Cardiff Economics Working Papers

Working Paper No. E2022/14

Tactical Refereeing and Signaling by Publishing

Sergey V. Popov

July 2022

ISSN 1749-6010

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

Tactical Refereeing and Signaling by Publishing*

Sergey V. Popov

July 20, 2022

Abstract

A peer review is used ubiquitously in hiring, promotional, and evaluation decisions, within academia and beyond. It is usually conducted to allocate limited resources, such as the budget of a funder or the pages of a journal. With limited capacity, a peer review may lead to negatively biased evaluations precisely because approving a peer's worthy project lowers the chance that a referee's own project will be approved. I show that limited capacity is inconsistent with a hypothesis that the decision-maker's policy is to stimulate efforts, and I discuss possible decision-maker motivations that could lead to a limited capacity policy.

Keywords: refereeing, peer review.

JEL: C78

In many environments, decisionmakers solicit peer advice to allocate a scarce resource to individuals. As an illustration, consider publishing a paper in a journal. Referees' incentives to report their true assessments are frequently aligned with the editors' incentives: referees, by being peers of the authors, are interested in being associated with solid work that provides meaningful impact. Strategic considerations often cloud this alignment, creating incentives for misreporting, and therefore these strategic interactions become a point of interest for both theoretical and empirical research.

In this paper, I show that, even if we suppress all channels that can induce strategic misreporting, *tactical* considerations remain: being a peer implies competing for the same scarce resource. This affects the referees' incentives to truthfully report whether the project under consideration is acceptable: approving a project means jeopardizing the opportunity to win the resource for themselves. I show that the editor who maximizes the quality of

*Popov: Cardiff University, PopovS@cardiff.ac.uk. The usual disclaimer applies.

the output will never choose a combination of capacity and quality thresholds that lead to dishonest referee reports in the equilibrium. I conclude by illustrating that one rationalization for dishonest replies from referees in equilibrium is the heterogeneity of authors. If the editor is interested in improving their journal's function as a signaling device, the editor can artificially suppress the capacity, to induce referees' misreporting, and lower the marginal benefit of effort. This discourages authors from exerting effort. Innate ability difference becomes relatively more important in the acceptance outcomes, and acceptance therefore becomes a better signal of innate ability.

This study's findings apply to other contexts of peer evaluation, for instance to students grading the work of their peers in classes graded by the curve, or office workers providing feedback about ideas of their colleagues. As long as capacity matters, peer evaluation becomes less reliable, which may or may not be a problem, depending on the decision-maker's motivations.

Literature

Many papers deal with refereeing from the perspective of the author. Frequently, the referee is not too strategic, by relatively straightforwardly reporting his or her signal to the editor. Engers and Gans (1998) present a model where referees decide whether to provide a review or not; paper studies the effects of monetary incentives on publication delays. Ellison (2002) builds a model of authors revising their papers, but the referees do not take any actions. Baghestanian and Popov (2018) study the choice of effort by authors of different abilities. They find that different incentives for authors of different abilities on submissions strategies lead to different consequences of changes in the refereeing black box on effort choice of different abilities, subsequently resulting in to publication strategies being informative about authors' abilities. However, in their paper referees merely report their signal to the editor. Gerardi et al. (2021) model the effort choice of the politician whose idea might be embraced or denounced by a regulator of varying impartiality, with implications on the politician's efforts due to this impartiality, however, the reason for this impartiality is not discussed. Bayar and Chemmanur (2021) provide a model of editor's choice between biased referees; again, it is not clear why there is a pre-existing bias.

Some empirical literature on soliciting advice from experts indicates that there are factors that might impede the referee's impartiality. Hale et al. (2021) report that "attractive" authors seem to collect more citations and publish better than "less attractive" colleagues. Hengel (2020) shows a difference in writing quality exhibited in published work of writers of different gender; Alexander et al. (2021) shows that female economists spend more time revising papers, and referees spend more time reviewing papers written by female academics in economics. Card and DellaVigna (2020) argue that editors go along with referee recommendations, and that they value citations; indeed, it appears that more enthusiastic references improve a paper's chances to receive more citations, while referees of differing skills provide a similar refereeing quality. Brogaard et al. (2014) find that editors in finance field are able to attract good papers from their colleagues; Colussi (2018) finds a similar result in economics. Publications appear to be informative in the refereeing process: Heckman and Moktan (2018) finds that publishing two top-5 papers in the economics field is significantly different from publishing one on a tenure decision in competitive US institutions.

Generically, referees are beneficial for the publishing process: Hadavand et al. (2020) report that referees improve papers (at least in the first round of refereeing), which is in line with Card and DellaVigna (2020), who find that referees improve citations.

The Symmetric Model

I propose a symmetric single-period global game where each active agent has two roles, one of an *author* and another of a *referee*. An *editor*'s role is to assign papers written by authors to referees so that authors don't referee their own papers. I concentrate on a single-period game to assume away strategic dishonesty, such as nepotism, or "tit-for-tat strategies"; all these concerns can be included with the interaction that we are studying here in a more sophisticated model.

There is a number of authors, $N > 2$, working in the field. There is one journal in the field. The publication process is modeled as follows:

- The editor chooses capacity $\Gamma \in \{2, .., N\}$ and paper quality threshold $\bar{q}$;
- Each author gets a paper idea with a known innate quality θ. Each author spends effort e, obtaining the paper of quality $q = \theta + e$; costs of effort are $c(e)$, which is for the sake of brevity concave and thrice differentiable. Authors send their papers to the editor;
- The editor sends these papers to referees for an evaluation of whether their quality is above $\bar{q}$. Referees are authors too, however each author gets a paper from a different author to referee. The process is true double blind, so referees cannot learn the identity of authors and use it in *quid pro quo* or other strategic interactions;

- Referees effortlessly evaluate the true paper's quality, $q + \varepsilon$, where ε is distributed on $\mathbb{R}$ with full support. Referees compare the paper's quality with $\bar{q}$ and respond to the editor. They might report the truth, or they might misrepresent their findings;

- The editor obtains the referees' recommendations. If the number of papers recommended for publication by referees exceeds Γ, the editor accepts at random Γ out of the number accepted by the referees, and rejects the rest;[1]

- Authors whose papers are accepted get a *reputational payoff* $W(G, B)$, which is a function of *good* and *bad* papers (with qualities above and below $\bar{q}$, respectively) published by the journal; all authors, published or not, bear the costs of their efforts.

For now, we will limit ourselves to a symmetric model: we will assume that all authors have the same level of ability θ; by symmetry, their strategies will be identical. The equilibrium in the symmetric model is a collection of:

- The probability that each of the referees misrepresent their signal when $q + \varepsilon > \bar{q}$, denoted by $\bar{s}_G \in [0, 1]$, consistent with referees' incentives;

- The probability that each of the referees misrepresent their signal when $q + \varepsilon < \bar{q}$, denoted by $\bar{s}_B \in [0, 1]$, consistent with referees' incentives;

- And e^*, the choice of effort by the authors, who maximize their utility:

$$\max_e E\left[W(G, B)|\bar{s}_G, \bar{s}_B\right] P[\text{pass referees}|\bar{s}_G, \bar{s}_B] P[\text{publish}|e] - c(e).$$

When refereeing, authors care about the field's reputation, which is a function of the number of good papers published and the number of bad papers published. If the journal publishes a good paper, this improves every published author's reputational payoff, while if the field has published a bad paper, it worsens every published author's payoff. The referee can only harvest the reputational payoff if their own paper is published:

$$u(s, \bar{s}) = \mathbf{E}\left[I(\text{publish yourself}) W(G, B)|s, \bar{s}\right],$$

[1] There are multiple examples of economics editors admitting that a lot of excellent papers are being rejected in their journals, in personal communications, and in public presentations (example). Even Nobel prize winners' papers are rejected, as documented by Gans and Shepherd (1994). At one time, a medical journal editor decided to reject some papers post-acceptance in order to deal with a publication lag. The summary of this case can be found here: `https://www.wame.org/post-acceptance-rejection-of-a-manuscript`.

where the chance for the author to publish and the distribution of counts of good and bad papers is determined by the propensity to deliver misleading reports in the profession $\bar{s}$ and the author's own propensity to be misleading s.

All authors are equally perfectly capable of evaluating the quality of the paper they receive. The refereeing strategy is therefore $\{s_G, s_B\}$, which are the probabilities to mislead about the signal to the editor if a referee sees a good or a bad paper, respectively.

When Capacity Does Not Matter Assume that the editor sets $\Gamma = N$. Then the referee's decision to accept another author's paper does not affect the chances of the referee's own paper to get accepted.

Since the referee's choices while refereeing do not affect the referee's outcomes regarding their own paper, the payoff can be rewritten as

$$u(s, \bar{s}) = \mathbf{E}\left[I(\text{publish})|\bar{s}\right]\mathbf{E}\left[W(G, B)|s, \bar{s}\right].$$

What happens to the reputational payoff? Observe that G and B are random variables of counts of good and bad papers, correspondingly, that get accepted by other referees. Then:

- If the referee sees a good paper, suggesting that the editor publishes it can only improve the referee's payoff: the reputational payoff's distribution improves in the first-order stochastic dominance sense because of the monotonicity of the payoff function.

- If the referee sees a bad paper, suggesting that the editor publishes it can only worsen the referee's payoff: the reputational payoff's distribution worsens in the first-order stochastic dominance sense because of the monotonicity of the payoff function.

Therefore, by first order stochastic dominance, no matter what other referees do, the only optimal behavior for referees is $s_B = 0$, $s_G = 0$, and only good papers are published. This immediately extends to the case when referees are only getting a precise enough signal about the quality of the paper under their consideration, as provided by the monotonicity of the payoff with respect to the signal under a wide family of signal distributions; the only caveat is that some bad papers end up getting published if referees obtain an unfortunate signal.

When Capacity Matters Assume that the editor imposes an upper bound on the number of papers they can accept for publication. If the total number of papers deemed acceptable by the referees is above $\Gamma \geq 2$, the editor picks Γ papers among those recommended for publication at random.[2] This creates a downside for referees to accept good papers for the

[2]The editor could also evaluate papers by reading and then discarding acceptable papers based on their

publication: even if the additional paper improves the distribution of reputation outcomes, it creates competition for capacity, which jeopardizes author's own chances of publishing by sheer chance of rejection by the editor. Therefore:

- If the referee sees a bad paper, suggesting that the editor publishes it can only worsen the referee's payoff: the reputational payoff's distribution worsens in the first-order stochastic dominance sense because of the monotonicity of the payoff function, and the chance of publishing of the author's own paper decreases, so $s_B = 0$, and therefore in equilibrium $\bar{s}_B = 0$;

- If the referee sees a good paper, suggesting that the editor publishes it can either increase or decrease the referee's payoff:

  - it can stochastically improve the reputational outcome: in cases when less than $\Gamma-1$ papers are deemed good by other referees, the distribution of payoffs improves in the first order stochastic dominance sense;

  - but in cases when at least Γ papers are deemed good by the other referees, advising a seemingly good paper to be published *lowers* the chance to publish the referee's own paper without affecting the distribution of payoffs: after all, there is at least Γ of other good papers, since bad papers do not survive the refereeing process by dominance.

$\bar{s}_G = 1$ cannot be a part of a meaningful equilibrium outcome: the other agent's behavior does not allow any papers to be published and either leads to the author going against the equilibrium or no papers being published at all, and neither of these outcomes are interesting. Can $s_G = \bar{s}_G = 0$ be an equilibrium outcome? Yes if the decrease of a chance to publish is less than the benefit from potentially improving the publication chances. This is possible if Γ is large enough relative to N: it stops to matter. If Γ matters, we face a mixed equilibrium, where some good papers might not get a favorable response from the referees.

Since $s_B = \bar{s}_B = 0$, I suppress the subscript of s_G where unnecessary.

Equilibrium in the Refereeing Subgame

EXAMPLE Let there be $N = 5$ authors, the chance of writing a good paper for each of them be $p = 1 - F(\bar{q} - \theta - e^*)$, and the journal capacity Γ be set at 2. Then, there are three

own assessment. From the perspective of the referee, this is observationally equivalent to randomization, at least in the symmetric case, where all authors, and therefore all papers, are identical.

possible publication outcomes: first, no papers published; second, one good paper published; and third, two good papers published. In the equilibrium, $\bar{s}$ must be such that the author is indifferent between misreporting and saying the truth about the good paper that they are refereeing. The author might expect the editor to have between 0 and 3 papers recommended for publication if their own paper is not recommended, and between 1 and 4 if their own paper is recommended.

Assume the reputational payoff from publishing G good papers and B bad papers is $W(G, B) = G - B$. Let $P_i(\bar{s})$ denote the probability that there is i good papers recommended for publication conditional on the equilibrium refereeing strategy $\bar{s}$. Since no one misrepresents bad papers, the overall payoff is

$$\underbrace{p(1-\bar{s})}_{\text{Publish}}\left(P_0(\bar{s})2 + P_1(\bar{s})\frac{2}{3}2 + P_2(\bar{s})\frac{2}{4}2 + P_3(\bar{s})\frac{2}{5}2\right)$$

if the referee tells the truth about the good paper they face, and

$$\underbrace{p(1-\bar{s})}_{\text{Publish}}\left(P_0(\bar{s})1 + P_1(\bar{s})2 + P_2(\bar{s})\frac{2}{3}2 + P_3(\bar{s})\frac{2}{4}2\right)$$

if the referee misleads. The difference between these two, omitting $p(1-\bar{s})$, is

$$P_0(\bar{s}) - P_1(\bar{s})\frac{1}{3} - P_2(\bar{s})\frac{1}{3} - P_3(\bar{s})\frac{1}{5},$$

which must be equal to zero if agents are indifferent between misreporting and telling the truth.

$$P_0(\bar{s}) = \overbrace{(1-p)^3}^{\text{all bad}} + \overbrace{3(1-p)^2 p\bar{s}}^{\text{one good}} + \overbrace{3(1-p)p^2\bar{s}^2}^{\text{two good}} + \overbrace{p^3\bar{s}^3}^{\text{all good}}.$$

$$P_1(\bar{s}) = \overbrace{3(1-p)^2 p(1-\bar{s})}^{\text{one good}} + \overbrace{3(1-p)p^2 \cdot 2\bar{s}(1-\bar{s})}^{\text{two good}} + \overbrace{p^3 \cdot 3(1-\bar{s})\bar{s}^2}^{\text{all good}}.$$

$$P_2(\bar{s}) = \overbrace{3(1-p)p^2(1-\bar{s})^2}^{\text{two good}} + \overbrace{p^3 \cdot 3(1-\bar{s})^2\bar{s}}^{\text{all good}}, \quad P_3(\bar{s}) = p^3(1-\bar{s})^3.$$

The indifference equation becomes a cubic equation in $\bar{s}$ with three solutions. One can notice that $p = \gamma/(1-\bar{s})$ for some γ, so that the chance to publish, $p(1-\bar{s})$, is constant with respect to p; this simplifies the indifference characterizing equation to $11\gamma^3 - 30\gamma^2 + 25\gamma - 5 = 0$, yielding a result of $\gamma = 0.2905$, making $\bar{s} = 1 - 0.2905/p$. It is an increasing function of p on $[0.2905, 1]$. If $p = 1$, $\bar{s} = 0.7095$: when everyone except the editor knows that all five papers are good, there is still a 30% chance at best to obtain a positive referee report, and after that the editor can discard the paper because of the capacity concerns.

General results The constancy of $p(1-\bar{s})$, as long as $\bar{s}$ is above zero, is a general result: both payoffs are the probability to publish $p(1-\bar{s})$ times the expected payoff from publishing. This is an expectation over payoffs that are not functions of p or s, distributed over states of the world that describe how many other papers are deemed worthy of publication by other referees, which is a negative binomial distribution with probability of success equal to $p(1-\bar{s})$.

Denote $\tilde{p} = p(1-\bar{s})$, and let $\xi(\pi, \Omega)$ be a random variable distributed with binomial distribution, with probability of success π and quantity of trials Ω.

Proposition 1. *Assume authors are risk-neutral: $dW/dG = a > 0$. In mixed strategy equilibria, $\tilde{p} = p(1-\bar{s})$ solves*

$$aP(\xi(\tilde{p}, N-2) \leq \Gamma - 2) = P(\xi(\tilde{p}, N) \geq \Gamma + 1)\frac{\Gamma W(\Gamma)}{N(N-1)\tilde{p}^2}. \tag{$\star$}$$

Proof. From the perspective of each referee, there are three papers: their own paper, the paper that they are refereeing, and the rest of the papers, the quantity of which is equal to $N-2$. There is nothing in the referee's powers that can affect the rest of the papers; the number of the rest of the papers that are going to get a positive referee report is distributed as $\xi(\tilde{p}, N-2)$. Since the referee only gets a payoff if their own paper is accepted, all payoffs could be evaluated conditional on the referees' own paper acceptance. The payoff in case the referee decides to reject the paper they review is:

$$EU(reject) = \sum_{i=0}^{N-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} W(\min(i+1, \Gamma)) P(\text{yours picked out of } i+1) =$$

$$= \sum_{i=0}^{\Gamma-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} W(i+1) + \sum_{i=\Gamma-1}^{N-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} W(\Gamma) \frac{\Gamma}{i+1}.$$

On the other hand, acceptance means there is one more paper that competes for the opportunity to be published:

$$EU(accept) = \sum_{i=0}^{N-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} W(\min(i+2, \Gamma)) P(\text{yours picked out of } i+2) =$$

$$= \sum_{i=0}^{\Gamma-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} W(i+2) + \sum_{i=\Gamma-1}^{N-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} W(\Gamma) \frac{\Gamma}{i+2}.$$

In the mixed strategy equilibrium, expected payoff of acceptance must be equal to the expected payoff of rejection, so the difference must be equal to zero:

$$\underbrace{\sum_{i=0}^{\Gamma-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!}}_{P(\xi(\tilde{p},N-2)\leq\Gamma-2)} \underbrace{(W(i+2)-W(i+1))}_{a} = \sum_{i=\Gamma-1}^{N-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} \frac{\Gamma W(\Gamma)}{(i+1)(i+2)}.$$

Rewrite the right-hand side sum as

$$\sum_{i=\Gamma-1}^{N-2} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} \frac{\Gamma W(\Gamma)}{(i+1)(i+2)} = \frac{\Gamma W(\Gamma)}{N(N-1)\tilde{p}^2} \sum_{i=\Gamma-1}^{N-2} \tilde{p}^{i+2} (1-\tilde{p})^{N-2-i} \frac{(N-2)!(N-1)N}{(i+2)!(N-2-i)!}.$$

Let $i+2=j$, then the right part transforms to

$$\frac{\Gamma W(\Gamma)}{N(N-1)\tilde{p}^2} \sum_{j=\Gamma+1}^{N} \tilde{p}^j (1-\tilde{p})^{N-j} \frac{N!}{j!(N-j)!} = \frac{\Gamma W(\Gamma)}{N(N-1)\tilde{p}^2} P(\xi(\tilde{p},N) \geq \Gamma+1).$$

Combining two sides yields the result in the proposition statement. □

Proposition 2. *Subgame equilibrium s^* exists.*

Proof. For a general increasing $W(G)$, $(\star)$ looks like

$$\sum_{i=0}^{\Gamma-2} \overbrace{[W(i+2)-W(i+1)]}^{>0} \tilde{p}^i (1-\tilde{p})^{N-2-i} \frac{(N-2)!}{i!(N-2-i)!} = P(\xi(\tilde{p},N) \geq \Gamma+1) \frac{\Gamma W(\Gamma)}{N(N-1)\tilde{p}^2}.$$

The left-hand side is positive at $\tilde{p} < 1$, zero at $\tilde{p} = 1$, and continuous in $\tilde{p}$. The right-hand side is zero at $\tilde{p} = 0$, positive at $\tilde{p} = 1$, and continuous in $\tilde{p}$. They must intersect at least once. □

The uniqueness is harder to establish: while the left-hand side is reliably decreasing in $\tilde{p}$, the right-hand side is generically not monotone.

Corollary 1. *If there is π chance that the referee will not consider tactical misleading, and will just convey their own signal, the chance of misleading for those who can mislead increases.*

Proof. If fraction π of referees always reply honestly, the chance of getting the paper accepted by referees is

$$\tilde{p} = p[\pi + (1-\pi)(1-\bar{s})].$$

The equation in terms of $\tilde{p}$ remains to be $(\star)$. When π increases, $\tilde{p}$ does not change, but that means that $1-\bar{s}$ must decrease. $\square$

This may rationalize why some authors find it easier to publish in journals than others, whilst simultaneously overburdened with refereeing requests. Editors may protect authors who they know to be producing good papers from random elimination at the publication decision stage in order to warrant knowing that these authors do not need to distort their reports about other authors. The cost is more competitiveness from the other authors.

Corollary 2. *If there is a publication incentive in an author's remuneration, $W(G) = b + aG$. Assume the subgame equilibrium is unique. Holding p fixed*[3]*, if b increases, $\tilde{p}$ decreases, which means $\bar{s}$ increases.*

Proof. An increase in unconditional payment b leads to no change in the marginal payment increase before the capacity matters, where the utility of accepting the reviewed paper is better than rejecting. However, when capacity matters, the loss due to more competing papers (the right-hand side of $(\star)$) includes the unconditional payment (multiplied by the difference in probability to not get published when rejecting and when accepting the paper which the referee reviews), so the right-hand side of $(\star)$ increases for every $\tilde{p}$. By monotonicity of the left-hand side of $(\star)$, $\tilde{p}$ must decrease. $\square$

Obviously, a change in $\bar{s}$ must lead to a change in incentives at the effort choice level of the game, canceling out some, if not all, of the changes in efforts due to a change in b. If we want to study how policy choices affect $\bar{s}$, we need to include efforts of authors into the analysis.

Authoring-Refereeing Game Equilibrium

The first stage of this game is studied extensively in Baghestanian and Popov (2018). Authors face a refereeing subgame where they know the equilibrium leads to a chance that the referee will misrepresent the paper with a chance $\bar{s}^*$ and expected reputational payoff

$$\omega = E\left[W(G,B)|\bar{s}^*, e^*, \Gamma, \text{author is accepted}\right],$$

[3]Efforts should increase due to better incentives to publish.

where e^* is the equilibrium level of effort[4]. Each author has ability θ, spends effort e, faces a quality threshold $\bar{q}$, and therefore solves

$$e^* = \arg\max_e \omega \times \overbrace{(1-\bar{s})}^{P[\text{pass referee}|\text{good}]} \times \underbrace{(1 - F(\bar{q} - \theta - e))}_{P[\theta+e+\varepsilon>\bar{q}]} - c(e). \tag{$\dagger$}$$

$F(\cdot)$ is the cdf of ε; assume it has a pdf $f(\cdot)$ which is single-peaked, with a mode at 0, and positive everywhere on $\mathbb{R}$. Meanwhile, the effort cost function $c(e)$ is a strictly convex thrice differentiable cost function with $c(0) = 0$, $c'(0) = 0$, and $c''(\cdot) > 0$.

Maximizing this utility function yields us, in a symmetric equilibrium, e^* as a function of $\bar{q}$ and $\bar{s}^*$. Denote $p^* = 1 - F(\bar{q} - \theta - e^*)$. After authors have made their choice of efforts, the refereeing subgame follows. Referees face a chance of seeing a good enough paper of p^* and a capacity of Γ, and the equilibrium of this subgame, according to Proposition 2, is $\bar{s}^* = \max(0, 1 - \tilde{p}/p^*)$ where $\tilde{p}$ is a function of Γ and N.

We therefore have a probability that the paper is good p^* as a function of the probability that the referee is going to be honest $1 - \bar{s}^*$ in the authoring stage, governed by ($\dagger$), and vice versa, $1 - \bar{s}^*$ as a function of p^* in the refereeing stage, governed by ($\star$). Manipulating Γ and $\bar{q}$ move the responses around, as depicted on Figure 1. Increase in Γ shifts the blue line in north-east direction (lowering $\bar{s}$ for every p), while an increase in $\bar{q}$ is decreasing p^* (Proposition 1 of Baghestanian and Popov (2018)), shifting the green line to the left. Why would the publication economy exhibit the outcome as on Figure 1 (solid lines intersection), where $\bar{s} \neq 0$?

Proposition 3. *Taking ω as fixed, if the editor is interested in maximizing the effort, they choose (i) Γ large enough so that $\bar{s} = 0$ and (ii) $\bar{q}$ is such that $\bar{q} - \theta - e = \arg\max_x F'(x)$ (such as the intersection of dashed lines on Figure 1).*

Proof. For every choice of $\bar{q}$, increasing Γ lowers $\bar{s}$ which improves the effort, because marginal benefit of effort improves. Lemma 1 of Baghestanian and Popov (2018) delivers the second part: effort is maximized when $\bar{q} - \theta - e$ is at maximum of density, which is at zero by normalization of $f(\cdot)$ so that the mode of ε distribution is at zero. □

[4] The author's level of effort does not matter for ω, because the author gets a reputational payoff only if their paper is accepted, and the expectation is over the quantity of other accepted papers.

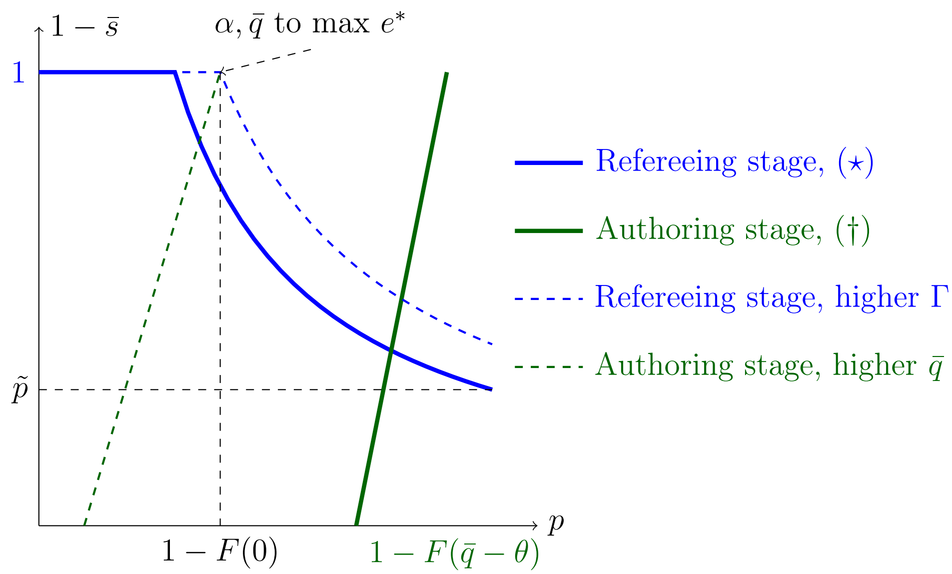

Figure 1: Endogenous p and $1-\bar{s}$ as a function of Γ and $\bar{q}$.

Discussion

So why then do editors choose lower capacity and lower quality thresholds than is optimal for effort stimulation? While higher thresholds decrease authors' utilities, the truthfulness of referees will make their utilities higher, even though the efforts will increase. Editors would likely appreciate more effort from authors, higher standards, and more good papers submitted for publication. What motivates the equilibrium where good papers go unpublished? Why not increase the capacity, especially taking into account that this can go hand-in-hand with increasing the standards?

Exclusivity Higher capacity might lead to lower payoff per publication ω, which may interest the editor in maximizing instead of effort. This goes against the refereeing subgame ideology: if more good papers lead to lower reputational payoff, referees are *never* interested in approving good papers, no matter the capacity. This penalty for publishing too many papers might work via other channels, such as citing; for example, if future generations choose one paper to cite, more published papers might lead to the lower payoff of published authors. However, a larger publishing capacity will lead to *more* papers published in the future, which means more opportunities to be cited now. Overall, it is not clear why a higher capacity, especially combined with higher standards, would lower the publication payoff; if anything, increasing the capacity increases the quantity of good papers associated with the field, so the contrary event likely applies. As long as capacity constraint matters, it can be relaxed, and, if anything, this should improve the effort choice of authors, and therefore the quality of the output.

Signaling In the field of economics, as well as other fields, publications in journals are used for hiring, tenure, and promotional decisions, among other things. This means that my basic model without heterogeneity in abilities can be extended to multiple ability levels. Nonetheless, it is arguable that these abilities are difficult to observe, which is possibly why publication statistics are used by decision-makers.

Asymptotic Asymmetric Model

In this section, I study the limit of the behavior of the agents as the number of authors increases, so that the refereeing subgame does not depend on the ability of the removed author.[5] Mathematically, assume $N \to +\infty$ and $\Gamma_N/N \to \gamma \in (0,1)$; this means that every author faces the same structure of referees, and therefore the same $\bar{s}$. This allows an easier characterization an asymptotic equilibrium with multiple types. This must also have implications with regard to the reputational payoff function.

Proposition 4. *Consider a symmetric subgame equilibrium problem where authors are risk-neutral: $W(G,B) = (aG - B)/N$. Let $\tilde{p}_N$ be the solution of that problem. As $N \to +\infty$, the solution $\tilde{p}_N = \gamma + o\left({}^1\!/\!\sqrt{N}\right)$.*

Proof. In a mixed strategy equilibria, $\tilde{p}_N = p(1 - \bar{s}_N)$ solves

$$\frac{a}{N} P(\xi(\tilde{p}_N, N-2) \leq \Gamma_N - 2) = P(\xi(\tilde{p}_N, N) \geq \Gamma + 1) \frac{\Gamma_N a \Gamma_N / N}{N(N-1)\tilde{p}^2}.$$

Multiply both sides by N, cancel out a, and rewrite probabilities in a way that could be used to apply the central limit theorem:

$$P\left(\overbrace{\sqrt{N}\frac{\frac{\xi(\tilde{p}_N, N-2)}{N} - \tilde{p}_N}{\sqrt{\tilde{p}_N(1-\tilde{p}_N)}}}^{\to N(0,1)} \leq \sqrt{N}\frac{\frac{\overbrace{\Gamma_N - 2}^{\to\gamma}}{N} - \tilde{p}_N}{\sqrt{\tilde{p}_(1-\tilde{p}_)}} \right) = \frac{\overbrace{\Gamma_N(\Gamma_N+1)/(N(N-1))}^{\to\gamma^2}}{\tilde{p}_N^2} \times$$

[5] See Olszewski and Siegel (2016) for more on this approach. The idea is to approximate the equilibrium of the finite N game, which—as in my case—might be hard to characterize in closed form with a solution of the asymptotic game.

$$\times P\left(\overbrace{\sqrt{N}\frac{\frac{\xi(\tilde{p}_N,N)}{N}-\tilde{p}_N}{\sqrt{\tilde{p}_N(1-\tilde{p}_N)}}}^{\to N(0,1)} \geq \sqrt{N}\frac{\overbrace{\frac{\Gamma_N+1}{N}}^{\to\gamma}-\tilde{p}_N}{\sqrt{\tilde{p}_N(1-\tilde{p}_N)}}\right).$$

p_N is a sequence bounded in $[0,p]$, so it must have a convergent subsequence. Take a limit point of that subsequence. If it is $\gamma+c$, and $c>0$, then for that subsequence, the limit of this equation is

$$P(\xi<-\infty)=\frac{\gamma^2}{(\gamma+c)^2}P(\xi>-\infty)\Rightarrow 0=\frac{\gamma^2}{(\gamma+c)^2},$$

where with a slight abuse of notation $\xi\sim N(0,1)$.

A similar argument can be made that c cannot be negative. Same would hold if $\tilde{p}_N=\gamma+O\left(N^{1/2+\epsilon}\right)$ for an arbitrarily small but positive ϵ. However, if $\tilde{p}_N=\gamma+\sqrt{\tilde{p}_N(1-\tilde{p}_N)}\frac{d}{\sqrt{N}}+o\left(\frac{1}{\sqrt{N}}\right)$, the limit of that sequence of equilibrium equations is

$$\Phi(d)=\frac{\gamma^2}{\gamma^2}(1-\Phi(d)),$$

which is clearly satisfied when $d=0$. Since all limit points of that bounded sequence are the same, $\tilde{p}_N\to\gamma$. □

Corollary 3. *If $W(G,B)=b+a(G-B)/N$, $\tilde{p}_N=\gamma+O\left(\frac{1}{\sqrt{N}}\right)$, so $\tilde{p}\to\gamma$, but in a way consistent with Corollary 2.*

Proof. As in Proposition 4, obtain a sequence of p_N, observe that it is bounded, and rule out $\gamma+c$ as limit points. Now let $\tilde{p}_N=\gamma-\sqrt{\tilde{p}_N(1-\tilde{p}_N)}\frac{d}{\sqrt{N}}+o\left(\frac{1}{\sqrt{N}}\right)$. Observe that the equilibrium condition converges to

$$\Phi(d)=\frac{\gamma(b/a+\gamma)}{\gamma^2}(1-\Phi(d)).$$

If b increases, the right-hand side increases, leading to an increase in d, as in Corollary 2. □

The asymptotic model is easier to analyze in an asymmetric case because author types do not influence either the refereeing pool or the paper that one expects to referee. If we had two ability levels, a high ability author would face relatively more referees whose own papers are less likely to be published, and vice versa; in an asymptotic model, we simply need to know the relative mass of authors of different abilities to decide on $\tilde{p}$ that would lead to the same $\bar{s}$ for different authors. We will maintain the assumption that referees of different authorship skills are equally capable, which seems to be consistent with the findings of Hadavand et al. (2020) and Card and DellaVigna (2020). Generically, one could assume a continuous distribution of θ, as in Baghestanian and Popov (2018). I will focus on two levels of ability, $\theta_H > \theta_L$, and pick $\lambda \in (0,1)$ to designate $P(\theta = \theta_H)$.

The asymmetric asymptotic model equilibrium is a collection of:

- The probability that each of the referees misrepresent their signal when $q + \varepsilon > \bar{q}$, denoted by $\bar{s}_G \in [0,1]$, consistent with referees' incentives;
- The probability that each of the referees misrepresent their signal when $q + \varepsilon < \bar{q}$, denoted by $\bar{s}_B \in [0,1]$, consistent with referees' incentives;
- e_H^* and e_L^*, the choice of effort by the authors, who maximize their utility:

$$\max_e E\left[W(G,B)|\bar{s}_G, \bar{s}_B\right] P[\text{pass referees}|\bar{s}_G, \bar{s}_B] P(\overbrace{\theta + e + \varepsilon > \bar{q}}^{\text{Paper is good}}) - c(e),$$

  where e_H^* solves the problem of an author with ability θ_H, and e_L^* solves the problem of an author with $\theta = \theta_L$;
- p_H^* is $P(\theta_H + e_H^* + \varepsilon > \bar{q})$ and p_L^* is $P(\theta_L + e_L^* + \varepsilon > \bar{q})$;
- Referees face $p = \lambda p_H^* + (1-\lambda) p_L^*$.

Corollary 4. *In asymmetric asymptotic model, $\bar{s}$ solves*

$$\gamma = \left(\lambda p_H^* + (1-\lambda) p_L^*\right)(1 - \bar{s}).$$

Proof. Identical to Proposition 4 approach. □

If the editor wants to maximize the signaling value of the publication, they want to consider

$$P(\theta = \theta_H|\text{published}) = \frac{\lambda p_H^*}{\lambda p_H^* + (1-\lambda) p_L^*}.$$

We know that a decrease in γ leads to an increase in $\bar{s}$. Take derivatives with respect to $\bar{s}$:

$$\frac{d}{d\bar{s}} P(\theta = \theta_H|\text{published}) = \frac{\lambda(1-\lambda)\left[\frac{dp_H^*}{d\bar{s}} p_L^* - \frac{dp_L^*}{d\bar{s}} p_H^*\right]}{\left(\lambda p_H^* + (1-\lambda) p_L^*\right)^2},$$

which is positive if $\frac{dp^*_H}{d\bar{s}}p^*_L > \frac{dp^*_L}{d\bar{s}}p^*_H$. $p^*_H > p^*_L$ by Lemma 1 in Baghestanian and Popov (2018). Both $\frac{dp^*_L}{d\bar{s}}$ and $\frac{dp^*_H}{d\bar{s}}$ are negative; by the optimality condition of authors of both types, they are equal to $c'(e^*_H)/(1-\bar{s})\frac{de^*_H}{d\bar{s}} - p^*_H$ and $c'(e^*_L)/(1-\bar{s})\frac{de^*_H L}{d\bar{s}} - p^*_L$. Altogether, the positivity of the original derivative is simplified to

$$c'(e^*_H)p_L\frac{de^*_H}{d\bar{s}} > c'(e^*_L)p_H\frac{de^*_L}{d\bar{s}}. \tag{$\diamond$}$$

Use the author's first order condition to obtain

$$(1-\bar{s})\omega f(\bar{q}-\theta-e) = c'(e) \Rightarrow \frac{de}{d\bar{s}} = \underbrace{\frac{-\overbrace{\omega f(\bar{q}-\theta-e)}^{c'(e)/(1-s)}}{c''(e)+\omega f'(\bar{q}-\theta-e)(1-\bar{s})}}_{>0 \text{ because it's author's SOC}}.$$

Substituting back to the ($\diamond$) and canceling $(1-\bar{s})$ from both sides obtains:

$$\frac{(c'(e^*_H))^2 p_L}{c''(e_H)+f'(\bar{q}-\theta_H-e^*_H)(1-\bar{s})} < \frac{(c'(e^*_L))^2 p_H}{c''(e_L)+f'(\bar{q}-\theta_L-e^*_L)(1-\bar{s})}.$$

To obtain the desired example where lowering of the capacity and therefore increasing of $\bar{s}$ would lead to a better signaling by the journal, take $c''(e) = const$, a single-peaked distribution of ε with a peak at 0, and pick $\bar{q}$, θ_H and θ_L so that $\bar{q}-\theta_H-e^*_H > 0 > \bar{q}-\theta_L-e^*_L$ while $e^*_L > e^*_H$. Clearly, this example is quite generic, as the same result obtains if $c'''(e) < 0$.

The editor might pursue multiple potentially contradicting aims, balancing as and when necessary. The key point here is that it is possible to rationalize the stifling of capacity by allowing one of these aims to be improving the value of their journal as a signaling device.

Conclusion

I provide a model of authors working as referees to show that, under conditions of limited capacity, even the most motivated referees might mislead an editor, exactly because the editor might run out of publishing capacity. I argue that limiting the capacity cannot be consistent with pure effort stimulation, precisely because efforts are discouraged when referees tend to misrepresent your paper, and I rationalize this behavior in line with the desire to make the journal into a better signaling device. Other explanations are possible, and should be analyzed and tested, and further data should become available to do so in the future. A future study could test the mechanism I describe here on data used in the paper of

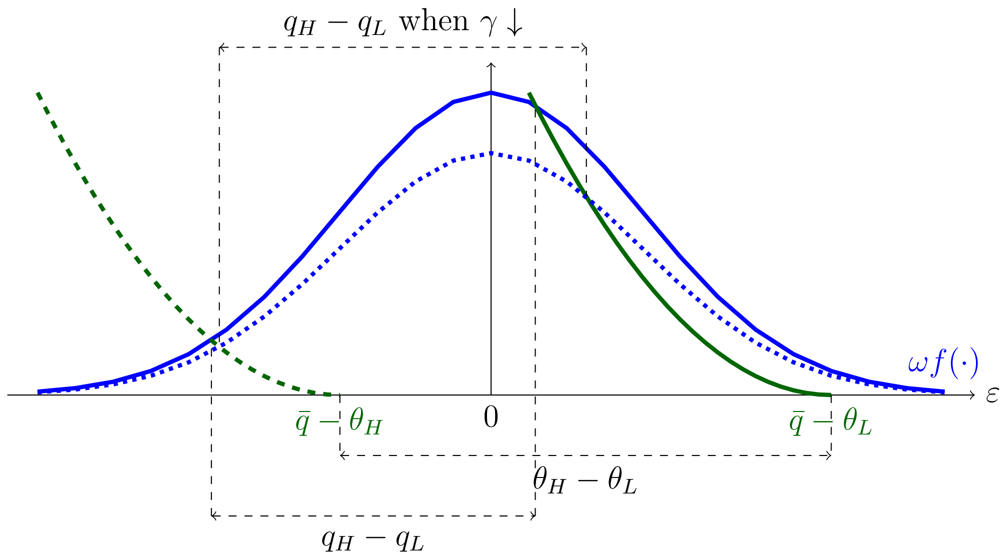

Figure 2: Example: $\varepsilon \sim N(0,1)$, $c'(e) = {}^{e^2}\!/_4$. Incentives for authors before (solid blue) and after (dotted blue) a decrease in capacity; green curves represent marginal cost graphs for efforts of high (dashed green) and low (solid green) ability authors, they start at the $\bar{q} - \theta_i$ for the respective types $i \in \{H, L\}$ where effort is equal to zero, and determine the effort choice e_i of each type at the intersection with the marginal benefit of the effort curve $\omega f(q - \theta_i - e_i)$. Note that the difference in the quality of papers, $q_H - q_L$, increased as a response to a decrease in capacity; this would lead to an increase in the quality of signaling by publishing in the journal, in the sense that now the chance of author to be of higher ability increases.

Card and DellaVigna (2020), but with finer data. Knowing the distance between the referee and the author, for instance using textual analysis of their papers a lá Onder et al. (2021), my mechanism would create an incentive for authors to become more stringent when the refereeing papers of their direct competitors. However, due to current privacy surrounding such data, the only way this analysis could be executed is by means of collaboration across multiple journals.

Strategic refereeing, such as *tit-for-tat* strategies, hurting own competitors, or helping your own narrower field to prosper in wider literature deserve their own studies. I intentionally omit these considerations from this study because I want to underline that competition for limited capacity undermines the efficiency of refereeing process, even if the refereeing process is perfect and costless. I leave career considerations, reputation building, network building and other strategic considerations for future studies.

References

Alexander, D., Gorelkina, O., and Hengel, E. (2021). Gender and the time cost of peer review.

Baghestanian, S. and Popov, S. V. (2018). On publication, refereeing, and working hard. *Canadian Journal of Economics*, 51(4):1419–1459.

Bayar, O. and Chemmanur, T. J. (2021). A model of the editorial process in academic journals. *Research Policy*, 50(9):104339.

Brogaard, J., Engelberg, J., and Parsons, C. A. (2014). Networks and productivity: causal evidence from editor rotations. *Journal of Financial Economics*, 111(1):251–270.

Card, D. and DellaVigna, S. (2020). What do editors maximize? Evidence from four economics journals. *Review of Economics and Statistics*, 102(1):195–217.

Colussi, T. (2018). Social ties in academia: a friend is a treasure. *Review of Economics and Statistics*, 100(1):45–50.

Ellison, G. (2002). Evolving standards for academic publishing: A q-r theory. *Journal of Political Economy*, 110(5):994–1034.

Engers, M. and Gans, J. S. (1998). Why referees are not paid (enough). *American Economic Review*, 88(5):1341–1349.

Gans, J. S. and Shepherd, G. B. (1994). How are the mighty fallen: Rejected classic articles by leading economists. *Journal of Economic Perspectives*, 8(1):165–179.

Gerardi, D., Grillo, E., and Monzón, I. (2021). The perils of friendly oversight. Technical report.

Hadavand, A., Hamermesh, D. S., and Wilson, W. W. (2020). Is scholarly refereeing productive (at the margin)? Technical report, National Bureau of Economic Research.

Hale, G., Regev, T., and Rubinstein, Y. (2021). Do looks matter for an academic career in economics?

Heckman, J. J. and Moktan, S. (2018). Publishing and promotion in economics: the tyranny of the top five. Technical report, National Bureau of Economic Research.

Hengel, E. (2020). Publishing while female.

Olszewski, W. and Siegel, R. (2016). Large contests. *Econometrica*, 84(2):835–854.

Onder, A. S., Popov, S. V., and Schweitzer, S. (2021). Leadership in scholarship: Editors' influence on the profession's narrative. Technical report, Cardiff University, Cardiff Business School, Economics Section.