Supplementary Material of "Can transparency undermine peer review? A simulation model of scientist behavior under open peer review"

This appendix includes results from comprehensive computational tests on the effect of initial conditions on simulation outcomes reported in the manuscript. All results here were averaged over 200 realizations of 3,000-run simulations.

1 Number of scientists

In order to examine whether the population size could affect outcomes, we performed a set of tests by varying the number of scientists (in the manuscript, we reported results from $n = 240$). We ran simulations by testing the impact of initializing a population of $n = 50, 100, 200, 500$ on the main outcomes.

Figure 1 shows that we did not find any relevant qualitative differences of the size of the population of scientists on *publication bias* across scenarios. We only found minimal differences between scenarios for $n \geq 100$, where the distribution of average publication bias over model realizations was narrower, thereby providing more precise mean estimates.

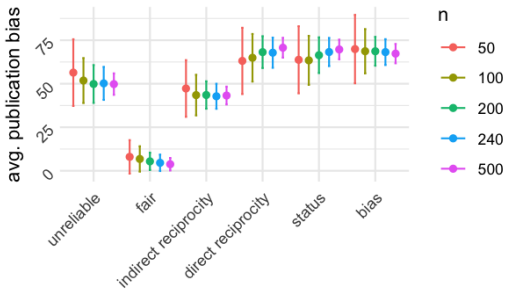

Figure 1: Sensitivity analysis of publication bias to N across different peer review models and scenarios.

Figure 2 shows similar results in terms of differences in *reviewing expenses* across scenarios. By varying n, we did not find any remarkable qualitative difference compared to the simulation results reported in the manuscript ($n = 240$). With the increase of n, the variance of the average review expenses over model realizations was reduced, thereby generating clearer differences across scenarios.

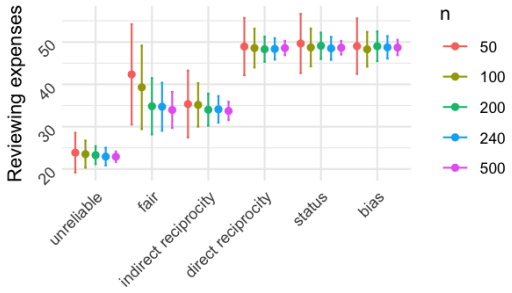

Figure 2: Sensitivity analysis of reviewing expenses to N across different peer review models and scenarios.

Figure 3 shows no remarkable effects of the size of the population on the *average quality of publications*. Differences between scenarios were more pronounced when $n \geq 100$.

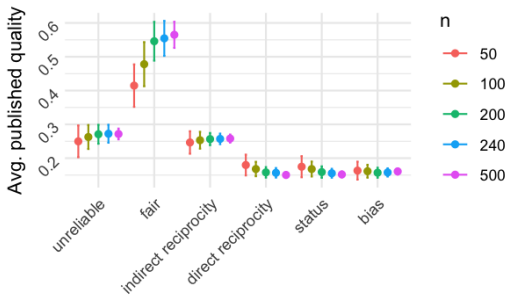

Figure 3: Sensitivity analysis of the average quality of publications to N across different peer review models and scenarios.

2 Gini index

Figure 4 shows the association between the size of the population and inequality in resource distribution. We measured the Gini index of each outcomes and estimated the effect of n across scenarios. While under confidential peer review, a larger population of scientists increased inequality of

outcomes, the opposite occurred with open peer review, where resource inequality decreased with any increase of the population. We found the most remarkable difference in the *fair* scenario under confidential peer review.

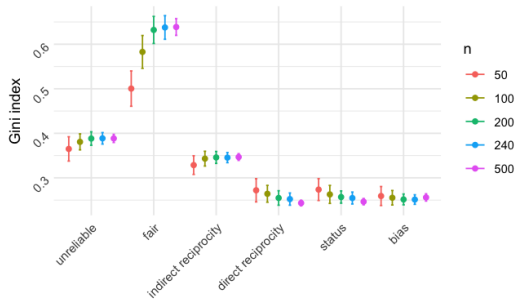

Figure 4: Gini index across different scenarios and values of n.

3 Velocity

Figure 7 shows that simulation results on *publication bias* are not sensitive to changes in v, as differences across values within scenarios were negligible.

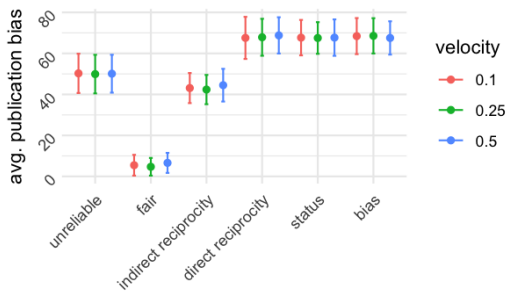

Figure 5: Sensitivity analysis of publication bias to v across different peer review models and scenarios.

Figure 6 shows that there were no qualitative differences in terms of *reviewing expenses* when varying initial values of v.

Figure 7 shows that simulation results on *average quality of publications* are not sensitive to changes in v, as differences across values within scenarios were negligible.

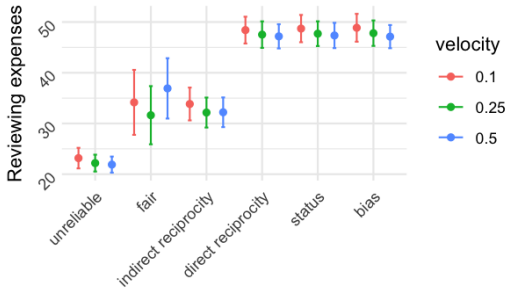

Figure 6: Sensitivity analysis of reviewing expenses to v across different peer review models and scenarios.

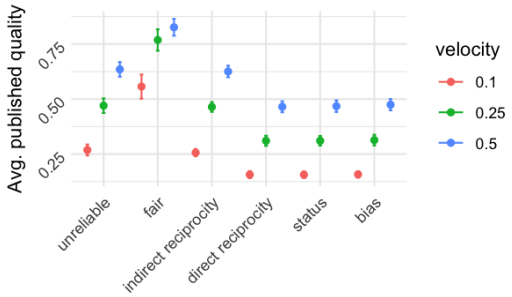

Figure 7: Sensitivity analysis of average quality of publications to v across different peer review models and scenarios.

4 Author bias factor

Figure 8 shows that varying initial values of B did not determine any qualitative difference on *publication bias* outcomes. The only remarkable difference was found under confidential peer review and *fair* and *indirect reciprocity* behaviour when B increased.

Figure 9 shows that varying initial values of B did not determine any relevant differences on *reviewing expenses* outcomes.

Figure 10 shows that varying B did not affect the *average quality of publications* across all scenarios.

5 Proportion of accepted publications

Figure 11 shows that *publication bias* simulation outcomes were not sensitive to the variation of P. We found remarkable differences between scenarios when P increased, as the distribution of mean values over model realizations were reduced, thereby providing more precise estimates. While within-scenario differences across P values were minimal, under confidential

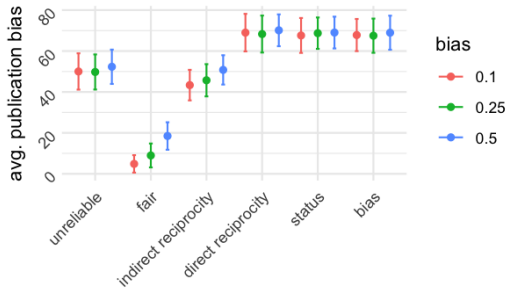

Figure 8: Sensitivity analysis of publication bias to B across different peer review models and scenarios.

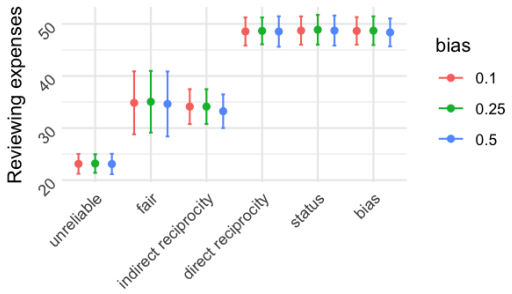

Figure 9: Sensitivity analysis of reviewing expenses to B across different peer review models and scenarios.

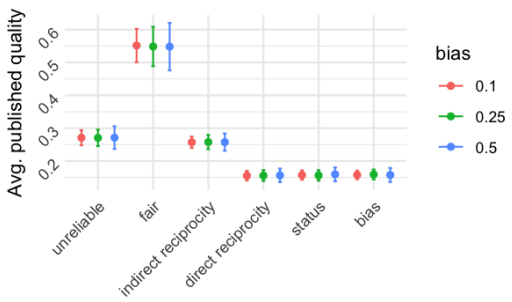

Figure 10: Sensitivity analysis of average quality of publications to B across different peer review models and scenarios.

peer review and in the two baseline scenarios (*fair* and *unreliable*), increasing the value of P (i.e., publishing a higher rate of manuscripts) tended to decrease *publication bias* within the other scenarios, especially if $P = 0.5$ (i.e., 50% of submitted manuscripts were accepted for publication at the end of each simulation tick).

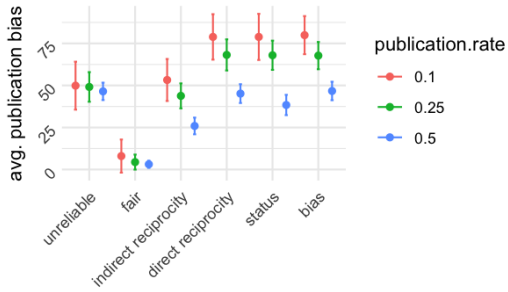

Figure 11: Sensitivity analysis of the publication bias to P across different peer review models and scenarios.

Figure 9 shows that *reviewing expenses* outcomes were not sensitive to variations in the initial values of P, except for the *fair* scenario, where $P = 0.1$ reduced the value of *reviewing expenses* the most.

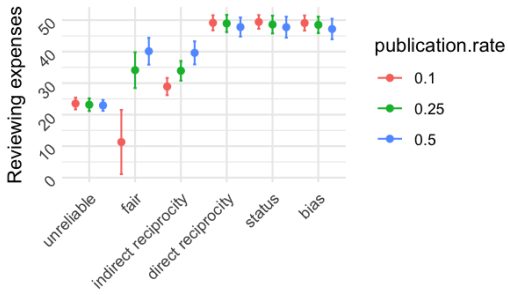

Figure 12: Sensitivity analysis of the reviewing expenses to P across different peer review models and scenarios.

Figure 10 shows that varying initial values of P did not affect the *average quality of publications* across all scenarios.

6 Discount factor on resources for unreliable reviewers

Figures 14, 15, and 16 show varying the initial values of d did not affect *publication bias*, *reviewing expenses*, and *average quality of publications* across scenarios.

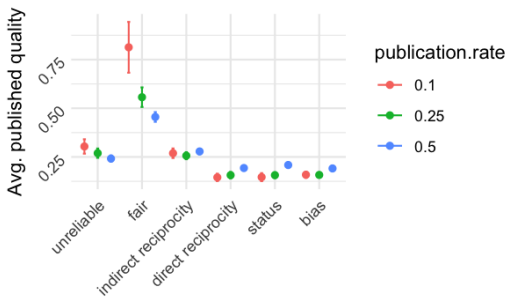

Figure 13: Sensitivity analysis of average quality of publications to P across different peer review models and scenarios.

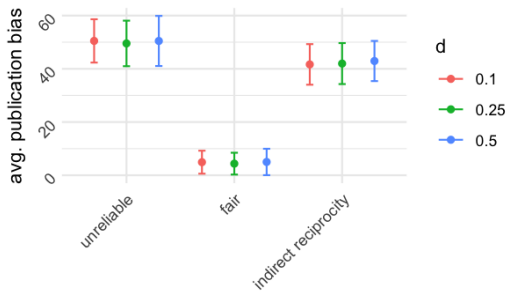

Figure 14: Sensitivity analysis of the publication bias to d across different scenarios in confidential peer review.

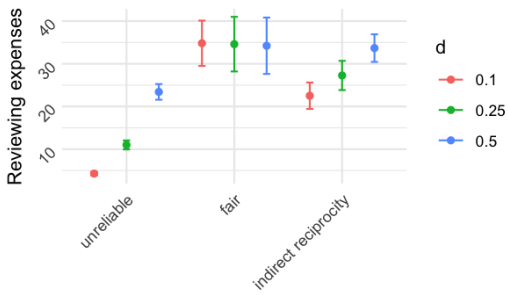

Figure 15: Sensitivity analysis of reviewing expenses to d across different scenarios in confidential peer review.

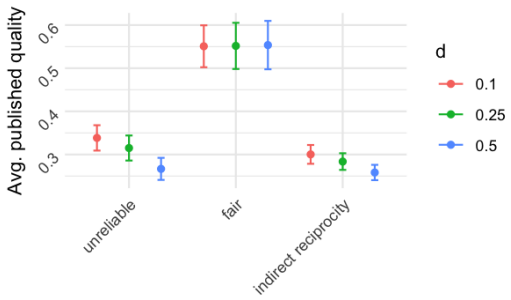

Figure 16: Sensitivity analysis of average quality of publications to d across different scenarios in confidential peer review.