THE

Teoría e Historia Económica
Working Paper Series

Rejected: Career concerns in the refereeing process

Ascensión Andina-Díaz, José A. García-Martínez, and Nektaria Glynia

WP 2024-03
June 2024

Departamento de Teoría e Historia Económica
Facultad de Ciencias Económicas y Empresariales
Universidad de Málaga
ISSN 1989-6908

Rejected: Career concerns in the refereeing process[*]

Ascensión Andina-Díaz,[†] José A. García-Martínez,[‡]
and Nektaria Glynia [§]

June 20, 2024

Abstract

We analyze the effect of career concerns on the refereeing process. We consider a journal editor and two referees who may differ in reputation and ability. A referee's reputation is public information, while a referee's ability is private information. We identify an incentive for low-ability referees to reject good papers —a phenomenon we call over-rejection— and find that this incentive increases with the referee's reputation. We show that over-rejection decreases with competition, referee homogeneity, and the anonymity of the refereeing process. In contrast to low-ability experts, high-ability referees are truthful in equilibrium. Since a referee with a higher reputation is ex-ante more likely to be high-ability, our results suggest that the probability of rejection is inverted U-shaped in the referee's reputation. We empirically test this result. We use data from Card and DellaVigna (2020) for submissions to four top economic journals in the period 2003-2013 and use the referee's publication record as a proxy for the referee's reputation. We find that the probability of sending a negative recommendation increases with the referee's reputation in the early stages of the career and decreases thereafter, suggesting an inverted U-shape form in line with our theoretical results.

Keywords: Career concerns: refereeing process; reputation; ability; endogenous transparency; competition; information transmission
JEL: C72; D83

[*]We are very grateful to Miguel A. Meléndez-Jiménez for helpful conversations and comments. We also thank Stefano DellaVigna, Gilat Levy, Nagore Iriberri, Aniol Llorente-Saguer, Maria Marino, Pilar Sorribas, Nikolas Tsakas, and Dimitrios Xefteris, as well as seminar participants at the University of Málaga, University of Cyprus, Oligo 2023 Padova, and the EPC Conference 2024 Vienna. We gratefully acknowledge the financial support from the Ministerio de Ciencia e Innovación through projects PID2021-127736NB-I00 and PID2022-137211NB-I00, and the Junta de Andalucía through project P18-FR-3840. The usual disclaimer applies. Declaration of interests: none.

[†]Dpto. Teoría e Historia Económica, Universidad de Málaga, Málaga, Spain. Email: aandina@uma.es
[‡]Dpto. Estudios Económicos y Financieros, Universidad Miguel Hernández, Elche, Spain. Email:jose.garciam@umh.es
[§]Dep. Economics, University of Cyprus, Nicosia, Cyprus. Email: glynia.nektaria-spyridoula@ucy.ac.cy

1 Introduction

The frequency of rejection decisions in the editorial process is high in many disciplines and particularly pronounced in social sciences (Zuckerman and Merton 1971, Hargens 1988).[1] In economics, Card and DellaVigna (2013) found that between 1970 and 2012, annual submissions to the top-five journals nearly doubled and the total number of articles published declined, resulting in acceptance rates falling from 15 percent to 6 percent. Beyond the top-five journals, the situation is much the same, with acceptance rates moving from 6.25 percentage points in the *Journal of Development Economics* in 2021 to 10 and 14 percentage points in the *European Economic Review* and the *Journal of Economic Theory*, respectively. While there are clear arguments in favour of a rejection decision —the advance of science should be based on sound grounds and not on potentially erroneous ideas— there are concerns that high rejection rates may discourage scientists, slow down the progress of science and open avenues for predatory journals.

There are a number of possible explanations for the high rejection rates, including the aforementioned increased competition for space in journals, the increased complexity of economics papers, or the democratization of the publishing process. Some of these arguments have also been proposed as explanations for the slowdown in the publication process (Ellisons 2002). Within this debate, Berk et al. (2017) maintain: *"We believe that part of the explanation [...] is that referees feel the need to demonstrate their intelligence or industriousness to editors."*

We share with Berk et al. (2017) the view that the publication process is a signalling game in which referees send recommendations, presumably under the desire to be perceived by the editor as highly competent professionals. Although this impression may be shared by many academics, we are not aware of many papers that approach the refereeing process from this perspective. The exception is Hirshleifer (2015), who proposes a model in which a referee can signal his ability by detecting flaws and blemishes in a paper. He finds that, in equilibrium, the referee will recommend fixing both mere blemishes and significant flaws, since signal-jamming gives him credit with the editor. Since some blemishes might be difficult to remove, those born under a lucky star will have their publication process delayed. Born under an unlucky one, the paper remains unpublished.

We build on the view of the refereeing process as a signaling game to propose a model of career concerns that, however, puts forward a different mechanism. We argue that a referee (he), by means of his recommendation, can affect how much the editor (she) learns about his ability. The idea is that a rejection recommendation might probably yield a rejection decision, which stops the learning process of the editor about the quality of the paper; hence, about the referee's ability. In contrast to this, a recommendation to revise and resubmit or to accept a paper might probably end up in further revisions and so further knowledge about both the paper's quality and the referee's ability. If the referee has career concerns, i.e., he cares about his reputation, then recommending rejection can be the mechanism to preclude further learning. We contribute to the literature on the publication process in two strands. On the one hand, we propose a new mechanism that explains high rejection rates and study how the mechanism works in the

[1] Zuckerman and Merton (1971) document rejection rates of 20 to 40 percent in the physical sciences, and of 70 to 90 percent in the social sciences. Hargens (1988) found similar results for the period 1960-1980, documenting acceptance rates of 16-17 percent in the *American Economic Review*, and of 79-83 percent in the *Physical Review*.

presence of competition and heterogenous referees. On the other hand, we provide empirical support for the strategic effect of reputation on referee recommendations.

The model has the following structure. There are two referees, each of whom receives a signal about the quality of a paper. The paper can be either good or bad. Each referee can be either high-ability or low-ability. The difference is that high-ability referees can perfectly judge the quality of the paper, whereas low-ability referees receive an imperfect —though informative— signal about the quality of the paper. The type or ability of a referee is the referee's private information. The two referees may also differ in their reputation, which is public information. A referee's reputation describes the perception that others have about the probability the referee is high-ability. In the real world, this is given by the referee's publication record, cv, and so on. The ability and reputation of a referee are positively correlated, so a referee with a higher reputation is ex-ante more likely to be of high ability; however, referees of high reputation can well be low-ability. Referees observe their reputation —which is public information— and their type and signal —which are private information— and write a recommendation to the editor seeking to maximize their reputation as high-ability referees. The editor updates her beliefs about the quality of the paper and chooses whether to reject the paper or to keep it in the refereeing process. A key ingredient of the model is that if a paper is kept in the refereeing process, the editor will gather more information about the quality of the paper, which helps the editor better assess the quality of the referee's recommendations. However, if a paper is rejected, further learning is precluded.

We show that, in equilibrium, high-ability referees send truthful recommendations. In contrast, we identify an incentive for low-ability referees to reject good papers. We call this phenomenon over-rejection. The phenomenon of over-rejection is detrimental to scientific welfare and is entirely due to strategic motives. We show that the incentive to over-reject increases with referee reputation and decreases with referee ability, suggesting that referee expertise in the field of the paper is more important for guaranteeing accurate and non-strategic recommendations than referee reputation. We introduce competition between referees and find that competition reduces over-rejection by low-ability referees, though it does not eliminate it. We also study the effect of referees' heterogeneity (in reputation) and obtain that the more homogeneous referees are, the less they over-reject. This result suggests that an anonymous refereeing process, where referees assign positive probability to the opponent being like them, may reduce over-rejection. We show that the insight is correct. Finally, we discuss variations of the model that allow us to explain, among others, why more unorthodox papers and top journal have higher rejection rates.

We complement our theoretical predictions with an empirical analysis that uses data from Card and DellaVigna (2020) on the evaluation process of nearly 30,000 submissions to four leading economic journals —the *Journal of the European Economics Association*, the *Quarterly Journal of Economics*, the *Review of Economics and Statistics*, and the *Review of Economic Studies*— for the period 2003-2013. For each submission, we have information on the year of the submission, the number of coauthors and their recent prominent-publication records, the editorial decision, the recommendation of each referee, the number of recommendations previously sent by each referee, and the recent prominent-publication record of each referee. We use the data to test our predictions about the incentive of a referee to recommend rejection.[2]

[2]Unfortunately, given the nature of the refereeing process —referees do not know the identity or number of referees in the

Following Card et al. (2020), we construct an index of referee recommendation and also estimate the probability that a referee rejects a paper. As a proxy for the referee reputation, we use the referee's record of publications. Unlike reputation, the ability of a referee is the referee's private information. Hence, taking our results to the data requires rewriting them in terms of reputation, rather than ability. Since referees of higher reputation are more likely to be high-ability —who are truthful in equilibrium— and less likely to be low-ability —who over-reject more often the higher their reputation— our results suggest that the probability that a referee sends a rejection recommendation is inverted U-shaped in the referee's reputation. In line with this result, in the data we find a non-monotonic effect of reputation on rejection recommendations, with rejection recommendations increasing in the referee's reputation in early stages of the referee's career (up to 3-4 prominent-publications in the past 5 years) and decreasing afterwards. This result is robust to different model specifications, including OLS models with quadratic terms and Logit models, and to the inclusion of controls and journal fixed effects.

The paper has the following structure. In Section 2 we describe the related literature. In Section 3 we describe the theoretical model and in Section 4 we analyze the model and present the results. Section 5 contains the empirical analysis, where we describe the data and present the empirical findings. We conclude in Section 6. The paper contains three appendices. Appendix A contains the proofs of all the results in the text. Appendix B supports the empirical section, presenting additional description of the data and empirical results. Finally, Appendix C complements the theoretical part of the paper, presenting extensions of the model and robustness checks, including model variations in the opportunity costs of publishing a paper, the editor's learning process and decision rule, and asymmetries in the referee's costs of actions.

2 Related literature

Our paper is part of the literature on the scientific publication process that analyzes its strengths and weaknesses. The theoretical research has focused attention on the incentives underlying the publication process, either from the journal's or editor's side, the referee's side, or the author's side. Atal (2010) focuses attention on the journal and, in particular, on the decision about the quality standard to establish in the presence of competing journals. Baghestanian and Popov (2018) focus on the decisions of authors about their level of effort. The analysis of referees' behavior is the focus in Bayar and Chemmanur (2021), Popov (2022), and Hirshleifer (2015). Bayar and Chemmanur (2021) consider referees with biases for or against a submission, and an editor who trades-off ability and bias in choosing referees. Popov (2022) analyzes the strategic incentives of referees who are also authors of papers when journals have capacity constraints. Closer to our work, Hirshleifer (2015) considers a referee who wants to signal his ability to detect blemishes and flaws in a paper. He finds that, in equilibrium, the referee spots relevant flaws but also insignificant blemishes, as it gives him credit with the editor. Unlike us, this paper does not consider competition (hence referee heterogeneity) or endogenous feedback.

The empirical literature on the topic is probably more abundant, including some well-known papers.

process— we cannot test our predictions about the effect of competition and anonymity. See footnote 12 for further discussion.

Card and DellaVigna (2020) collect data from paper submissions to four top journals during the period 2003-2013 and study referee recommendations and editorial decisions. They find that referee recommendations are good predictors of the paper's future citations and editors follow recommendations closely. They also find that editors give more weight to referees with a strong record of publications, though their recommendations are equally predictive of future citations.[3] Card et al. (2020) build on the journal sample collected by Card and DellaVigna (2020) adding information on the gender of authors and referees. They find that referee gender has no effect on the relative assessment of female versus male authored papers. They also find no differences in the informativeness of female versus male referees, or in the weight that editors place on female versus male referee recommendations. However, they find that holding constant referee evaluations, female-authored papers receive more citations which suggests a stricter standard for female-authored papers relative to a citation-maximizing benchmark. Cherkashin et al. (2009) use data from the *Journal of International Economics* to analyze factors that affect the probability of publishing in the journal and the cost of desk rejections, among others. A few other papers include Welch (2014), who focuses on the strength of agreement between referees, and Chetty et al. (2014), who investigate different measures to increase pro-social behavior in the refereeing process.

Methodologically, our paper belongs to the literature on information transmission in the presence of career concerns. We contribute to this literature by combining in one model two ingredients that have not yet been considered together: competition among heterogeneous agents and endogenous feedback. By endogenous feedback, we refer to the ability of an expert to influence the probability that the principal will verify an underlying state of the world. Existing work in the literature has considered different types of endogenous feedback. Canes-Wrone et al. (2001) study the incentive of an executive to act in the public interest when different policies have different uncertainty resolutions. Levy (2005) identifies an incentive for judges to be "creative" and contradict previous decisions when they can affect the probability that the decision will be appealed. Leaver (2009) shows how the desire to avoid criticism leads bureaucrats to be "generous" to third parties. Camara and Dupuis (2019) and Andina-Díaz and García-Martínez (2020) enrich the analysis by adding competition among experts and show that competition moderates "conservative" behavior. Mariano (2012) obtains similar effects of competition in the presence of a popular belief. Our paper goes a step further by considering that experts may differ in ability and reputation, and they are uncertain about who the opponent is.

3 The model

An editor (she) of a journal receives a paper for evaluation. The editor has to decide on whether to keep the paper in the journal —revise and resubmit (R&R) verdict hereafter— or to reject the paper. To make an informed decision, the editor consults two referees (he).

The paper: The paper has quality $\theta \in \{G, B\}$, where G stands for a *good paper* and B for a *bad paper*. We assume that the editor and the referees share a common balanced prior about the quality of the paper,

[3] The result that editors give more weight to referees with a strong record of publications is supported by our theoretical model. See the discussion after Proposition 1 and footnote 21. Our model also provides support for their finding that referees' recommendations are equally good, provided that referees in our model have sufficient ability. See Propositions 3 and 4.

$P(\theta = G) = 1/2$. This describes the most neutral scenario, where no referee or editor has a bias for or against the submission.[4]

A good paper has value K to the journal and we normalize the value of a bad paper to 0. Additionally, asking for R&R implies a cost c to the journal that illustrates the time cost and opportunity cost of further processing the paper and eventually publishing it. This cost may vary from one paper submission to another —the length of the review process varies between submissions— and so we consider it is a random variable with uniform distribution in the interval $[0, K]$. Since $0 \leq c \leq K$, asking for R&R of a good paper yields payoff $K - c > 0$ to the journal, whereas asking for R&R of a bad paper yields payoff $-c < 0$.[5] Rejecting the paper yields payoff 0.

The referees: There is a pool of referees. Each referee is characterized by two variables: reputation and ability. The reputation of a referee describes his academic records, e.g., record of publications and cv, and it is public information. The ability of a referee describes his expertise to evaluate and judge the quality of the paper and it is private information of the referee. We think that in the real world both dimensions are positively related and so we assume that the higher the reputation of a referee, the higher the *expected* ability of the referee is.

To formalize these ideas, we consider a simple structure where each of two referees, 1 and 2, can be either of two types, high-ability H and low-ability L. Types differ in the ability or expertise to judge the quality of the paper. Let $\gamma_t = P(s = \theta \mid \theta)$ be the *ability* of a referee of type $t \in \{H, L\}$, with $s \in \{G, B\}$ being the signal the referee receives after reading the paper. We consider that high-ability referees receive a perfect signal about the quality of the paper, i.e., $\gamma_H = 1$, whereas low-ability referees observe a signal that is correct with probability $\gamma_L \in (\frac{1}{2}, 1)$. Since $\gamma_H = 1$, hereafter we simplify notation and denote γ_L by γ. Each referee knows his ability, which is the referee's private information. The only information that the other players have about the ability of referee $i \in \{1, 2\}$ is his *reputation* α_i, which denotes the prior probability that the referee is high-ability. The abilities of the two referees are i.i.d. and their signals are i.i.d. conditional on the state.

Upon reading the paper and observing a signal, each referee $i \in \{1, 2\}$ sends a recommendation $m_i \in \{G, B\}$ to the editor, where G stands for *"this is a good paper"* and B stands for *"this is a bad paper"*. We interpret message B as a recommendation to reject the paper and message G as a recommendation to R&R the paper.

The editor: The editor observes the pair of referee recommendations $\mathbf{m} = (m_1, m_2)$, updates her beliefs about the quality of the paper, and decides whether to reject the submission or give it a R&R verdict.

A crucial assumption we make is that the probability that the editor learns the true quality of the paper (i.e., the state of the world) is not the same for papers with a R&R verdict or a rejection verdict. The idea is that a paper that receives a R&R verdict will receive more scrutiny within the journal —e.g., from reviewers in subsequent rounds, the editor, or even future readers— than a rejected paper —for which it may be difficult to assess the quality of the original submission, since even if the paper is later published in another journal, the final version may be an improved version of the original submission.

[4]Note that this is not a limitation but a way to control for other lying incentives, i.e., herding and anti-herding incentives.

[5]See Appendix C for an extension where the cost c may vary between journals and/or disciplines.

This asymmetry in the feedback of the two verdicts implies that the amount of information the editor has to judge the quality of the referee's work depends indirectly on the referee's recommendations (as they influence the editor's verdict). In other words, referees can affect the quality of the monitoring process. For simplicity, in the main body of the paper we assume that if the editor gives a rejection verdict, the quality of the paper remains unknown to the editor. In contrast, if the paper receives a R&R verdict, the editor learns the quality of the paper. We relax this assumption in Appendix C.

Let $X \in \{G, B, \varnothing\}$ denote this additional information about the quality of the paper that the editor can observe. We use $X = G$ (alternatively B) to denote a situation in which the editor learns that the paper is good (alternatively bad); and $X = \varnothing$ to denote a situation in which the paper is rejected and the editor does not observe its quality.

Payoffs, strategies, and equilibrium concept: The editor chooses her verdict seeking to maximize the expected payoff of the journal.[6] The journal, hence the editor, receives payoff $K - c$ when a good paper is kept the refereeing process, $-c$ when a bad paper is kept, and 0 when the paper is rejected, with $K - c \geq 0 \geq -c$.

We assume that referees have career concerns and maximize the editor's posterior that they are high-ability. To simplify the analysis, we next assume that high-ability referees are honest and make truthful recommendations. This is a reasonable assumption that allows us to focus the following analysis on the low-ability referees' attempt to manipulate the editor to appear high-ability. However, this assumption is later relaxed and shown not to affect the results. See Lemma 1 in Appendix A, where we prove that our results are robust to the consideration of strategic high-ability referees.

For a low-ability referee i, let $\sigma_s^i \in [0, 1]$ be the probability that he recommends rejection after observing the signal s_i, i.e., $\sigma_s^i = P(B \mid s_i)$. If $\sigma_G^i > 0$, we say that the low-ability referee i *over-rejects*. Note that this type of rejection occurs for strategic reasons, not for quality reasons, and is therefore welfare-detrimental. Finally, given the profile of referee recommendations $\mathbf{m}$ and feedback X, let $\hat{\alpha}_i(\mathbf{m}, X)$ be the editor's posterior belief that referee $i \in \{1, 2\}$ is high-ability (type H). This posterior describes the payoff function of referee i.

We analyze the one-shot version of the game. The equilibrium concept is perfect Bayesian equilibrium.

4 Analysis

The game we analyze has two stages. Stage 1 is the referees' stage and Stage 2 is the editor's stage. In stage 1 the referees make their recommendations and in stage 2 the editor gives her verdict. We solve the game by backward induction.

In the equilibrium of the stage 2, the editor listens to the referees' recommendations, updates her belief about the quality of the paper, and gives her verdict: to R&R or reject the paper. Let $P(G \mid \mathbf{m})$ and $P(B \mid \mathbf{m})$ be the editor's posterior beliefs about the paper being good- and bad-quality, respectively, given

[6]This assumption is supported by the empirical literature (see Card and DellaVigna 2020 and Sobel 2020). It is also a common assumption in the theoretical literature. See Atal (2010), Baghestanian and Popov (2018), Bayar and Chemmanur (2021), Bertomeu (2020), Hirshleifer (2015), and Popov (2022).

the vector of referees' recommendations $\mathbf{m}$. Let $\mu_{\mathbf{m}}$ be the probability the editor rejects the paper given $\mathbf{m}$. We obtain the following result.

Proposition 1. *Given the profile of recommendations $\boldsymbol{m}$, in equilibrium the editor rejects the paper with probability $\mu^*_{\boldsymbol{m}} = P(B|\boldsymbol{m})$.*

The result says that, in equilibrium, the editor rejects the paper with the posterior probability that she assigns to the paper being bad. The referees anticipate this behavior and each chooses his recommendation so as to maximize the editor's posterior about him being high-ability. Prior to analyzing the behavior of referees under competition, we first describe the equilibrium behavior of a referee.[7]

Proposition 2. *With one referee, there exists thresholds $\bar{\alpha}_1(\gamma), \bar{\alpha}_2(\gamma) \in (0,1)$, with $\bar{\alpha}_1(\gamma) < \bar{\alpha}_2(\gamma)$, such that in equilibrium $\sigma^*_B = 1$ always. Additionally,*

1. *$\sigma^*_G = 0$ if and only if $\alpha \leq \bar{\alpha}_1(\gamma)$,*
2. *$\sigma^*_G \in (0,1)$ if and only if $\alpha \in (\bar{\alpha}_1(\gamma), \bar{\alpha}_2(\gamma))$,*
3. *$\sigma^*_G = 1$ otherwise.*

This result shows that a referee always recommends to reject after a bad signal and after a good signal he recommends to reject with positive probability. The latter phenomenon is referred to as over-rejection. Proposition 2 also shows that over-rejection increases in the referee's reputation, suggesting that referees of low-ability but high reputation reject more often than referees of lower reputation. The reason is that referees with higher reputation have small room for positive update if proven right but large room for negative update if proven wrong. In contrast, referees of lower reputation can gain a lot if proven right and do not risk that much if proven wrong. This asymmetry in the reputation at stake makes referees with higher reputation be "timid" to give R&R recommendations and so recommend to reject too often.

It is easy to observe that the over-rejection phenomenon appears for strategic motives and it is welfare-detrimental. In fact, since the referee's signal is informative and the editor uses the referee's recommendation to give the verdict about the paper, we can show that the strategy of the referee that maximizes the probability the editor gives the correct verdict —R&R for a good paper and Rejection for a bad paper— is the truthful strategy.[8] Accordingly, in the following we say that an equilibrium is *efficient* when the referee/s plays the truthful strategy $(\sigma^i_G, \sigma^i_B) = (0,1)$ for $i = 1,2$ in period 1, and the editor rejects the paper with probability $\mu^*_{\mathbf{m}} = P(B|\mathbf{m})$ in period 2. Hereafter, we focus our attention on the conditions for the existence of an efficient equilibrium. With one referee, we obtain the following result.

Corollary 1. *With one referee, an efficient equilibrium exists if and only if $\alpha \leq \bar{\alpha}_1(\gamma)$. In this equilibrium, $\mu^*_G = (1-\alpha)(1-\gamma)$ and $\mu^*_B = \alpha + (1-\alpha)\gamma$. The cutoff function $\bar{\alpha}_1(\gamma)$ is increasing in γ.*

[7] The expressions of thresholds $\bar{\alpha}_1(\gamma)$ and $\bar{\alpha}_2(\gamma)$ are in the proof of the result in Appendix A.

[8] We can show that the probability that the editor gives the correct verdict —given by $\sum_{k \in \{G,B\}} P(w_k)\left(P(a_k/m_k)P(m_k|\omega_k) + P(a_k/m_{-k})P(m_{-k}|\omega_k)\right)$, where a stands for the verdict (action) of the editor, m is the recommendation (message) of the referee, and ω is the quality (state of the world) of the paper— is decreasing in σ_G. We can also show that this probability is increasing in α and γ.

We observe that the higher the ability of the referee γ, the higher the probability the referee writes a truthful recommendation. This result highlights the relevance of the referee' ability in the field of the paper to guarantee not only a correct understanding of the paper (a first prerequisite for a good recommendation) but also the writing of a non-strategic (free from career-concerns) recommendation. We also observe that the higher the reputation of the referee, the higher the probability that the editor follows the referee's recommendation, i.e., $\frac{\partial \mu_B^*}{\partial \alpha} > 0$ and $\frac{\partial \mu_G^*}{\partial \alpha} < 0$. This is consistent with the findings in Card and DellaVigna (2020) (c.f. footnote 21).

Next, we analyze the behavior of referees with competition. We proceed in two steps. First, we consider two referees with known identities and then, we analyze an anonymous refereeing process. This procedure allows us to discern the effects of competition and heterogeneity on the one hand, and anonymity on the other hand. In general, we denote the referees by $i \in \{1,2\}$. When the referees have different reputations, we sometimes refer to them as the *senior* and the *junior* referees and use scripts S and J, respectively, with $\alpha_S > \alpha_J$.

4.1 Refereeing process with known identities

Let us start considering that the identities of the two referees are known to each other. This is a common feature in the evaluation of grants and fellowship applications, where the experts of a committee give an individual assessment about the merits of the applicant knowing who the other experts in the committee are.[9] We obtain the following result.

Proposition 3. *With two referees with known identities, there exist functions $\hat{\alpha}(\alpha_J,\gamma) \in (0,1)$ and $\hat{\gamma}(\alpha_J) \in (1/2,1)$ such that an efficient equilibrium exists if and only if $\alpha_S \leq \hat{\alpha}(\alpha_J,\gamma)$ and $\gamma > \hat{\gamma}(\alpha_J)$. The equilibrium probabilities $\mu_{\boldsymbol{m}}^*$, with $\boldsymbol{m} \in \{G,B\}^2$ are given by expressions (3)-(6).*

A first comment is that also with competition, the ability of the referees γ is key to guarantee truthful recommendations. The second comment refers to the reputation of the referees α and it says that the incentive to over-reject persists the introduction of competition. Figure 1 illustrates the results of Proposition 3. We observe that for a given α_J and γ, only when the senior referee has reputation smaller than cutoff $\bar{\alpha}(\alpha_J,\gamma)$, there is an equilibrium where the two referees are truthful. We also observe that higher ability γ sustains information transmission with more senior experts. A third and final comment refers to the effect of referee heterogeneity on their equilibrium behavior. Corollary 2 presents this result.

Corollary 2. *The cutoff function $\hat{\alpha}(\alpha_J,\gamma)$ is increasing in α_J. Hence, the more similar the referees are, the higher the region where the efficient equilibrium exists. In the limit, when the two referees have the same reputation α, low-ability referees are truthful if and only if $\gamma > \hat{\gamma}(\alpha)$. The cutoff function $\hat{\gamma}(\alpha)$ has an inverted-U shaped form in α.*

Two ideas are worth discussing. The first one is that the more similar referees are in their reputations, the softer the constraint on the referee's reputation and the higher the likelihood there is an efficient

[9] It is also the rule in the evaluation of accreditations to hold a position in the academia in countries such as Spain, France, or Italy. The accreditation is an official approval given by a national entity that certifies that the applicant has achieved the standard to promote in the academia.

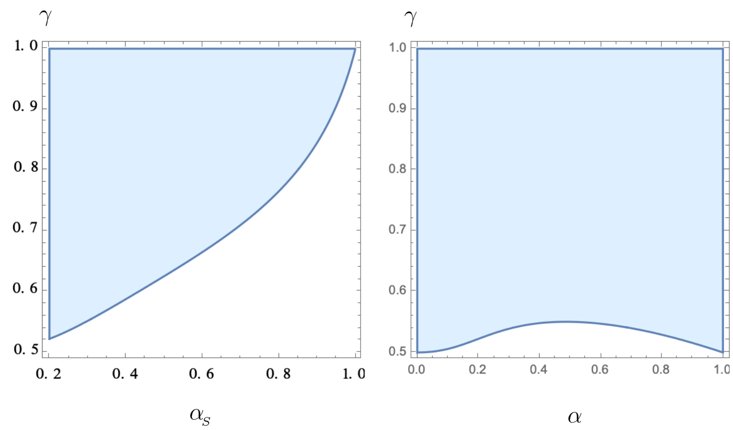

Figure 1: In blue, we represent the region where the efficient equilibrium exists. The left-hand side panel considers $\alpha_J = 0.2$ and describes the region where two low-ability referees are truthful, as a function of α_S. The right-hand side panel represents this region for the case of homogeneous referees, i.e., $\alpha_1 = \alpha_2 = \alpha$.

equilibrium. In the limit, when $\alpha_S \to \alpha_J$, the two referees are truthful conditioned only on them having sufficient ability, i.e., $\gamma > \hat{\gamma}(\alpha)$. The left-hand side panel of Figure 1 illustrates this result. The right-hand side panel shows how the cutoff function $\hat{\gamma}(\alpha)$ changes with α, when $\alpha_S \to \alpha_J = \alpha$. We observe that the effect if non-monotonic. This result suggests that with *homogeneous referees*, truthful recommendations are more likely to come from junior and senior referees and less likely to come from more "intermediate" referees.[10]

4.2 Refereeing process with unknown identities: Anonymous refereeing

In this section, we consider an anonymous refereeing process, i.e., referees who do not know the identity of the other referee, and thus do not know the opponent's reputation. This new source of uncertainty about the reputation of the other referee (the probability that he is high-ability, i.e. α_j) adds to the two sources of uncertainty already at work: uncertainty about the opponent's ability γ and uncertainty about the quality of the paper (or state of the world) ω. To keep the model tractable, we assume that the pool of referees can be divided into two groups by reputation: the group of senior referees and the group of junior referees. Senior referees have reputation α_S and junior referees have reputation α_J, and this is common-knowledge. Let λ_S be the probability that a senior referee competes with another senior referee, so $1 - \lambda_S$ is the probability that he competes with a junior referee. Similarly, let α_J be the probability that

[10]Though this result suggests an inverted-U shaped form, this is not the result we later test in the empirical analysis. Note that this result refers to an effect we obtain with homogeneous referees, whereas the result we test does not require it. Next, we give an intution for the non-monotonic result when $\alpha_1 = \alpha_2$. It helps to identify the two forces that underneath our results. The first one, already discussed, is that by rejecting a paper the referee reduces the probability the editor can assess his ability. This effect induces low-ability referees of higher reputation to over-reject and low-ability referees of lower reputation to recommend R&R. The second one is that with homogeneous referees, the ability of a referee is a perfect signal of the ability of the other referee. This effect induces low-ability senior referees to accept good papers (as the opponent is likely to be high-ability) and junior ones to reject them. For the low-ability senior referees, the second effect dominates the first one and so they are truthful. For the low-ability junior referees, the first effect dominates the second one and so they are truthful too. The referees with more "intermediate" levels of reputation are less affected by these disciplining forces and so they "misbehave" more often.

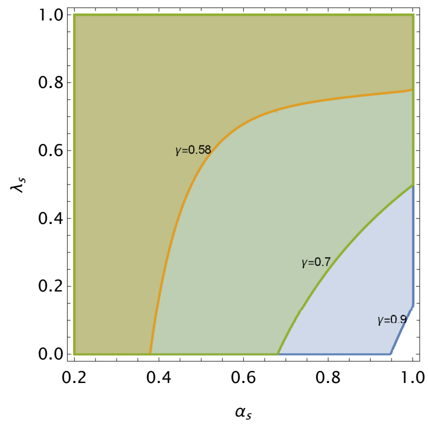

Figure 2: In color, we represent the region where the efficient equilibrium exists. We consider $\alpha_J = 0.2$ and represent this region as a function of α_S and λ_S, for different values of γ.

a junior referee competes with another junior referee, and let $1 - \lambda_J$ be the probability that he competes with a senior referee. Proposition 4 describes the equilibrium behavior of the referees in this case. It restricts attention to the case $\gamma > \hat{\gamma}_m$ (*A.1*).[11]

Proposition 4. *Under (A.1), with two referees with unknown identities, there exists cutoff $\hat{\lambda}_S \in (0,1)$ such that an efficient equilibrium exists if and only if either $\alpha_S \leq \hat{\alpha}(\alpha_J, \gamma)$ and $\gamma \geq \hat{\gamma}(\alpha_J)$, or $\lambda_S > \hat{\lambda}_S$. The equilibrium probabilities $\mu^*_{\boldsymbol{m}}$, with $\boldsymbol{m} \in \{G, B\}^2$ are given by expressions (3)-(6).*

This result says that with anonymous referees, there is room for an efficient equilibrium. It requires the referees to have sufficient ability and not very high reputation (the cutoff functions $\hat{\alpha}(\alpha_J, \gamma)$ and $\hat{\gamma}(\alpha_J)$ are those of Proposition 3) or, alternatively and perhaps more interesting, a high and known likelihood that the two referees can well be senior. This result suggests that a blind peer-review process is a good design of the reviewing process provided that (senior) referees believe that the other referee is very likely to be senior. Figure 2 illustrates the results of Proposition 4. We observe that for a given α_J and γ, the higher the belief λ_S, the higher the region where referees are truthful. We also observe that for a sufficiently high belief, i.e., $\lambda_S > \hat{\lambda}_S$, the efficient equilibrium always exists.

4.3 Discussion on the results

In this section, we summarize the theoretical results and derive the hypothesis that we test with data in the next section. Our model yields three main results describing the behavior of a referee. The first result characterizes the behavior of a low-ability referee and shows that the higher the referee's reputation, the higher his incentive to over-reject a paper. The second and third results show that the incentive of low-ability referees to over-reject decreases with competition, referees' homogeneity and with the anonymity of the refereeing process. With respect to high-ability referees, Lemma 1 in Appendix A shows that they are truthful in equilibrium.

[11] When $\gamma < \hat{\gamma}_m$, with $\hat{\gamma}_m \sim 0.551$, the analysis is less clear and many possibilities arise. The results for this case are available from the authors upon request.

Unfortunately, the nature of the refereeing process —referees do not know the identity or number of referees in the process— does not allow us to test our predictions about the effects of competition and anonymity.[12] Thus, in the empirical analysis that follows, we focus our attention on the effect of a referee's reputation on his recommendation.

However, testing the effect of reputation in the data is challenging. Our theoretical results show that the effect of reputation depends on the ability of the referee, but this is the referee's private information. Consequently, we cannot observe this information in the data. Unlike ability, a referee's reputation is public information, and our dataset contains information about the referee's publication record, which we believe is a good measure of a referee's reputation. Therefore, before testing our results in the data, we need to rewrite them in terms of reputation rather than ability. To do so, we take into account two important results. The first is that while low-ability referees reject more often the higher their reputation, high-ability referees are always truthful. The second is that the higher a referee's reputation, the more likely the referee is high-ability and the less likely he is low-ability. Combining these two effects, the probability that a referee rejects a paper is $\alpha\frac{1}{2} + (1-\alpha)(\frac{1}{2}\sigma_B^* + \frac{1}{2}\sigma_G^*)$, where the first term describes the probability that a high-ability referee rejects the paper and the second term the probability that a low-ability referee rejects it, with $\sigma_B^* = 1$ and σ_G^* given by the equilibrium value of Proposition 2. This probability defines an inverted U-shaped curve that describes the hypothesis that we will test in the empirical section.

Hypothesis. *The probability that a referee sends a rejection recommendation is inverted U-shaped in the referee's reputation.*

Figure 3 depicts this probability for the case $\gamma = 0.6$. The left-hand side panel represents in orange the equilibrium probability that a low-ability referee rejects the paper after a good signal, σ_G^*, and in blue the probability that the referee is low-ability, $1-\alpha$. The probability that a high-ability referee rejects a paper is constant in reputation. The right-hand side panel combines these effects, together with the fact that after a bad signal the low-ability referee always rejects the paper, to represent the probability that a referee random rejects the paper as a function of the referee's reputation.

[12] Although we do not present an empirical test of these results, we would like to briefly discuss a plausible approach to testing the effect of competition. Our conjecture is that the number a referee receives in the refereeing process may be informative (albeit an imperfect signal) about the degree of competition in the process. Namely, we conjecture that referees with higher referee numbers, e.g., referee numbers 3, 4, or higher, may perceive competition as more likely than referees with lower referee numbers, e.g., referee numbers 1 or 2. Accordingly, we would predict lower rejection rates from referees with higher referee numbers and higher rejection rates from referees with lower referee numbers. The analysis of the data shows that this is the case (even after controlling for the referee's publication record and other covariates). Although the result is consistent with our theoretical prediction, we are cautious about its validity. We acknowledge that a majority of referees may submit their reports without either knowing or paying attention to their referee number, so the statistical significance of this variable may mean something else, e.g., it may capture some unobservable characteristic of a referee's reputation that the publication record does not fully capture.

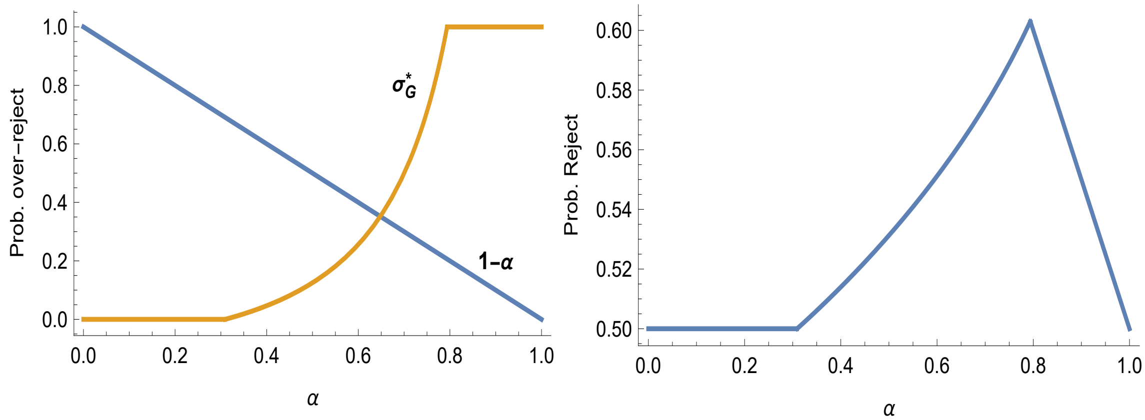

Figure 3: In the left-hand side panel we represent, in orange, the equilibrium probability that a low-ability referee over-rejects (σ_G^*) and, in blue, the probability that the referee is low-ability ($1-\alpha$). The probability that a high-ability rejects is constant in reputation. The right-hand side panel represents the probability that a referee rejects a paper. This probabilily is given by $\alpha\frac{1}{2} + (1-\alpha)(\frac{1}{2}\sigma_B^* + \frac{1}{2}\sigma_G^*)$, where $\sigma_B^* = 1$ and σ_G^* is defined by the equilibrium values of Proposition 2.

5 Empirical evidence

In this section, we use data from referees' recommendations at four leading economic journals —the *Quarterly Journal of Economics* (QJE), the *Review of Economic Studies* (REStud), the *Journal of the European Economic Association* (JEEA), and the *Review of Economics and Statistics* (REStat)— to test the prediction of an inverted U-shaped form between referee reputation and the recommendation to reject. Notice that due to the nature of the refereeing process —referees do not know the identity or number of referees in the process— we cannot test our theoretical predictions about the effect of competition, heterogeniety, and anonymity (c.f. footnote 12). Despite this limitation, the empirical analysis of this section provides a practical perspective that complements and helps support our theoretical insights.

5.1 Data

We make use of microdata used and well documented in Card and DellaVigna (2020) (CDV hereafter).[13] The dataset is anonymized and it contains information on all the submissions to the QJE, REStud, JEEA, and REStat during the period 2003-2013. For each submission, the dataset has information on the year of the submission, the number of coauthors and their recent prominent-publication records, the editorial decision, the recommendation of each referee, the number of recommendations previously done by each referee to the journal, and the recent prominent-publication record of each referee, among others. It makes 29,872 submissions, distributed as follows: QJE (N=10,824), REStud (N=8,335), JEEA (N=4,946), and REStat (N=5,767).

Our focus is on the referees' recommendations. Hence, we just consider non-desk-rejected papers.[14]

[13]The data is freely available at https://dataverse.harvard.edu/

[14]In Appendix B, we consider the whole sample of papers and perform a Heckman correction model.

This amounts a total of 17,441 submissions.[15] For non-desk-rejected papers, the referee's recommendation is one of the following: "Reject", "Definitively Reject (DefReject)", "Accept", "Revise and Resubmit (R&R)", "Strong R&R", "Weak R&R", and "No Recommendation (NoRec)". Non-desk-rejected papers have a minimum of 1 referee and a maximum of 10 referees. The mode of the distribution is 2 referees.[16] We focus on the recommendations made by referee numbers 1 to 6, which makes a total of 41,153 observations (the observation level is here the referee's recommendation) that are distributed as follows: Referee 1's recommendation (N=15,839), Referee 2's recommendation (N=13,602), Referee 3's recommendation (N=8,338), Referee 4's recommendation (N=2,800), Referee 5's recommendation (N=527), and Referee 6's recommendation (N=48). Referee numbers 7 to 10 are rare and just make 9 observations.

Table 1 presents summary statistics of the data. The first column presents descriptive statistics of our sample of referees 1 to 6. The other columns present statistics for each referee number. The upper panel presents information on the referee recommendations. For all the samples, we observe that the mode is "Reject", which accounts for about 50% of the referees' recommendations. The distribution of recommendations is quite stable across the samples.

[15] In the sample, out of 29,872 papers, 12,431 were desk-rejected. The others went into the refereeing process with 15,071 papers being rejected in the first round and 2,370 papers receiving R&R.

[16] The distribution of the number of referees of a paper is: 1 referee (N=2,990), 2 referees (N=7,238), 3 referees (N=5,414), 4 referees (N=1,567), 5 referees (N=217), 6 referees (N=14), 8 referees (N=1), and 10 referees (N=1).

Table 1. Summary Statistics for Referee Recommendations

	All		Ref 1		Ref 2		Ref 3		Ref 4		Ref 5		Ref 6	
	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent
Referee Recommendations														
Accept	1450	3.52	519	3.28	467	3.43	314	3.77	114	4.07	32	6.07	4	8.33
DefReject	5237	12.73	2520	15.91	1582	11.63	881	10.57	210	7.50	39	7.40	5	10.42
NoRec	2302	5.59	955	6.03	789	5.80	387	4.64	144	5.14	26	4.93	1	2.08
RR	4428	10.76	1466	9.26	1476	10.85	1016	12.19	380	13.57	82	15.56	8	16.67
Reject	21583	52.45	8244	52.06	7222	53.10	4361	52.30	1473	52.61	258	48.96	25	52.08
StrongRR	1815	4.41	608	3.84	613	4.51	414	4.97	155	5.54	23	4.36	2	4.17
WeakRR	4338	10.54	1526	9.63	1453	10.68	965	11.57	324	11.57	67	12.71	3	6.25
Referee Number of Publications in Past 5 Years in Prominent Journals														
0 publ	9664	23.48	3717	23.47	3140	23.08	1982	23.77	671	23.96	140	26.57	14	29.17
1 publ	6756	16.42	2604	16.45	2248	16.53	1371	16.44	444	15.86	83	15.75	6	12.50
2 publ	6182	15.02	2388	15.08	2061	15.15	1251	15.00	408	14.57	67	12.71	7	14.58
3 publ	5492	13.35	2133	13.47	1850	13.60	1077	12.92	348	12.43	80	15.18	4	8.33
4 publ	4246	10.32	1642	10.37	1412	10.38	848	10.17	293	10.46	46	8.73	5	10.42
5 publ	3213	7.81	1233	7.78	1037	7.62	648	7.77	245	8.75	46	8.73	4	8.33
6 publ	5600	13.61	2121	13.39	1854	13.63	1161	13.92	391	13.96	65	12.33	8	16.67
Total	41153	100.00	15839	100.00	13602	100.00	8338	100.00	2800	100.00	527	100.00	48	100.00
Referee Rejection Recommendations over Recommendations as a function of the Referee Number of Publications														
0 publ	5974	65,28	2338	66,82	1943	65,80	1204	62,97	399	62,44	83	61,48	7	50,00
1 publ	4346	68,18	1731	71,18	1446	68,34	852	64,69	272	64,15	40	50,63	5	83,33
2 publ	4097	70,01	1665	73,90	1347	69,08	797	66,64	245	64,47	38	56,72	5	83,33
3 publ	3710	71,47	1510	74,83	1235	70,73	698	68,84	224	67,47	42	54,55	1	25,00
4 publ	2882	71,96	1188	76,74	949	71,09	545	68,04	164	60,07	33	76,74	3	60,00
5 publ	2139	70,22	872	74,72	696	70,88	398	64,30	146	62,39	23	57,50	4	100,00
6 publ	3672	70,20	1460	74,26	1188	68,63	748	68,50	233	62,30	38	63,33	5	62,50
Total	26820		10764		8804		5242		1683		297		30	

Table 1: The table presents information on the frequencies and percentages of referee recommendations (upper panel), the referee number of prominent publications in the past 5 years (middle panel), and the negative recommendations of the referees as a function of the referee's record of publications (bottom panel). Negative recommendations include "Reject" and "DefReject" recommendations. The share of negative recommendations is obtained over the total number of referee recommendations, where we exclude "NoRec" recommendation. The data refers to submissions to the four leading journals *Quarterly Journal of Economics*, *Review of Economics and Statistics*, *Journal of the European Economic Association*, and *Review of Economic Studies*, during the period 2003-2013.

The middle panel of Table 1 presents descriptive statistics of the referees' record of publications in the past 5 years in prominent journals —top-5 and top-field journals.[17] We use this variable as the proxy of the referees' reputation. The variable ranges from 0 and is top-coded at 6 publications. We observe that the mode is 0 publications, which accounts for about 23% of the observations. We also observe that the fraction of referees with 6 publications accounts for about 14% of the observations. The numbers are quite stable across the samples.

Finally, the bottom panel of Table 1 presents summary statistics of the share of referee negative recommendations over total recommendations. We define a negative recommendation as a recommendation to either "Reject" or "DefReject". The columns of frequencies present, for each sample, the total number of negative recommendations. Since the mode of publications is 0, we observe that the higher number of negative recommendations comes from referees with 0 publications. The columns of percentages present, for each sample, the share of referee negative recommendations over total recommendations, where we exclude the "NoRec" recommendation.[18] We observe that the referees who reject less often are those with 0 publications. We also observe that the share of negative recommendations increases in the record of publications of the referee, attaining the maximum at the level of 3-4 publications and slightly decreasing afterwards. This suggests a non-monotonic effect of the referee reputation on the probability to send a negative recommendation. Figures 4 and 5 in Appendix B provide graphical support for this observation.

5.2 Results

Data observation suggests an inverted U-shaped relationship between the share of negative recommendations of a referee and the referee's record of publications. It suggests that for referees with lower records of publications (in prominent journals in the past 5 years), an increase in the number of publications increases the share of negative recommendations. However, when the record of publications of the referee is higher, having more publications reduces the share of negative recommendations. Note that the number of observations is consistently high across all the categories "number of publications", which implies that our results are not driven by a small number of anomalous observations in a particular category.

To investigate the suggested non-monotonic effect of reputation, we use two main measures of referee recommendation. The first measure summarizes the seven categories of referee recommendations into an index. From "Accept" to "DefReject", the index takes values 0 to 7.[19] Following Card et al. (2020) (CDVFI hereafter), we also construct a second index based on the predicted *asinh(GScitations)* coefficients of each recommendation category. The *asinh(GScitations)* variable is defined in CDV as the "inverse hyperbolic sine (asinh) of the citation count" and has the property to be well defined at zero —an important feature

[17]For the list of journals categorized as prominent, see Table 1 in the Online Appendix in Card and DellaVigna (2020). The list contains 35 journals including the top-5 and top-field journals such as *Journal of Finance, Journal of Econometrics, Journal of Economic Growth, Journal of Labor Economics, Journal of Public Economics*, and *Journal of Economic Theory*, among others.

[18]The share of referee negative recommendations is defined as the share of "Reject" + "DefReject" recommendations over "Reject" + "DefReject" + "Accept" + "R&R" + "WeakR&R" + "StrongR&R" recommendations, times 100. The exclusion of "NoRec"recommendations has not substantial effects on the results. See posterior discussion.

[19]Categories (in order) are "Accept", "Strong R&R", "R&R, "Weak R&R", "NoRec", "Reject", and "DefReject", taking values 0 to 7. We have also constructed an alternative index with 6 categories —without "NoRec" recommendation— and results maintain.

as about 30% of submitted papers have no citations. Using the coefficients of the cites model of CDV (Table II, column (4)), CDVFI propose an index that takes values 0, 0.67, 1.01, 1.47, 1.92, 2.27, 2.33 from "DefReject" to "Accept". To apply this index to our analysis, where the focus is on negative recommendations, we turn the positive values into negative values (from -2.33 to 0), for the index to increase from a positive to a negative recommendation. The second measure of referee recommendation we use considers the share of negative recommendations —"Reject" and "Definitively Reject"— over the total number of recommendations of the referees, where we exclude "NoRec" recommendations. Excluding "NoRec", total recommendations amount to 38,851 observations that constitutes our sample in the Logit analyses that we present next. Results for the Logit regressions do not change with the inclusion of "NoRec" recommendation —significance of coefficients do not change, only minor changes in point estimates.

Table 2 presents a series of OLS models for the index of referee recommendations. Models (1) to (4) consider the index taking values 0 to 7, from "Accept" to "DefReject". Models (5) and (6) consider the CDVFI index taking values -2.33 to 0, from "Accept" to "DefReject". The estimates reveal a clear negative relationship between the number of publications referees have and the recommendations they give, exhibiting a non-monotonic effect that is captured by the quadratic term. The coefficient of this term is always negative and statistically significant, suggesting an inverted U-shaped form in the referee's reputation. This effect remains robust across various model specifications. Models (1) to (3) differ in the inclusion of additional explanatory variables and controls for journal and field, and model (4) includes paper fixed effects. Models (5) and (6) replicate models (3) and (4) with the CDVFI index. Along with the effect of reputation, we observe that negative recommendations have also increased in time. Conversely, the number of publications by the authors positively influences recommendation scores, likely reflecting a prestige effect or acting as a proxy for the paper's quality. Additionally, the variable representing the inverse hyperbolic sine transformation of citations is positive and significant, indicating that, as expected, papers with more citations are likely to receive more positive recommendations. Last, the coefficients for the number of previous reports by the referee are negative and significant, suggesting that referees with more report-writing experience in a journal are more critical. Since a higher number of past reports could indicate a higher reputation acknowledged by the editor, capturing some unobservable that the referee's number of publications does not capture, we repeat our analysis running the OLS regression models (1)-(6) but proposing the quadratic form at the number of past reports instead. Results are presented in Table 5 in Appendix B. Aligned with previous results, we also observe an inverted-U shaped curve in the number of past reports, suggesting that this variable also captures some sort of reputation. The curvature of the inverted U-shaped form in the new estimations of Table 5 is however consistently smaller than in the estimations of Table 2, suggesting that the non-monotonic effect of reputation is stronger in the record of publications.[20]

To provide additional support for the non-monotonic effect of reputation on referee recommendations, we use a second simpler measure that considers the share of negative recommendations —"Reject" and "Definitively Reject"— over the total number of recommendations of the referees. Table 3 presents a

[20] The correlation between the variables "record of publications" and "number of past reports" is lower than what we expected: 0.0882.

Table 2. OLS models for index of referee recommendations

Variables	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6
Ref. no. publications	0.118***	0.124***	0.108***	0.148***	0.0444***	0.0612***
	(0.0131)	(0.0126)	(0.0125)	(0.0208)	(0.00513)	(0.00848)
Ref. no. publications^2	-0.0150***	-0.0146***	-0.0134***	-0.0169***	-0.00527***	-0.00679***
	(0.00215)	(0.00208)	(0.00205)	(0.00341)	(0.000846)	(0.00139)
Ref. no. past reports		0.0298***	0.0151***	0.0219***	0.00605***	0.00838***
		(0.00190)	(0.00179)	(0.00316)	(0.000755)	(0.00131)
No. ref. did not decline		-0.182***	-0.243***		-0.109***	
		(0.00914)	(0.00971)		(0.00402)	
Asinh(GScitations)		-0.161***	-0.171***		-0.0730***	
		(0.00517)	(0.00516)		(0.00213)	
Author no. publications		-0.0530***	-0.0629***		-0.0267***	
		(0.00430)	(0.00425)		(0.00174)	
Year submission		0.0101***	0.0119***		0.00425***	
		(0.00357)	(0.00364)		(0.00150)	
Paper fixed effects	No	No	No	Yes	No	Yes
Control for journal	No	No	Yes	–	Yes	–
Control for field	No	No	Yes	–	Yes	–
Constant	4.049***	-15.29**	-19.05***	3.955***	-9.208***	-1.047***
	(0.0159)	(7.168)	(7.316)	(0.0212)	(3.014)	(0.00861)
N	41,153	41,153	41,153	41,153	41,153	41,153
R^2	0.003	0.080	0.108	0.521	0.115	0.535

Table 2: The index of referee recommendation in columns (1)-(4) take values 0 to 7, from "Accept" to "DefReject". The index of columns (5) and (6) uses the CDVFI index taking (negative) values -2.33, -2.27, -1.92, -1.47, -1.01, -0.67, 0, from "Accept" to "DefReject". Robust standard errors, clustered at the paper level, in parentheses: $*** \, p < 0.01$, $** \, p < 0.05$, $* \, p < 0.1$.

series of Logit models for the probability to send a negative recommendation. The results also reveal a clear non-monotonic relationship between the probability to send a negative recommendation and the record of publications of the referee. Specifically, referees with one publication send more negative recommendations compared to those with zero publications (base category), and this effect intensifies with an increasing number of publications, reaching its peak at four publications and decreasing afterwards. This pattern remains robust across various model specifications. The inclusion of other explanatory variables, control variables, and fixed effects consistently produce similar results, albeit with some variability in the magnitude of the estimates.

Although the coefficients of the categorical variable "referee number of publications" in models (1)-(4) of Table 3 suggest a statistically significant effect of increasing number of publications on the probability of sending a negative recommendation (with respect to the base category: 0 publications), they do not necessarily imply that an increase in one publication has a statistically significant effect on the probability to send a negative recommendation. To address this question, next we obtain the (predicted) marginal effects of each category of publications and then perform contrast tests of the marginal effects. We perform the analysis for Model 3 of Table 3, which performs the best, and present the results in Table 4. The first column and row provides the marginal effect of each category of publications. Marginal effects are in brackets along with their significance. We observe that all the marginal effects are positive and statistically different from zero. The information inside the table provides the results from the contrast test of the marginal effects. Specifically, we compare the marginal effects of having X and $X + k$ publications on the probability to send a negative recommendation, with $X \in \{0, 1, 2, 3, 4, 5\}$ and $k \in \{1, 2, 3, 4, 5, 6\}$. The null hypothesis H_0 is *Marginal effect X+k publ - Marginal effect X publ = 0*. The diagonal of the matrix provides the results for an increase in the probability to send a negative recommendation when the referee has one more publication. We observe that the signs of the diagonal align with our theoretical prediction, albeit the effects are only statistically significant for the first and second steps, i.e., 0 to 1 and 1 to 2 publications.

6 Conclusion

This paper postulates a model of career concerns in the refereeing process in which referees seek to build a reputation for high ability. This reputational incentive introduces strategic considerations into the refereeing process that affect referee recommendations and, hence, editor decisions. We identify an incentive for low-ability referees to reject good papers, a phenomenon we call over-rejection. The incentive to over-reject a paper increases with referee reputation and decreases with referee ability. Competition in the refereeing process mitigates this incentive, as do the homogeneity of the referee pool and the anonymity of the refereeing process. These results suggest that common features of the current refereeing process in most scientific journals, such as anonymity and peer review, have beneficial properties. However, they also show that incentives to over-reject due to career concerns still persist in these scenarios.

We complement the theoretical analysis with an empirical validation of part of our results. Using data from Card and DellaVigna (2020) on paper submissions to the journals QJE, REStud, JEEA, and

Table 3. Logit models for probability of negative recommendation

Variables	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Ref. no. publications				
1 pub.	0.131***	0.135***	0.111***	0.193***
	(0.0347)	(0.0361)	(0.0368)	(0.0511)
2 pub.	0.217***	0.247***	0.215***	0.344***
	(0.0362)	(0.0376)	(0.0383)	(0.0521)
3 pub.	0.287***	0.303***	0.268***	0.474***
	(0.0380)	(0.0395)	(0.0403)	(0.0559)
4 pub.	0.311***	0.344***	0.304***	0.488***
	(0.0416)	(0.0431)	(0.0439)	(0.0608)
5 pub.	0.227***	0.286***	0.251***	0.375***
	(0.0460)	(0.0479)	(0.0488)	(0.0669)
6 pub.	0.226***	0.277***	0.211***	0.437***
	(0.0378)	(0.0398)	(0.0405)	(0.0562)
Referee no. past reports		0.0605***	0.0323***	0.0474***
		(0.00501)	(0.00463)	(0.00695)
No. ref. did not decline		-0.246***	-0.354***	
		(0.0137)	(0.0155)	
Asinh(GScitations)		-0.224***	-0.247***	
		(0.00809)	(0.00846)	
Author no. publ.		-0.0633***	-0.0814***	
		(0.00616)	(0.00631)	
Year submission		-0.00670	0.00283	
		(0.00562)	(0.00595)	
Paper fixed effects	No	No	No	Yes
Control for journal	No	No	Yes	–
Control for field	No	No	Yes	–
Constant	0.631***	15.53	-3.856	
	(0.0228)	(11.29)	(11.94)	
N	38,851	38,851	38,851	18,006
Pseudo R^2	0.00209	0.0574	0.0805	0.0172

Table 3: Robust standard errors, clustered at the paper level, in parentheses: $*** \; p < 0.01$, $** \; p < 0.05$, $* \; p < 0.1$.

REStat for the period 2003-2013, we analyze referee recommendations. Based on our theoretical analysis, we hypothesize that a referee's recommendation to reject is inverted U-shaped in the referee's reputation. Our empirical results support this hypothesis, showing a non-monotonic effect of reputation that is robust to different model specifications and the inclusion of controls and paper fixed effects.

We believe that the results in this paper have broader implications that go beyond the refereeing process and speak to other scenarios. Any scenario in which the implementation of a project reveals information about the experts' abilities, and the probability of implementation depends on the experts' decisions, can be addressed under the parameters of this model. Examples include financial advisors competing to influence an investor, or policy experts advising a politician on whether or not to implement a project. For these scenarios, the results in this paper lead to some relevant policy implications. The first is that the selection of experts should be based on their ability (or expertise) in the field, not on their reputation. Indeed, ability leads to more accurate reporting, whereas reputation leads to strategic motives. The second policy implication is that it is a good policy to let experts compete and to make them aware that they are competing. The third and final implication is that revealing the identity of experts in the process is a good policy only if experts are homogeneous in reputation. Otherwise, if the pool of experts is very heterogeneous, an anonymous process is better.

A Appendix: Proofs

In Appendix A we present the proofs of the results of the text. This appendix also contains the proofs of Propositions 5 and 6 that can be found in Appendix C.

Proof of Proposition 1

Proof. Given $\mathbf{m}$, the expected payoff to the editor from rejecting the paper is 0 and that of keeping the paper in is $P(G \mid \mathbf{m})(K-c) + (1 - P(G \mid \mathbf{m}))(-c)$. The editor rejects the paper whenever $P(G \mid \mathbf{m})(K - c) + (1 - P(G \mid \mathbf{m}))(-c) \leq 0$. This condition simplifies to $c \geq P(G \mid \mathbf{m})K$. Since $c \sim U[0, K]$, the probability that the editor rejects the paper is $1 - F\left(P(G \mid \mathbf{m})K\right) = P(B \mid \mathbf{m})$. □

Proof of Proposition 2

Proof. Since there is an only referee, we skip subindex i. Given the strategy profile of the referee (σ_G, σ_B), for a given $m \in \{G, B\}$, the editor's consistent beliefs about the quality of the paper being bad $P(\omega|m)$, with $\omega \in \{G, B\}$, are $P(B \mid G) = \frac{(1-\alpha)(\gamma(1-\sigma_B)+(1-\gamma)(1-\sigma_G))}{\alpha+(1-\alpha)(2-\sigma_B-\sigma_G)}$ and $P(B \mid B) = \frac{\alpha+(1-\alpha)(\gamma\sigma_B+(1-\gamma)\sigma_G)}{\alpha+(1-\alpha)(\sigma_B+\sigma_G)}$. From Proposition 1, these posteriors define μ_G and μ_B, respectively.

Additionally, for a given $m \in \{G, B\}$ and $X \in \{G, B, \varnothing\}$, the editor's consistent beliefs about the type of the referee $\hat{\alpha}(m, X)$ are $\hat{\alpha}(G, \varnothing) = \frac{\alpha}{\alpha+(1-\alpha)((1-\sigma_G)+(1-\sigma_B))}$, $\hat{\alpha}(B, \varnothing) = \frac{\alpha}{\alpha+(1-\alpha)(\sigma_B+\sigma_G)}$, $\hat{\alpha}(G, G) = \frac{\alpha}{\alpha+(1-\alpha)(\gamma(1-\sigma_G)+(1-\gamma)(1-\sigma_B))}$, $\hat{\alpha}(B, B) = \frac{\alpha}{\alpha+(1-\alpha)(\gamma\sigma_B+(1-\gamma)\sigma_G)}$, and $\hat{\alpha}(B, G) = \hat{\alpha}(G, B) = 0$.

In equilibrium, given signal $s \in \{G, B\}$, the low-ability referee anticipates the editor's behavior and computes the expected gain for sending message B rather than G, $\Delta_s = EU(B \mid s) - EU(G \mid s)$. These expressions are:

$$\Delta_G = \mu_B \hat{\alpha}(B,\varnothing) + (1-\mu_B)(\gamma\hat{\alpha}(B,G) + (1-\gamma)\hat{\alpha}(B,B))$$
$$-\mu_G\hat{\alpha}(G,\varnothing) - (1-\mu_G)(\gamma\hat{\alpha}(G,G) + (1-\gamma)\hat{\alpha}(G,B)), \tag{1}$$
$$\Delta_B = \mu_B \hat{\alpha}(B,\varnothing) + (1-\mu_B)(\gamma\hat{\alpha}(B,B) + (1-\gamma)\hat{\alpha}(B,G))$$
$$-\mu_G\hat{\alpha}(G,\varnothing) - (1-\mu_G)(\gamma\hat{\alpha}(G,B) + (1-\gamma)\hat{\alpha}(G,G)). \tag{2}$$

After some algebra we can show that $\Delta_B > \Delta_G$ always. See the Supplementary Appendix for the details. This result implies that if $\Delta_G \geq 0$, in which case $\sigma_G^* \in (0,1]$, then $\Delta_B > 0$ and so, in equilibrium, $\sigma_B^* = 1$. Otherwise, i.e., if $\Delta_G < 0$, then it can be that $\Delta_B \leq 0$, in which case, in equilibrium, $\sigma_G^* = 0$ and $\sigma_B^* \in [0,1)$. However, for the latter to occur we need that $\Delta_B \leq 0$ when $\sigma_G = 0$. After some algebra we can show that $\Delta_B|_{\sigma_G=0} > 0$. This is a contradiction. Hence, in equilibrium, $\sigma_B^* = 1$. In the rest of the proof we assume this result.

The expression of Δ_G evaluated at $\sigma_B = 1$ is:

$$\Delta_G|_{\sigma_B=1} = \alpha\left(\frac{1-\gamma}{\gamma+\alpha(1-\gamma)(1-\sigma_G)+\sigma_G(1-\gamma)} + \frac{\alpha-\alpha\gamma}{(1-(1-\alpha)\sigma_G)^2} + \frac{\alpha(1-\gamma)+2\gamma-1}{(1+\sigma_G-\alpha\sigma_G)^2} + \frac{1}{\sigma_G(1-\alpha)-1}\right).$$

We can show that $\frac{\partial \Delta_G|_{\sigma_B=1}}{\partial \sigma_G} < 0$, which implies that the equilibrium value σ_G^* is unique. We can also show that:

$$\Delta_G|_{\sigma_B=1,\sigma_G=0} < 0 \leftrightarrow [1/2 < \gamma < 1 \wedge 0 < \alpha < \bar{\alpha}_1(\gamma)]$$
$$\Delta_G|_{\sigma_B=1,\sigma_G=1} > 0 \leftrightarrow [1/2 < \gamma < 1 \wedge \bar{\alpha}_2(\gamma) < \alpha < 1],$$

with $\bar{\alpha}_1(\gamma) = \frac{2\gamma-1}{2(\gamma-1)} + \frac{1}{2}\sqrt{\frac{2\gamma-1}{(\gamma)^2-1}}$, and $\bar{\alpha}_2(\gamma)$ being the first real root of the following degree-3 polynomial in a:

$$f_{\alpha_2}(a;\gamma) = (-1+\gamma)a^3 + (3-2\gamma)a^2 - (3+2\gamma)a + 4\gamma.$$

The thresholds satisfy $\bar{\alpha}_1(\gamma) \in (0,1)$, $\bar{\alpha}_2(\gamma) \in (0,1)$, and $\bar{\alpha}_2(\gamma) > \bar{\alpha}_1(\gamma)$ for all $\gamma \in (1/2,1)$. Finally, in the non-empty region $\bar{\alpha}_1(\gamma) < \alpha < \bar{\alpha}_2(\gamma)$, we can show that the equilibrium value $\sigma_G^* \in (0,1)$ is given by the second real root of the degree-4 polynomial in a:

$$f_{\sigma_G}(a;\alpha,\gamma) = \sum_{k=0}^{4} \zeta_k(\alpha,\gamma) \cdot a^k,$$

with all the coefficients of $f_{\sigma_G}(a;\alpha,\gamma)$ being themselves polynomials in α and γ (hence, continuous functions) that for space reasons we omit. See the Supplementary Appendix. This completes the proof. □

Proof of Corollary 1

Proof. Given the referee's strategy profile $(\sigma_G, \sigma_B) = (0,1)$, the editor's consistent beliefs about the quality of the paper simplify to $P(B \mid G) = (1-\alpha)(1-\gamma)$ and $P(B \mid B) = \alpha + (1-\alpha)\gamma$, with $\mu_G = (1-\alpha)(1-\gamma)$

and $\mu_B = \alpha + (1-\alpha)\gamma$.

Similarly, her posterior beliefs about the type of the referee simplify to $\hat{\alpha}(m, \varnothing) = \alpha$, $\hat{\alpha}(B,B) = \hat{\alpha}(G,G) = \frac{\alpha}{\alpha+(1-\alpha)\gamma}$, and $\hat{\alpha}(B,G) = \hat{\alpha}(G,B) = 0$, with $m \in \{G, B\}\}$.

From Proposition 2 we know that, in equilibrium, $\sigma_B^* = 1$. Hence, we next evaluate expression (1) at the conjectured strategy profile $(\sigma_G, \sigma_B) = (0, 1)$. After some algebra, we obtain:

$$\Delta_G = -\frac{\alpha(\gamma-1)\left(-2\alpha^2 + 2(\alpha-1)^2\gamma + 2\alpha - 1\right)}{\alpha(\gamma-1) - \gamma},$$

We can show that $\Delta_G \leq 0$ if and only if $\alpha \leq \bar{\alpha}_1$, with $\bar{\alpha}_1 = \frac{2\gamma-1}{2(\gamma-1)} + \frac{1}{2}\sqrt{\frac{2\gamma-1}{(\gamma-1)^2}}$ as defined in the proof of Proposition 2. In addition, it is straightforward to show that $\partial\bar{\alpha}_1(\gamma)\partial\gamma > 0$. This completes the proof. $\square$

Proof of Proposition 3

Proof. Given the referees' strategy profile $(\sigma_G^1, \sigma_B^1; \sigma_G^2, \sigma_B^2) = (0,1;0,1)$, for a given $\mathbf{m} = (m_i, m_j)$ with $m_i, m_j \in \{G, B\}$ and $i, j \in \{1, 2\}$, the editor's consistent beliefs about the quality of the paper being bad $P(\omega|\mathbf{m})$, which from Proposition 1 is $\mu_{\mathbf{m}}$, are:

$$\mu_{G_i,G_j} = P(B \mid G_i, G_j) = \frac{(1-\alpha_i)(1-\alpha_j)(1-\gamma)(1-\gamma)}{(1-\alpha_i)(1-\alpha_j)(1-\gamma)(1-\gamma) + ((1-\alpha_i)\gamma+\alpha_i)((1-\alpha_j)\gamma+\alpha_j)}, \tag{3}$$

$$\mu_{B_i,B_j} = P(B \mid B_i, B_j) = \frac{((1-\alpha_i)\gamma+\alpha_i)((1-\alpha_j)\gamma+\alpha_j)}{(1-\alpha_i)(1-\alpha_j)(1-\gamma)(1-\gamma) + ((1-\alpha_i)\gamma+\alpha_i)((1-\alpha_j)\gamma+\alpha_j)}, \tag{4}$$

$$\mu_{G_i,B_j} = P(B \mid G_i, B_j) = \frac{(1-\alpha_i)(1-\gamma)((1-\alpha_j)\gamma+\alpha_j)}{(1-\alpha_j)(1-\gamma)((1-\alpha_i)\gamma+\alpha_i) + (1-\alpha_i)(1-\gamma)((1-\alpha_j)\gamma+\alpha_j)}, \tag{5}$$

$$\mu_{B_i,G_j} = P(B \mid B_i, G_j) = \frac{(1-\alpha_j)(1-\gamma)((1-\alpha_i)\gamma+\alpha_i)}{(1-\alpha_j)(1-\gamma)((1-\alpha_i)\gamma+\alpha_i) + (1-\alpha_i)(1-\gamma)((1-\alpha_j)\gamma+\alpha_j)}, \tag{6}$$

with $P(B \mid B_i, G_j) > P(B \mid G_i, B_j) \Leftrightarrow \alpha_i > \alpha_j$.[21]

Additionally, given $(\sigma_G^i, \sigma_B^i; \sigma_G^j, \sigma_B^j) = (0,1;0,1)$, for a given $\mathbf{m} = (m_i, m_j)$ and $X \in \{G, B, \varnothing\}$, the editor's consistent beliefs about the type of referee $i \in \{1, 2\}$, $\hat{\alpha}_i(m_i, m_j, X)$, are:

$$\hat{\alpha}_i(G_i, m_j, G) = \hat{\alpha}_i(B_i, m_j, B) = \frac{\alpha_i}{\alpha_i + (1-\alpha_i)\gamma},$$

$$\hat{\alpha}_i(G_i, G_j, \varnothing) = \hat{\alpha}_i(B_i, B_j, \varnothing) = \frac{\alpha_i((1-\alpha_j)\gamma+\alpha_j)}{\alpha_i((1-\alpha_j)\gamma+\alpha_j) + (1-\alpha_i)(\gamma((1-\alpha_j)\gamma+\alpha_j) + (1-\gamma)(1-\alpha_j)(1-\gamma))},$$

$$\hat{\alpha}_i(G_i, B_j, \varnothing) = \hat{\alpha}_i(B_i, G_j, \varnothing) = \frac{\alpha_i(1-\alpha_j)(1-\gamma)}{\alpha_i(1-\alpha_j)(1-\gamma) + (1-\alpha_i)((1-\alpha_j)(1-\gamma)\gamma + (1-\gamma)((1-\alpha_j)\gamma+\alpha_j))},$$

and $\hat{\alpha}_i(G_i, m_j, B) = \hat{\alpha}_i(B_i, m_j, G) = 0$, with $m_j \in \{G, B\}$.

Analogously to expressions (1)-(2), with two referees, the expected gain to the (low-ability) referee i from sending message B_i rather than G_i after observing signal $s_i \in \{G_i, B_i\}$ is $\Delta_s^i = EU_i(B_i \mid s_i) - EU_i(G_i \mid$

[21]This result supports the finding in Card and DellaVigna (2020) that editors give more weight to referees with stronger record of publications.

s_i), with:

$$EU_i(m_i \mid s_i) = P_i(B_j \mid m_i, s_i) EU_i(m_i, B_j, X \mid s_i) + P_i(G_j \mid m_i, s_i) EU_i(m_i, G_j, X \mid s_i) \tag{7}$$

and $m_i \in \{G_i, B_i\}$. The probabilities satisfy $P_i(m_j \mid m_i, s_i) = P_i(m_j \mid s_i)$, with $P_i(G_j \mid G_i) = \gamma(\alpha_j + (1-\alpha_j)\gamma) + (1-\gamma)(1-\alpha_j)(1-\gamma)$ and $P_i(G_j \mid B_i) = \gamma(1-\alpha_j)(1-\gamma) + (1-\gamma)(\alpha_j + (1-\alpha_j)\gamma)$.

The expected payoff $EU_i(\mathbf{m}, X \mid s_i)$, given the messages profile $\mathbf{m} = (m_i, m_j)$, is:

$$EU_i(\mathbf{m}, X \mid s_i) = \mu_{\mathbf{m}} \hat{\alpha}_i(\mathbf{m}, \varnothing) + (1-\mu_{\mathbf{m}}) \sum_{\omega} P_i(\omega \mid s_i, m_j) \hat{\alpha}_i(\mathbf{m}, \omega),$$

for $\omega \in \{G, B\}$, with:

$$P_i(B \mid s_i, B_j) = \frac{P_i(s_i \mid B)(\alpha_j + (1-\alpha_j)\gamma)}{P_i(s_i \mid B)(\alpha_j + (1-\alpha_j)\gamma) + P_i(s_i \mid G)(1-\alpha_j)(1-\gamma)},$$
$$P_i(B \mid s_i, G_j) = \frac{P_i(s_i \mid B)(1-\alpha_j)(1-\gamma)}{P_i(s_i \mid B)(1-\alpha_j)(1-\gamma) + P_i(s_i \mid G)(\alpha_j + (1-\alpha_j)\gamma)},$$

and $P_i(G \mid s_i, m_j) = 1 - P_i(B \mid s_i, m_j)$, for $s_i \in \{G_i, B_i\}$ and $m_j \in \{G_j, B_j\}$.

We are now ready to show the conditions under which the strategy profile $(\sigma_G^1, \sigma_B^1; \sigma_G^2, \sigma_B^2) = (0,1;0,1)$ is an equilibrium. It requires $\Delta_G^i \leq 0$ and $\Delta_B^i \geq 0$, for all $i = 1, 2$. Hereafter, we assume $\alpha_i \geq \alpha_j$, i.e., i denotes the senior referee and j denotes the junior referee.

First, we consider referee i, with $\alpha_i \geq \alpha_j$. Evaluated at $(\sigma_G^1, \sigma_B^1; \sigma_G^2, \sigma_B^2) = (0,1;0,1)$, after some algebra we can show that (see the Supplementary Appendix):

$$\Delta_B^i > 0 \ \text{ always}, \tag{8}$$
$$\Delta_G^i \leq 0 \Leftrightarrow [\hat{\gamma}_1(\alpha_j) < \gamma \leq \hat{\gamma}_2(\alpha_j) \wedge \alpha_i \leq \hat{\alpha}_1(\alpha_j, \gamma)] \vee [\gamma > \hat{\gamma}_2(\alpha_j) \wedge \alpha_i \leq \hat{\alpha}_2(\alpha_j, \gamma)], \tag{9}$$

where $\hat{\gamma}_1(\alpha_j)$ is the first real root of the degree-5 polynomial in a:

$$f_{\gamma_1}(a; \alpha_j) = \sum_{k=0}^{5} \lambda_k(\alpha_j) \cdot a^k,$$

with all the coefficients of $f_{\gamma_1}(a; \alpha_j)$ being themselves polynomials in α_j (hence, continuous functions), given by the following expressions:

$$\lambda_0(\alpha_j) = -3 + 12\alpha_j - 20\alpha_j^2 + 16\alpha_j^3 - 8\alpha_j^4,$$
$$\lambda_1(\alpha_j) = 18 - 80\alpha_j + 148\alpha_j^2 - 144\alpha_j^3 + 64\alpha_j^4,$$
$$\lambda_2(\alpha_j) = -48 + 236\alpha_j - 476\alpha_j^2 + 464\alpha_j^3 - 176\alpha_j^4,$$
$$\lambda_3(\alpha_j) = 76 - 392\alpha_j + 780\alpha_j^2 - 688\alpha_j^3 + 224\alpha_j^4,$$
$$\lambda_4(\alpha_j) = -72 + 352\alpha_j - 624\alpha_j^2 + 480\alpha_j^3 - 136\alpha_j^4,$$
$$\lambda_5(\alpha_j) = 32 - 128\alpha_j + 192\alpha_j^2 - 128\alpha_j^3 + 32\alpha_j^4.$$

Similarly, $\hat{\gamma}_2(\alpha_j)$ is the first real root of the degree-3 polynomial in a:

$$f_{\gamma_2}(a;\alpha_j) = \sum_{k=0}^{3} \eta_k(\alpha_j) \cdot a^k,$$

with all the coefficients of $f_{\gamma_2}(a;\alpha_j)$ being themselves polynomials in α_j (hence, continuous functions), given by the following expressions:

$$\begin{aligned}
\eta_0(\alpha_j) &= -1 + 2\alpha_j - 2\alpha_j^2,\\
\eta_1(\alpha_j) &= 4 - 14\alpha_j + 12\alpha_j^2,\\
\eta_2(\alpha_j) &= -10 + 28\alpha_j - 18\alpha_j^2,\\
\eta_3(\alpha_j) &= 8 - 16\alpha_j + 8\alpha_j^2.
\end{aligned}$$

Finally, $\hat{\alpha}_1(\alpha_j,\gamma)$ and $\hat{\alpha}_2(\alpha_j,\gamma)$ are the second and third highest real roots, respectively, of the degree-4 polynomial in a:

$$f_{\alpha}(a;\alpha_j,\gamma) = \sum_{k=0}^{4} \xi_k(\alpha_j,\gamma) \cdot a^k,$$

with all the coefficients of $f_{\alpha}(a;\alpha_j,\gamma)$ being themselves polynomials in (α_j,γ). They are very long polynomials and for space reasons, we omit them. See the Supplementary Appendix for the details.

Second, we consider referee j, with $\alpha_i \geq \alpha_j$. After some algebra we can show that:

$$\Delta_G^j \leq 0 \Leftrightarrow [\gamma \leq \hat{\gamma}_3(\alpha_j) \wedge \alpha_i \geq \hat{\alpha}_3(\alpha_j,\gamma)] \vee [\gamma > \hat{\gamma}_3(\alpha_j)], \tag{10}$$

where $\hat{\gamma}_3(\alpha_j) = \hat{\gamma}_1(\alpha_j)$ and $\hat{\alpha}_3(\alpha_j,\gamma)$ is the third highest real root of a degree-4 polynomial (that for space reasons we omit). From (9), we know that $\gamma > \hat{\gamma}_1(\alpha_j)$ is a necessary condition for $\Delta_G^i < 0$ to hold. Since, from (10), $\Delta_G^j < 0$ when $\gamma > \hat{\gamma}_1(\alpha_j)$, we conclude that after signal G, if referee i is truthful, referee j is truthful too. It implies that (10) does not restrict the equilibrium conditions given by (9).

Finally, we show that referee j lies after signal B when:

$$\Delta_B^j \leq 0 \Leftrightarrow [\gamma \leq \hat{\gamma}_4(\alpha_j) \wedge \hat{\alpha}_4(\alpha_j,\gamma) < \alpha_i \leq \hat{\alpha}_5(\alpha_j,\gamma)], \tag{11}$$

where $\gamma_4(\alpha_j)$ is the first real root of a degree-4 polynomial and $\hat{\alpha}_4(\alpha_j,\gamma)$ and $\hat{\alpha}_5(\alpha_j,\gamma)$ are the third and fourth highest real roots, respectively, of a degree-4 polynomial (that for space reasons we omit). We compare these conditions with (9) and show that $\hat{\gamma}_4(\alpha_j) < \hat{\gamma}_1(\alpha_j)$ if and only if $\alpha_j > \hat{\alpha}_j$, with $\hat{\alpha}_j$ being the scalar $\simeq 0.83$. This means that $\forall \alpha_j > \hat{\alpha}_j$, if referee i is truthful after signal G, referee j is truthful after signal B. Last, $\forall \alpha_j < \hat{\alpha}_j$, $\hat{\gamma}_1(\alpha_j) < \hat{\gamma}_4(\alpha_j)$. In this case we can show that $\hat{\alpha}_4(\alpha_j,\gamma) > \hat{\alpha}_1(\alpha_j,\gamma), \hat{\alpha}_2(\alpha_j,\gamma)$. Then, if referee i is truthful after signal G, referee j is truthful after signal B. Hence, we conclude that (11) does not affect the global equilibrium conditions.

To conclude, note that since we consider $\alpha_i > \alpha_j$, subscripts j and J both refer to the junior referee. We denote:

$$\hat{\alpha}(\alpha_J,\gamma)=\begin{cases}\hat{\alpha}_1(\alpha_j,\gamma) \text{ when } \hat{\gamma}_1(\alpha_j)<\gamma\leq\hat{\gamma}_2(\alpha_j), \text{ in which case } \hat{\gamma}(\alpha_J)=\hat{\gamma}_1(\alpha_j),\\ \hat{\alpha}_2(\alpha_j,\gamma) \text{ when } \gamma>\hat{\gamma}_2(\alpha_j), \text{ in which case } \hat{\gamma}(\alpha_J)=\hat{\gamma}_2(\alpha_j).\end{cases} \tag{12}$$

We prove that the piecewise function $\hat{\alpha}(\alpha_J,\gamma)$ is well defined in its domain. In this respect, it can be checked that $\hat{\alpha}_1(\alpha_j,\gamma)$ is defined for all $\hat{\gamma}_1(\alpha_j)<\gamma\leq\hat{\gamma}_2(\alpha_j)$ and function $\hat{\alpha}_2(\alpha_j,\gamma)$ is defined for all $\gamma>\hat{\gamma}_2(\alpha_j)$. We also prove that function $\hat{\alpha}(\alpha_J,\gamma)$ is continuous at the kink $\gamma=\hat{\gamma}_2(\alpha_j)$. It can be checked that $Limit_{\gamma\to\hat{\gamma}_2(\alpha_j)^-}\hat{\alpha}_1(\alpha_j,\gamma)=Limit_{\gamma\to\hat{\gamma}_2(\alpha_j)^+}\hat{\alpha}_2(\alpha_j,\gamma)$. Finally, we can prove that $\hat{\alpha}(\alpha_J,\gamma)$ is increasing in α_J. In this respect, it can be checked that $\frac{\partial\hat{\alpha}_1(\alpha_j,\gamma)}{\partial\alpha_j}>0\ \forall\hat{\gamma}_1(\alpha_j)<\gamma\leq\hat{\gamma}_2(\alpha_j)$ and $\frac{\partial\hat{\alpha}_2(\alpha_j,\gamma)}{\partial\alpha_j}>0\ \forall\gamma>\hat{\gamma}_2(\alpha_j)$.

We also prove that $\hat{\gamma}_1(\alpha_J)$ is inverted U-shaped in α_J, with $\frac{\partial\hat{\gamma}_1(\alpha_J)}{\partial\alpha_J}>0\ \forall\alpha_J\leq\alpha_{\gamma_1}$ and $\frac{\partial\hat{\gamma}_1(\alpha_J)}{\partial\alpha_J}<0$ $\forall\alpha_J>\alpha_{\gamma_1}$, with $\alpha_{\gamma_1}\sim 0.487$. It can be checked that in the limit it takes values $Limit_{\alpha_J\to 0}\hat{\gamma}_1(\alpha_J)=1/2$ and $Limit_{\alpha_J\to 1}\hat{\gamma}_1(\alpha_J)=1/2$. Finally, we prove that $\hat{\gamma}_2(\alpha_J)$ is decreasing in α_J and in the limit it takes values $Limit_{\alpha_J\to 0}\hat{\gamma}_2(\alpha_J)=0.829$ and $Limit_{\alpha_J\to 1}\hat{\gamma}_2(\alpha_J)=1/2$. See the Supplementary Appendix for the details. This completes the proof. □

Proof of Corollary 2

Proof. From the proof of Proposition 3, we know that $\hat{\alpha}(\alpha_J,\gamma)$ is increasing in α_J.

Additionally, when $\alpha_i=\alpha_j=\alpha$ and $(\sigma_G^1,\sigma_B^1;\sigma_G^2,\sigma_B^2)=(0,1;0,1)$, we can show that:

$$\Delta_B^i>0 \text{ always},$$
$$\Delta_G^i\leq 0\Leftrightarrow\gamma\geq\hat{\gamma}(\alpha),$$

where $\hat{\gamma}(\alpha)=\hat{\gamma}_1(\alpha_j)$ of the proof of Proposition 3. This proof shows the inverted-U shaped form of $\hat{\gamma}_1(\alpha_j)$, hence of $\hat{\gamma}(\alpha)$. □

Proof of Proposition 4

Proof. When the refereeing process is anonymous, the expected gain to the (low-ability) referee i for sending B_i rather than G_i after signal $s_i\in\{G_i,B_i\}$ is:

$$\Delta_s^i(\lambda) = \lambda_S\Delta_s^i|_{\alpha_i=\alpha_j}+(1-\lambda_S)\,\Delta_s^i|_{\alpha_i\neq\alpha_j} \text{ if } i \text{ is senior},$$
$$\Delta_s^i(\lambda) = \lambda_J\Delta_s^i|_{\alpha_i=\alpha_j}+(1-\lambda_J)\,\Delta_s^i|_{\alpha_i\neq\alpha_j} \text{ if } i \text{ is junior},$$

where $i,j\in\{1,2\}$, $\lambda_S,\lambda_J\in(0,1)$, $\Delta_s^i|_{\alpha_i=\alpha_j}$ is the expected gain when $\alpha_i=\alpha_j$, and $\Delta_s^i|_{\alpha_i\neq\alpha_j}$ is the expected gain when $\alpha_i\neq\alpha_j$.

Since $\Delta_s^i(\lambda)$ is a convex combination of $\Delta_s^i|_{\alpha_i=\alpha_j}$ and $\Delta_s^i|_{\alpha_i\neq\alpha_j}$, the strategy profile $(\sigma_G^1,\sigma_B^1;\sigma_G^2,\sigma_B^2)=(0,1;0,1)$ is an equilibrium if and only if evaluated at $(\sigma_G^1,\sigma_B^1;\sigma_G^2,\sigma_B^2)=(0,1;0,1)$, we have $\Delta_B^i|_{\alpha_i=\alpha_j}\geq 0$, $\Delta_B^i|_{\alpha_i\neq\alpha_j}\geq 0$, $\Delta_G^i|_{\alpha_i=\alpha_j}\leq 0$, and $\Delta_G^i|_{\alpha_i\neq\alpha_j}\leq 0$, for $i=1,2$. Next we demonstrate when this occurs. Hereafter, we assume $\alpha_i\geq\alpha_j$.

Consider first referee i and his expected gain after signal B, Δ_B^i. From (8) we know that $\Delta_B^i > 0$ for all $\alpha_i \geq \alpha_j$. It implies $\Delta_B|_{\alpha_i=\alpha_j} > 0$ and $\Delta_B|_{\alpha_i \neq \alpha_j} > 0$; hence $\Delta_B^i(\lambda) > 0$.

Consider now referee i and his expected gain after signal G, Δ_G^i. From (9) and the general conditions in Proposition 3, we know that for all $\alpha_i > \alpha_j$, $\Delta_G^i \leq 0 \Leftrightarrow \gamma > \hat{\gamma}(\alpha_j) \wedge \alpha_i \leq \hat{\alpha}(\alpha_j, \gamma)$. Additionally, from Corollary 2, we know that when $\alpha_i = \alpha_j = \alpha$, these conditions simplify to $\gamma > \hat{\gamma}(\alpha)$, with $\hat{\gamma}(\alpha)$ being inverted-U shaped. Let $\alpha_m = argmax\ \hat{\gamma}(\alpha)$ and let $\hat{\gamma}_m = \hat{\gamma}(\alpha_m)$, with $\hat{\gamma}_m \sim 0.551$. Then, $\gamma > \hat{\gamma}_m$ (A.1) is a sufficient condition for $\Delta_G^i|_{\alpha_i=\alpha_j} \leq 0$. Hence, if $\gamma > \hat{\gamma}(\alpha_j)$ and $\alpha_i \leq \hat{\alpha}(\alpha_j, \gamma)$ hold, then $\Delta_G^i(\lambda) \leq 0$ always. Otherwise, $\Delta_G^i|_{\alpha_i=\alpha_j} \leq 0$ but $\Delta_G^i|_{\alpha_i \neq \alpha_j} \geq 0$. In the latter case, because of the continuity of Δ_G^i in λ, we can assert there always exists $\hat{\lambda}_S < 1$ such that for all $\lambda \geq \hat{\lambda}_S$, $\Delta_G^i(\lambda) \leq 0$.

Consider next referee j and his expected gain after signal G. From condition (10) of Proposition 3, we know that if $\gamma > \hat{\gamma}_3(\alpha_j)$, then $\Delta_G^j \leq 0$. Additionally, in this case, $\hat{\gamma}_3(\alpha_j) = \hat{\gamma}_1(\alpha_j)$. From Corollary 2, $\hat{\gamma}_1(\alpha_j) = \hat{\gamma}(\alpha)$ and from the definition of $\hat{\gamma}_m$, $\hat{\gamma}_m = \max \hat{\gamma}(\alpha)$. Hence, under (A.1), $\Delta_G^j \leq 0$ always.

Finally, consider referee j and his expected gain after signal B. Condition (11) of Proposition 3 states the conditions under which this referee lies after signal B. From (11), if $\gamma > \hat{\gamma}_4(\alpha_j)$, then $\Delta_B^j > 0$. We can show that $\hat{\gamma}_4(\alpha_j)$ is increasing in α_j, with $\hat{\gamma}_4(\alpha_j)|_{\alpha_j=1} < \hat{\gamma}_m$. This completes the proof. □

Proof of Proposition 5

Proof. Let $\mu = \mu_{G_i,G_j} = \mu_{B_i,G_j} = \mu_{G_i,B_j} = \mu_{B_i,B_j}$. Evaluated at $(\sigma_G^1, \sigma_B^1; \sigma_G^2, \sigma_B^2) = (0,1;0,1)$, we can show that for all $i \in \{1,2\}$, $\Delta_B^i > 0$ and $\Delta_G^i < 0$ always. □

Proof of Proposition 6

Proof. Let χ_g be the ex-ante probability that any referee $i \in \{1,2\}$ affects the editor's decision when the latter uses decision rule $g \in \{D.1, Bayes\}$ and the other referee is truthful:

$$\chi_g = |\mu^*_{B_i,m_j} - \mu^*_{G_i,m_j}| = |1/2(\mu_{B_i,G_j} - \mu_{G_i,G_j}) + 1/2(\mu_{B_i,B_j} - \mu_{G_i,B_j})|.$$

Then, $\chi_{D.1} = 1/2$ and χ_{Bayes} is given by expressions (3)-(6). It can be shown that $\chi_{D.1} - \chi_{Bayesian} > 0 \Leftrightarrow \alpha_i > \tilde{\alpha}(\alpha_j, \gamma)$, with $\tilde{\alpha}(\alpha_j, \gamma) \in (0,1)$. The expression of the cutoff function is given by:

$$\tilde{\alpha}(\alpha_j, \gamma) = \frac{1}{2}\left(\frac{2(\alpha_j - 1)}{(2\alpha_j(\gamma-1)-2\gamma+1)^2} + \frac{2-2\alpha_j}{2\alpha_j(\gamma-1)-2\gamma+1} + \frac{1}{\gamma-1} + 2\right) +$$
$$+\frac{1}{2}\left(\sqrt{\frac{16\alpha_j^4(\gamma-1)^4 - 32\alpha_j^3(2\gamma-1)(\gamma-1)^3 + 4\alpha_j^2(24(\gamma-1)\gamma+5)(\gamma-1)^2 - 4\alpha_j(2\gamma-1)(8(\gamma-1)\gamma+1)(\gamma-1) + 4\gamma(1-2\gamma)^2(\gamma-1)+1}{(\gamma-1)^2(2\alpha_j(\gamma-1)-2\gamma+1)^4}}\right).$$

We can also show that $\frac{\partial\tilde{\alpha}(\alpha_j,\gamma)}{\partial\alpha_j} > 0$ and $\frac{\partial\tilde{\alpha}(\alpha_j,\gamma)}{\partial\gamma} > 0$. This completes the proof. □

The following lemma allows us to focus our analysis only on the equilibrium behavior of the low-ability referees.

Lemma 1. *In an efficient equilibrium, high-ability referees always use the truthful strategy.*

Proof. In an efficient equilibrium, low-ability referees use a truthful strategy, i.e., $(\sigma_G^1, \sigma_B^1; \sigma_G^2, \sigma_B^2) = (0, 1; 0, 1)$ —where the absence of a script refereing to the type denotes low-ability referees. Given this behavior, in the proof we show that the strategy profile $(\sigma_G^{H,1}, \sigma_B^{H,1}; \sigma_G^{H,2}, \sigma_B^{H,2}) = (0, 1; 0, 1)$ for the high-ability referees H is an equilibrium strategy profile.

We proceed like in the proof of Proposition 3. First, note that from the point of view of the editor, nothing changes. Then, for a given $\mathbf{m} = (m_i, m_j)$ with $m_i, m_j \in \{G, B\}$, the editor's consistent beliefs about the quality of the paper being bad $P(\omega|\mathbf{m})$ are given by equations (3) - (6). Similarly, the editor's consistent beliefs about the type of referee i, $\hat{\alpha}_i(m_i, m_j, X)$, are the same as in the proof of Proposition 3.

From the point of view of a high-ability referee i, her expected gain from sending message B_i rather than G_i after observing signal $s_i \in \{G_i, B_i\}$ is now $\Delta_s^{H,i} = EU_i^H(B_i \mid s_i) - EU_i^H(G_i \mid s_i)$, with:

$$EU_i^H(m_i \mid s_i) = P_i^H(B_j \mid m_i, s_i) EU_i^H(m_i, B_j, X \mid s_i) + P_i^H(G_j \mid m_i, s_i) EU_i^H(m_i, G_j, X \mid s_i)$$

and $m_i \in \{G_i, B_i\}$. The corresponding probabilities for a high-ability referee i are now $P_i^H(G_j \mid G_i) = \alpha_j + (1 - \alpha_j)\gamma$ and $P_i^H(G_j \mid B_i) = (1 - \alpha_j)(1 - \gamma)$. The expected payoff $EU_i^H(\mathbf{m}, X \mid s_i)$, given the messages profile $\mathbf{m} = (m_i, m_j)$, is:

$$EU_i^H(\mathbf{m}, X \mid s_i) = \mu_{\mathbf{m}} \hat{\alpha}_i(\mathbf{m}, \varnothing) + (1 - \mu_{\mathbf{m}}) \sum_{\omega, s_i, m_i, m_j \in \{G, B\}} P_i^H(\omega \mid s_i, m_j) \hat{\alpha}_i(\mathbf{m}, \omega),$$

with:

$$P_i^H(B \mid B_i, m_j) = P_i^H(B \mid B_i, m_j) = 1$$
$$P_i^H(B \mid G_i, m_j) = P_i^H(G \mid B_i, m_j) = 0$$

with $m_j \in \{G_j, B_j\}$. We can show that $\Delta_G^{H,i} < 0$ and $\Delta_B^{H,i} > 0$ always.

□

B Appendix: Further data description and empirical results

In Appendix B we provide further description of the data and a complementary empirical analysis.

B.1 Additional data description

Figures 4 and 5 provide graphical support for the information contained in the upper and bottom panels, respectively, of Table 1. For each samples of the referee numbers Ref 1 to Ref 6, Figure 4 represents the probability distribution of the variable "referee recommendation" and Figure 5 represents the share of negative and positive recommendations over total recommendations as a function of the variable "referee number of publications".

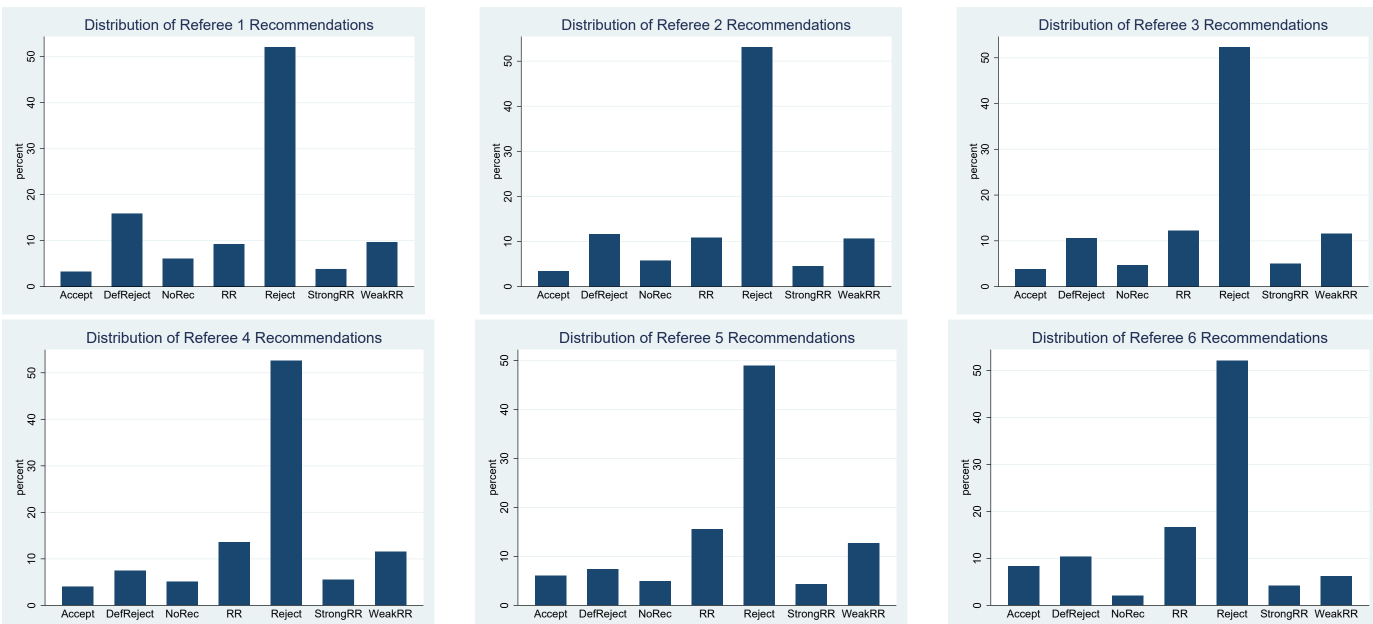

Figure 4: We represent the share of referee recommendations of each category. Panels 1 to 6 display the information for referee numbers Ref 1 to Ref 6.

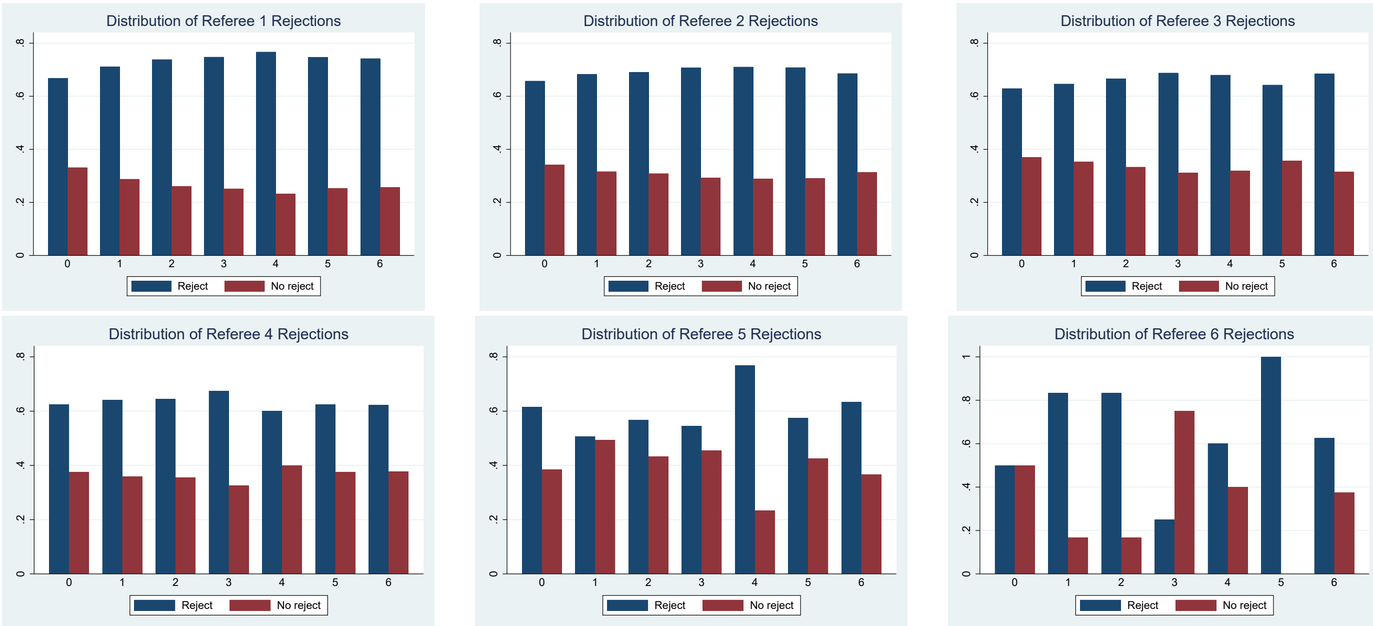

Figure 5: We represent the share of negative and positive recommendations over total recommendations —“Reject” and “No reject” respectively— as a function of the referee number of publications. We exclude “NoRec” recommendations. Panels 1 to 6 display the information for referee numbers Ref 1 to Ref 6.

Figure 6 represents the share of negative and positive recommendations over total recommendations as a function of the variable “referee number of past reports”, i.e., variable of interest in Table 5. We exclude “NoRec” recommendations. The representation corresponds to the full sample of referees.

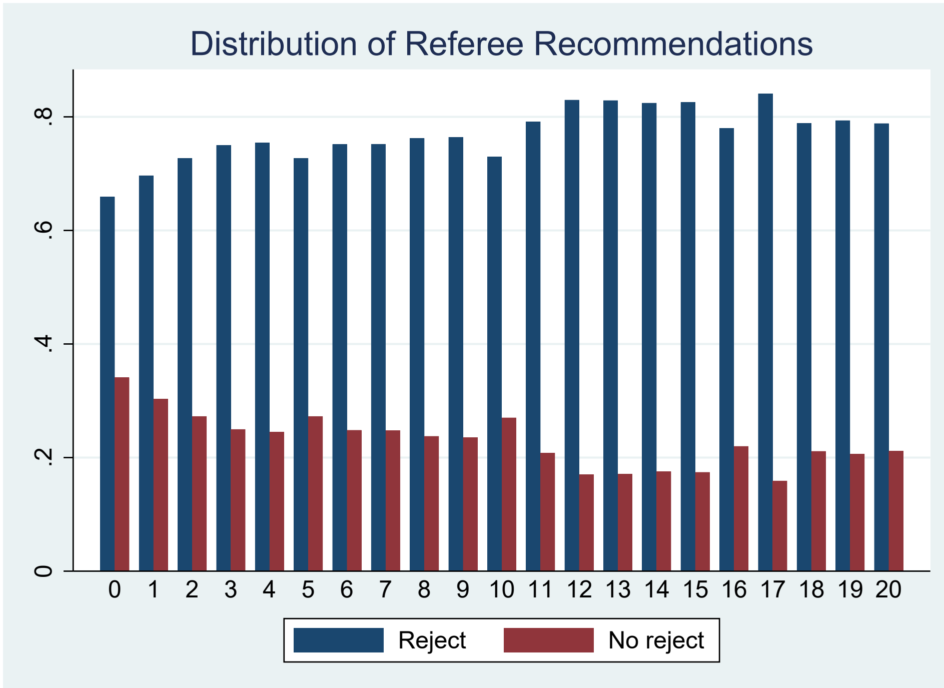

Figure 6: We represent the share of negative and positive recommendations over total recommendations —"Reject" and "No reject" respectively— as a function of the referee number of past reports. We exclude "NoRec" recommendations.

Table 5 presents the results for OLS models (1)-(6) that test the hypothesis of an inverted-U shaped form between the referee recommendation and the number of past reports of the referee, as suggested by Figure 6.

B.2 Additional empirical analysis

In this section we provide a robust test of the empirical results. We employ the Heckman two-step correction to address the potential selection bias arising from the non-random desk rejection process in the analysis of the referees' recommendations. This method is particularly well-suited for situations where the selection process —"Desk rejection" in our case— is not random, potentially leading to biased estimates if uncorrected. The first step of the Heckman correction models the probability of a paper being non-desk-rejected, accounting for observable characteristics that influence this initial selection phase. The second step then analyzes the referees' recommendations, incorporating the inverse Mills ratio (IMR) derived from the first step to correct for selection bias. This two-step approach enables us to make inferences about the entire submission process as if all papers, including those desk-rejected, could be observed in the final analysis. The method not only enhances the credibility of our findings by controlling for selection bias but also allows for a nuanced understanding of the factors influencing editor's decision at the first stage and the referees' recommendations. The models estimated are as follows:

Step 1. Selection Equation.— Probit model for desk rejection

$$P(NonDeskRejection_j = 1) = \Phi(\alpha_0 + \boldsymbol{\alpha_1 R_i} + \boldsymbol{\alpha_2 S_i} + \varepsilon_i) \quad (13)$$

where $P(NonDeskRejection_j = 1)$ is the probability that paper j is not desk rejected by the editor, Φ represents the cumulative distribution function of the normal distribution, indicative of a probit model, $\boldsymbol{R_j}$ and $\boldsymbol{S_j}$ are vectors representing the referee and submission characteristics, respectively, and ε_i is the error term.

Step 2. Main Equation.— Probit model for referee i's rejection recommendation

$$P(Rejection_{ij} = 1) = \Phi(\beta_0 + \beta_1 \text{IMR}_j + \boldsymbol{\beta_2 X_{ij}} + \boldsymbol{\beta_3 R_j} + \boldsymbol{\beta_4 S_j} + \mu_i) \quad (14)$$

where $P(Rejection_{ij} = 1)$ is the probability that referee i recommends rejection ("Reject" and "DefReject") of paper j. Φ again represents the CDF of the normal distribution, IMR_i is the inverse Mills ratio from the first step, which corrects for selection bias. $\boldsymbol{X_{ij}}$, $\boldsymbol{R_j}$, and $\boldsymbol{S_j}$ represent the characteristics of the paper, referee, and submission-related variables, respectively, and μ_i is the error term.

Table 6 presents the estimated coefficients from the two stages of the Heckman correction model. We observe that the new coefficients of the main equation are statistically significant which shows that our results are robust to this specification. The *athrho* value and the Wald test for independence of equations provide statistical tests for the presence of selection bias and the appropriateness of applying Heckman's correction. The *athrho* coefficient and standard value (not statistically different from 0) and a non-significant Wald test suggests that selection bias seems not to be a concern in our case.

C Appendix: Extensions

In Appendix C we discuss extensions of the theoretical model that we present in Section 3 and the robustness of our results to these variations. We consider four extensions: changes in the opportunity costs of publishing a paper, an alternative approximation to the editor's learning process, a different decision rule of the editor, and asymmetries in the referee's costs of actions.

Opportunity cost: The time cost and opportunity cost of keeping a paper in the refereeing process may vary across journals, fields of research, and methodologies. For example, it is reasonable to consider that top journals face a higher cost of processing a paper than non-top journals, as the opportunity cost is higher in top-journals. Similarly, interdisciplinary papers, papers using new methodologies, and papers speaking to new fields of research may have a higher opportunity cost. We can model differences in the cost of processing a paper by changing the support of the cost. In particular, we now consider $c \sim U[0, \rho K]$, with $\rho \in (0, 2)$, where a higher ρ stands for a higher cost of processing a paper.[22]

The analysis of this scenario yields two new insights. The first insight refers to the probability that

[22]We assume $\rho < 2$, so that the expected return from keeping a good paper in the process is always positive, i.e., $K - E(c) = K - \frac{\rho K}{2} > 0$.

the editor rejects the paper, which is now given by $1 - \frac{1-P(B|\mathbf{m})}{\rho}$. This result suggests that, ceteris paribus, the higher ρ, the higher the likelihood that the editor rejects a paper. This result provides a rationale to explain why top journals have higher rejection rates. It also helps rationalize why rejection rates are also usually higher for unconventional and unorthodox papers, as Gans and Shepherd (1994) document.[23]

The second insight refers to the power of a referee to affect the editor's decision. Since $\mu_{\mathbf{m}}$ is the probability the editor rejects a paper, $|\mu^*_{m_i,m_j} - \mu^*_{m'_i,m_j}|$ with $m_i \neq m'_i \in \{G, B\}$ measures the power of referee i to affect the editor's decision. We obtain that $|\mu^*_{B_i,m_j} - \mu^*_{G_i,m_j}| = |\frac{P(B|B_i,m_j)-P(B|G_i,m_j)}{\rho}|$, with this ratio decreasing in ρ. This result suggests that the higher the time cost or opportunity cost of keeping a paper in the refereeing process, the smaller the power of the referee to affect the editor's final decision.

Feedback: Rather than considering that the editor never observes the quality of the paper after rejection and she perfectly does it otherwise, we may consider a more intermediate situation in which learning occurs with positive probability after either decision. For simplicity, suppose that the likelihood the editor learns about the quality of the paper is the same independently of her decision. Let v be this probability. We obtain the following result.

Proposition 5. *For any $v \in (0, 1)$, there is always an equilibrium where the two referees are truthful.*

Note that when the probability of feedback is symmetric across the two decisions, the referees have no predisposition for taking either of the two actions, since the two actions are no longer different. Furthermore, with symmetric feedback, the power of the referee to affect the editor's decision vanishes. In this set-up, the best a referee can do is to follow her informative signal. This result suggests that the more similar the two decisions are (in terms of consequences), the smaller the distortion that career concerns introduce and the smaller the incentive to reject a good paper.

Editor's decision rule: Beyond considering a bayesian editor that maximizes the expected payoff of the journal taking into account all information she grasps from the referees, we may consider an editor who is ex-ante committed to a pre-specified and publicly known decision rule. For example, suppose an editor who is known to reject a paper unless both referees recommend keeping it in the process (*D.1*). It is straightforward to show that the power of referee i to affect the editor's decision, measured as the ex-ante probability he affects the editor's decision rule when the other referee is truthful, is $|\mu^*_{B_i,m_j} - \mu^*_{G_i,m_j}| = |\sum_{m\in\{G,B\}} \frac{1}{2}(\mu_{B_i,m_j} - \mu_{G_i,m_j})|$, which is 1/2 under *D.1*. In contrast, if the editor is bayesian, this ex-ante unconditional probability is given by expressions (3)-(6) in the Appendix. A comparison of the two decision rules yields the following result.

Proposition 6. *There exists $\tilde{\alpha}(\alpha_j, \gamma) \in (0, 1)$ such that for all $\alpha_i > \tilde{\alpha}(\alpha_j, \gamma)$, the ex-ante power of referee $i \in \{1, 2\}$ to change the editor's decision is higher when the editor is bayesian. Otherwise, his power to affect the editor's*

[23] The authors write: "*Until the 1970s, editors regularly rejected articles because they contained technical mathematics. The dominant editorial orthodoxy emphasized intuition, and viewed sophisticated mathematics as arid and irrelevant. Early papers by Tinbergen, Friedman, Hotelling, Debreu, and Lucas were all rejected for excess mathematics. In the 1970s, the technical tide rolled in. Leading journals filled with theorems and equations. Articles that contained only clear ideas in clear prose began to be rejected because they contained insufficient mathematics. Examples include the Akerlof and Arthur articles.*"

decision is higher when the latter uses decision rule D.1.

The result suggests that referees with high reputation have (relatively) more power under a bayesian editor, whereas referees with low reputation have (relatively) more power under decision rule *D.1*. Since a higher power increases the distortion of the referee's recommendations (towards rejection), and the distortion is higher for referees with high reputation, this result casts doubts on the optimality of having bayesian editors who use all information when referees have career concerns.[24]

Asymmetric cost of actions: Finally, similarly to considering that rejecting a paper has zero cost to the editor whereas keeping it in the process has cost c, we may argue that the cost to a referee of writing a rejection recommendation is different than the cost of asking for a revise and resubmit (alternatively, an acceptance recommendation). Without lose of generality, let us consider that recommending rejection has a smaller cost (e.g. due to lower effort). If we model the difference in cost between the two decisions as a fix cost, then it is easy to show that it has an effect on the region where referees recommend rejection, which indeed increases.

References

Andina-Díaz, Ascensión and José A García-Martínez (2020), 'Reputation and news suppression in the media industry', *Games and Economic Behavior* **123**, 240–271.

Atal, Vidya (2010), 'Do journals accept too many papers?', *Economics Letters* **107**(2), 229–232.

Baghestanian, Sascha and Sergey V Popov (2018), 'On publication, refereeing and working hard', *Canadian Journal of Economics/Revue canadienne d'économique* **51**(4), 1419–1459.

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

Berk, Jonathan B, Campbell R Harvey and David Hirshleifer (2017), 'How to write an effective referee report and improve the scientific review process', *Journal of Economic Perspectives* **31**(1), 231–44.

Bertomeu, Jeremy (2020), 'The editor's problem', *Available at SSRN 3710261* .

Camara, Fanny and Nicolas Dupuis (2019), 'Avoiding judgement by recommending inaction: Beliefs manipulation and reputational concerns', *Available at SSRN 3496638* .

Canes-Wrone, Brandice, Michael C Herron and Kenneth W Shotts (2001), 'Leadership and pandering: A theory of executive policymaking', *American Journal of Political Science* pp. 532–550.

[24] A different possibility is to consider a bayesian editor who wants to have "impact on the profession" and so might be inclined to support the topics she considers more relevant. We can model this disposition towards a topic as an extra payoff $h > 0$ that the editor receives when keeping a paper in this topic in the refereeing process. It is straightforward to show that the probability the editor rejects the paper is now $P(B \mid \mathbf{m}) - \frac{h}{K}$, which is smaller than before. This result suggests that the editor's disposition towards certain topics affects the journal's rejection rates, though it does not affect the power of a referee to affect the final decision.

Card, David and Stefano DellaVigna (2013), 'Nine facts about top journals in economics', *Journal of Economic Literature* **51**(1), 144–161.

Card, David and Stefano DellaVigna (2020), 'What do editors maximize? evidence from four economics journals', *Review of Economics and Statistics* **102**(1), 195–217.

Card, David, Stefano DellaVigna, Patricia Funk and Nagore Iriberri (2020), 'Are referees and editors in economics gender neutral?', *The Quarterly Journal of Economics* **135**(1), 269–327.

Cherkashin, Ivan, Svetlana Demidova, Susumu Imai and Kala Krishna (2009), 'The inside scoop: Acceptance and rejection at the journal of international economics', *Journal of International Economics* **77**(1), 120–132.

Chetty, Raj, Emmanuel Saez and László Sándor (2014), 'What policies increase prosocial behavior? an experiment with referees at the Journal of Public Economics', *Journal of Economic Perspectives* **28**(3), 169–88.

Ellisons, Glenn (2002), 'The slowdown of the economics publishing process', *Journal of Political Economy* **110**(5), 947–993.

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

Hargens, Lowell L. (1988), 'Scholarly consensus and journal rejection rates', *American Sociological Review* **53**(1), 139–151.

Hirshleifer, David (2015), 'Cosmetic surgery in the academic review process', *The Review of Financial Studies* **28**(3), 637–649.

Leaver, Clare (2009), 'Bureaucratic minimal squawk behavior: Theory and evidence from regulatory agencies', *American Economic Review* **99**(3), 572–607.

Levy, Gilat (2005), 'Careerist judges and the appeals process', *RAND journal of Economics* pp. 275–297.

Mariano, Beatriz (2012), 'Market power and reputational concerns in the ratings industry', *Journal of banking & Finance* **36**(6), 1616–1626.

Popov, Sergey V (2022), 'Tactical refereeing and signaling by publishing', *Available at SSRN4166875* .

Sobel, Joel (2020), 'Mourning edition', *UC San Diego WP* .

Welch, Ivo (2014), 'Referee recommendations', *The Review of Financial Studies* **27**(9), 2773–2804.

Zuckerman, Harriet and Robert K. Merton (1971), 'Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system', *Minerva* **9**, 66–100.

Table 4. Matrix of Contrast Test of Marginal Effects

	Margin 0 pub. [0.657***]	Margin 1 pub. [0.680***]	Margin 2 pub. [0.700***]	Margin 3 pub. [0.710***]	Margin 4 pub. [0.717***]	Margin 5 pub. [0.707***]
Margin 0 pub. [0.657***]						
Margin 1 pub. [0.680***]	0.022*** (0.007)					
Margin 2 pub. [0.700***]	0.042*** (0.007)	0.020** (0.008)				
Margin 3 pub. [0.710***]	0.052*** (0.008)	0.030*** (0.008)	0.010 (0.008)			
Margin 4 pub. [0.717***]	0.059*** (0.008)	0.037*** (0.009)	0.017* (0.009)	0.007 (0.009)		
Margin 5 pub. [0.707***]	0.049*** (0.009)	0.027*** (0.010)	0.007 (0.010)	−0.003 (0.010)	−0.010 (0.010)	
Margin 6 pub. [0.699***]	0.042*** (0.008)	0.019** (0.008)	−0.001 (0.008)	−0.011 (0.0089)	−0.017* (0.009)	−0.007 (0.010)

Table 4: Contrast tests of marginal effects of the categorical variable "referee number of publications" on the probability to send a negative recommendation. We use Model 3 of Table 3 where a negative recommendation is equal to 1. The marginal effects of each category and the significance levels are in brackets. Contrast test checks the difference between the probability to reject with $X + k$ and X publications (column minus row); level of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5. OLS models for index of referee recommendations: no. past reports

Variables	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6
Ref. no. past report	0.0458***	0.0529***	0.0299***	0.0411***	0.0124***	0.0163***
	(0.00303)	(0.00289)	(0.00288)	(0.00546)	(0.00121)	(0.00225)
Ref. no. past report^2	-0.000721***	-0.000812***	-0.000502***	-0.000573***	-0.000216***	-0.000235***
	(7.66e-05)	(7.54e-05)	(6.76e-05)	(0.000130)	(2.80e-05)	(5.30e-05)
Ref. no. publications		0.0355***	0.0286***	0.0470***	0.0130***	0.0204***
		(0.00378)	(0.00375)	(0.00638)	(0.00154)	(0.00260)
No. ref. did not decline		-0.186***	-0.243***		-0.109***	
		(0.00912)	(0.00970)		(0.00402)	
Asinh(GScitations)		-0.161***	-0.171***		-0.0730***	
		(0.00517)	(0.00516)		(0.00213)	
Author no. publications		-0.0529***	-0.0627***		-0.0266***	
		(0.00430)	(0.00425)		(0.00174)	
Year submission		0.00828**	0.0104***		0.00363**	
		(0.00357)	(0.00365)		(0.00150)	
Paper fixed effects	No	No	No	Yes	No	Yes
Controls for journal	No	No	Yes	–	Yes	–
Controls for field	No	No	Yes	–	Yes	–
Constant	4.119***	-11.49	-16.05**	4.008***	-7.931***	-1.026***
	(0.0101)	(7.167)	(7.323)	(0.0164)	(3.018)	(0.00667)
N	41,153	41,153	41,153	41,153	41,153	41,153
R^2	0.006	0.081	0.108	0.521	0.115	0.535

Table 5: The index of referee recommendation in columns (1)-(4) take values 0 to 7, from "Accept" to "DefReject". The index of columns (5) and (6) uses CDVFI index taking (negative) values -2.33, -2.27, -1.92, -1.47, -1.01, -0.67, 0, from "Accept" to "DefReject". Robust standard errors in parentheses: $***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$.

Table 6: Heckman correction model

Variables	(1) Stage 2: Main eq.	(2) Stage 1: Selection eq.
Ref. no. publications		
1 pub.	0.0841***	
	(0.0217)	
2 pub.	0.152***	
	(0.0226)	
3 pub.	0.187***	
	(0.0237)	
4 pub.	0.210***	
	(0.0258)	
5 pub.	0.179***	
	(0.0287)	
6 pub.	0.171***	
	(0.0239)	
Ref. no. past reports	0.0327***	
	(0.00288)	
No. ref. did not decline	-0.152***	
	(0.00841)	
Asinh(GScitations)	-0.134***	0.225***
	(0.00548)	(0.00529)
Author no. publications	-0.0391***	0.191***
	(0.00450)	(0.00521)
Year submission	-0.00289	0.0110***
	(0.00336)	(0.00360)
Control for journal		0.874***
		(0.0300)
Control for field		0.269***
		(0.0487)
Constant	7.066	-23.68***
	(6.754)	(7.227)
N	113,437	113,437

Table 6: Coefficient of correlation parameter (*athrho*) and standard error (in parenthesis)= -0.00253 (0.0248). Wald test of independent equations (rho = 0): chi2(1)=0.01, Prob > chi2 = 0.9187. Robust standard errors in parentheses: $***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$.