Enter the Editor: Assessing editorial decisions' impact in Double Blind and Open Review systems via evolutionary game logit-simulations

Mantas Radzvilas, Francesco De Pretis, William Peden and Daniele Tortoli

Department of Biomedical Sciences and Public Health, Marche Polytechnic University, Italy
Department of Communication and Economics, University of Modena and Reggio Emilia, Italy
Erasmus Institute for Philosophy and Economics, Erasmus University Rotterdam, The Netherlands

Corresponding author: francesco.depretis@unimore.it

Background

Science has been tremendously successful in producing enlightening and useful knowledge. Throughout science's history, its institutions have been examined and reformed. The Replication Crisis, in which research in many sciences fails to replicate at the expected rate (Ioannidis, 2005; Schooler, 2014) suggests that scientific institutions still have much room for (Bohannon, 2016). Economics can make a particular contribution by studying the incentive structure of scientific institutions in a systematic way. One of the most prominent incentives for scientists is the pressure to publish, i.e. "publish or perish". Consequently, the publication system in general and especially peer-review is a target of suggested reforms. However, despite numerous controversies about peer-review systems, plus the clearly complex incentives affecting the decisions of authors and reviewers, they have received very little systematic and strategic analysis (Rogers, 2002).

Study rationale, objectives and methods.

In previous work (Radzvilas et al. 2020), we started addressing this feature of the peer-review literature by applying the tools of game theory. We used simulations to develop an evolutionary model based around a game played by authors and reviewers, before exploring some of its tendencies. In particular, we examined the relative impact of double-blind peer-review and open review on incentivising reviewer effort under a variety of parameters. We also compared (a) the impact of one review system versus another with (b) other alterations, such as higher costs of reviewing. One idealization in that study was that editors were not part of our model. In this work, we add editors as additional players in this game and we assess the impact of their decisions over the quality of manuscripts. For methods, we use game theory, but we do not assume perfect rationality; we use an evolutionary approach and logit-simulation to uncover the behavioral tendencies of agents with imperfect rationality and stochastic unpredictability. The incentives, strategies, and interactive effects of editors in the scientific publication system has not previously been studied using this methodology, so our objective is to enrich the economics of science in a novel and practically relevant way.

Main results and findings

As in our previous work, we find that is no reliable difference between peer-review systems in our model. Furthermore, under some conditions, higher payoffs for good reviewing can lead to less (not more) author effort under open review. Finally, compared to the other parameters that we vary, it is the exogenous utility of author effort that makes an important and reliable difference in our model.

Table 1: Examples of payoff matrices for $\epsilon = 0.3$. Matrix **(a)** is defined for a Double Blind system and it is characterized by having 6 Nash equilibria in pure strategies, whereas matrix **(b)** is calculated for an Open Review mechanism and owns a unique Nash equilibrium in pure strategies.

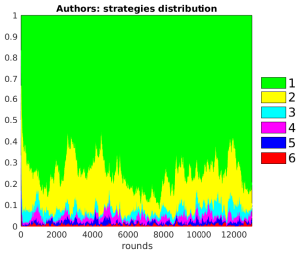
Author \ Reviewer	$S1$	$S2$	$S3$	$S4$	$S5$	$S6$
$S1$	(1.13, 1.00)	(1.13, 1.00)	(−0.03, 1.00)	(−0.03, 1.00)	(−0.03, 1.00)	(−0.03, 1.00)
$S2$	(1.14, 1.00)	(1.14, 1.00)	(1.14, 1.00)	(−0.04, 1.00)	(−0.04, 1.00)	(−0.04, 1.00)
$S3$	(1.15, 1.00)	(1.15, 1.00)	(1.15, 1.00)	(1.15, 1.00)	(−0.05, 1.00)	(−0.05, 1.00)
$S4$	(1.16, 1.00)	(1.16, 1.00)	(1.16, 1.00)	(1.16, 1.00)	(1.16, 1.00)	(−0.06, 1.00)
$S5$	(1.17, 1.00)	(1.17, 1.00)	(1.17, 1.00)	(1.17, 1.00)	(1.17, 1.00)	(1.17, 1.00)
$S6$	(1.18, 1.00)	(1.18, 1.00)	(1.18, 1.00)	(1.18, 1.00)	(1.18, 1.00)	(1.18, 1.00)

(a) Payoff matrix computed for $(\delta, \mu) = (0, 0)$

Author \ Reviewer	$S1$	$S2$	$S3$	$S4$	$S5$	$S6$
$S1$	(1.13, 1.03)	(1.13, 1.00)	(−0.03, 1.00)	(−0.03, 1.00)	(−0.03, 1.00)	(−0.03, 1.00)
$S2$	(1.14, 1.06)	(1.14, 1.03)	(1.14, 1.00)	(−0.04, 1.00)	(−0.04, 1.00)	(−0.04, 1.00)
$S3$	(1.15, 1.17)	(1.15, 1.14)	(1.15, 1.11)	(1.23, 1.08)	(−0.05, 1.00)	(−0.05, 1.00)
$S4$	(1.16, 1.36)	(1.16, 1.33))	(1.16, 1.30))	(1.24, 1.27))	(1.40, 1.24)	(−0.06, 1.00)
$S5$	(1.17, 1.55)	(1.17, 1.52)	(1.17, 1.49)	(1.25, 1.46)	(1.41, 1.43)	(1.57, 1.40)
$S6$	(1.18, 1.74)	(1.18, 1.71)	(1.18, 1.68)	(1.26, 1.65)	(1.42, 1.62)	(1.58, 1.59)

(b) Payoff matrix computed for $(\delta, \mu) = (0.3, 1.6)$

Figure 1: A strategy distribution for authors. Here we give an example of a strategies distribution among authors. In panel **(1a)** strategies are graphically represented with green color marked by number 1 being the highest effort level and red color marked by number 6 being the lowest one, whereas in panel **(1b)** a table based on the same data is shown

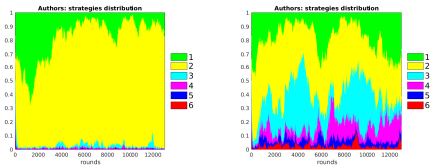


(a) Graph for an Open Review system for $\epsilon = 0.3$ with $(\delta, \mu) = (0.3, 1.6)$ ruled by matrix **(b)** shown in **table (1)**. In this picture, strategies percentages (Y-axis) are plotted against rounds (X-axis)

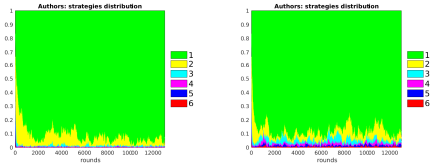
	Strategies							
Data	1	2	3	4	5	6	$\overline{s}$	$\mathbb{E}(\overline{s})$
Last 3000 rounds	71.74 ± 6.72	17.26 ± 5.31	5.31 ± 2.37	2.85 ± 1.63	1.73 ± 1.09	1.12 ± 0.71	1.490	0.751
Entire history	73.61 ± 7.84	17.15 ± 6.72	3.93 ± 2.19	2.65 ± 1.72	1.62 ± 1.18	1.05 ± 0.76	1.447	0.755

(b) Table presenting mean values and standard deviations (in percentage terms) computed for each strategy over time (rounds) as shown in panel **(1a)**. Last two columns report a weighted average for strategies and the expected value of authors effort levels. Both measures are presented for the full set of data and for the last 3000 rounds

Figure 2: Strategy distributions for authors. On left panels, figures representing interactions under Double Blind system are shown, whereas on right panels examples stemming from Open Review mechanism are reported. Strategies percentages (Y-axis) are pictured against rounds (X-axis). At the same time, in the upper panels (a) and (b) we display authors' strategy distributions for $\epsilon = 0.1$ and in lower panels (c) and (d) for $\epsilon = 0.4$

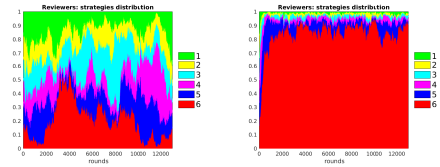


(a) $(\delta, \mu) = (0, 0)$ **(b)** $(\delta, \mu) = (0.3, 1.6)$

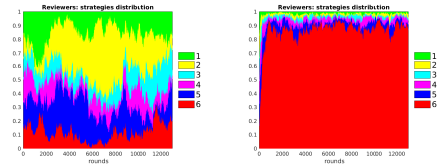


(c) $(\delta, \mu) = (0, 0)$ **(d)** $(\delta, \mu) = (0.3, 1.6)$

Figure 3: Strategy distributions for reviewers. Left and right panels are organized mirroring for reviewers what is presented for authors in **figure (2)**. Similarly, in upper panels (a) and (b) we display reviewers' strategy distributions for $\epsilon = 0.1$ and in lower panels (c) and (d) for $\epsilon = 0.4$



(a) $(\delta, \mu) = (0, 0)$ **(b)** $(\delta, \mu) = (0.3, 1.6)$



(c) $(\delta, \mu) = (0, 0)$ **(d)** $(\delta, \mu) = (0.3, 1.6)$

Conclusions

It is hard to determine a reliable positive impact for any peer-review system over another. Different systems and associated incentives can have surprising and negative effects. Despite the prominence of peer-review systems in contemporary debates about scientific institutions, other incentives for author effort might be better targets for change.

References

1. Bohannon, J. (2016). About 40% of economics experiments fail replication survey. Science.
2. Ioannidis, J. P. A. (2005). Why most published research findings are false. PLoS Medicine 2(8), e124.
3. Radzvilas, M., De Pretis F., Peden W., Tortoli D., and Osimani B. (2020). Double blind vs. open review: an evolutionary game logit-simulating the behavior of authors and reviewers. Annals of Operations Research. 24 pages, under review.
4. Rogers, L. F. (2002, December). Manuscripts: To blind, or not to blind, that is the question. American Journal of Roentgenology 179(6), 1373–1373.
5. Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. Nature 515(7525), 9–9.