


ORIGINAL PAPER

# Open science, closed peer review?

Gary Charness<sup>1</sup>, Anna Dreber<sup>2</sup>, Daniel Evans<sup>3</sup>, Adam Gill<sup>4</sup> and Séverine Toussaert<sup>5</sup> 

<sup>1</sup>Department of Economics, University of California, Santa Barbara, California, United States

<sup>2</sup>Department of Economics, Stockholm School of Economics, Stockholm, Sweden

<sup>3</sup>Department of Economics, University of Bonn, Bonn, Germany

<sup>4</sup>Department of Economics, Uppsala University, Uppsala, Sweden

<sup>5</sup>Department of Economics, University of Oxford, Oxford, UK

**Corresponding author:** Séverine Toussaert; Email: [severine.toussaert@economics.ox.ac.uk](mailto:severine.toussaert@economics.ox.ac.uk)

(Received 24 December 2024; revised 19 July 2025; accepted 22 January 2026)

## Abstract

Open science initiatives have gained traction in recent years. However, open *peer-review* practices, i.e., reforms that (i) modify the identifiability of stakeholders and (ii) establish channels for the open communication of information between stakeholders, have seen very little adoption in economics. In this paper, we explore the feasibility and desirability of such reforms. We present insights derived from survey data documenting the attitudes of 802 experimental/behavioral economists, a conceptual framework, a literature review, and cross-disciplinary data on current journal practices. On (i), most respondents support preserving anonymity for referees, but views about anonymity for authors and associate editors are mixed. On (ii), most respondents are open to publishing anonymized referee reports, sharing reports between referees, and allowing authors to appeal editorial decisions. Active reviewers, editors, and respondents from the US/Canada are generally less open to transparency reforms.

**Keywords:** editorial policies; journals; open science; peer review; transparency

**JEL Codes:** A11; A14; C83

## 1. Introduction

The open science movement in economics has seen success in recent years with the adoption of transparency initiatives on the production side of research, e.g., pre-analysis plans (Ferguson et al., 2023). However, little attention has been paid to transparency on the evaluation side, i.e., open *peer-review* practices. These practices can be divided into (i) those that modify the identifiability of authors, referees, and editors and (ii) those that facilitate the open communication of information between stakeholders. Assuming that journals aim to maximize the quality of the papers they publish, the optimal level of transparency is uncertain without knowing more about the (perceived) costs and benefits of various policies. In this paper, we leverage a survey of experimental/behavioral economists to learn how stakeholders reason about key trade-offs and whether the growing support for open science practices extends to peer review. We complement the survey with a conceptual framework, a literature review, and cross-disciplinary data on journal practices.

The take-up of open peer-review practices has been slow across disciplines (Wolfram et al., 2020), including economics. Among the top-5 economics and Economic Science Association (ESA) journals (Table 1), some disclose metadata and handling editor identities, but none reveal referee identities

Table 1 Transparency policies

Identifiability	ESA journals		Top-5 journals					Other disciplines		
	<i>Exp Econ</i>	<i>JESA</i>	<i>AER</i>	<i>ECMA</i>	<i>JPE</i>	<i>QJE</i>	<i>REStud</i>	<i>BMJ</i>	<i>Frontiers</i> <sup>a</sup>	<i>Nature</i> <sup>b</sup>
Blinded review	Single	Single	Single	Single	Single	Single	Single	Open	Single	Opt-in
Referee identity disclosed	–	–	–	–	–	–	–	✓	✓	Opt-in
Handling editor disclosed	–	–	✓	✓	✓	–	✓	✓	✓	–
Journal publishes list of referees	No	No	Full	Full	Full	Partial	Partial	Full	No	No
Authors can nominate referees	–	–	–	–	–	–	–	✓	✓	✓
Authors can oppose referees	–	–	–	–	–	–	–	✓	✓	✓
Conflict-of-interest policy for editors	✓	✓	✓	✓	✓	–	✓	✓	✓	✓
Conflict-of-interest policy for referees	✓	✓	–	–	✓	–	–	✓	✓	✓
<b>Disclosure</b>										
Turnaround time statistics	–	–	✓	✓	✓	✓	✓	✓	✓	✓
Acceptance/rejection statistics	–	–	✓	✓	✓	✓	–	✓	✓	✓
Manuscript received/decision dates	✓	✓	–	✓	–	✓	✓	✓	✓	✓
Public decision letters	–	–	–	–	–	–	–	✓	–	✓
Public referee reports	–	–	–	–	–	–	–	✓	–	✓
Signed public referee reports	–	–	–	–	–	–	–	✓	–	Opt-in
Prior manuscript versions	–	–	–	–	–	–	–	✓	–	–
<b>Communication channels</b>										
Policy on author appeals	✓	✓	–	✓	✓	–	–	✓	✓	✓
Interactive review	–	–	–	–	–	–	–	–	✓	–

Notes: For journal abbreviations and further information, please see Online Resource Appendix A.4. The sources for many journal policies are available in Online Resource Table A.6. <sup>a</sup>After acceptance, *Frontiers* publishes the names of referees who have recommended the manuscript for publication. <sup>b</sup>*Nature* allows authors to opt in to double-anonymity, while authors and referees must both agree for referee identities to be publicly shared. All *Nature* publications must now feature a peer-review file, which includes signed or unsigned reports and author responses, but only some files feature decision letters.

or publish referee reports. One notable exception to the pattern of slow take-up comes from the *Nature* journals. On identifiability, *Nature* began disclosing reviewer identities in 2016, subject to the consent of the authors and referees (Nature, 2019). Around 80% of papers identified at least one referee during the trial period, suggesting some openness to these practices. On open communication, *Nature Communications* began publishing referee reports in 2016, subject to author consent (Nature Communications, 2015). The opt-in rate rose from 60% in 2016 to 70% in 2021, after which the journal decided to publish all reports without exception (Nature Communications, 2016; Nature Communications, 2022). In 2025, *Nature* announced a similar full disclosure policy (Nature, 2025).

Given the limited evidence from economics, we surveyed economists on their views about open peer-review practices as part of a larger survey on peer review (Charness et al., 2022). In this article, we use this data to examine the main arguments for and against open peer review. We focus on the attitudes of experimental/behavioral economists ( $N = 802$ ), many of whom are ESA members and from whom we recorded a high response rate.

We document three main findings. First, on identifiability, most respondents support preserving anonymity for referees, but views about authors and associate editors are mixed. However, there is clear support for giving authors the option to request, ex-ante, that certain reviewers not be contacted. Second, regarding open communication, most respondents support publishing anonymized referee reports, sharing reports between referees, and opening a channel for authors to appeal editorial decisions. Third, in a latent class analysis (LCA), we find that responses cluster in three groups, which we label as Traditionalists (33% of respondents), Incrementalists (38%), and Reformists (28%). Respondents from the US/Canada, active reviewers, top-5 reviewers, and editors are more likely to be classified as Traditionalists, while female respondents are less likely to be.

These attitudes could serve as inputs into journal decision-making in two ways. First, they signal which policies are ripe for implementation (e.g., sharing reports) and which might provoke backlash (e.g., referee identifiability). Second, they may proxy for beliefs about policy efficacy, which could help journals identify policies in need of further experimentation (e.g., double-blind anonymity).

Our paper complements recent surveys on peer review that have been conducted in cross-disciplinary settings (Publishing Research Consortium, 2016; Ross-Hellauer et al., 2017) and within economics (Altonji et al., 2025). Compared to the cross-disciplinary surveys, we cover ground especially relevant to journals in economics and take an economic perspective on issues common across disciplines. This distinction is relevant because findings from other disciplines may not generalize to economics. Economics is a relatively small discipline<sup>1</sup> and the effects of identifiability may differ when networks are small. Additionally, peer review in economics features long publication lags (Hadavand et al., 2024), which incentivizes authors to disseminate their preprints widely, thereby complicating anonymity policies. Finally, economics is a social science, which may entail more subjectivity in how we evaluate research findings and imply a greater need for policies that protect the ability of referees to dissent. As compared to Altonji et al. (2025), who survey economists from all fields, we focus on the views of experimental/behavioral economists, whose attitudes may be more relevant to the readership of this journal. Moreover, we investigate several policies that they do not, including referee identifiability and the disclosure of review histories.

Our efforts are divided into six sections. We introduce our survey design and dataset in Section 2. Section 3 features a conceptual framework that formalizes how journals may reason about policy adoption. Next, we explore policies that affect the identifiability of stakeholders (Section 4) and open communication in peer review (Section 5). Section 6 investigates potential heterogeneity and clustering in participant responses. Finally, Section 7 discusses the limitations of our study and possible steps forward.

<sup>1</sup>For example, the average top-5 journal in economics published 241 articles in 2019–2021, compared to an average of 1,718 articles at top medicine journals in the same timeframe (namely, *The Journal of the American Medical Association*, *The Lancet*, and *The New England Journal of Medicine*). These patterns hold when looking at alternative measures of journal reach.

## 2. Survey on peer review

We conducted an anonymous survey of economists between July 2020 and January 2021. Researchers could participate if, over the prior two years, they had (i) written at least one referee report and (ii) received reports on at least one journal submission. We inquired about their experiences with the system from their dual perspectives as authors and reviewers, their opinions on its current performance, and their attitudes toward certain proposals for reform.<sup>2</sup> The full questionnaire is in Online Resource Appendix C and the dataset is on our OSF page (<https://doi.org/10.17605/OSF.IO/GJ586>).

This paper focuses on respondents' attitudes about eight proposed policies. For each policy, respondents indicated either how favorable they are or how useful they find it for improving referee reports on a 1–5 Likert scale. The policies include publishing referee reports and author responses (either anonymously or signed), sharing reports between reviewers, de-anonymizing senior referees and associate editors, implementing double-blind review, and allowing authors to oppose the assignment of specific reviewers or to appeal manuscript rejections. When relevant, we also report results from other parts of the survey, including perceptions of referee report quality and conflicts of interest.

Our recruitment strategy (detailed in Online Resource Appendix A.1) leveraged direct emails, advertising through professional networks (e.g., CESifo, CEPR), discussion forum posts (e.g., ESA), and social media. The first wave of emails (July–October 2020) targeted experimental/behavioral economists, yielding 606 responses from 1,802 emails (34%). The second wave (November 2020–January 2021) expanded to other subfields, yielding 269 responses from 3,618 emails (7%). In total, we obtained 1,459 complete responses. Given the interests of this journal's readership, we restrict our presentation of the data to the experimental/behavioral subsample ( $N = 802$ , 54% of respondents).

Basic descriptive statistics for this subsample are presented in Online Resource Table A.3. To assess potential selection bias, we compared the characteristics of this subsample to those of 2020 ESA members (Online Resource Table A.4) and those of our full sample to the weighted sample in Andre and Falk (2021) and the full sample in Altonji et al. (2025) (Online Resource Appendix A.3). Relative to these benchmarks, we over-sample researchers in Europe and, in the full sample, we also over-sample experimental/behavioral economists.

## 3. Conceptual framework

A natural objective for journals is to maximize the expected quality of the papers they publish (Card & DellaVigna, 2020). Consider a journal that receives a manuscript submission with unobserved quality  $q \sim \mathcal{N}(\bar{q}, \sigma_q^2)$ . The journal cannot observe  $q$  directly; instead, it assigns the paper to a representative reviewer who receives a private signal  $s$ :

$$s = q + b(r, t) + \varepsilon,$$

where  $b(r, t)$  is a bias term dependent on the author's reputation ( $r \in \mathbb{R}$ ) and the reviewer's taste ( $t \in \mathbb{R}$ ).  $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2(e))$  is an error term.  $\frac{\partial b(r, t)}{\partial r} > 0$  implies more favorable reviews for influential authors due to, e.g., status bias (Huber et al., 2022) or expected reciprocity/retaliation. The parameter  $t$  represents a reviewer's perception of a manuscript that is related to factors other than  $r$  or  $q$ , e.g., personal relationships, conflicts of interest, and/or idiosyncratic preferences. The variance of the noise depends on reviewer effort  $e \geq 0$ , with  $\frac{\partial \sigma_\varepsilon^2(e)}{\partial e} < 0$ . Effort also improves the final quality of the published paper via revisions, i.e.:

$$q^{\text{published}} = q + \Delta(e), \quad \text{with } \Delta'(e) > 0.$$

<sup>2</sup>We block-randomized the order of the author and reviewer question blocks, but found no evidence of order effects in the responses.

Effort is costly, but referees may choose to exert positive effort due to the potential for reputational gain:

$$U = -c(e) + \eta e,$$

where  $c(e)$  is increasing and convex and  $\eta$  captures the salience of reputational gains. Each reviewer has a reservation utility below which they would reject the request; in these cases, the journal will rely on the signal and suggested revisions of an alternative reviewer. Critically,  $\eta$  may depend on the journal's policy regime, e.g., referee identifiability. The journal learns about the signal  $s$  and must form a posterior belief over  $q$ . If the journal is myopic about  $b(r, t)$ , it treats  $s$  as follows:

$$s = q + \varepsilon', \quad \text{with } \varepsilon' \sim \mathcal{N}(0, \sigma_s^2) \text{ and } \sigma_s^2 = \sigma_\varepsilon^2(e) + \sigma_{b(r,t)}^2$$

The journal's posterior belief about  $q$  is therefore:

$$\mathbb{E}[q \mid s] = \lambda s + (1 - \lambda)\bar{q}, \quad \text{where } \lambda = \frac{\sigma_q^2}{\sigma_q^2 + \sigma_s^2}.$$

The journal offers an R&R if  $\mathbb{E}[q^{\text{published}} \mid s] = \mathbb{E}[q \mid s] + \Delta(e)$  exceeds a publication threshold  $\tau$  (accounting for space constraints). Both Type I errors (accepted despite  $q^{\text{published}} < \tau$ ) and Type II errors (rejected despite  $q^{\text{published}} > \tau$ ) are possible. We use this framework throughout the article to discuss policy adoption.

## 4. Identifiability

Peer-review systems differ in terms of *identifiability*. Single-blind review, where authors are blinded to reviewer identities, is the norm in economics (Table 1), but other possibilities include double-blind review (blinding both authors and referees) and open identifiability (COPE Council, 2017). This section considers various modifications to the timing and conditions of identity revelation.

### 4.1. Author identifiability

#### *The case for double-blind review*

Single-blind review is currently ubiquitous in economics, but this has not always been the case. In fact, the *American Economic Review* (AER) and the *Quarterly Journal of Economics* (QJE) only abandoned double anonymity 10–20 years ago (Hengel, 2022). Furthermore, some prominent journals that publish research in economics or psychology continue to use it, including *Management Science*, *Economic Inquiry*, and *Psychological Bulletin* (Online Resource Table A.6).<sup>3</sup>

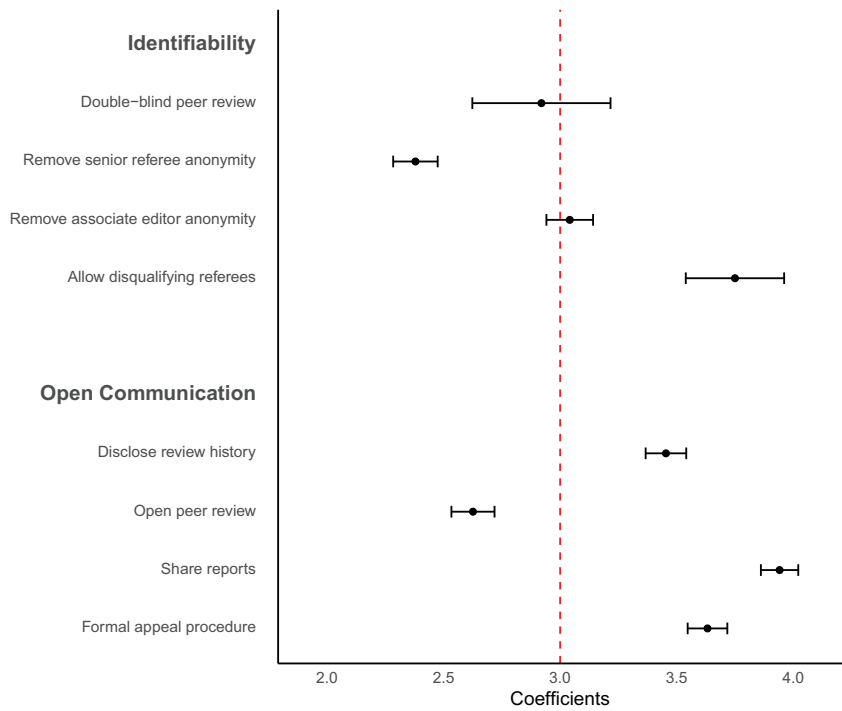
Double anonymity, if implemented effectively, protects against identity-based discrimination, i.e., reviewers favoring prominent authors due to status bias (Huber et al., 2022), taste-based discrimination, or statistical discrimination (inferring quality  $q$  from author reputation  $r$ ). As such, we expect heterogeneous effects: double anonymity should decrease the frequency of Type I errors for low- $q$  papers from high- $r$  authors and Type II errors for high- $q$  papers from low- $r$  authors.

#### *Scope for enabling double-blind review*

Our respondents' attitudes toward double-blind review are divided: 39% view it favorably (rating  $\geq 4$ ) and 46% unfavorably (rating  $\leq 2$ ).<sup>4</sup> The average rating is 2.92 out of 5 ( $p = 0.59$  against 3; see Fig. 1).

<sup>3</sup>Another possibility is a "triple-blind" system with blinded editors (Jung et al., 2016), as done by some philosophy journals like *Philosophical Review* (Online Resource Table A.6), but this is more difficult to implement and prevents editors from considering referee conflicts of interests.

<sup>4</sup>The question on double-blind review received only 112 responses, as it was introduced in a later survey wave.



**Fig. 1** Mean response relative to the midpoint value of 3  
*Notes:* This figure plots the means and 95% confidence intervals of the respondents' ratings (1–5) for each of the main eight policies we discuss in the paper. The number of observations is  $N = 802$  for all questions except those on double-blind peer review and disqualifying referees, which have  $N = 112$  respondents each.

Qualitative responses reveal concerns about reviewer objectivity, with “double” and “blind” appearing frequently in comments related to transparency (Online Resource Fig. B.6).<sup>5</sup>

These mixed views appear in line with existing empirical evidence on double-blind review, with some studies finding bias reduction and others reporting null effects (Blank, 1991; Laband & Piette, 1994; Snodgrass, 2006; Tomkins et al., 2017; Kolev et al., 2019; Ersoy & Pate, 2023). This heterogeneity could relate to field- and time-specific factors that affect the feasibility of anonymizing authors. Anonymity is likely difficult to preserve in small disciplines where authors disseminate preprints widely, like economics, especially in the age of search engines (Goldberg, 2012). Despite these concerns, Altonji et al. (2025) recommend that journals attempt to implement double-blind review, since it is fairly costless and ideal in principle. One possible implementation could be an opt-in model similar to *Nature* (Table 1) and *Nature Human Behaviour* (Online Resource Table A.6). While authors who opt in may signal low  $r$ , thereby undermining potential benefits, the evidence about this policy remains limited and experimentation could be valuable.

4.2. Referee and editor identifiability

*The case for open identifiability*

Instead of keeping key stakeholders anonymous, one could consider the opposite. In economics, the identities of referees and associate editors are generally concealed to authors and readers, while

<sup>5</sup> A sizable share of respondents (268/802, 33%) left at least one comment about peer-review policies, but those who select into commenting are unlikely to be representative. Out of all 506 comments, 113 (22%) discuss transparency-related policies (Online Resource Fig. B.6).

handling editors are known to the authors and sometimes publicly revealed (Table 1). However, referee identifiability is becoming increasingly common in other fields (Wolfram et al., 2020),<sup>6</sup> which could make an examination of this policy worthwhile.

Identifiability is likely to increase the scrutiny faced by referees and editors, as authors can link reports to identities, and the research community may view identified referees and editors as offering “seals of approval” to manuscripts. The implications of this heightened scrutiny are ambiguous. Bias could increase if referees become more vulnerable to retaliation, favor-swapping, and social pressure. If so, this would raise the likelihood of Type I errors for low- $q$  papers from high- $r$  authors. At the same time, greater scrutiny could enhance the salience of reputational benefits from reviewing  $\eta$  and thus increase effort  $e$ . This could improve review quality, yielding more precise signals  $\sigma_\varepsilon^2(e)$ , more useful revisions  $\Delta(e)$ , and fewer Type I and Type II errors. Low-quality reports are a common complaint under the status quo (Silbiger & Stubler, 2019).<sup>7</sup> Publicly identifying referees alongside manuscripts could also help referees earn credit for peer-review contributions in tenure, promotion, and hiring decisions, and allow the community to observe the Type I error rates of editors and referees. However, this visibility might make editors and referees more risk averse, leading them to demand more extensive revisions or to shy away from bold research. Finally, identifiability could trigger selection effects if some referees prefer to decline review requests rather than be named, potentially altering the taste  $t$  or precision  $\sigma_\varepsilon^2(e)$  of those who opt in.

### Scope for increasing identifiability

Respondents are generally skeptical about disclosing the identities of senior referees: only 23% consider it useful, while 60% do not. The average rating is 2.38 ( $p < 0.001$  against 3; see Fig. 1). Views on associate editor anonymity are more divided, with roughly equal shares finding disclosure useful (42%) vs. not useful (40%). The average rating is 3.04 ( $p = 0.42$  against 3). Overall, attitudes toward referee identifiability are more skeptical than those reported in other surveys (Publishing Research Consortium, 2016; Ross-Hellauer et al., 2017).

The empirical evidence on the effects of identifiability is generally mixed. In terms of bias, McNutt et al. (1990), Walsh et al. (2000), and Fox (2021) find that identified or voluntary signers tend to give more favorable recommendations, though van Rooyen et al. (1999) find no differences. The impact on report quality is similarly inconclusive: Walsh et al. (2000) find that identifiability increases effort (time spent) as well as the politeness and quality of reports, but van Rooyen et al. (1999) and Vinther et al. (2012) find no significant effects. Non-experimental evidence from McNutt et al. (1990) suggests that voluntarily signed reports are more constructive, courteous, and fair. Finally, van Rooyen et al. (1999) report that referees are less likely to accept review invitations under open identifiability, and Fox (2021) finds significant (though non-causal) differences in the demographic characteristics of signers versus non-signers.

### Alternative or complementary approaches

Given respondents' negative views on full identifiability, we considered alternative policies. One option is to allow referees to opt in to revealing their identity, as practiced by *Nature* (Table 1) and *Nature Human Behaviour* (Online Resource Table A.6). Referees could use this as a commitment device to produce higher-quality reports and it might help shift norms over time, though concerns about favor trading would likely persist.

<sup>6</sup>Prominent adopters include *The BMJ* and journals in the *Frontiers* family, including *Frontiers Behavioral Economics* and *Frontiers Environmental Economics* under certain conditions (Table 1).

<sup>7</sup>Our survey respondents are moderately satisfied with the reports they receive: on average, they judge 42% to be high quality and 27% to be low quality. Unsatisfactory reports most often contain “inaccurate statements about what the paper does” (76% of respondents; Online Resource Fig. B.1), “vague and unconstructive comments” (67%), and either “an aggressive tone” or “personal insults” (42% select at least one of these).



Another option is for journals to periodically acknowledge the contributions of their referees (Tennant et al., 2017), as some economics journals already do (Table 1). This would raise the visibility of refereeing and allow for the assessment of the diversity of referee pools. A variation of this policy, in which referee turnaround times were also published, was tested experimentally by Chetty et al. (2014). The main effect, relative to a no-visibility group, was to lower the acceptance rate of review requests.

We also consider whether exposing unrecorded conflicts of interest, a purported benefit of identifiability (Benos et al., 2007), can be achieved without piercing anonymity. Conflicts can create either positive or negative bias, depending on whether they stem from personal relationships, competing papers, or ideological differences. Regarding positive bias, network proximity and social ties are predictive of positive recommendations and decisions (Brogaard et al., 2014; Colussi, 2018; Carrell et al., 2024). While this may partly reflect other factors, such as lower search costs, fairness concerns remain (Altonji et al., 2025). Among our respondents, 16% reported rejecting or considering rejecting a referee request due to a conflict over the prior two years. We also asked them about “people refereeing papers by co-authors or friends.” Most answered either “this should never happen” (40%) or “this should happen as little as possible but cannot be avoided sometimes” (44%). The latter answer may reflect a potential trade-off between bias  $b(r, t)$  and variance  $\sigma_\epsilon^2(e)$ : the most knowledgeable referees are often connected to authors working in the same area. Altonji et al. (2025) propose several policies to help navigate this trade-off, including requiring reviewers and authors to disclose potential conflicts and offering clearer guidelines about what constitutes a conflict. Nine of the ten journals listed in Table 1 have conflict policies for editors, while six have them for referees.

To minimize negative bias from conflicts, one idea is to give authors the option to proactively oppose the appointment of certain referees. This idea has broad support: 61% of respondents report favorable attitudes, 12% report unfavorable ones, and the average response is 3.75 ( $p < 0.001$  against 3, Fig. 1). Attitudes are similarly favorable in Altonji et al. (2025). While authors may attempt to use this policy strategically (Moore et al., 2011; Teixeira da Silva & Alkhatib, 2018), the option would be available to everyone. Moreover, editors may find complaints about certain referees more credible when they are communicated before manuscript decisions, rather than ex-post. This option exists at neither the ESA journals nor the top-5 journals (Table 1).

## 5. Open communication

We now examine two other dimensions of peer review: the *publication* of peer-review documents and the *mediation* of interactions between stakeholders (COPE Council, 2017). COPE identifies three publication policies: keeping reports confidential, publishing unsigned reports, and publishing signed reports. For mediation, they distinguish between editor-mediated interactions, open interaction between reviewers, and open interaction between reviewers and authors. Unpublished reports and editor-mediated interactions are uniformly embraced by the top-5 and ESA journals (Table 1). We consider policies that deviate from these norms.

### 5.1. Publication of peer-review documents and metadata

#### *The case for publishing peer-review data*

Most journals keep referee reports and decision letters confidential by default, as sharing them would likely require the informed consent of authors, referees, and editors (COPE Council, 2017). Yet, some exceptions outside economics exist, including *The BMJ* and *Nature* (Table 1). Journals are also increasingly sharing statistics and metadata, such as average turnaround times.

Publishing peer-review documents is likely to increase the scrutiny faced by reviewers and editors because such materials often contain identifying elements (even if unsigned) and because low-quality reports and letters can harm the reputation of journals. Fully open peer review, i.e., publishing signed



reports, would amplify this further by removing any doubts about identities. As discussed previously, the net effect of enhanced scrutiny depends on the relative strength of the effort and bias channels.

Making peer-review documents public (with author responses and prior manuscript versions) could also help readers assess a paper's intrinsic quality  $q$ , the informativeness of referee feedback (reflected in  $\sigma_e^2(e)$  and  $b(r, t)$ ), and the added value of revisions  $\Delta(e)$ . Beyond individual papers, such documents could shed light on how peer review shapes research by providing stronger evidence on whether  $\Delta(e)$  varies with the number of revision rounds or the nature of revision requests (Malički et al., 2022). This may be particularly relevant in experimental economics, where requests for additional experiments can be especially costly.

### *Scope for increasing data sharing*

Our respondents are generally supportive of sharing anonymized referee reports and author responses: 55% hold favorable views, 22% unfavorable views, and the average rating is 3.45 ( $p < 0.001$  against 3; Fig. 1). However, respondents appear less convinced of the impact of this policy on review quality: only 38% expect it to be useful, while 41% expect otherwise; the average rating is 2.94 ( $p = 0.21$  against 3). The divergence between favorability and perceived usefulness – moderate at the individual level (correlation  $\rho = 0.59$  between responses) – may reflect the view that publishing these documents is informative but unlikely to change reviewer behavior.

To go one step further, we also measured attitudes toward fully open peer review. Respondents were fairly skeptical, with 50% holding unfavorable attitudes, 27% holding favorable attitudes, and an average response of 2.63 ( $p < 0.001$  against 3, Fig. 1). Interestingly, attitudes are less positive than those toward publishing unsigned reports, but more positive than toward referee identifiability with confidential reports.

Our respondents' views appear more cautious than those expressed in a cross-disciplinary survey on open peer review (Ross-Hellauer et al., 2017). Cautious attitudes could be expected given the very limited evidence available on the impact of such policies. The two main studies on the topic are Bravo et al. (2019), who observationally examine how publishing unsigned referee reports affects behavior at five journals, and van Rooyen et al. (2010), who randomly informed some referees that their signed reviews might be published online. Bravo et al. (2019) find no overall effect on key outcomes, including willingness to review, time spent, report quality (proxied by objective language), or referee recommendations. However, younger and non-academic referees increased their willingness to review, showed greater objectivity, and gave more positive recommendations, suggesting that less influential referees may be more sensitive to scrutiny. Results differ somewhat in van Rooyen et al. (2010), where treated referees spent more time on reports and produced higher-quality reviews for accepted papers, though they were no more likely to recommend acceptance. Suggestively, Zong et al. (2020) find that papers with voluntarily disclosed review histories are cited more often.

### *Alternative or complementary approaches*

Short of sharing full review histories, journals could publish summaries of how manuscripts changed from submission to publication, as done by *F1000Research* (Online Resource Table A.6). This would provide standardized, low-burden evidence on the effects of peer review. As proposed by Altonji et al. (2025), journals could also standardize how they report metadata, such as turnaround times and acceptance rates, to enable comparisons across journals. More granular information, such as submission dates or the number of revision rounds, could further enhance transparency, promote accountability, and help authors make more informed submission decisions without compromising confidentiality. Finally, Altonji et al. (2025) suggest that journals could make author contribution statements mandatory for papers with many authors, as done in other disciplines; this would allow all readers, including those on tenure, hiring, and promotion committees, to understand the distribution of work and the contributions of specific co-authors.

## 5.2. Open channels of communication

### *The case for opening communication*

Interactions between stakeholders during the review process in economics tend to be limited and asynchronous. In particular, editors interact directly with authors and reviewers in sequential stages and mediate interactions between these stakeholders via referee reports and author responses. As such, under the status quo, the editor is responsible for aggregating the demands of reviewers and giving direction to authors. Such efforts appear essential, as a near majority of respondents (46%) complained about “unrealistic” and/or “inconsistent” demands from reviewers. However, many of our respondents also expressed discontent regarding the extent of editorial direction (Charness et al., 2022) and the word “editor” appeared frequently in their comments, as shown in Online Resource Fig. B.5. In principle, a more directed and dynamic review process could allow stakeholders to discuss and resolve these issues promptly. This could also substantially speed up turnaround times, especially if these discussions replace multi-month revision rounds, an issue that distinguishes economics from other disciplines (Hamermesh, 1994; Hadavand et al., 2024). These benefits must be balanced against communication costs, as a more interactive review process might require more intense effort from editors and referees at specific times. In addition, increasing the volume of interactions makes de-anonymization more likely and could create opportunities for well-connected authors to apply pressure on reviewers.

### *Scope for increasing communication channels*

The most ambitious version of increased communication would allow real-time interactions between authors and referees, as practiced by *Frontiers* (Table 1). Very few journals currently use this format, and empirical evidence on mediation policies remains limited (Ross-Hellauer et al., 2017), making it unlikely to be adopted in disciplines with more conservative peer-review practices, such as economics.

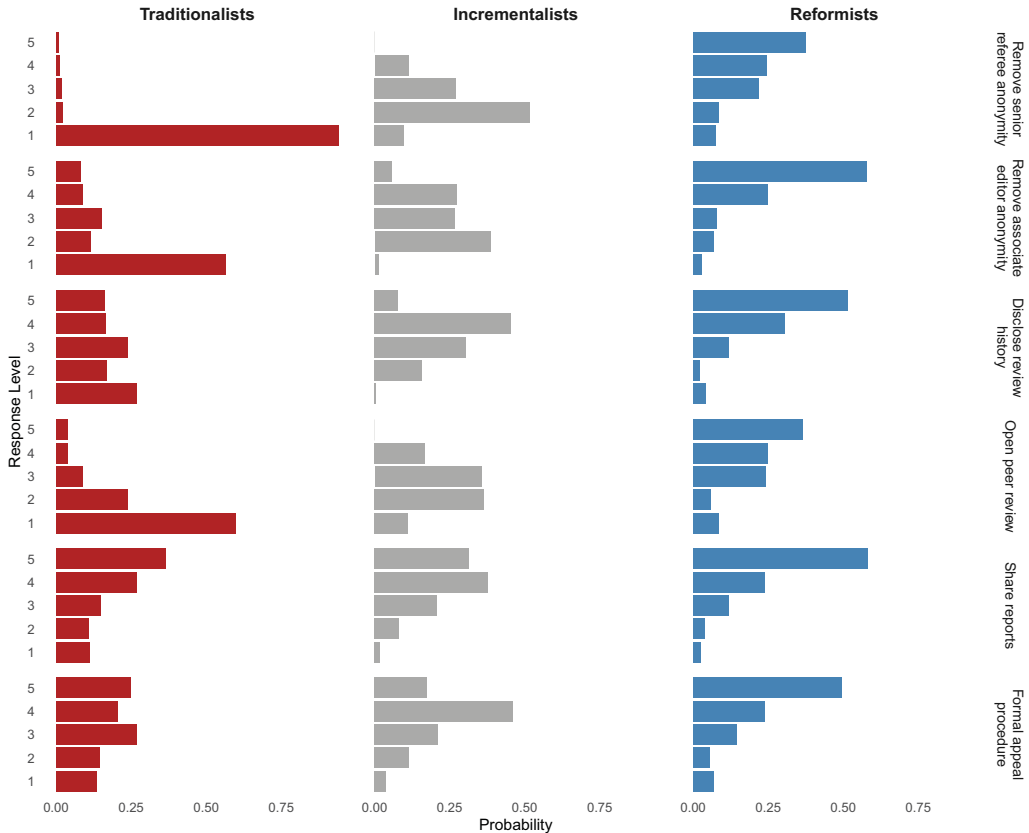
Short of a fully interactive system, journals could open new channels between selected stakeholders. One option is to establish an official procedure for authors to respond to editors and referees after receiving a rejection. This would allow authors to plead their case without imposing significant additional work on editors or referees in most cases. Formal appeals are standard in computer science conferences (Shah, 2022) but are not always available in economics (Table 1). Anecdotal evidence suggests that informal appeals occur in practice; if such appeals disproportionately favor well-connected authors, a formal procedure could help create a more level playing field. We asked respondents for their views on such a policy: over 60% were favorable, 19% were unfavorable, and the average response was 3.63 ( $p < 0.001$  against 3; Fig. 1). Support was similar among editors ( $-0.01$ ,  $p = 0.85$ ) but somewhat lower among active reviewers ( $-0.10$ ,  $p < 0.001$ ).

Another option is to ask authors to submit a revision proposal to editors and referees before proceeding with an R&R. This could clarify key issues before costly revisions are undertaken and reduce wasted effort on revising manuscripts that are ultimately rejected. The downside is the additional upfront work required.

Finally, journals could share reports between reviewers more systematically. This policy is popular among respondents: 71% find it useful and 13% do not, with an average response of 3.9 ( $p < 0.001$  against 3; Fig. 1). This is typically done *ex post* but could also occur at an intermediate stage to help referees with their evaluations. Some editors, however, may prefer to preserve signal independence (Lorenz et al., 2011).

## 6. Evaluating policies jointly

The previous sections examined support for each policy separately. To identify possible complementarities, we now consider whether support for transparency in one dimension predicts support in others.

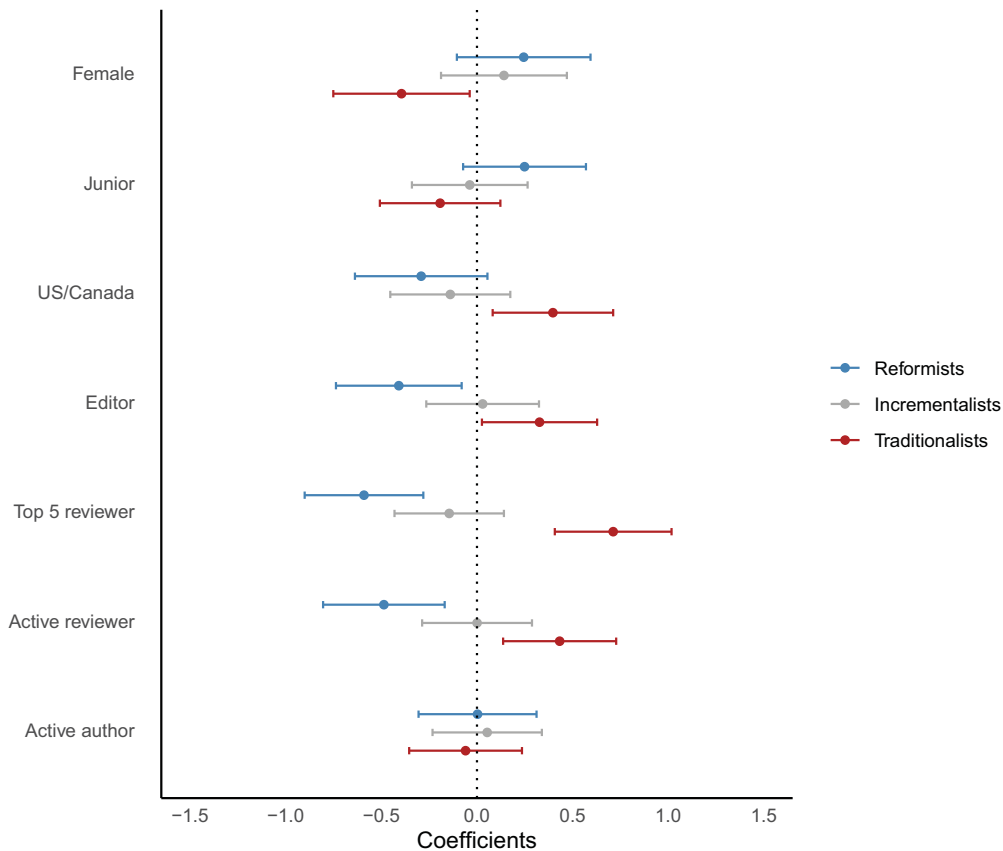


**Fig. 2** LCA classification of responses

*Notes:* This figure shows the results of an LCA, which aims to identify clusters of responses to six survey questions ( $N = 802$  for each question). Responses to the questions on “disqualifying referees” and “double-blind peer review” (both  $N = 112$ ) are excluded since the LCA requires non-missing values for all respondents and questions. The number of classes was chosen to minimize the Bayesian information criterion. There are 18 subplots in the figure, each of which shows the probability (bottom axis) that an individual in that class (top axis) would give a specific response (left axis) to a given question (right axis). Our chosen names for each class can be seen on the top axis.

Correlations between respondents’ support for different policies are generally weak, with most below 0.30 (Online Resource Fig. B.2). The strongest is between de-anonymizing associate editors and senior referees ( $\rho = 0.68$ ). Moderate correlations also appear between support for fully open peer review and both de-anonymizing senior referees ( $\rho = 0.58$ ) and disclosing review histories ( $\rho = 0.54$ ). No other correlation exceeds  $\rho = 0.50$ .

Since correlations only capture linear, pairwise relationships, we also use LCA to identify distinct respondent types with similar attitudes. Using the six questions answered by the full sample, we group respondents based on their rating patterns, selecting the number of classes using the Bayesian Information Criterion. As shown in Fig. 2, the LCA identifies three groups differing in their openness to reform. The first group (red), which we label *Traditionalists*, generally favors the status quo and holds the most negative views on open peer review. They have the highest likelihood of giving each policy a rating of 1. At the other end, the third group (blue), the *Reformists*, shows the highest probability of rating each policy as 5. In between, the second group (gray), the *Incrementalists*, gives responses closer to 3 for each question. No group forms a majority: 38% are Incrementalists, 33% are Traditionalists, and 28% are Reformists (Online Resource Fig. B.3).



**Fig. 3** Demographic predictors of LCA classification  
*Notes:* This figure plots point estimates and 95% confidence intervals for potential demographic predictors of class membership. Each coefficient was obtained from a separate linear probability model, regressing class membership on the listed demographic characteristics (left axis).

To explore potential heterogeneity, we regressed indicators for class membership on a range of demographic characteristics (Fig. 3).<sup>8</sup> Respondents from the US/Canada, editors, top-5 reviewers, and active reviewers, possible “gatekeepers” of the system, are more likely to be classified as Traditionalists, while female respondents are less likely. Editors, top-5 reviewers, and active reviewers are less likely to be classified as Reformists. No other significant differences emerge.<sup>9</sup> This classification exercise reveals that the lack of consensus on many policies may reflect the existence of distinct factions in the transparency debate.

Our heterogeneity findings align with recent surveys suggesting that seniority is negatively associated with support for certain open science practices, such as pre-analysis plans (Ross-Hellauer et al., 2017; Logg & Dorison, 2021; Spitzer & Mueller, 2023; Imai et al., 2025). For example, Imai et al. (2025) find that full professors are less favorable toward pre-analysis plans, policies recommending or mandating their use, registered reports, and results-blind review. By contrast, Ferguson et al. (2023) find no difference in support for open science practices between published authors and PhD students.

<sup>8</sup>For heterogeneity analysis at the individual policy level, see Online Resource Table B.1.  
<sup>9</sup>The share of respondents within each class who belong to each demographic group is shown in Online Resource Fig. B.4.

## 7. Looking forward

In this article, we investigated our respondents' attitudes toward transparency reforms in peer review. Overall, we find some degree of openness toward making changes, but views differ based on the exact policy considered and respondent characteristics. Starting with identifiability, there is no clear majority in favor of changing the status quo for senior referees, associate editors, or authors, though opinions about the latter two are more ambivalent. In contrast, there is strong support for allowing authors to oppose the assignment of specific reviewers. On open communication, respondents express broad support for publishing unsigned referee reports and author responses, sharing reports among reviewers, and establishing a formal appeal process, while remaining skeptical of publishing signed reports. Looking across policies, we identify three respondent types: *Traditionalists* (33%), *Incrementalists* (38%), and *Reformists* (28%). Membership in the Traditionalist class is somewhat predicted by seniority (editor status) and network centrality (US/Canada, top-5 reviewer) and negatively predicted by whether a respondent is female.

Our findings come with three key policy implications. First, it is likely that a handful of policies are already possible to implement globally with minimal risk of backlash, including sharing reports between reviewers and allowing author appeals. Second, given possible heterogeneous effects and divided opinions about policies like anonymity for authors, some degree of horizontal differentiation between journals may make sense. For example, *Econometrica* recently opened an additional channel for manuscript readers to engage in post-publication discussion about papers (Online Resource Table A.6). Policies like these allow stakeholders to sort into journals whose transparency norms are consistent with their preferences and create data that helps other journals make informed decisions about policy adoption. Third, given that respondent views differ substantially based on demographic characteristics, recruiting diverse editorial boards could help ensure that policy adoption decisions are made with input from multiple perspectives.

Our findings also come with caveats. First, even through the lens of the single objective function that we examine, most policies have ambiguous effects on the quality of peer review due to competing forces. Furthermore, different journals likely have different objective functions, with some placing more weight on the experiences of authors (e.g., ensuring fast and fair decisions) and others prioritizing the interests of referees, editors, and readers. These factors may help explain the lack of consensus among our respondents on some policies. Second, while we discuss a broad range of policies, we necessarily do not cover every topic. Certain omitted topics, like the role of large language models in peer review, have received a great deal of attention recently (Altonji et al., 2025; Rao et al., 2025). Third, among the policies that we do consider, our survey results focus on measuring attitudes rather than the determinants of these attitudes. Relatedly, our elicitations are unincentivized, subject to measurement error, and might depend on whether respondents answer from their perspective as an author, referee, reader, or editor.<sup>10</sup> Finally, we acknowledge limitations in how actionable our findings are, given the minimal causal evidence that exists to assist journals with policy adoption. To address this, we encourage journals and other institutions to run their own experiments and surveys in a systematic manner, as outlined by Bendiscioli et al. (2022).

This further research could take many directions. In particular, future surveys could investigate how beliefs about policy impact map into overall policy support, measure support for alternative policies, and track how views evolve over time. Additional experimental studies, particularly within economics, would be useful for filling the current gap in causal evidence. Since real-world trials are difficult to implement, lab or online experiments that simulate peer review could serve as a first step. Studies of opt-in policies could also shed light on take-up rates and help shift norms over time.

<sup>10</sup>For what it is worth, regarding our unincentivized elicitation, some papers suggest that the effect of incentives on respondent behavior is minimal, e.g., Enke et al. (2023). Regarding framing effects, we find no evidence that policy ratings differ based on whether the author or reviewer question block appeared first.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/esa.2026.10030>.

**Data availability statement.** The replication material for the study is available at <https://doi.org/10.17605/OSF.IO/GJ586>.

**Acknowledgements.** This article represents an adaptation of the transparency section of our earlier [report](#) on peer review in economics. The analyses presented in this paper do not follow a pre-analysis plan and, as such, are exploratory. We would like to thank all of the associations and networks that helped us advertise the survey to their members, including CEPR, CESifo, the ESA, and the EEA. Our appreciation also goes out to the editors at various journals for the statistics provided and to Peter Andre and Fulya Ersoy for the statistics and datasets they shared with us. This project benefited from the wonderful research assistance of Raffaele Blasone, Ingrid Löfman, Jessica Milligan, Rafael Suchy, Alena Wabitsch, and Le Wu.

**Funding statement.** Anna Dreber acknowledges funding support from the Jan Wallander and Tom Hedelius Foundation (Svenska Handelsbankens Forskningsstiftelser), the Knut and Alice Wallenberg Foundation, the Marianne and Marcus Wallenberg Foundation (in the form of a Wallenberg Scholar grant awarded to AD), and the Austrian Science Fund (FWF, SFB F63). Daniel Evans acknowledges funding support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/2 – 390838866. Adam Gill acknowledges funding support from the Jan Wallander and Tom Hedelius Foundation (H22-0007). Séverine Toussaert acknowledges funding support from the Economic and Social Research Council (ES/V003461/1).

**Ethical standards.** The survey received ethics approval from the University of Oxford (Approval # ECONCIA21-21-20).

## References

- Altonji, J., Imbens, G., Lang K., Luttmer, E., Rasul, I., Stantcheva, S., & Wacziarg, R. (2025). *Report on improving the publication process in economics* (Technical Report). Ad-hoc Joint AEA-EEA-ES-RES Committee.
- Andre, P., & Falk, A. (2021). *What's worth knowing? Economists' opinions about economics* (Technical Report IZA Discussion Paper No. 14527). IZA Institute of Labor Economics.
- Bendiscoli, S., Firpo, T., Bravo-Biosca A., Czibor E., Garfinkel M., Stafford T., Wilsdon J., & Buckley Woods H. (2022). *The experimental research funder's handbook* (Report).
- Benos, D. J., Bashari, E., Chaves, J. M., Gaggari, A., Kapoor, N., LaFrance, M., Mans, R., Mayhew, D., McGowan, S., Polter, A., Qadri, Y., Sarfare, S., Schultz, K., Splittgerber, R., Stephenson, J., Tower, C., Walton, R. C., & Zotov, A. (2007). The ups and downs of peer review. *Advances in Physiology Education*, 31(2), 145–152.
- Blank, R. M. (1991). The effects of double-blind versus single-blind reviewing: Experimental evidence from the American Economic Review. *American Economic Review*, 81(5), 1041–1067.
- Bravo, G., Grimaldo, F., Lopez-Inesta, E., Mehmani, B., & Squazzoni, F. (2019). The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature Communications*, 10(1), 322.
- Brogaard, J., Engelberg, J., & Parsons, C. A. (2014). Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, 111(1), 251–270.
- Card, D., & Della Vigna, S. (2020). What do editors maximize? Evidence from four economics journals. *The Review of Economics and Statistics*, 102(1), 195–217.
- Carrell, S., Figlio, D., & Lusher, L. (2024). Clubs and networks in economics reviewing. *Journal of Political Economy*, 132(9), 2999–3024.
- Charness, G., Dreber, A., Evans, D., Gill, A., & Toussaert, S. (2022). *Improving peer review in economics: Stocktaking and proposals* (Technical Report).
- Chetty, R., Saez, E., & Sandor, L. (2014). What policies increase prosocial behavior? An experiment with referees at the Journal of Public Economics. *Journal of Economic Perspectives*, 28(3), 169–88.
- Colussi, T. (2018). Social ties in academia: A friend is a treasure. *The Review of Economics and Statistics*, 100(1), 45–50.
- COPE Council. (2017). Ethical guidelines for peer reviewers. <https://publicationethics.org/guidance/guideline/ethical-guidelines-peer-reviewers>.
- COPE Council. (2017). Who 'owns' peer reviews? <https://publicationethics.org/guidance/discussion-document/who-owns-peer-reviews>.
- Enke, B., Gneezy, U., Hall, B., Martin, D., Nelidov, V., Offerman, T., & van de Ven, J. (2023). Cognitive biases: Mistakes or missing stakes? *The Review of Economics and Statistics*, 105(4), 818–832.
- Ersoy, F. Y., & Pate, J. (2023). Invisible hurdles: Gender and institutional differences in the evaluation of economics papers. *Economic Inquiry*, 61(4), 777–797.
- Ferguson, J., Littman, R., Christensen, G., Paluck, E. L., Swanson, N., Wang, Z., Miguel, E., Birke, D., & Pezzuto, J. -H. (2023). Survey of open science practices and attitudes in the social sciences. *Nature Communications*, 14(1), 5401.

- Fox, C. W. (2021). Which peer reviewers voluntarily reveal their identity to authors? Insights into the consequences of open-identities peer review. *Proceedings of the Royal Society B: Biological Sciences*, 288(1961), 20211399.
- Goldberg, P. K. (2012). Report of the Editor: American Economic Review. *American Economic Review*, 102(3), 653–65.
- Hadavand, A., Hamermesh, D. S., & Wilson, W. W. (2024). Publishing economics: How slow? Why slow? Is slow productive? How to fix slow? *Journal of Economic Literature*, 62(1), 269–93.
- Hamermesh, D. S. (1994). Facts and myths about refereeing. *Journal of Economic Perspectives*, 8(1), 153–163.
- Hengel, E. (2022). Publishing while female: Are women held to higher standards? Evidence from peer review. *The Economic Journal*, 132(648), 2951–2991.
- Huber, J., Inoua, S., Kerschbamer, R., König-Kersting, C., Palan, S., & Smith, V. L. (2022). Nobel and novice: Author prominence affects peer review. *Proceedings of the National Academy of Sciences*, 119(41), e2205779119.
- Imai, T., Toussaert, S., Baillon, A., Ertaç, S., Johannesson, M., Neyse, L., & Villeval, M.-C. (2025). *Pre-registration and pre-analysis plans in experimental economics* (No. 220). I4R Discussion Paper Series.
- Jung, J., Kim, J.-I., & Yoon, J. W. (2016). A practical approach to constructing triple-blind review process with maximal anonymity and fairness. In *International workshop on information security applications* (pp. 198–209). Springer International Publishing.
- Kolev, J., Fuentes-Medel, Y., & Murray, F. (2019). *Is blinded review enough? How gendered outcomes arise even under anonymous evaluation* (No. w25759). National Bureau of Economic Research.
- Laband, D. N., & Piette, M. J. (1994). Does the “blindness” of peer review influence manuscript selection efficiency? *Southern Economic Journal*, 60(4), 896.
- Logg, J. M., & Dorison, C. A. (2021). Pre-registration: Weighing costs and benefits for researchers. *Organizational Behavior and Human Decision Processes*, 167, 18–27.
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22), 9020–9025.
- Malički, M., Jerončić A., ter Riet G., Bouter L., Ioannidis J. P. A., Aalbersberg I. J. and Goodman S. N. (2022, September 14–16). Elucidating the effects of peer review: A living synthesis of studies on manuscript changes. In *PUBMET2022: The 9th Conference on Scholarly Communication in the Context of Open Science*, University of Zadar, Croatia.
- McNutt, R. A., Evans, A. T., Fletcher, R. H., & Fletcher, S. W. (1990). The effects of blinding on the quality of peer review: A randomized trial. *JAMA*, 263(10), 1371–1376.
- Moore, J. L., Neilson, E., & Siegel, V. (2011). Effect of recommendations from reviewers suggested or excluded by authors. *Journal of the American Society of Nephrology: JASN*, 22(9), 1598–602.
- Nature. (2019). Three-year trial shows support for recognizing peer reviewers. *Nature*, 568(7752), 275.
- Nature. (2025). Transparent peer review to be extended to all of Nature’s research papers. *Nature*, 642(8068), 542.
- Nature Communications. (2015). Transparent peer review at Nature Communications. *Nature Communications*, 6(1), 10277.
- Nature Communications. (2016). Transparent peer review one year on. *Nature Communications*, 7(1), 13626.
- Nature Communications. (2022). Transparent peer review for all. *Nature Communications*, 13(1), 6173.
- Publishing Research Consortium. (2016). Peer review survey 2015, Technical Report, Publishing Research Consortium May.
- Rao, V., Kumar A., Lakkaraju H., & Shah N. B. (2025). Detecting LLM-generated peer reviews. *PLoS One*, 20(9), e0331871.
- Ross-Hellauer, T., Deppe, A., & Schmidt, B. (2017). Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers. *PLoS One*, 12(12), e0189311.
- Shah, N. B. (2022). Challenges, experiments, and computational solutions in peer review. *Communications of the ACM*, 65(6), 76–87.
- Silbiger, N. J., & Stubler, A. D. (2019). Unprofessional peer reviews disproportionately harm underrepresented groups in STEM. *PeerJ*, 7, e8247.
- Snodgrass, R. (2006). Single- versus double-blind reviewing: an analysis of the literature. *SIGMOD Record*, 35(3), 8–21.
- Spitzer, L., & Mueller, S. (2023). Registered report: Survey on attitudes and experiences regarding preregistration in psychological research. *PLOS ONE*, 18(3), e0281086.
- Teixeira da Silva, J., & Alkhatib, A. (2018). Should authors be requested to suggest peer reviewers? *Science and Engineering Ethics*, 24(1), 275–285.
- Tennant, J., Dugan, J., Grazier, D., Jacques, D., Waldner, F., Mietchen, D., Elkhatib, Y., Collister, L., Pikas, C., Crick, T., Masuzzo, P., Caravaggi, A., Berg, D., Niemeyer, K., Ross-Hellauer, T., Mannheimer, S., Rigling, L., Katz, D. S., Tzovaras, B., & Colomb, J. (2017). A multi-disciplinary perspective on emergent and future innovations in peer review. *F1000Research*, 6, 111151.
- Tomkins, A., Zhang, M., & Heavlin, W. D. (2017). Reviewer bias in single-versus double-blind peer review. *Proceedings of the National Academy of Sciences*, 114(48), 12708–12713.
- van Rooyen, S., Godlee, F., Evans, S., Black, N., & Smith, R. (1999). Effect of open peer review on quality of reviews and on reviewers’ recommendations: A randomised trial. *BMJ*, 318(7175), 23–27.
- van Rooyen, S., and Delamothe, T., & Evans, S. J. W. (2010). Effect on peer review of telling reviewers that their signed reviews might be posted on the web: Randomised controlled trial. *BMJ*, 341, 7782.



- Vinther, S., Nielsen, O. H., Rosenberg, J., Keiding, N., & Schroeder, T. V. (2012). Same review quality in open versus blinded peer review in “Ugeskrift for Læger”. *Dan Med J*, 59(8), A4479.
- Walsh, E., Rooney, M., Appleby, L., & Wilkinson, G. (2000). Open peer review: A randomised controlled trial. *The British Journal of Psychiatry*, 176(1), 47–51.
- Wolfram, D., Wang, P., Hembree, A., & Park, H. (2020). Open peer review: promoting transparency in open science. *Scientometrics*, 125(2), 1–19.
- Zong, Q., Xie, Y., & Liang, J. (2020). Does open peer review improve citation count? Evidence from a propensity score matching analysis of PeerJ. *Scientometrics*, 125(1), 607–623.