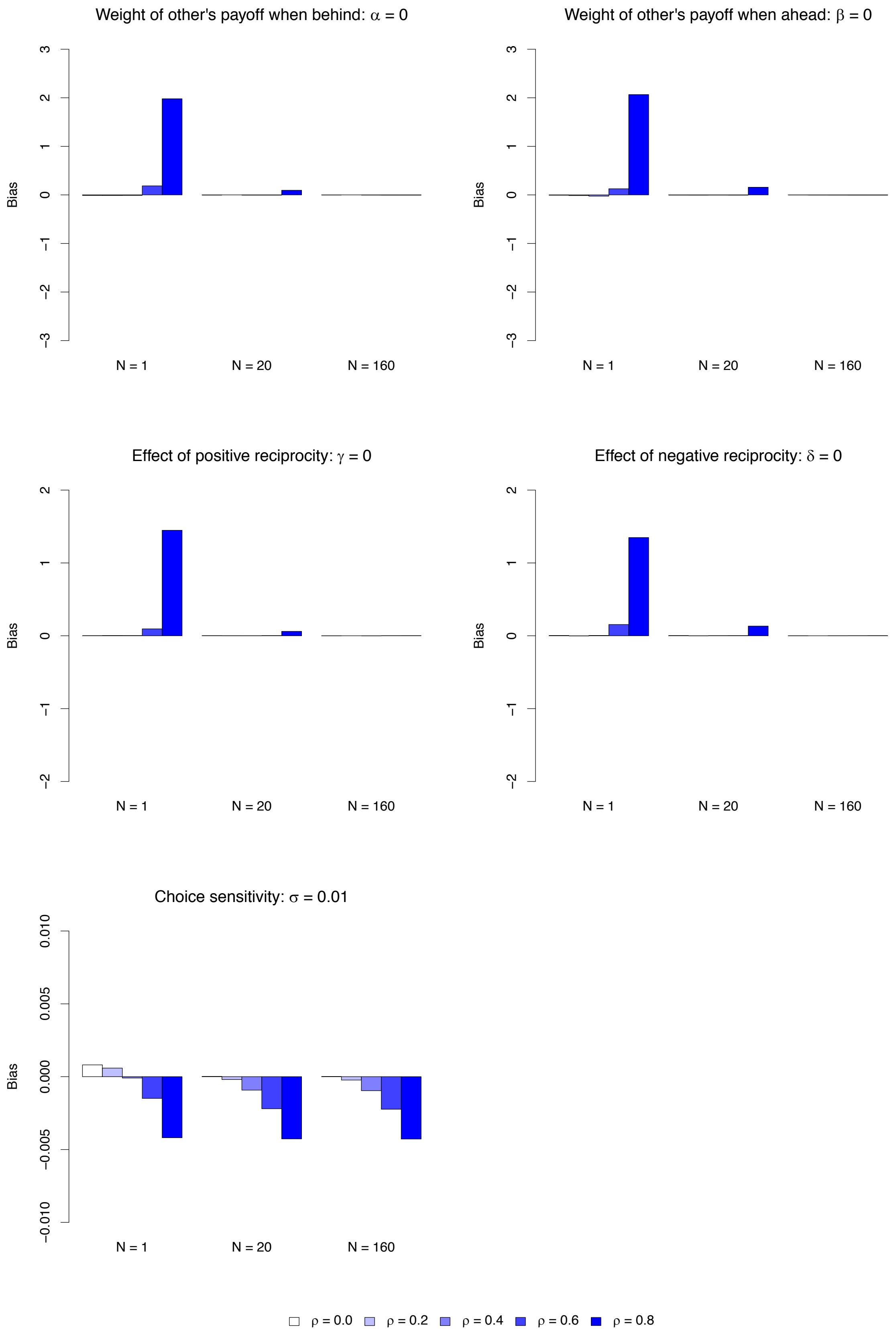

Figure B.1: Bias in parameter estimates in Simulation 1 with selfish subjects.

True Parameter	ρ	Bias $N = 1$	Bias $N = 20$	Bias $N = 160$	MSE $N = 1$	MSE $N = 20$	MSE $N = 160$
$\alpha = 0.20$	0.0	-0.003	-0.001	0.000	0.021	0.001	0.000
	0.2	-0.013	-0.001	0.000	0.022	0.001	0.000
	0.4	-0.010	-0.002	-0.001	0.026	0.001	0.000
	0.6	0.003	-0.002	-0.002	2.486	0.002	0.000
	0.8	0.053	0.021	0.001	40.397	0.295	0.000
$\beta = 0.50$	0.0	0.008	0.000	0.000	0.023	0.001	0.000
	0.2	-0.003	0.000	0.000	0.024	0.001	0.000
	0.4	0.003	0.001	-0.001	0.029	0.001	0.000
	0.6	-0.091	0.001	0.000	4.018	0.002	0.000
	0.8	-1.722	-0.054	0.003	62.936	1.600	0.000
$\gamma = 0.15$	0.0	-0.006	0.000	0.000	0.028	0.001	0.000
	0.2	0.011	0.001	0.000	0.029	0.001	0.000
	0.4	0.001	0.000	0.001	0.032	0.002	0.000
	0.6	-0.093	0.001	0.001	4.132	0.002	0.000
	0.8	-1.487	-0.048	-0.002	45.587	1.292	0.001
$\delta = -0.10$	0.0	-0.009	0.002	0.000	0.028	0.001	0.000
	0.2	0.006	0.001	0.000	0.030	0.001	0.000
	0.4	-0.006	0.000	0.000	0.036	0.002	0.000
	0.6	-0.031	-0.001	0.000	1.287	0.002	0.000
	0.8	0.276	0.023	-0.003	34.558	0.438	0.001
$\sigma = 0.01$	0.0	0.001	0.000	0.000	0.000	0.000	0.000
	0.2	0.000	0.000	0.000	0.000	0.000	0.000
	0.4	0.000	-0.001	-0.001	0.000	0.000	0.000
	0.6	-0.002	-0.002	-0.002	0.000	0.000	0.000
	0.8	-0.005	-0.004	-0.004	0.000	0.000	0.000

Biases and Mean Squared Errors (MSE) are calculated based on 1,000 simulation runs, each with $N \in \{1,20,160\}$ subjects. The subjects' simulated choices are based on the random utility model presented in section 3.1 of the paper, using the true parameters shown in the first column. To simulate the effect of serially correlated errors across games, the errors of each subject follow an AR(1) process with serial correlation ρ and type-1-extreme-value-distributed innovations. The simulation considers that games were presented in random order within the blocks of dictator and reciprocity games.

Table B.2: Results of Simulation 2 with strongly altruistic subjects.

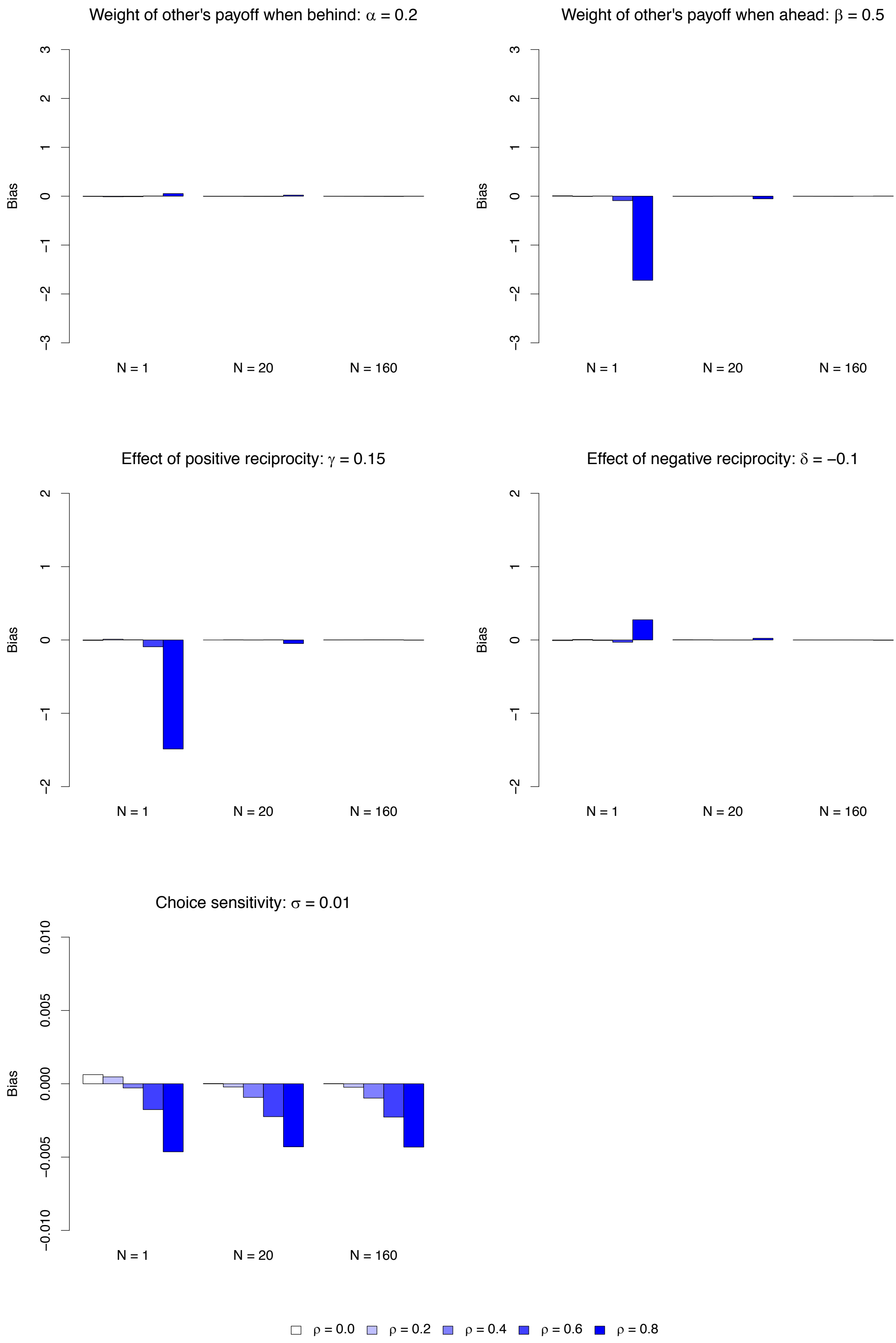

Figure B.2: Bias in parameter estimates in Simulation 2 with strongly altruistic subjects.

True Parameter	ρ	Bias $N = 1$	Bias $N = 20$	Bias $N = 160$	MSE $N = 1$	MSE $N = 20$	MSE $N = 160$
$\alpha = 0.05$	0.0	-0.008	-0.002	0.000	0.023	0.001	0.000
	0.2	-0.008	0.001	0.000	0.026	0.001	0.000
	0.4	-0.019	-0.004	-0.002	0.057	0.001	0.000
	0.6	0.145	-0.003	-0.003	4.159	0.001	0.000
	0.8	2.095	0.089	-0.003	49.951	1.589	0.000
$\beta = 0.10$	0.0	-0.009	-0.002	0.000	0.022	0.001	0.000
	0.2	-0.013	-0.001	-0.001	0.027	0.001	0.000
	0.4	-0.015	-0.001	-0.002	0.029	0.001	0.000
	0.6	0.197	-0.004	-0.003	4.890	0.002	0.000
	0.8	1.722	0.158	-0.002	47.662	4.096	0.000
$\gamma = 0.00$	0.0	-0.005	0.001	0.000	0.029	0.001	0.000
	0.2	-0.002	-0.001	0.000	0.035	0.002	0.000
	0.4	0.015	0.000	0.000	0.139	0.002	0.000
	0.6	0.194	0.001	0.000	3.900	0.002	0.000
	0.8	1.287	0.063	0.001	39.949	0.760	0.000
$\delta = -0.05$	0.0	0.002	0.002	0.000	0.029	0.001	0.000
	0.2	-0.004	-0.002	0.000	0.035	0.002	0.000
	0.4	-0.004	0.001	0.000	0.060	0.002	0.000
	0.6	0.077	0.001	0.000	3.361	0.002	0.000
	0.8	1.244	0.129	0.000	35.938	3.374	0.000
$\sigma = 0.01$	0.0	0.001	0.000	0.000	0.000	0.000	0.000
	0.2	0.001	0.000	0.000	0.000	0.000	0.000
	0.4	0.000	-0.001	-0.001	0.000	0.000	0.000
	0.6	-0.002	-0.002	-0.002	0.000	0.000	0.000
	0.8	-0.004	-0.004	-0.004	0.000	0.000	0.000

Biases and Mean Squared Errors (MSE) are calculated based on 1,000 simulation runs, each with $N \in \{1,20,160\}$ subjects. The subjects' simulated choices are based on the random utility model presented in section 3.1 of the paper, using the true parameters shown in the first column. To simulate the effect of serially correlated errors across games, the errors of each subject follow an AR(1) process with serial correlation ρ and type-1-extreme-value-distributed innovations. The simulation considers that games were presented in random order within the blocks of dictator and reciprocity games.

Table B.3: Results of Simulation 3 with moderately altruistic subjects.

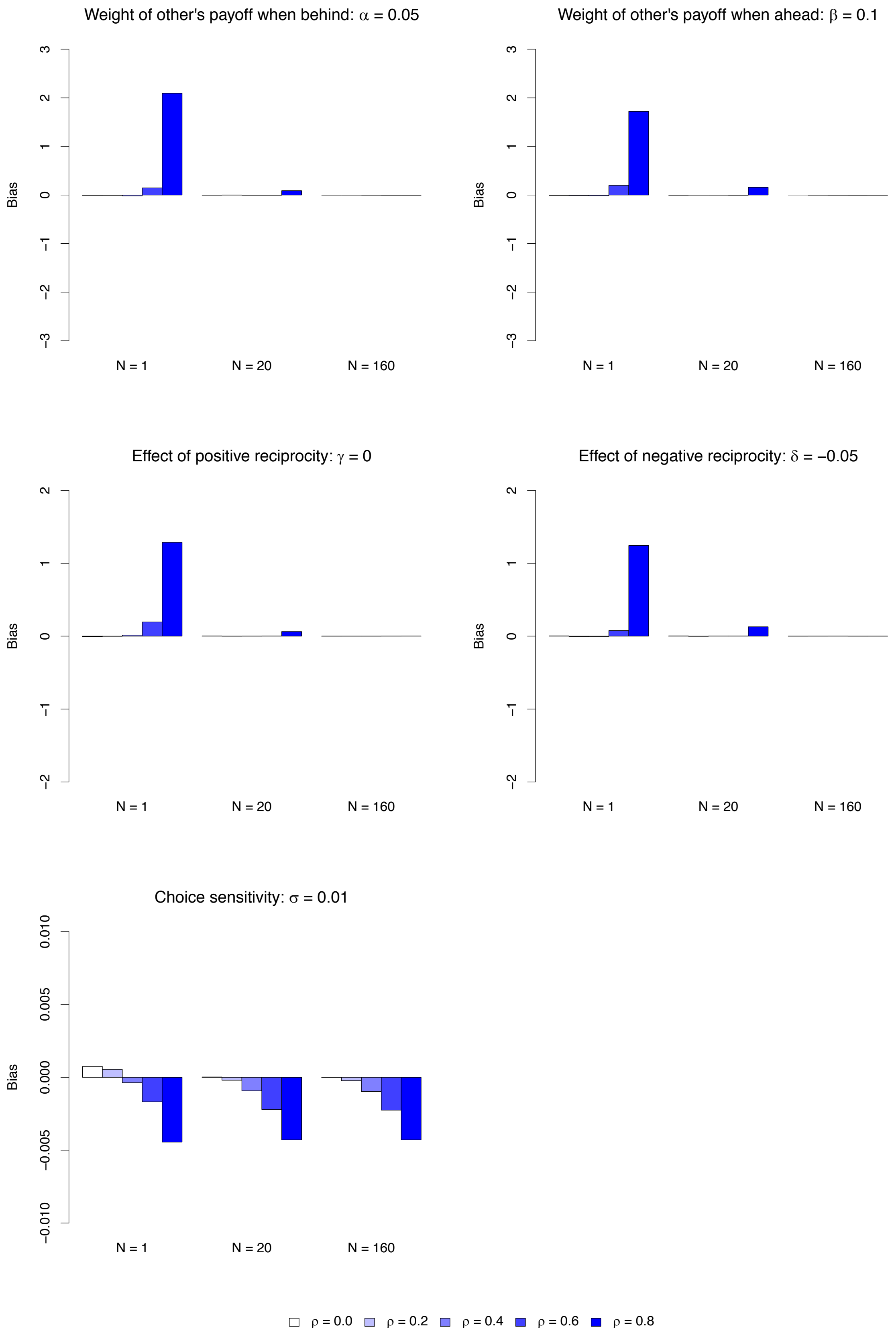

Figure B.3: Bias in parameter estimates in Simulation 3 with moderately altruistic subjects.

True Parameter	ρ	Bias			MSE		
		$N = 1$	$N = 20$	$N = 160$	$N = 1$	$N = 20$	$N = 160$
$\alpha = -0.35$	0.0	-0.016	-0.004	-0.001	0.036	0.002	0.000
	0.2	-0.020	0.000	0.000	0.035	0.002	0.000
	0.4	-0.024	-0.003	-0.001	0.044	0.002	0.000
	0.6	0.047	-0.004	-0.004	1.821	0.002	0.000
	0.8	2.745	-0.009	-0.008	59.811	0.005	0.001
$\beta = 0.00$	0.0	-0.010	-0.003	0.000	0.029	0.001	0.000
	0.2	-0.009	-0.002	-0.001	0.030	0.001	0.000
	0.4	-0.018	-0.001	-0.002	0.032	0.001	0.000
	0.6	0.106	-0.002	-0.003	2.058	0.002	0.000
	0.8	2.172	-0.006	-0.003	50.978	0.003	0.000
$\gamma = 0.10$	0.0	-0.002	0.002	0.000	0.042	0.002	0.000
	0.2	0.007	-0.001	0.000	0.040	0.002	0.000
	0.4	0.002	0.000	0.000	0.047	0.002	0.000
	0.6	0.103	0.001	0.001	1.604	0.003	0.000
	0.8	1.497	0.003	0.003	30.262	0.004	0.000
$\delta = -0.05$	0.0	-0.003	0.002	0.000	0.041	0.002	0.000
	0.2	-0.005	-0.002	0.000	0.045	0.002	0.000
	0.4	0.011	0.000	0.000	0.049	0.002	0.000
	0.6	0.057	-0.001	0.000	1.438	0.003	0.000
	0.8	1.600	0.002	0.002	32.746	0.004	0.001
$\sigma = 0.01$	0.0	0.001	0.000	0.000	0.000	0.000	0.000
	0.2	0.001	0.000	0.000	0.000	0.000	0.000
	0.4	0.000	-0.001	-0.001	0.000	0.000	0.000
	0.6	-0.001	-0.002	-0.002	0.000	0.000	0.000
	0.8	-0.004	-0.004	-0.004	0.000	0.000	0.000

Biases and Mean Squared Errors (MSE) are calculated based on 1,000 simulation runs, each with $N \in \{1,20,160\}$ subjects. The subjects' simulated choices are based on the random utility model presented in section 3.1 of the paper, using the true parameters shown in the first column. To simulate the effect of serially correlated errors across games, the errors of each subject follow an AR(1) process with serial correlation ρ and type-1-extreme-value-distributed innovations. The simulation considers that games were presented in random order within the blocks of dictator and reciprocity games.

Table B.4: Results of Simulation 4 with behindness averse subjects.

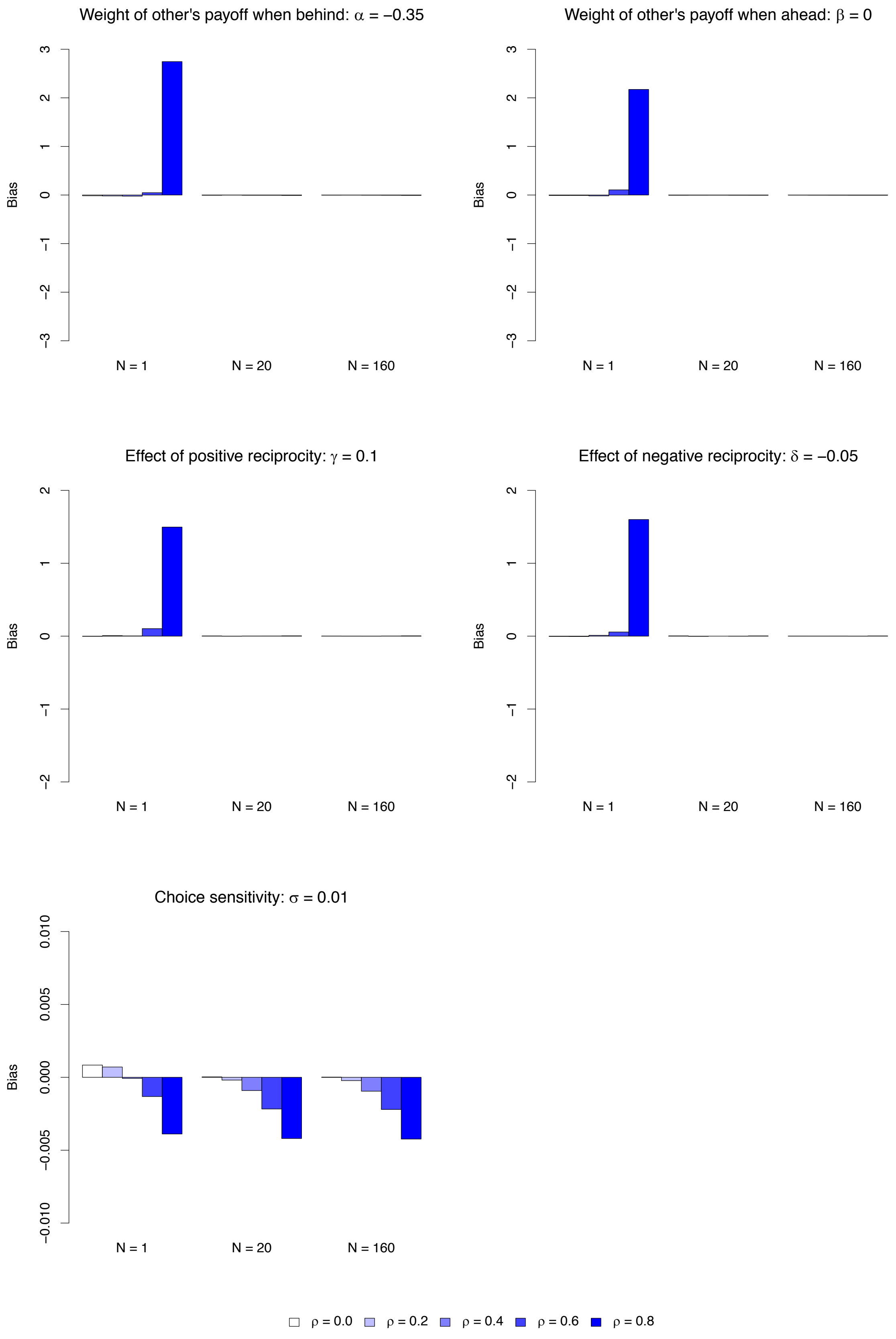

Figure B.4: Bias in parameter estimates in Simulation 4 with behindness averse subjects.

B.2 Monte Carlo simulations: Parameter recovery and impact of serially correlated errors in the finite mixture model

This section presents two Monte Carlo simulations in the context of the finite mixture model with $K = 3$ types. As in section B.1 of this online supplement, the main goals are twofold: (i) to examine whether the finite mixture model reliably recovers the true parameters – and in particular, whether it discriminates well between selfish and moderately altruistic behavior – and (ii) to assess the impact of serially correlated errors.

B.2.1 Set-up

The two Monte Carlo simulations have a similar set-up as the ones described in section B.1.1 of this online supplement. In each of the $R = 1{,}000$ simulation runs, we first simulate the subjects' choices in the dictator and reciprocity games based on the random utility model described in section 3.1 of the paper. Subsequently, we estimate a finite mixture model with $K = 3$ types. After completing all simulation runs, we calculate the bias and the overall accuracy (i.e. MSE) of each parameter.

However, in contrast to the simulations shown in section B.1 of this online supplement, we now consider heterogeneous populations that consist of the following three preference types:

			True parameters				
Simulation	Preference Types	N	α_k	β_k	γ_k	δ_k	σ_k
5	Moderately Altruistic (MA)	80	0.05	0.10	0.00	-0.05	0.03
	Strongly Altruistic (SA)	60	0.20	0.50	0.15	-0.05	0.02
	Behindness Averse (BA)	20	-0.15	-0.40	0.20	-0.10	0.01
6	Selfish (SF)	80	0.00	0.00	0.00	0.00	0.00
	Strongly Altruistic (SA)	60	0.20	0.50	0.15	-0.05	0.02
	Behindness Averse (BA)	20	-0.15	-0.40	0.20	-0.10	0.01

In Simulation 5, the MA, SA, and BA types' sizes and parameters closely resemble the estimates of the finite mixture model in Session 1 (see Table 2 in the paper). This allows us to check whether the finite mixture model can recover the parameters of a sample that is similar to the one elicited in our experiment. In Simulation 6, the MA type is replaced by a selfish (SF) type. This enables us to check whether the finite mixture model could also reliably pick up a selfish type.

We also consider the impact of serially correlated errors. Hence, the simulated random errors in a subject's utility follow an AR(1) process across games with serial correlation $\rho \in \{0, 0.5, 0.8\}$ and type-1-extreme-value-distributed innovations. As in the simulations before, the order of games within the blocks of dictator and reciprocity games is randomly reshuffled in each simulation run. This random reshuffling takes into account that in the experiment, the games were presented in random order within the blocks of dictator and reciprocity games (see section 2.3 in the paper).

B.2.2 Results

Tables B.5 and B.6 show the results for Simulation 5 and 6, respectively.

The results are qualitatively similar to the ones of the aggregate model with $N = 160$ subjects. If there is no or only moderate serial correlation in the errors, i.e. $\rho \in \{0,0.5\}$, the finite mixture models' estimators are virtually unbiased and exhibit great overall accuracy. If there is strong serial correlation in the errors, i.e. $\rho = 0.8$, the estimators tend to be slightly biased and their overall accuracy is somewhat lower. However, relative to the size of the true parameters the bias is negligible. In sum, these results highlight the finite mixture model's efficiency in taking the prevalent heterogeneity into account.

Moreover, Table B.6 also reveals that the finite mixture model could reliably uncover a group of selfish subjects, even if the errors in the subjects' utility are serially correlated. The estimators for the simulated SF-type are virtually unbiased and overall very accurate. Consequently, if there were a distinct type of selfish subjects in our experimental data, we would have most likely detected it – in particular once we increased the number of types from $K = 3$ to $K = 4$.

ρ	Preference Type	Bias $\hat{\pi}_k$	Bias $\hat{\alpha}_k$	Bias $\hat{\beta}_k$	Bias $\hat{\gamma}_k$	Bias $\hat{\delta}_k$	Bias $\hat{\sigma}_k$	MSE $\hat{\pi}_k$	MSE $\hat{\alpha}_k$	MSE $\hat{\beta}_k$	MSE $\hat{\gamma}_k$	MSE $\hat{\delta}_k$	MSE $\hat{\sigma}_k$
0.0	MA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	SA	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	BA	–	-0.003	0.000	0.000	0.000	0.000	–	0.002	0.001	0.002	0.002	0.000
0.5	MA	0.000	-0.001	-0.001	0.000	0.001	-0.004	0.000	0.000	0.000	0.000	0.000	0.000
	SA	0.000	-0.002	0.000	0.000	-0.001	-0.003	0.000	0.000	0.000	0.000	0.000	0.000
	BA	–	-0.004	-0.004	0.001	0.001	-0.001	–	0.002	0.002	0.003	0.003	0.000
0.8	MA	0.004	-0.003	-0.003	0.000	0.001	-0.012	0.000	0.000	0.000	0.000	0.000	0.000
	SA	-0.004	-0.003	0.003	0.002	-0.002	-0.008	0.000	0.000	0.000	0.001	0.001	0.000
	BA	–	-0.007	-0.003	0.007	-0.001	-0.004	–	0.007	0.004	0.005	0.005	0.000

Biases and Mean Squared Errors (MSE) are calculated based on 1,000 simulation runs. The subjects' simulated choices are based on the random utility model presented in section 3.1 of the paper, using the true parameters shown in Section B.2.1 of this online supplement. To simulate the effect of serially correlated errors across games, the errors of each subject follow an AR(1) process with serial correlation ρ and type-1-extreme-value-distributed innovations. The simulation considers that games were presented in random order within the blocks of dictator and reciprocity games. The estimator of the mixing proportion of the BA-type, $\hat{\pi}_{BA}$, has no bias and MSE because it is not a separately estimated parameter and given by $\hat{\pi}_{BA} = 1 - \hat{\pi}_{MA} - \hat{\pi}_{SA}$.

Table B.5: Results of Simulation 5 with a mixture of 80 MA-, 60 SA-, and 20 BA-types.

ρ	Preference Type	Bias $\hat{\pi}_k$	Bias $\hat{\alpha}_k$	Bias $\hat{\beta}_k$	Bias $\hat{\gamma}_k$	Bias $\hat{\delta}_k$	Bias $\hat{\sigma}_k$	MSE $\hat{\pi}_k$	MSE $\hat{\alpha}_k$	MSE $\hat{\beta}_k$	MSE $\hat{\gamma}_k$	MSE $\hat{\delta}_k$	MSE $\hat{\sigma}_k$
0.0	SF	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	SA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	BA	–	-0.002	-0.001	0.001	0.000	0.000	–	0.002	0.001	0.002	0.002	0.000
0.5	SF	0.000	-0.001	-0.001	0.001	0.000	-0.004	0.000	0.000	0.000	0.000	0.000	0.000
	SA	0.000	-0.002	0.001	0.002	0.000	-0.003	0.000	0.000	0.000	0.000	0.000	0.000
	BA	–	-0.003	-0.004	-0.001	0.002	-0.001	–	0.003	0.002	0.003	0.003	0.000
0.8	SF	0.000	-0.004	-0.001	0.000	-0.001	-0.012	0.000	0.000	0.000	0.000	0.000	0.000
	SA	-0.002	-0.003	0.002	0.002	-0.003	-0.008	0.000	0.000	0.000	0.001	0.000	0.000
	BA	–	-0.004	-0.004	0.001	0.001	-0.004	–	0.007	0.004	0.005	0.005	0.000

Biases and Mean Squared Errors (MSE) are calculated based on 1,000 simulation runs. The subjects' simulated choices are based on the random utility model presented in section 3.1 of the paper, using the true parameters shown in Section B.2.1 of this online supplement. To simulate the effect of serially correlated errors across games, the errors of each subject follow an AR(1) process with serial correlation ρ and type-1-extreme-value-distributed innovations. The simulation considers that games were presented in random order within the blocks of dictator and reciprocity games. The estimator of the mixing proportion of the BA-type, $\hat{\pi}_{BA}$, has no bias and MSE because it is not a separately estimated parameter and given by $\hat{\pi}_{BA} = 1 - \hat{\pi}_{SF} - \hat{\pi}_{SA}$.

Table B.6: Results of Simulation 6 with a mixture of 80 SF-, 60 SA-, and 20 BA-types.

B.3 Behavioral model with a CES utility function

This section considers a more general specification of the behavioral model, which is inspired by Fisman et al. (2007). In this specification, player A's deterministic utility,

$$\widetilde{U}^A = \left[(1 - \alpha s - \beta r - \gamma q - \delta v) * (\Pi^A)^\zeta + (\alpha s + \beta r + \gamma q + \delta v) * (\Pi^B)^\zeta\right]^{1/\zeta}, \tag{B.3}$$

exhibits constant elasticity of substitution (CES) between her own and player B's payoff, where $\zeta \leq 1$ represents the curvature of the indifference curves, and $\eta = (\zeta - 1)^{-1}$ is the elasticity of substitution.

The behavioral model with CES utility nests the one with piecewise linear utility as a special case, since for $\zeta = 1$, $\widetilde{U}^A = U^A$ (see also equation (2) in the paper). Moreover, for $\zeta < 1$, the indifference curves are strictly convex, and the behavioral model with CES utility could also rationalize an interior solution if A had to choose Π^A and Π^B from a continuous budget set.

B.3.1 Monte Carlo simulation

This subsection applies a Monte Carlo simulation to examine how well the behavioral model with CES utility performs relative to the one with piecewise linear utility in our experimental data. This is an important question, as on the one hand the model with CES utility is more general, but the on the other hand, the binary dictator and reciprocity games are specifically designed to reliably identify the weights player A puts on B's payoff in the context of piecewise linear utility.

B.3.1.1 Set-up

The Monte Carlo simulation has a set-up similar to the one described in section B.1.1 of this online supplement. In each simulation run, we first simulate the choices of $N \in \{1,20,160\}$ subjects in the dictator and reciprocity games. We apply a random utility model with CES utility and parameters similar to the ones obtained in the aggregate estimations – i.e. $\alpha = 0.10$, $\beta = 0.25$, $\gamma = 0.05$, $\delta = -0.05$, $\sigma = 0.01$ – and a curvature of the indifference curves corresponding to $\zeta \in \{1,0.5,0.1,-1,-5\}$. Subsequently, we try to recover the true parameters by estimating two versions of the random utility model on the simulated choices: one with CES utility $\widetilde{U}^A$ and the other with piecewise linear utility U^A.

B.3.1.2 Results

Tables B.7 and B.8 present the results in terms of bias and MSE, respectively. In both tables, the first set of columns shows the results for the version of the random utility model with CES utility $\widetilde{U}^A$, while the second set of columns shows the results of the version with piecewise linear utility U^A.

If the simulated choices stem from a piecewise linear utility function, i.e. $\zeta = 1$, the performance of both versions of the random utility models is nearly identical. Except for $N = 1$, the estimators are both virtually unbiased and with MSEs close to zero also very accurate overall. This is not surprising, as the version of the random utility model with piecewise linear utility is correctly specified and the more general version with CES utility nests $\zeta = 1$ as a special case.

However, if the simulated choices stem from a CES utility function with strictly convex indifference curves, i.e. $\zeta \in \{0.5,0.1,-1,-5\}$, there are important differences in the performance of the two versions of the model. In terms of bias, shown in Table B.7, the differences in performance can be summarized as follows:

- The version of the random utility model with CES utility yields estimators of the behavioral parameters with small to negligible bias, unless $N = 1$ and/or ζ is very small. However, the estimator of ζ tends to be downward-biased, even if $N = 20$. Hence, the model has problems to recover the curvature of the indifference curves from the binary choices in the dictator and reciprocity games.

- The version of the random utility model with piecewise linear utility yields estimators of the distributional parameters α and β that are biased regardless of the number of subjects. However, the absolute size of the bias is relatively small if the indifference curves are not overly convex. Only if $\zeta < 0.5$, the absolute size of the bias becomes severe. Interestingly, the estimators of γ and δ that measure how the weight on B's payoff changes under positive and negative reciprocity seem to be unbiased.

Hence, in case we only consider the bias in the estimators – in particular those of α and β – the general version of the random utility model with CES utility outperforms the version with piecewise linear utility if $\zeta < 1$. However, this result changes in case we also consider the estimators' overall accuracy, i.e. their MSEs as shown in Table B.8:

- The version of the random utility model with CES utility yields estimators of the distributional parameters α and β that are inaccurate – especially if ζ is small. If ζ drops below 1, the estimators remain (nearly) unbiased but their variance quickly increases, leading to poor overall accuracy.

- In contrast, the version of the random utility model with piecewise linear utility yields estimators that, with some exceptions, are equally accurate or sometimes even more accurate overall than the ones of the model with CES utility. If ζ drops below 1, the estimators of the model with piecewise linear utility are biased but their variance remains relatively low.

In conclusion, these results imply that, in case we consider the overall accuracy of the estimators, the general version of the random utility model with CES utility does not offer much benefit over the more parsimonious model with piecewise linear utility. Note that this conclusion is probably highly specific to our data, which stems from an experiment designed to elicit how subjects weight another player's payoff with a series of easy-to-grasp, binary dictator and reciprocity games. We expect the general model with CES utility to perform much better in data stemming from experiments that are specifically designed to estimate a CES utility function and allow the subjects to also make interior choices, such as the one presented in Fisman et al. (2007). However, as the general version of the random utility model with CES utility has no advantage in our data, we stick to the more parsimonious version with piecewise linear utility.

True Parameter	ζ	Bias: CES utility $\widetilde{U}^A$			Bias: piecewise linear utility U^A		
		$N = 1$	$N = 20$	$N = 160$	$N = 1$	$N = 20$	$N = 160$
$\alpha = 0.10$	1	0.008	0.000	0.000	0.008	0.000	0.000
	0.5	0.068	0.029	0.003	-0.041	-0.048	-0.047
	0.1	0.013	0.026	0.003	-0.080	-0.081	-0.081
	-1	-0.950	0.041	0.007	-0.154	-0.150	-0.148
	-5	-8.413	-2.424	-0.111	-0.183	-0.181	-0.181
$\beta = 0.25$	1	-0.009	0.001	0.000	-0.009	0.001	0.000
	0.5	-0.049	-0.015	-0.002	0.098	0.101	0.101
	0.1	-0.022	-0.012	-0.002	0.186	0.194	0.193
	-1	0.402	-0.015	-0.002	0.424	0.421	0.422
	-5	3.521	0.679	0.025	0.670	0.665	0.662
$\gamma = 0.05$	1	-0.005	-0.002	0.000	-0.005	-0.002	0.000
	0.5	-0.007	-0.004	-0.001	-0.002	-0.002	-0.002
	0.1	0.003	-0.003	0.000	-0.006	-0.004	-0.005
	-1	0.143	-0.003	-0.001	-0.016	-0.017	-0.018
	-5	0.637	0.103	0.002	-0.033	-0.034	-0.033
$\delta = -0.05$	1	0.005	-0.001	0.000	0.005	-0.001	0.000
	0.5	0.010	0.005	0.000	0.001	0.002	0.001
	0.1	0.023	0.003	0.001	0.006	0.001	0.002
	-1	0.044	0.005	0.000	0.018	0.013	0.012
	-5	-0.100	-0.088	-0.005	0.024	0.030	0.031
ζ	1	0.000	0.000	0.000	–	–	–
	0.5	-0.963	-0.175	-0.014	–	–	–
	0.1	-1.050	-0.155	-0.016	–	–	–
	-1	-1.737	-0.294	-0.032	–	–	–
	-5	-3.434	-3.110	-0.883	–	–	–
$\sigma = 0.01$	1	0.001	0.000	0.000	0.001	0.000	0.000
	0.5	0.001	0.000	0.000	0.001	0.000	0.000
	0.1	0.001	0.000	0.000	0.001	0.000	0.000
	-1	0.002	0.000	0.000	0.002	0.001	0.001
	-5	0.007	0.001	0.000	0.001	0.000	0.000

Biases in the estimators of the random models with CES utility and piecewise linear utility are calculated based on 1,000 simulation runs, each with $N \in \{1,20,160\}$ subjects. The subjects' simulated choices are based on a random utility model with CES utility with parameters as shown in the first column and a varying curvature of the indifference curves $\zeta \in \{1, 0.5, 0.1, -1, -5\}$.

Table B.7: Bias in estimators of random utility models with CES and piecewise linear utility.

True Parameter	ζ	MSE: CES utility $\widetilde{U}^A$ $N=1$	$N=20$	$N=160$	MSE: piecewise linear utility U^A $N=1$	$N=20$	$N=160$
$\alpha = 0.10$	1	0.018	0.001	0.000	0.018	0.001	0.000
	0.5	0.131	0.011	0.001	0.022	0.004	0.002
	0.1	1.038	0.012	0.001	0.026	0.008	0.007
	-1	85.827	0.035	0.003	0.045	0.023	0.022
	-5	538.259	64.138	3.156	0.059	0.034	0.033
$\beta = 0.25$	1	0.021	0.001	0.000	0.021	0.001	0.000
	0.5	0.039	0.009	0.001	0.029	0.011	0.010
	0.1	0.158	0.009	0.001	0.054	0.038	0.037
	-1	10.109	0.014	0.002	0.199	0.178	0.178
	-5	103.064	5.896	0.203	0.479	0.443	0.438
$\gamma = 0.05$	1	0.026	0.001	0.000	0.026	0.001	0.000
	0.5	0.042	0.001	0.000	0.028	0.001	0.000
	0.1	0.136	0.001	0.000	0.024	0.001	0.000
	-1	12.141	0.002	0.000	0.025	0.002	0.000
	-5	74.525	0.671	0.004	0.034	0.003	0.001
$\delta = -0.05$	1	0.025	0.001	0.000	0.025	0.001	0.000
	0.5	0.029	0.001	0.000	0.028	0.001	0.000
	0.1	0.110	0.001	0.000	0.025	0.001	0.000
	-1	10.160	0.002	0.000	0.027	0.001	0.000
	-5	48.832	0.363	0.014	0.033	0.002	0.001
ζ	1	0.000	0.000	0.000	–	–	–
	0.5	3.630	0.426	0.024	–	–	–
	0.1	4.356	0.398	0.026	–	–	–
	-1	9.381	0.965	0.057	–	–	–
	-5	96.690	57.138	7.619	–	–	–
$\sigma = 0.01$	1	0.000	0.000	0.000	0.000	0.000	0.000
	0.5	0.000	0.000	0.000	0.000	0.000	0.000
	0.1	0.000	0.000	0.000	0.000	0.000	0.000
	-1	0.000	0.000	0.000	0.000	0.000	0.000
	-5	0.000	0.000	0.000	0.000	0.000	0.000

MSEs in the estimators of the random models with CES utility and piecewise linear utility are calculated based on 1,000 simulation runs, each with $N \in \{1,20,160\}$ subjects. The subjects' simulated choices are based on a random utility model with CES utility with parameters as shown in the first column and a varying curvature of the indifference curves $\zeta \in \{1,0.5,0.1,-1,-5\}$.

Table B.8: Mean Squared Errors (MSEs) in estimators of random utility models with CES and piecewise linear utility.

B.3.2 Aggregate estimation

We also estimated the aggregate parameters in Sessions 1 and 2 using the general version of the random utility model with CES utility. As this general version of the model nests the model with piecewise linear utility as a special case, we can test whether the subjects' choices reject the assumption of piecewise linearity.

Table B.9 shows the results. In both sessions, the estimates of ζ are very close to 1 and we cannot reject the null hypothesis that the subjects' utility is piecewise linear. As expected, considering the results of the previous Monte Carlo simulation (section B.3.1 of this online supplement), the other parameter estimates nearly coincide with the ones reported in Table 1 of the paper.

	Estimates of Session 1	*Estimates of Session 2*	*p-value of z-test with H_0: Session1=Session2*
α: Weight on other's payoff when behind	0.088*** (0.018)	0.103*** (0.014)	0.521
β: Weight on other's payoff when ahead	0.253*** (0.021)	0.237*** (0.020)	0.570
γ: Measure of positive reciprocity	0.071*** (0.014)	0.028*** (0.010)	0.011
δ: Measure of negative reciprocity	-0.042*** (0.011)	-0.044*** (0.009)	0.919
ζ: Curvature of indifference curves	0.958*** (0.085)	0.949*** (0.062)	0.933
σ: Choice sensitivity	0.016*** (0.001)	0.019*** (0.001)	0.006
P-value (H_0: ζ=1)	0.619	0.407	
# of observations	18,720	18,720	
# of subjects	160	160	
Log Likelihood	-5,472.20	-4,540.65	

Individual cluster robust standard errors in parentheses.
*** significant at 1%; ** significant at 5%; * significant at 10%

Table B.9: Preferences of the representative agent ($K = 1$) in Sessions 1 and 2 with CES utility.

B.4 Correlation of psychological and demographic variables with individual type-membership

Table B.10 shows that neither the psychological nor the demographic variables are significantly correlated with individual type-membership. It exhibits the result of a multinomial logit regression with individual type-membership as dependent variable and the various psychological and demographic variables as independent variables. The base category is membership in the MA-type. While the individual estimates indicate that studying law (being female) is negatively (positively) associated with membership in the BA- and SA-type, the likelihood ratio tests at the bottom of the table reveal that the estimated coefficients are jointly insignificant for the whole regression as well as for various subsets of variables.

Multinomial logit regression with dependent variable: type-membership base category: Moderately Altruistic (MA) type	Behindness Averse (BA) type	Strongly Altruistic (SA) type
Constant	-2.278 (3.437)	-1.515 (2.026)
Big 5: consciousness	-0.170 (0.101) *	-0.098 (0.065)
Big 5: openness	-0.003 (0.080)	-0.014 (0.050)
Big 5: extraversion	0.065 (0.080)	-0.009 (0.050)
Big 5: agreeableness	0.074 (0.104)	0.100 (0.071)
Big 5: neuroticism	0.081 (0.077)	0.080 (0.048) *
Cognitive ability score	-0.005 (0.122)	0.018 (0.079)
Field of study: natural sciences	-1.221 (0.978)	-0.566 (0.729)
Field of study: law	-2.422 (1.464) *	-2.482 (1.294) *
Field of study: social sciences	-1.725 (1.385)	-0.127 (0.828)
Field of study: medicine	-17.165 (1855.68)	-0.535 (0.861)
Monthly income (1,000 CHF)	-0.906 (0.962)	0.204 (0.414)
Age	-0.076 (0.107)	-0.038 (0.061)
Female	1.263 (0.682) *	0.847 (0.414) **
# of observations / subjects	160	
Log likelihood	-132.49	
P-values of likelihood ratio tests		
H_0: All coefficients jointly insignificant	0.181	
H_0: Big 5 measures jointly insignificant	0.374	
H_0: Field-of-study-dummies jointly insignificant	0.161	
H_0: M.inc., age, female jointly insignificant	0.221	

Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%.
As the model is non-linear, only the coefficients' signs and significance are directly interpretable: a positive (negative) sign indicates that the variable has a positive (negative) effect on the probability of being associated with the corresponding type.

Table B.10: Multinomial logit regression showing that individual type-membership is not correlated with Big 5 personality traits, cognitive ability, field of study, and other socio-economic characteristics.

B.5 Figures of the Paper in Color and Serif Free Font

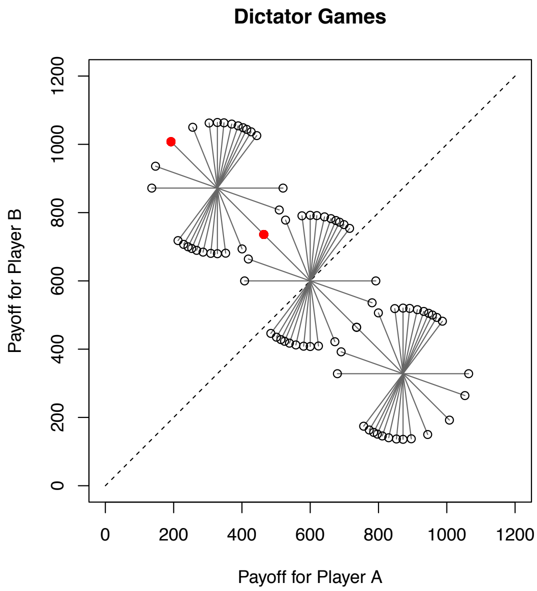

Figure B.5 (Analogue to Figure 1 in the paper): The dictator games. Each of the three circles contains 13 binary dictator games. Each game is represented by the two payoff allocations connected by a grey line. Player A can choose one of the extreme points on the line. For every game, the slope of the grey line indicates A's cost of altering player B's payoff. Allocations above (below) the dashed 45° line help identifying the weight player A puts on B's payoff under disadvantageous (advantageous) inequality.

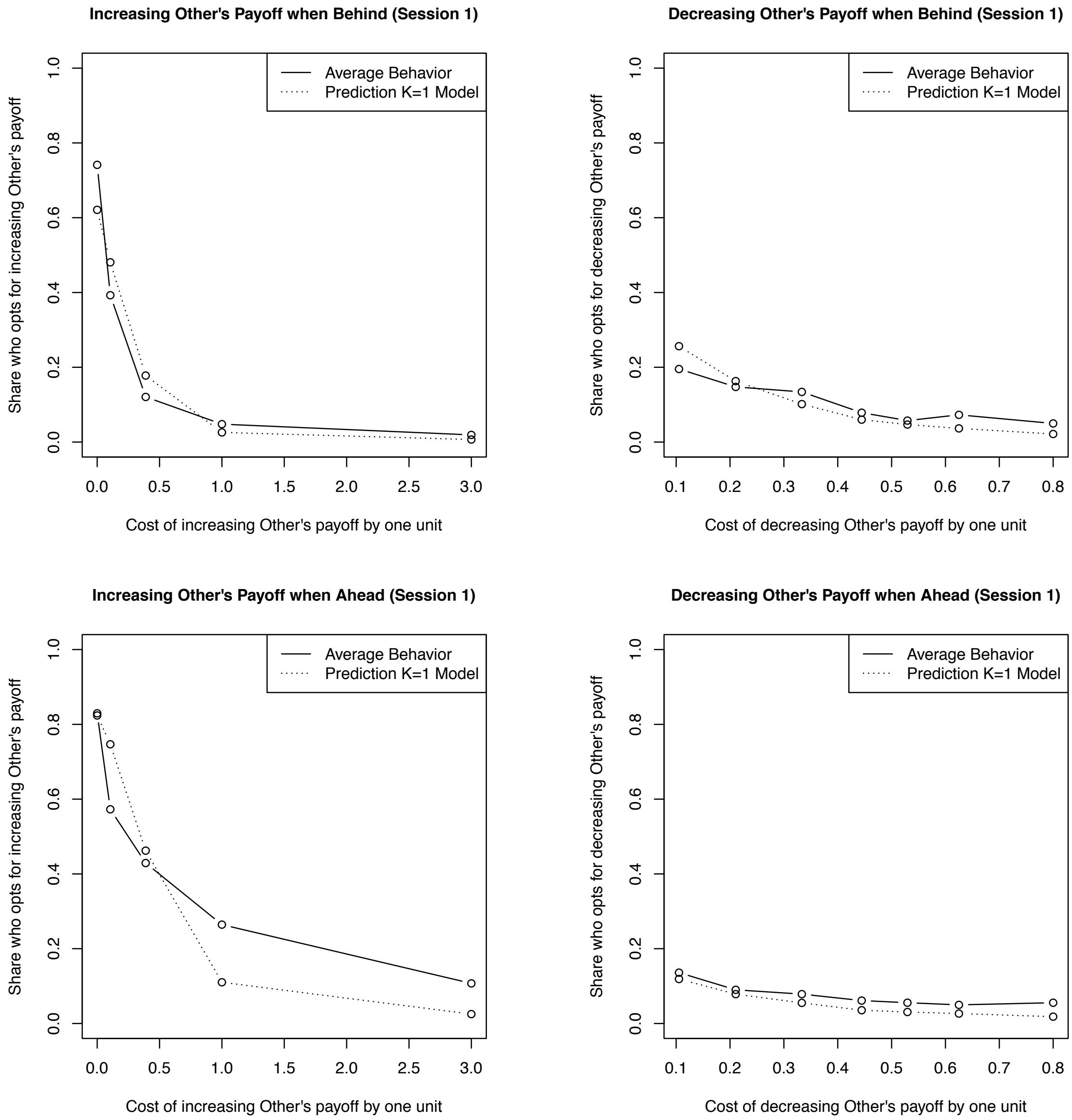

Figure B.6 (Analogue to Figure 2 in the paper): Representative agent's empirical and predicted willingness to change the other player's payoff across cost levels in Session 1. The empirical willingness corresponds to the fraction of subjects that chose to change the other player's payoff in the indicated direction. The predicted willingness corresponds to the predicted probability that the representative agent changes the other player's payoff in the indicated direction. It is based on the random utility model presented in Section 3.1 and uses the estimated aggregate parameters of Session 1 on all dictator and reciprocity games.

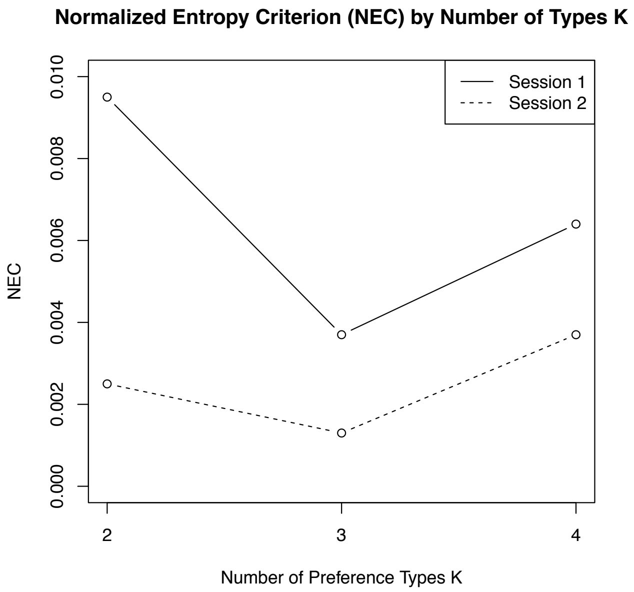

Figure B.7 (Analogue to Figure 3 in the paper): Normalized entropy criterion (NEC) for different numbers of preference types in Sessions 1 and 2. The NEC summarizes the ambiguity in the subjects' classification into types relative to the finite mixture model's improvement in fit compared to the representative agent model with K = 1 (see equations (8) and (9)). By minimizing the NEC, we can determine the optimal number of preference types K the finite mixture model should take into account.

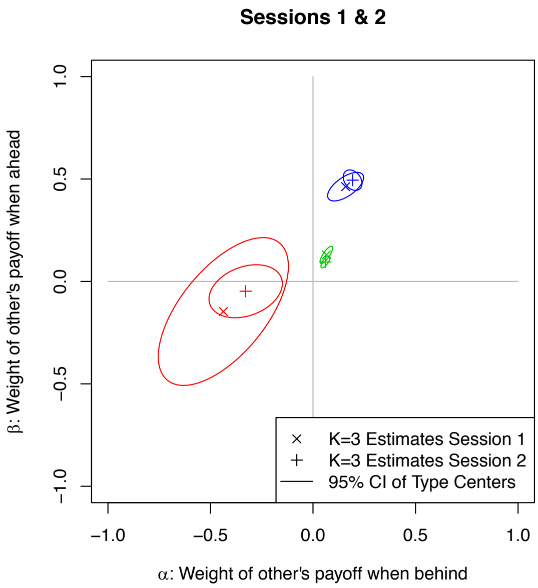

Figure B.8 (Analogue to Figure 4 in the paper): Temporal stability of the type-specific parameter estimates of the finite mixture models with $K = 3$ preference types. The type-specific parameter estimates are stable over time as their 95% confidence intervals overlap between Session 1 and 2.)). Moderately Altruistic (MA) Types are shown in green, Strongly Altruistic (SA) Types in blue, and Behindness Averse (BA) Types in red.

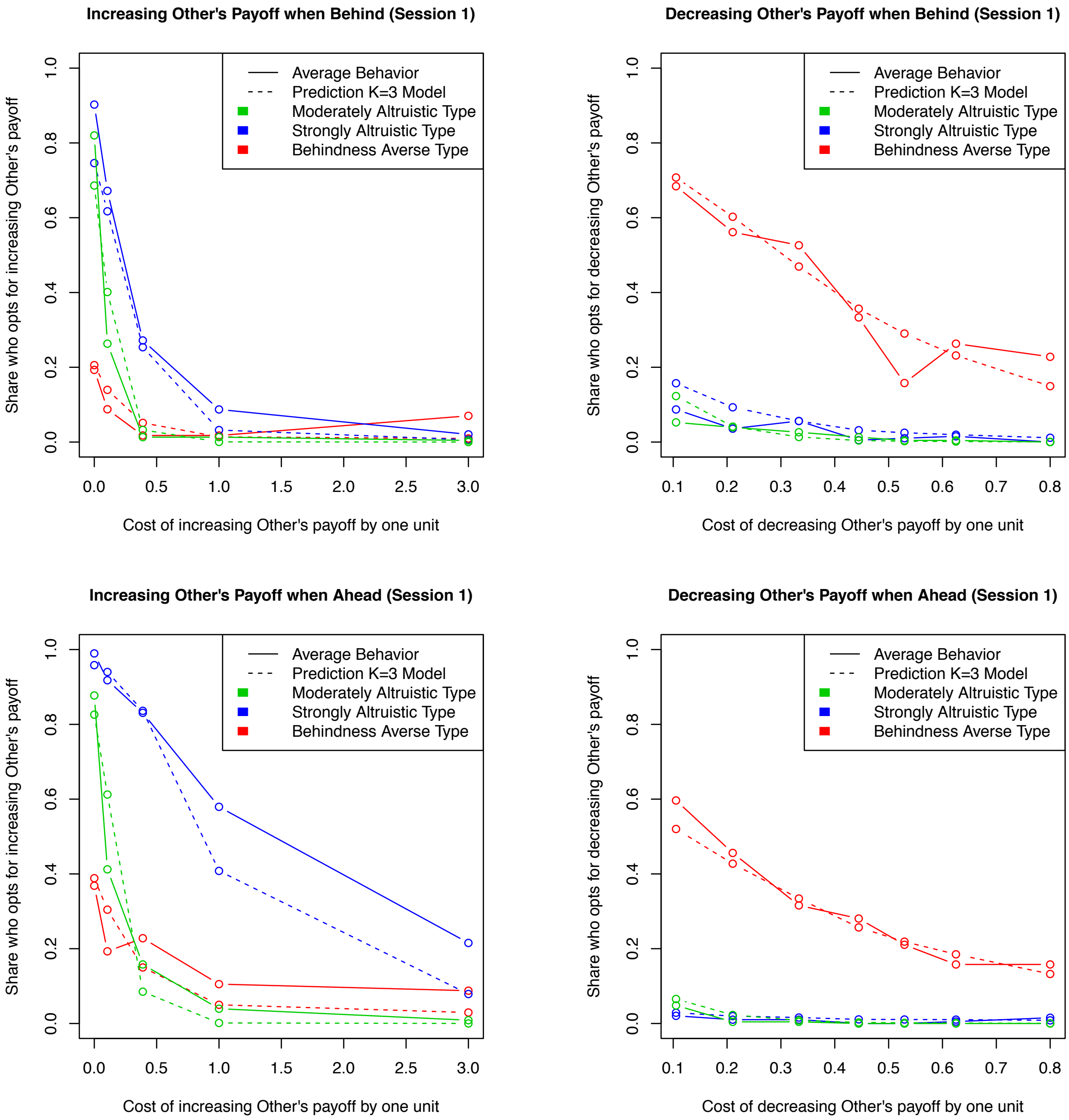

Figure B.9 (Analogue to Figure 5 in the paper): Empirical and predicted willingness to change the other player's payoff of the different preference types across cost levels in Session 1. The empirical willingness corresponds to the fraction of subjects of a given preference type that chose to change the other player's payoff in the indicated direction. The predicted willingness corresponds to the predicted probability that a given preference type changes the other player's payoff in the indicated direction. It is based on the random utility model presented in Section 3.1 and uses the estimated type-specific parameters of Session 1 on all dictator and reciprocity games.

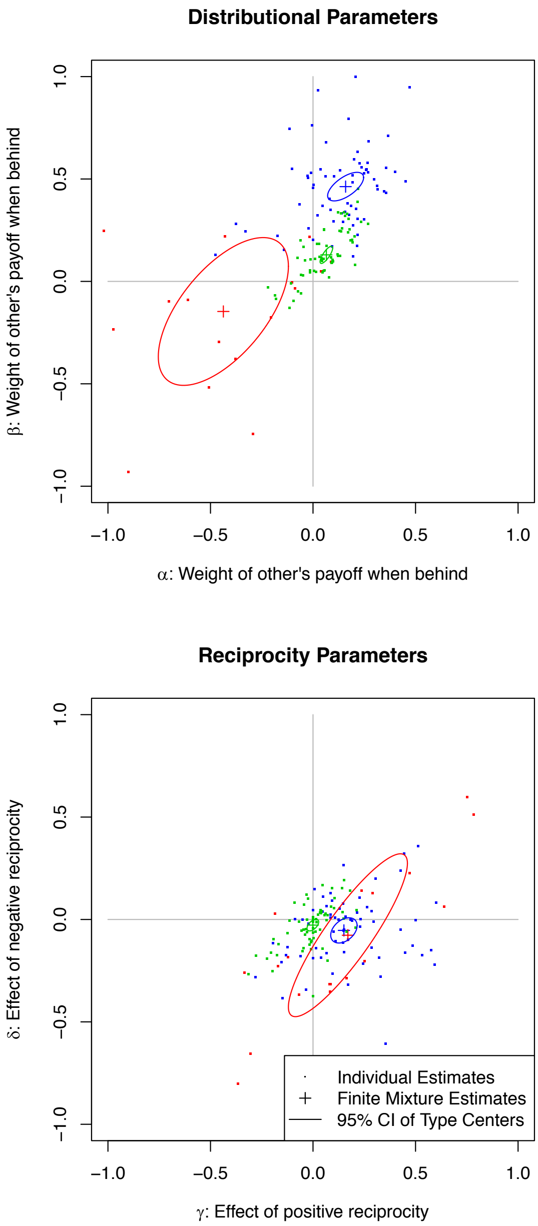

Figure B.10 (Analogue to Figure 6 in the paper): Distribution of individual-specific parameter estimates along with type-specific parameter estimates ($K = 3$ model) in Session 1. The colors of the individual-specific estimates indicate the underlying subjects' classification into preference types according to the individual posterior probabilities of type-membership (see equation (7)): Moderately Altruistic (MA) Types are shown in green, Strongly Altruistic (SA) Types in blue, and Behindness Averse (BA) Types in red.

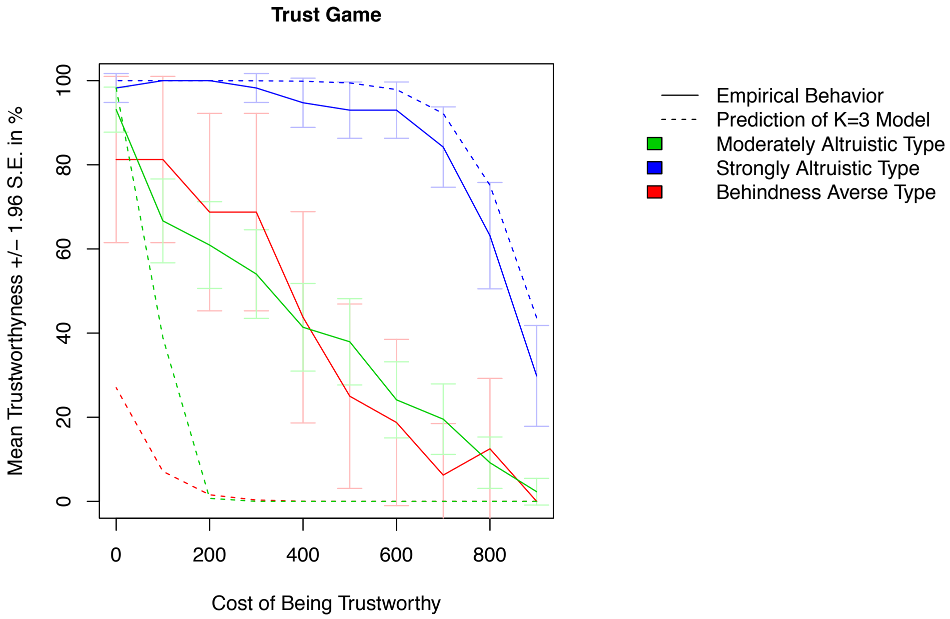

Figure B.11 (Analogue to Figure 8 in the paper): Empirical and predicted mean trustworthiness of the different preference types across cost levels (with 95% confidence intervals). The empirical mean trustworthiness corresponds to the fraction of subjects of a preference type that chose the trustworthy action at a given cost of being trustworthy. The predicted trustworthiness corresponds to the predicted probability that a subject of a given preference type chooses the trustworthy action at a given cost of being trustworthy. The predicted probability is based on the random utility model presented in Section 3.1 and uses the estimated type-specific parameters of Session 2.

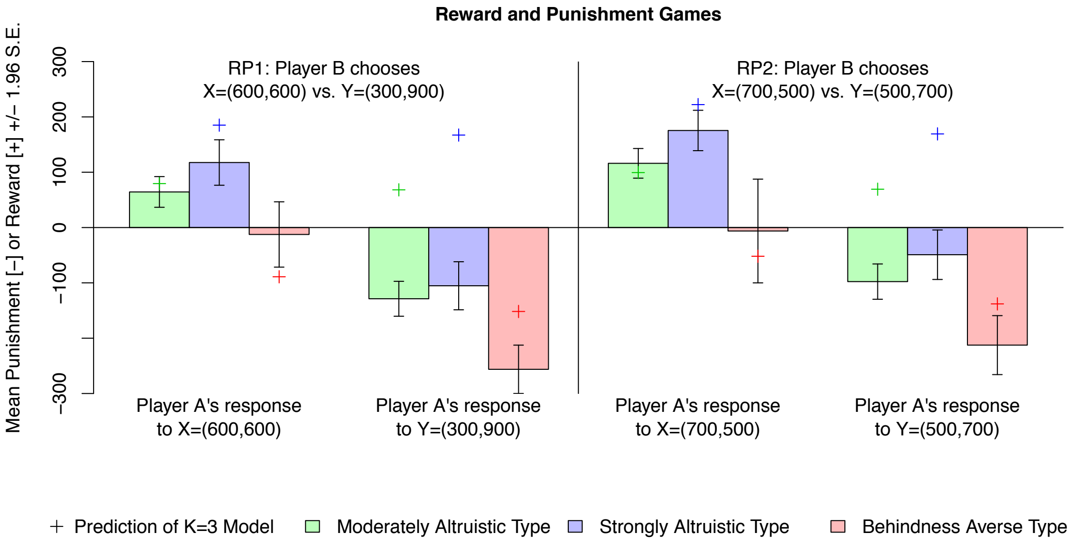

Figure B.12 (Analogue to Figure 9 in the paper): Empirical and predicted mean reward and punishment of player A in response to player B's choice (with 95% confidence intervals). The bars correspond to the mean reward or punishment level the subjects of a given preference type implement in response to player B's choice. The plus signs correspond to the reward or punishment level a subject of a given preference type is predicted to implement in response to player B's choice. The prediction is based on the random utility model presented in Section 3.1 and uses the estimated type-specific parameters of Session 2.

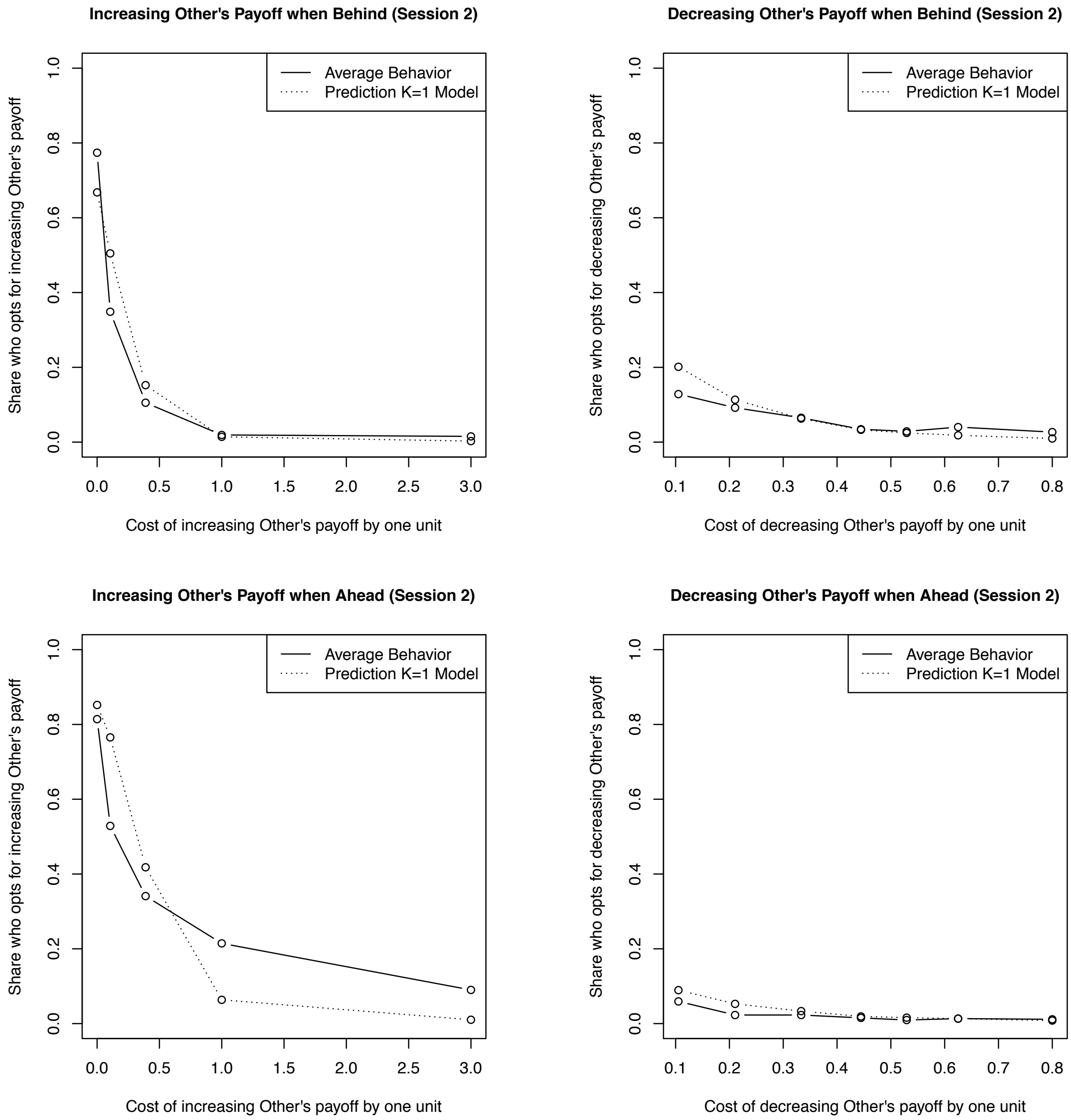

Figure B.13 (Analogue to Figure A.3 in the paper): Representative agent's empirical and predicted willingness to change the other player's payoff across cost levels in Session 2. The empirical willingness corresponds to the fraction of subjects that chose to change the other player's payoff in the indicated direction. The predicted willingness corresponds to the predicted probability that the representative agent choses to change the other player's payoff in the indicated direction. It is based on the random utility model presented in Section 3.1 and uses the estimated aggregate parameters of Session 2 on all dictator and reciprocity games.

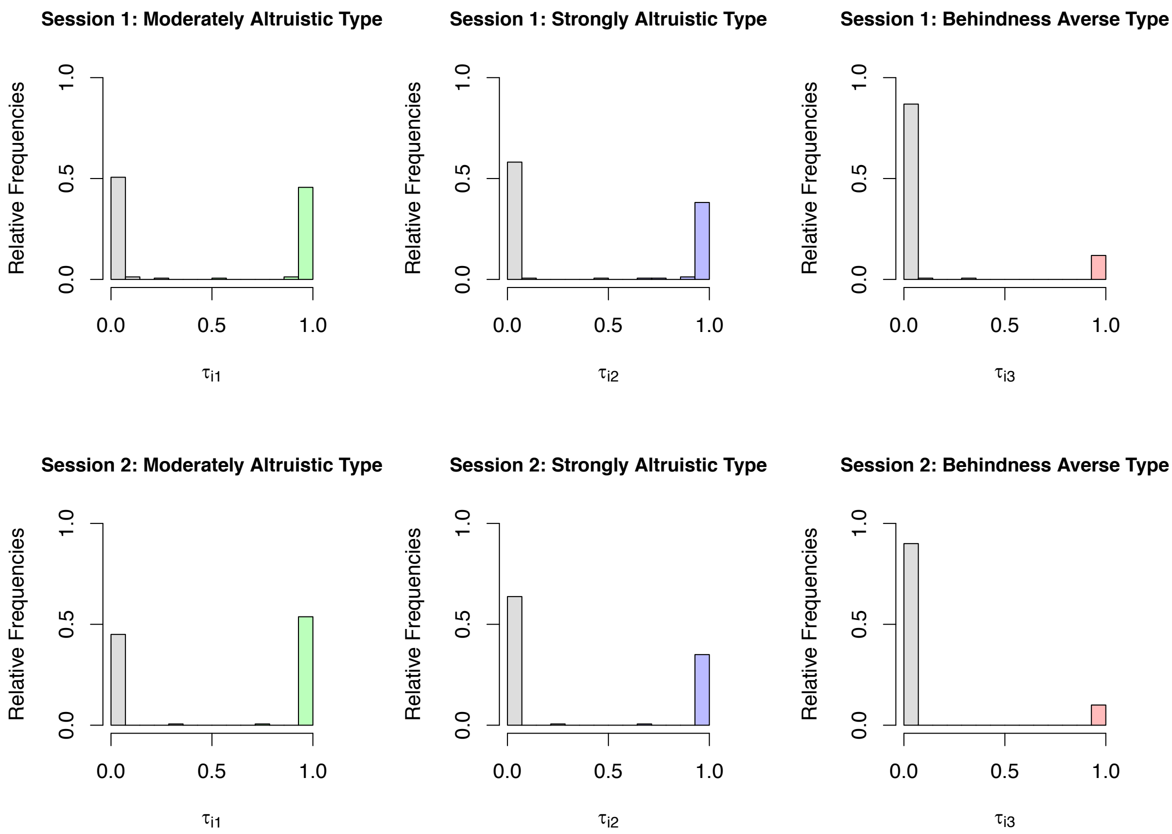

Figure B.14 (Analogue to Figure A.4 in the paper): Distribution of posterior probabilities of individual type-membership in Sessions 1 (upper row) and 2 (lower row).

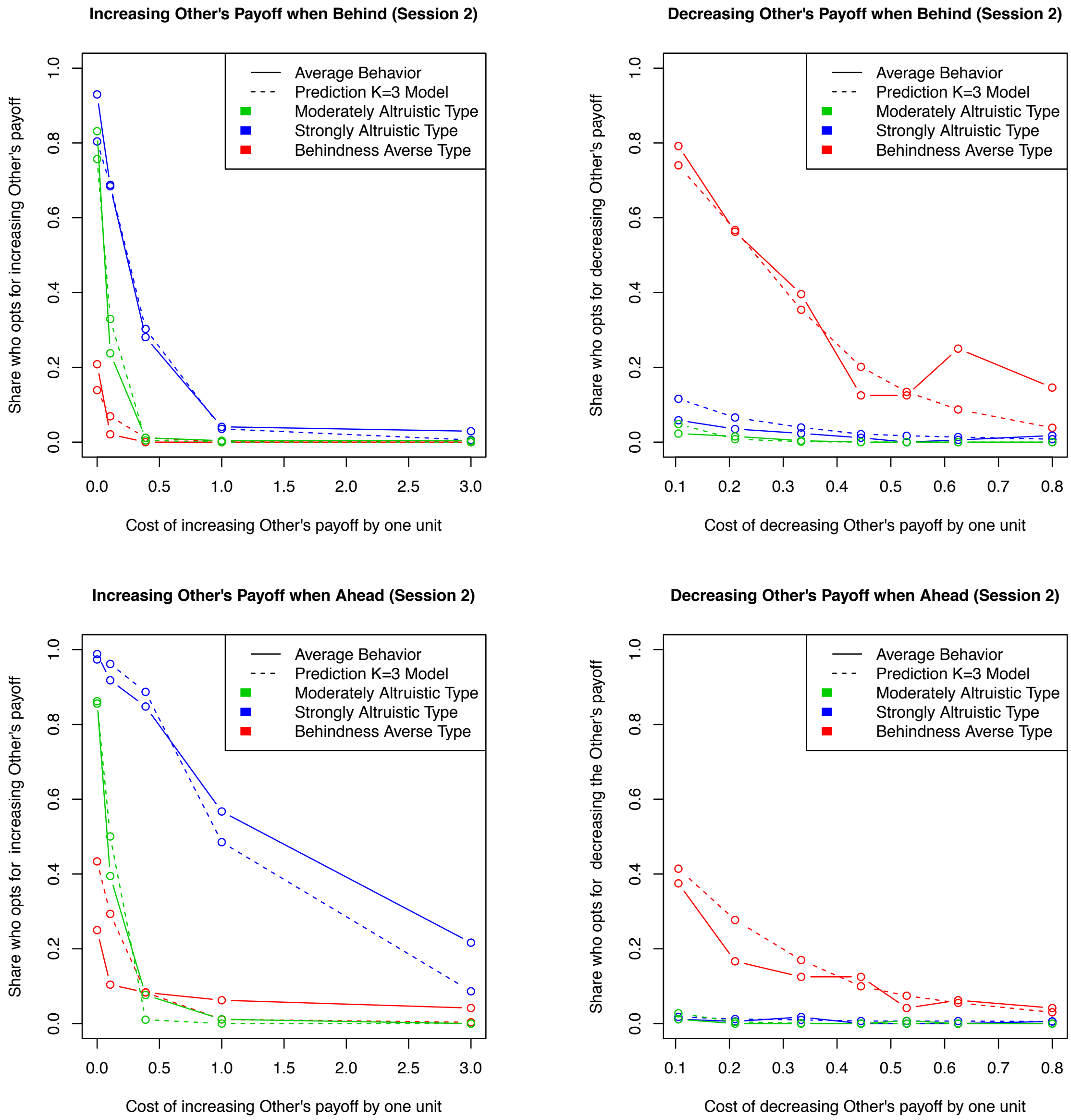

Figure B.15 (Analogue to Figure A.5 in the paper): Empirical and predicted willingness of the different preference types to change the other player's payoff across cost levels in Session 2. The empirical willingness corresponds to the fraction of subjects of a given preference type that chose to change the other player's payoff in the indicated direction. The predicted willingness corresponds to the predicted probability that a given preference type changes the other player's payoff in the indicated direction. It is based on the random utility model presented in Section 3.1 and uses the estimated type-specific parameters of Session 2 on all dictator and reciprocity games.

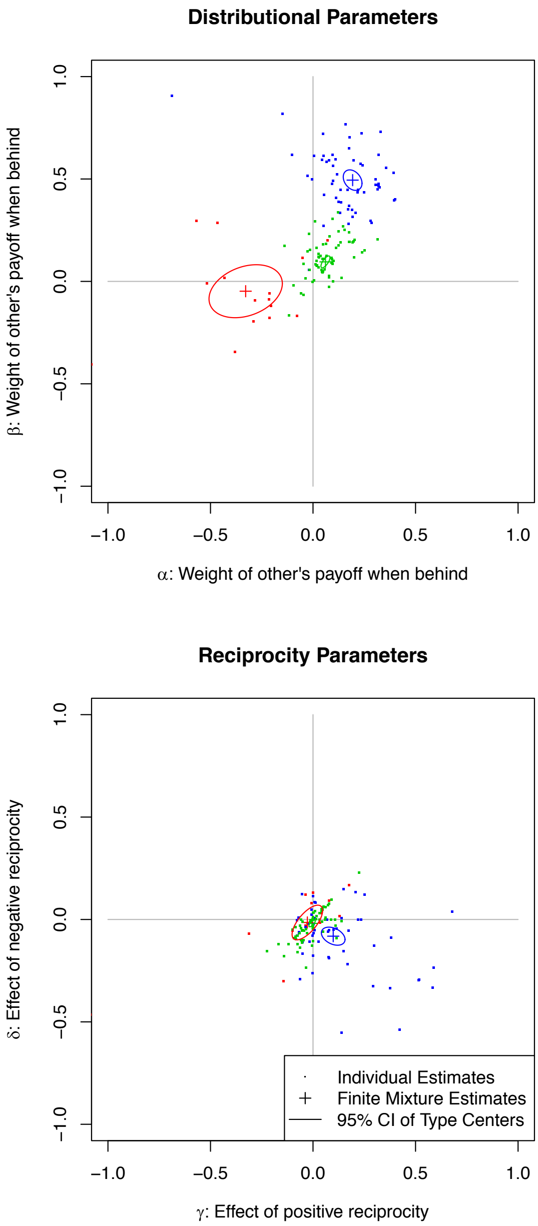

Figure B.16 (Analogue to Figure A.6 in the paper): Distribution of individual-specific parameter estimates along with type-specific parameter estimates ($K = 3$ model) in Session 2. The colors of the individual-specific estimates indicate the underlying subjects' classification into preference types according to the individual posterior probabilities of type-membership (see equation (7)): Moderately Altruistic (MA) Types are shown in green, Strongly Altruistic (SA) Types in blue, and Behindness Averse (BA) Types in red.

References

Fisman, R., S. Kariv and D. Markovits (2007): Individual Preferences for Giving. American Economic Review, 97, 1858-1876.