Supplementary material for Zollman, García, and Handfield, "Academic journals, incentives, and the quality of peer review: a model"

Accepted 25 March 2023

1 Mathematical results

There is a continuum of papers, indexed by a quality $q \in [0, 1]$. The journal chooses a publication strategy and this strategy is made public. Authors then observe the quality of their papers (without error) and the publication strategy of the journal. They then choose whether or not to submit their paper. The journal observes a noisy signal of quality of the paper and chooses to publish those papers according to its publication strategy. Then everyone receives their payoffs.

The journal makes two choices which constitute its publication strategy. They choose a minimum acceptable paper, Q_T, and an error rate ϵ. As a matter of notation, let $J(q, Q_T, \epsilon)$ represent the probability that a paper of quality q is accepted if the journals strategy is given by Q_T and ϵ. We treat this probability as fixed given the arguments, and it does not depend on the papers which are submitted. (This means we are not treating the peer reviewers as Bayesian agents.)

While we make particular functional assumptions for J later, for the purposes of our proofs we need only the following assumptions:

1. For all Q_T, q, ϵ, $0 < J(q, Q_T, \epsilon) < 1$ (No outcome is guaranteed.)

2. $J(q, Q_T, \epsilon)$ is increasing in q and decreasing in Q_T. (While noisy, peer review is detecting quality.)

3. $J(q, Q_T, \epsilon)$ is decreasing in ϵ when $q > Q_T$ and increasing in ϵ when $q < Q_T$. (This correspond to our assumption that the journal is paying for better, i.e. less noisy, peer review.

4. $J(q, Q_T, \epsilon)$ is continuous.

Suppose that the set of papers S submits. Two numbers will be important for the journal going forward. One is the average quality of published papers which is given by:

$$\bar{q} = \frac{\int_S J(q, Q_T, \epsilon) q dq}{\int_S J(q, Q_T, \epsilon) dq}$$

We will adopt the convention that if only one paper, q, is submitted then $\bar{q} = J(q, Q_T, \epsilon)q$ and if no papers are submitted $\bar{q} = 0$.

Second is a journal's rejection rate, this is the proportion of papers that are rejected by the journal.

$$r = \frac{\mu(R)}{\mu(S)} = \frac{1 - \int_S J(q, Q_T, \epsilon) dq}{\mu(S)}$$

where $\mu(S)$ is the measure of set S. In the case where only a single paper, q, is submitted we define r as $(1 - J(q, Q_T, \epsilon)$ We will assume that if no papers are submitted $r = 0$.

We assume that the journal pays a cost for ϵ, which we represent by $c_J(\epsilon)$ which is decreasing in ϵ. Our intended interpretation is that ϵ represents the error rate and lower error rates entail higher costs.

With this in hand, we can specify the two version of the journal. Those incentivized by quality:

$$u_R(Q_T, \epsilon) = r - c_J(\epsilon).$$

And those incentivized by selectivity:

$$u_Q(Q_T, \epsilon) = \bar{q} - c_J(\epsilon).$$

In the non-strategic version of the model, we assume that all authors submit, and so their payoff is irrelevant. In this setting we have the following theorem:

Theorem 1. *In the non-strategic model, when $c(\epsilon) > 0$ for all ϵ, the optimal strategy for the journal incentivized by rejection is to set $Q_T = 1$.*

Proof. By assumption 2, $J(q, Q_T, \epsilon)$ is decreasing in Q_T, so the rejection rate will increase for higher Q_T raising the payoff for the journal without any attendant costs. □

Now let us consider strategic authors who choose whether or not to submit. If the paper is accepted the authors are paid $p(q, \bar{q})$. If the paper is rejected the authors lose opportunity cost $-c(q)$ To make life easy, assume that authors submit when indifferent. We will assume both are continuous.

Let $s = 1$ if the author submits and $s = 0$ if they don't. the payoff for the author is

$$u(q, s) = s\big(J(q, Q_T, \epsilon)p(q, \bar{q}) - c(q)\big).$$

The next theorem establishes that when the cost to authors is constant, self-selection proceeds from the "bottom up" That is, the worst papers are the ones who choose not to submit.

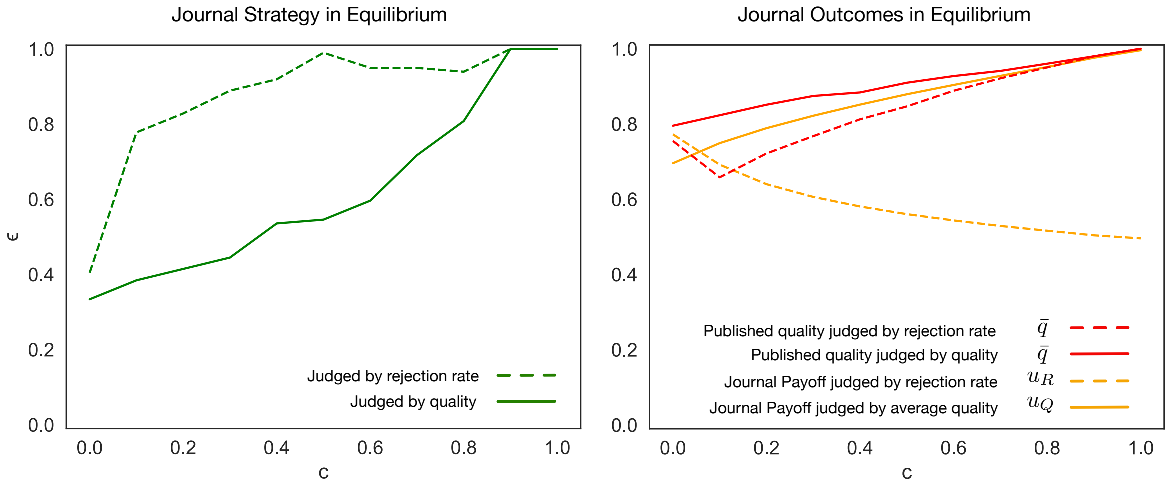

Figure 1: Journal strategy and outcomes in equilibrium for a version of the constant cost, variable benefit model.

Theorem 2. *Suppose there is a constant cost to rejection, $c(q) = c > 0$ and that Q_T, ϵ and c are such that a paper $q < 1$ submits. Then there exists a q^- such that $S = [q^-, 1]$.*

Proof. If all papers submit, $S = [0, 1]$ then the theorem is trivially true. So, consider the case where at least one paper q_S submits and one paper q_D does not. Since q_S submits, $u(q_S, 1) \geq 0$ and since q_D does not $u(q_D, 1) < 0$. By the intermediate value theorem, there must be a q_0 where $u(q_0, 1) = 0$. By assumption q_0 submits. Because of the assumption 2 for J, this means that for any $q < q_0$, $u(q, 1) < 0$ and for any $q > q_0$, $u(q, 1) > 0$. Therefore q_0 is q^-. □

1.1 Computational results for variable benefit

The central conclusions remain the same when considering a version of the model where the benefit is given by $\bar{q}$ (i.e. variable cost) instead of 1 (i.e. constant cost).

First, we see the same pattern in how submissions change as the costs increase. Compare figure **??** (constant cost, constant benefit) and figure 1 (constant cost, variable benefit). Both figures exhibit the “bottom up” pattern of self-selection, where the worst papers choose to not submit. Similarly, a comparison of figure **??** (variable cost, constant benefit) and figure 4 (variable cost, variable benefit), shows that in both of these models self-selection proceed from the “top down” or “middle out.”

Second the patterns of quality of reviewing, journal payoffs, and quality of published papers remain largely similar. Figure **??** (constant cost, constant benefit) and figure 1 (constant cost, variable benefit) look largely the same.

While in a large part the same is true for figure **??** (variable cost, constant benefit) and figure 3 (variable cost, variable benefit), there is one striking dissimilarity. In figure 3 the average quality of published papers, $\bar{q}$ drops to zero. This occurs

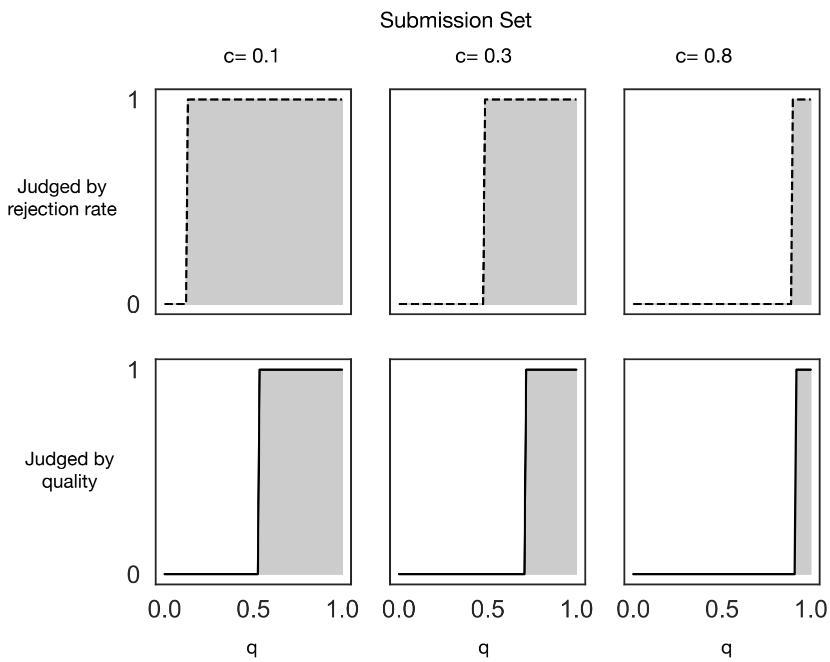

Figure 2: Submission set for a version of the constant cost, variable benefit model.

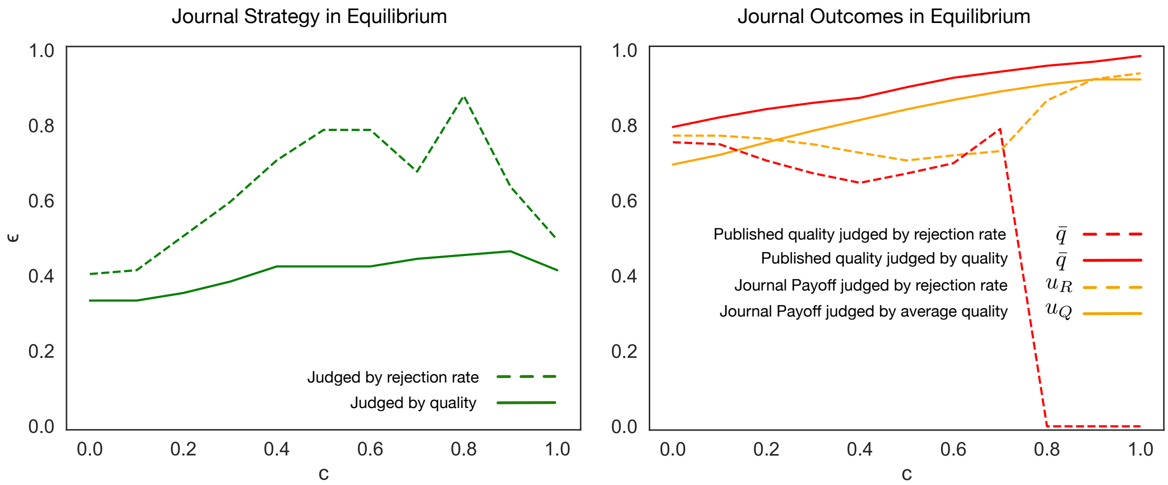

Figure 3: Journal strategy and outcomes in equilibrium for a version of the variable cost, variable benefit model.

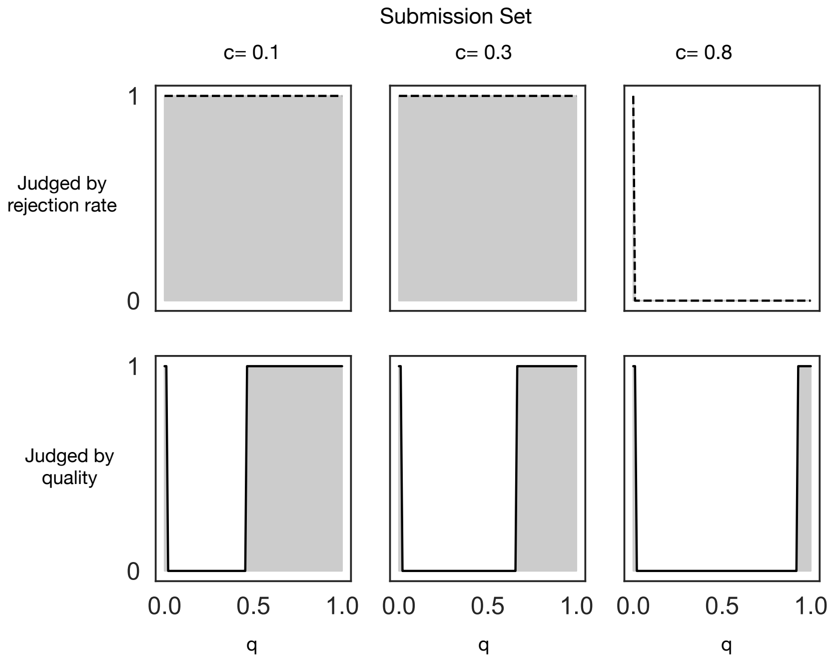

Figure 4: Submission set for a version of the variable cost, variable benefit model.

because only one paper (the worse paper, $q = 0$) submits and is always rejected. By convention, we assign this average quality, $\bar{q} = 0$. From the journal's perspective this is quite good, they maintain a rejection rate of 100%. Of course, from an epistemic standpoint this is worthless, because the journal is serving no purpose whatsoever.

2 Description of computational analysis

We pose a search problem over discretised bounded intervals. In particular, we discretise all possible values of $q \in [0, 1]$, all possible values of $\epsilon \in [0, 1]$ and all values of $Q_t \in [0, 1]$. The number of bins is parametrisable – with a default value of 100. We denote these discretise sets as q_d, ϵ_d and Q_{td}; for q, ϵ and Q_t respectively.

2.1 Non-strategic authors

For the non-strategic versions, we simply maximise the journal payoff over q_{td} and ϵ_d. That is, the algorithm finds a pair Q_T, ϵ that solves

$$\arg\max \Big(u(Q_T, \epsilon) = r - c_J(\epsilon) \Big)$$

or

$$\arg\max \Big(u(Q_T, \epsilon) = \bar{q} - c_J(\epsilon) \Big)$$

by searching exhaustively over discretised intervals.

2.2 Strategic authors

We now need to exhaustibly search submission sets, in addition to the journal's parameters. We do this by considering subsets Q_{td}, which are easy to enumerate.

Consider the following submission sets, $T_\lambda = \{q \in Q_{td} | q > \lambda\}$, $C_{\lambda_1,\lambda_2} = \{q \in Q_{td} | \lambda_1 < q < \lambda_2\}$; and their complements, $\bar{T_\lambda} = \{q \in Q_{td} | q <= \lambda\}$ and $C_{\bar{\lambda_1,\lambda_2}} = \{q \in Q_{td} | \lambda_1 >= q >= \lambda_2\}$. These define discrete sets where agents are submitting above a given threshold q, or only between certain quality values, as well as their complements.

We define our search space:

$$U = \{C \cup \bar{C} \cup T \cup \bar{T}\}$$

.

With $T = \{T_\lambda \forall \lambda \in Q_{td}\}$, $\bar{T} = \{\bar{T_\lambda} \forall \lambda \in Q_{td}\}$, $C = \{c_{\lambda_1,lambda_2} \forall \lambda_1 < \lambda_2 \in Q_{td}\}$, and $\bar{C} = \{c_{\lambda_1,lambda_2} \forall \lambda_1 > \lambda_2 \in Q_{td}\}$

The procedure that searches for equilibria is presented in Algorithm 1. It searches over the space of submission sets, quality thresholds and ϵ values, and exhaustively checks if every triplet is consistent with equilibrium behaviour. When many equilibria are produced, we picked the one that maximises the payoff of the journal.

Algorithm 1 Searching for equilibria with strategic authors. Will return a list of equilibria, with each element including a submission set, and a noise and threshold set by the journal.

```
function EnumerateEquilibria(U)
    E ← []
    for u ∈ U do                                ▷ Iterate over all possible submission sets
        for ε ∈ [0,1] do                        ▷ Iterate over all possible ε values
            for Q_t ∈ Q_td do                   ▷ Iterate over all possible Q_t thresholds
                if EquilibriumConsistent(u, ε, q_t) then
                    E ← (u, ε, Q_t)             ▷ add to list of equilibria
                end if
            end for
        end for
    end for
    return E                                    ▷ Return list of equilibria
end function
```

Consistency is checked by computing the counterfactual payoff for each scientist, and checking that the submission decision given by u cannot be unilaterally improved. This process is described in Algorithm 2.

Algorithm 2 Checks if the behaviour implied by u, ϵ, Q_t is consistent with equilibrium play

```
function EquilibriumConsistent(u, ϵ, Q_t)
    for q ∈ q_td do
        if q ∈ u then                    ▷ Scientist at q is submitting according to u
            q̄ = AverageQualityPublished(u, ϵ, Q_t)
            Π ← q̄ − c
            Π_c ← 0                      ▷ Counterfactual payoff is set to 0
        else                             ▷ Scientist at q is not submitting according to u
            q̄ = AverageQualityPublished(u_c, ϵ, Q_t)
                ▷ u_c is a counterfactual submission set assuming q does not submit
            Π ← 0
            Π_c ← q̄ − c                  ▷ Counterfactual payoff is set to submission payoff
        end if
        if Π ≤ Π_c then
            return FALSE
        end if
    end for
    return TRUE
end function
```