A Data discussions

This appendix is devoted to explaining aspects of our data collection and sample in greater detail and presenting additional figures and tables.

A.1 Survey recruitment

To increase participation in the survey, we conducted multiple stages of recruitment within different communities of academic economists. A timeline and summary of these efforts can be found in Table A.1. The first part (Wave 1) lasted from July 2020 to October 2020. In this phase, we targeted groups of experimental/behavioral economists. An initial pilot study was sent to a select group in order to gather feedback on the content and structure of the survey. After that, a link to the survey was posted on the Economic Science Association (ESA) discussion forum. Next, we created a database of email addresses of experimental/behavioral economists in order to reach a larger group of potential respondents (mail merge 1). Email addresses in the database came from a wide variety of sources, including lists of experimental/behavioral economists on the RePEc database and participation lists from various conferences and seminars, i.e., the ESA conference, the Early-Career Behavioral Economics Conference (ECBE), and the Stanford Institute for Theoretical Economics conference (SITE). A few emails were also added individually. The entire database was sent an initial email in early August, which was followed by a reminder email in late September, sent only to those who had not provided their email address after taking the survey (i.e., in order to be considered for the prize drawing). We sent emails out to 1,802 researchers, from whom we received 606 complete or partially complete responses (34%). A separate but similar email was also sent to participants in the Virtual East Asia Experimental and Behavioral Seminar (VEAEBES).

The second part (Wave 2) was conducted from November 2020 to January 2021. In this stage, we shifted our efforts to recruiting economists from outside of experimental/behavioral economics. We did this by reaching out to communities of economists that are not specific to any subfield, as well as conducting efforts targeted at some particular subfields. Posts on the European Economic Association (EEA) website and Twitter feed advertised the survey to general groups of economists. Emails sent to the CESifo and CEPR networks targeted similarly varied groups. Subfield-specific outreach efforts included a post on the Decision Theory (DT) forum and emails sent to Health Economics at Lancaster (HEAL) seminar series members. Finally, we constructed another database of email addresses targeted at non-experimental/behavioral economists, with a particular emphasis on reaching out to underrepresented fields like macroeconomics (mail merge 2). The database was partly constructed with participant lists of conferences hosted by various organizations, including the Society for Economic Dynamics (SED), the American Economic Association/Allied

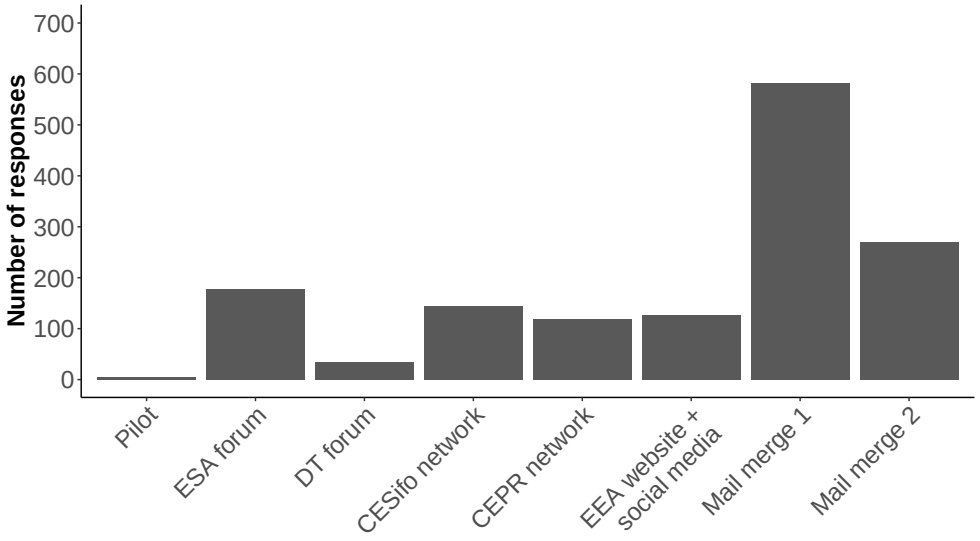
Social Science Associations (AEA/ASSA), the Society for Judgment and Decision Making (SJDM), and the Society for the Advancement of Economic Theory (SAET). We also included contact details collected from the NBER database and added some emails individually. The entire email database received a survey request in mid-December, followed by a reminder email in mid-January (once again to those we could not confirm took the survey). We sent emails out to 3,618 researchers, from whom we received 269 responses (7%). Informal recruitment efforts took place throughout the entire recruitment period, including via individual emails and social media posts sharing the survey link.

As mentioned elsewhere, 1,497 (1,459) individuals at least partially (fully) responded to our recruitment efforts. The median response time for fully completed surveys was 15.5 minutes. Most respondents were recruited from the two mail merges, from which we received 875 completed or partially completed responses (58% of our sample). Four other recruitment channels garnered $>$ 100 completed responses each (Figure A.1).[11] Characteristics of the sample broken down by recruitment channel can be found in Table A.2. Some clear demographic differences are worth noting. First, the CESifo and CEPR channels have particularly high percentages of respondents over 50 (44% and 39%, respectively), respondents who are full professors (57% and 59%), and respondents with editorial experience (50% and 56%). Additionally, the percentage of US/Canada-based respondents from mail merge 2 (63%) is much higher than the other channels, which feature more Europe-based researchers. Finally, while all channels have a fairly high percentage of researchers who have refereed for top-5 journals, this percentage is particularly high in the CEPR and mail merge 2 channels (84% and 71%, respectively).

After completing our primary analysis of the survey data, we sent out a follow-up survey in February 2022 ($N = 117$) to clarify our interpretation of the initial results and gather further evidence. Our additional inquiries included questions on what respondents consider to be reasonable report-writing activity, whether they feel pressured to write more referee reports due to publication concerns around their own manuscripts, and the percentage of their submissions from 2020-2021 that were desk rejected (28%). These results were included in our original report, but are excluded from the present article. Compared to the original survey, the follow-up survey sampled relatively few full (34%) and assistant (19%) professors, and relatively many postdoc/PhD candidates (19%). 33% of respondents reported that they are currently editors.

[11] A recruitment channel is based on the survey link used. Some survey links were used in multiple methods of recruitment; these are considered to be one recruitment channel.

Figure A.1: Distribution of responses across recruitment channels



Notes: This figure only includes respondents who fully completed the survey. $N = 1,459$.

Table A.1: Recruitment strategy summary

Wave	Date	Targeted group	Recruitment channel (survey link)
Wave 1 (Experimental/behavioral)	Early July 2020	Selected group for initial feedback	Pilot (individual emails)
	16 July 2020	ESA forum	ESA forum
	8 August 2020	RePEc/ESA/AEA-ASSA/ECBE/SITE	Mail merge 1
	18 September 2020	VEABES	Mail merge 1
Wave 2 (All fields)	13 November 2020	EEA	EEA website + social media
	15 November 2020	DT forum	DT forum
	24 November 2020	CESifo	CESifo network
	26 November 2020	CEPR	CEPR network
	7 December 2020	HEAL	EEA website + social media
	16 December 2020	NBER/SAET/SED/AEA-ASSA/SJDM	Mail merge 2
Informal outreach	Sporadically	Personal emails	EEA website + social media[a]
	Sporadically	Social media	EEA website + social media

Notes: See previous page for more information about the meaning of the various acronyms. ***a:*** Some personal emails were sent using different survey links, but the majority used the link for "EEA website + social media."

Table A.2: Characteristics across recruitment channels[a]

Characteristic	ESA forum	CESifo	CEPR	EEA/social media	Mail merge 1	Mail merge 2
Demographics						
Female	28%	18%	28%	29%	24%	26%
Age:						
Under 40	55%	30%	27%	51%	40%	49%
40–49	28%	26%	34%	36%	38%	26%
50–59	11%	26%	22%	10%	15%	16%
60 and over	6%	18%	17%	4%	7%	10%
Location:						
US/Canada	32%	27%	23%	29%	31%	63%
Europe	58%	69%	74%	61%	54%	31%
Asia/Oceania	7%	3%	3%	8%	14%	5%
Other countries	3%	1%	0%	2%	2%	2%
Position						
Full professor	23%	57%	59%	25%	39%	37%
Associate professor	26%	13%	22%	23%	26%	17%
Assistant professor	27%	16%	12%	29%	26%	34%
Postdoc/PhD candidate	18%	5%	0%	15%	5%	7%
Other position	7%	8%	8%	7%	4%	5%
Experience						
Editorial experience	29%	50%	56%	30%	38%	41%
Average number of publications[b]	21	40	27	25	28	20
Referee for top-5 journals	46%	55%	84%	43%	58%	71%
N[c]	190	145	119	128	606	269

Notes: ***a:*** This table only looks at the recruitment channels (i.e., survey links) that received > 100 partially- or fully-completed responses (= 97% of the total sample). ***b:*** Full range used (no top coding). ***c:*** These sample sizes refer to all partially- or fully-completed surveys for each recruitment channel. For the individual statistics, sample sizes may differ from the stated N as unanswered questions and "Prefer not to say" responses were removed from these calculations.

A.2 Sample descriptive statistics

Table A.3: Descriptive statistics for the main dimensions of heterogeneity

Variable	N	Share
Female	196	24%
Junior	268	33%
US/Canada	239	30%
Editor	297	37%
Top-5 reviewer	441	55%
Active reviewer	350	44%
Active author	367	46%
N	802	100%

Variable descriptions:

- `Female`: Binary variable = 1 if the respondent selected "Female," with the baseline being the combination of "Male" (71%) + "Prefer not to say" (4%).
- `Junior`: Binary variable = 1 if the respondent selected "PhD candidate" or "Post-doctoral researcher" or "Assistant professor" as their position.
- `US/Canada`: Binary variable = 1 if the respondent selected "United States" or "Canada" as the country in which their job is located.
- `Editor`: Binary variable = 1 if the respondent answered "Yes" to having held an editorial position.
- `Top-5 reviewer`: Binary variable = 1 if the respondent indicated that $> 0\%$ of their referee reports had been written for a top-5 journal.
- `Active reviewer`: Binary variable = 1 if the respondent wrote more reports annually than the median respondent in our sub-sample (> 8 reports).
- `Active author`: Binary variable = 1 if the respondent made more submissions over the designated timeframe than the median respondent in our sub-sample (> 6 submissions over two years).

A.3 Sample representativeness

Since there are no comprehensive statistics on the global population of economists, we assess potential selection bias by comparing the characteristics of our sample to statistics from relevant subpopulations and other surveys. We first benchmark our experimental/behavioral subsample against 2020 ESA members

Table A.4: Comparison between experimental/behavioral sample and ESA membership

Categories	**Our sample (%)**	**ESA 2020 members (%)**
Female	24%	38%
Postdoc/PhD	9%	26%
Europe	55%	40%
US/Canada	30%	35%
Asia/Oceania	12%	23%
Africa/Other Americas	2%	2%
N	802	994

Notes: The demographic characteristics of ESA 2020 members were obtained directly from the ESA.

(Table A.4) and our full sample against the weighted sample in Andre and Falk (2021) (Table A.5).[12] Relative to these groups, we somewhat under-sample women and over-sample researchers based in Europe. In addition, we significantly over-sample those in experimental/behavioral economics relative to Andre and Falk (2021). While the ESA provides a useful reference point, its membership may not fully represent the broader experimental/behavioral community. For additional context, we also compare our full sample with that from Altonji et al. (2025), though their survey includes limited demographic data; again, we observe an over-representation of European and experimental/behavioral researchers.

[12] Andre and Falk (2021) use a database containing nearly all active economists with EconLit publication data, weighted to reflect the underlying population.

Table A.5: Comparison between full sample and Andre & Falk (2021) benchmarks

Characteristic	Peer-review survey	EconLit population (AF)	Unweighted sample (AF)	Weighted sample (AF)
Demographics				
Female	24%	26%	23%	26%
Age:				
Under 40	43%	-	33%	35%
40-49	32%	-	33%	32%
50-59	16%	-	19%	18%
60 and over	9%	-	16%	16%
Location:				
US/Canada	35%	34%	24%	34%
Europe	54%	40%	54%	41%
Asia/Oceania	8%	21%	17%	21%
Other regions	2%	4%	5%	5%
Field (excl. Exp/Behav)[a]				
Microeconomics	27%	15%	18%	15%
Macroeconomics	13%	31%	24%	30%
Econometrics	9%	4%	3%	3%
Development	8%	7%	8%	8%
Labor	11%	9%	12%	10%
Industrial organization	6%	8%	7%	8%
Public economics	11%	4%	4%	4%
Other fields	15%	22%	23%	23%
Position				
Full professor	38%	-	41%	37%
Associate professor	22%	-	27%	28%
Assistant professor	26%	-	20%	22%
Postdoc/PhD candidate	8%	-	10%	10%
Other position	6%	-	2%	3%
Research output				
Average number of publications	25	17	18	16
N	1,459[b]	53,779	7,794	7,794

Notes: The statistics in the last three columns were either directly taken from Andre and Falk (2021) (AF) or derived for us by Peter Andre. ***a:*** For the field statistics, we removed the "Experimental/Behavioral" selection from our survey data to improve comparability with the Andre & Falk data for the other fields. As such, it is useful to remember that we report a conditional distribution. For example, the microeconomics category, consisting of respondents who selected either applied microeconomics, decision theory, game theory, or microeconomic theory as a field, accounts for 28% of all field selections that were not behavioral or experimental economics ($N = 2,668$ remaining selections). Experimental/Behavioral economics accounts for 33% of the total number of field selections (out of $N = 3,982$ selections made across all fields). ***b:*** The sample size reported for our peer-review survey is the number of completed surveys. However, for our statistics to be comparable with the data in AF, we needed to remove the respondents who selected "Prefer not to say" for `Age`, `Location`, and `Position`. This leads to smaller sample sizes for those variables (with $N = 1,381, 1,392,$ and $1,401$, respectively). For the `Female` statistic, responses of "Prefer not to say" were kept.

Below we elaborate on the definitions of certain variables and note any instances where variables had to be reformulated in order to ensure the comparability of our data with the external data.

- **Location:** From the Andre and Falk (2021) (AF) statistics, we combine the "Asia" and "Australia and

New Zealand" categories to create the "Asia/Oceania" category and we combine the "Latin America" and "Africa" categories to create the "Other regions" category.

- **Field of research:** The field categories from AF are based on the JEL codes:
  - Microeconomics = JEL D (Microeconomics)
  - Macroeconomics = JEL E (Macroeconomics and Monetary Economics) + JEL F (International Economics) + JEL G (Financial Economics)
  - Econometrics = JEL C (Mathematical and Quantitative Methods)
  - Development = JEL O (Growth and Development Economics)
  - Labor = JEL J (Labor and Demographic Economics)
  - Industrial Organization = JEL L (Industrial Organization)
  - Public Economics = JEL H (Public Economics)
  - Other fields = JEL Q (Agriculture and Environmental Economics) + Other fields
- **Position:** The "Full Professor" category combines the AF categories of "Professor" and "Emeritus"; the "Postdoc/PhD" category combines "Post-doc" and "Doctoral student"; and the "Other" category combines "Graduate student" and "Other."
- **Average number of publications:** The number of publications is capped at 200.

A.4 Construction of the journal transparency policies table

The abbreviations used in Table 1 are as follows:

- *Exp Econ* = *Experimental Economics*; *JESA* = *Journal of the Economic Science Association*; *AER* = *American Economic Review*; *ECMA* = *Econometrica*; *JPE* = *Journal of Political Economy*; *QJE* = *Quarterly Journal of Economics*; *REStud* = *Review of Economic Studies*

We made use of a wide range of data sources to collect the data presented in Table 1. We also made several subjective judgments about the best way to code the available information and data. In this section, we provide more information about the construction of this table:

- Our first judgment relates to which journals' transparency policies to include in the table. Due to the significant contribution of the Economic Science Association (ESA) and its membership to our survey results, we include the two journals of the ESA (i.e., *Experimental Economics* and *JESA*). Given their importance to the careers of economists (Heckman and Moktan, 2020), we also include each of the traditional "top-5" journals. For the other disciplines, we selectively chose three journals or journal families in order to highlight a number of unique or innovative policies for discussion.

- The policies in the **Identifiability** section of the table were obtained from a range of sources, which are listed in Table A.6.

    - The ***Blinded Review*** variable indicates whether the referees are blinded to the identity of the author during the review process ("Double") or not. Usually, when this is not the case, referee identities still remain obscured to authors ("Single"). However, in the case of *The BMJ*, author and referee identities are mutually known ("Open"). At *Nature*, authors can opt-in to double-blind review ("Opt-in"). Most journals mention their review policy in their guidelines. The ESA and top-5 journals are uniform in their application of single-blind refereeing, which we (the authors) can confirm through personal experience.

    - The ***Referee Identity Disclosed*** variable indicates whether the journal publicly discloses the identities of the referees who evaluated each manuscript after the review process is complete (✓) or not (–). At *Frontiers*, the identities of referees who recommend publication are disclosed after review is complete and conditional on acceptance. At *Nature*, referee identities are revealed and published alongside the manuscript post-review, as long as both the authors and referees agree to this. This information was obtained by reading journal policies and checking published manuscripts.

- The ***Handling Editor Disclosed*** variable indicates whether the journal publicly discloses the identity of the handling editor(s) who decided to accept each manuscript after the review process is complete (✓) or not (–). This information was obtained by inspecting a selection of articles on each journal website.
- The ***Journal Publishes List of Referees*** variable indicates whether the journal periodically publishes a full list of its referees ("Full"), a partial list ("Partial"), or no list at all ("No"). *QJE* and *REStud* limit their referee list to those who have met a certain threshold of quality/quantity contribution. These lists can usually be found in periodic journal reports or on journal websites.
- The ***Authors Can Nominate Referees*** and ***Authors Can Oppose Referees*** variables indicate whether the journal specifically allows or asks authors to make suggestions to include or exclude specific referees by name for each submission (✓) or not (–), respectively, as part of the regular submission process. This information was obtained by clicking through the submission portal of each journal.
- The ***Conflict-of-Interest Policy for Editors*** and ***Conflict-of-Interest Policy for Reviewers*** variables indicate whether the journal has public conflict-of-interest policies for editors and reviewers (✓) or not (–), respectively. Policies for editors usually state the conditions under which an editor should recuse themselves from handling a manuscript, while those for reviewers typically relate to reporting potential conflicts with the author for consideration by the editor. This information was obtained by reading journal policies.

- The statistics in the **Disclosure** section also come from a wide range of sources (Table A.6).
  - ***Turnaround Time Statistics*** and ***Acceptance/Rejection Statistics*** record whether each journal releases statistics on its turnaround time and its acceptance/rejection rates (✓) or not (–), respectively. This is usually done via a periodic report or on the journal website.
  - The ***Manuscript Received/Decision Dates*** variable indicates whether a journal discloses certain manuscript metadata alongside publications (✓) or not (–), in particular the date that the manuscript was received, the date of first response, the dates that subsequent versions were returned, and/or the acceptance date. This information was obtained by inspecting a selection of articles on each journal website.
  - The ***Public Decision Letters*** variable is an indicator for whether a journal releases editors' decision letters alongside published manuscripts (✓) or not (–). For *The BMJ*, this was verified by inspecting recently published articles. *Nature* states on its peer review page that decision

letters are released "in some cases."

- The ***Public Referee Reports*** variable is an indicator for whether a journal releases referees' reports alongside published manuscripts (✓) or not (–). The ***Signed Public Referee Reports*** is an additional indicator for whether the authors of each report are publicly specified. This was verified by inspecting recently published articles from *The BMJ* and *Nature*. We note once again that *Nature* policies are subject to referee and author consent.
- The ***Prior Manuscript Versions*** variable is an indicator for whether a journal releases the prior versions of published manuscripts alongside the final one (✓) or not (–). This was verified by inspecting recently published articles at *The BMJ*.

- The information in the **Communication** section comes from a smaller range of sources (Table A.6).
  - The ***Policy on Author Appeals*** variable records whether the journal explicitly states that authors can appeal manuscript decisions (✓) or not (–). Most journals in our sample do this, but the extent to which they have a formal procedure for doing so varies. This information comes from policies on journal websites.
  - The ***Interactive Review*** variable indicates whether the journal reports having established channels for real-time communication between reviewers and authors (✓) or not (–). *Frontiers* is the only journal in our sample that reports having a system of interactive review.

A.5 Journal sourcing table

We use Table A.6 to specify the sources for journal policies that appear in Table 1 and throughout the article.

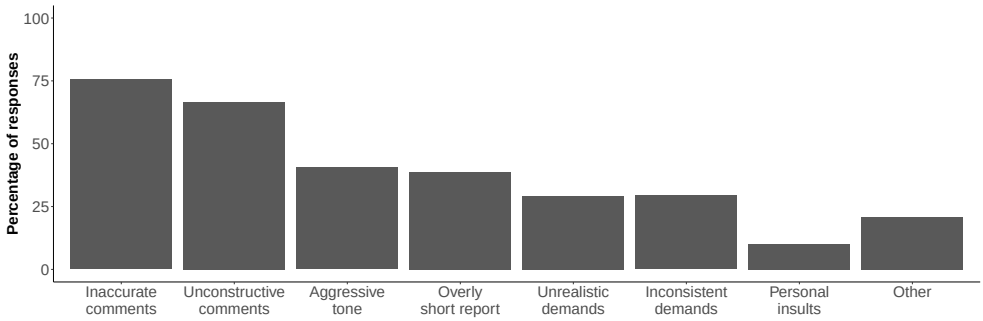
Table A.6: Sources for journal statistics and editorial policies

Journal	Source	Year	Link
American Economic Review	Report of the Editor *AER*	2025	AER-1
	Editorial Policy	2025	AER-2
The BMJ	Resources for reviewers	2025	BMJ-1
	The BMJ's reviewers 2013-2024	2025	BMJ-2
	Competing interest policy	2025	BMJ-3
	Publishing model	2025	BMJ-4
Econometrica	Annual Report 2023-2024 Referees	2025	ECMA-1
	Editorial Procedures	2025	ECMA-2
	Annual Report 2023-2024	2025	ECMA-3
	Bluesky post	2025	ECMA-4
Economic Inquiry	Journal Policies	2025	EI-1
Experimental Economics	Publishing ethics	2025	EX-1
F1000Research	How it Works	2025	F1000-1
Frontiers	Principles of peer review	2025	FRN-1
	Editorial policies and publication ethics	2025	FRN-2
	Annual Report	2025	FRN-3
	Peer Review	2025	FRN-4
JESA	Publishing ethics	2025	JESA-1
Journal of Political Economy	Recent Referees	2025	JPE-1
	ETHICS	2025	JPE-2
	JPE TURNAROUND TIMES	2025	JPE-3
Management Science	Submission Guidelines	2025	MS-1
Nature	Peer Review	2025	NAT-1
	Competing interests	2025	NAT-2
	Journal Metrics	2025	NAT-3
	Editorial criteria and processes	2025	NAT-4
Nature Human Behaviour	Peer Review	2025	NHB-1
Psychological Bulletin	Submission Guidelines	2025	PB-1
Philosophical Review	Submission Guidelines	2025	PR-1
QJE	Acknowledgment of Referees	2025	QJE-1
	Tweet	2025	QJE-2
REStud	Excellence in Refereeing Award	2025	RES-1
	Submissions	2025	RES-2
	Turnaround statistics	2025	RES-3

Journal abbreviations: *JESA = Journal of the Economic Science Association*; *QJE = Quarterly Journal of Economics*; *REStud = Review of Economic Studies.*

B Additional figures and tables referenced in text

Figure B.1: Characteristics of low-quality reports



Notes: Percentages do not add up to 100% as respondents could select multiple reasons (3.1 reasons selected on average). $N = 802$.

Table B.1: Relationships between demographic characteristics and policy attitudes

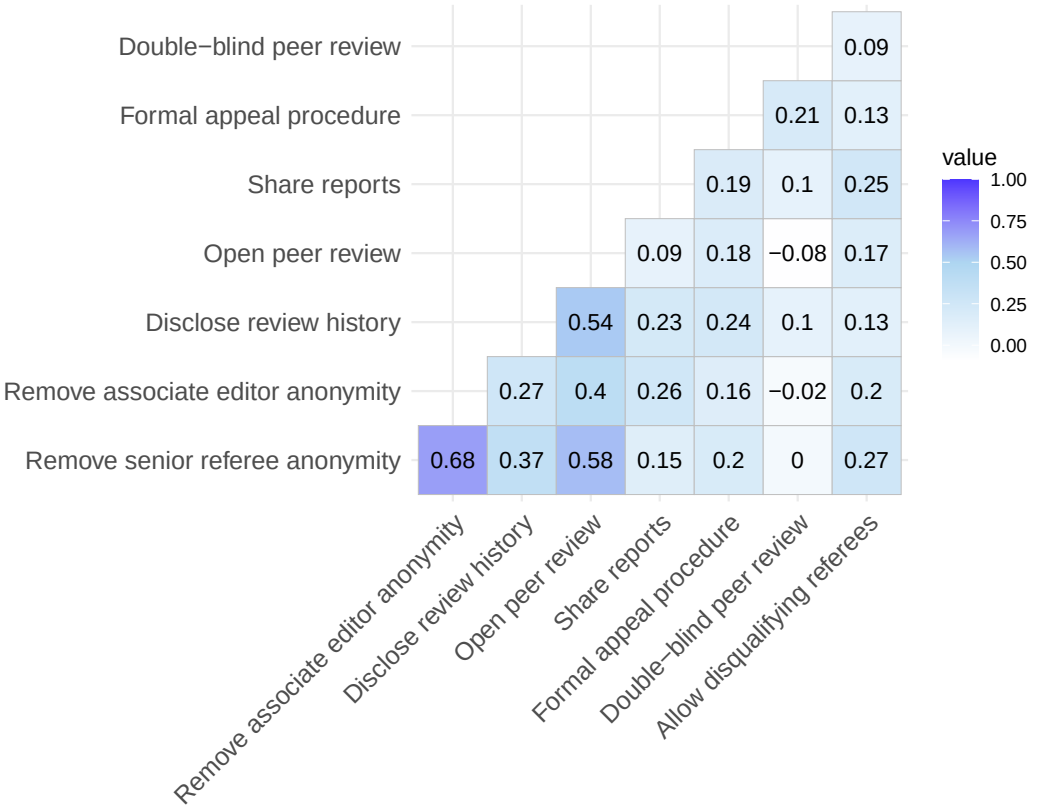
	Double-blind peer review	Remove senior referee anonymity	Remove associate editor anonymity	Allow disqualifying referees	Disclose review history	Open peer review	Share reports	Formal appeal procedure
Active author	−0.088	0.062	0.152	0.121	−0.077	0.050	−0.020	0.113
	(0.331)	(0.099)	(0.104)	(0.241)	(0.090)	(0.095)	(0.084)	(0.088)
Active reviewer	−0.541	−0.106	0.128	0.314	−0.110	−0.286***	0.061	−0.290***
	(0.357)	(0.113)	(0.119)	(0.260)	(0.103)	(0.108)	(0.096)	(0.101)
Top 5 reviewer	−0.276	−0.348***	−0.307***	−0.251	−0.356***	−0.369***	−0.010	−0.124
	(0.363)	(0.111)	(0.117)	(0.264)	(0.101)	(0.107)	(0.095)	(0.100)
Editor	0.026	−0.006	−0.114	−0.417	−0.117	0.003	−0.028	0.038
	(0.381)	(0.116)	(0.122)	(0.278)	(0.105)	(0.111)	(0.098)	(0.103)
US/Canada	−0.168	−0.034	−0.101	0.194	−0.119	−0.224**	−0.031	−0.033
	(0.326)	(0.109)	(0.115)	(0.237)	(0.099)	(0.104)	(0.093)	(0.097)
Junior	−0.010	0.085	0.063	−0.255	−0.034	−0.187*	0.101	0.109
	(0.399)	(0.116)	(0.123)	(0.290)	(0.106)	(0.111)	(0.099)	(0.104)
Female	0.888**	0.183	0.158	0.234	−0.064	0.148	0.167*	0.056
	(0.342)	(0.114)	(0.121)	(0.249)	(0.104)	(0.109)	(0.097)	(0.102)
Constant	3.175***	2.527***	3.097***	3.879***	3.839***	3.023***	3.874***	3.721***
	(0.404)	(0.116)	(0.118)	(0.294)	(0.101)	(0.107)	(0.095)	(0.100)
N	112	802	802	112	802	802	802	802

Notes: This table presents coefficients from a series of multivariate linear regressions of policy attitudes on demographic characteristics.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the main text, we look at heterogeneity in terms of class membership. In this table, we examine whether there is heterogeneity in support for each individual policy using multivariate OLS regressions. We find that the extent of heterogeneity is modest overall. Top-5 reviewers are less favorable to removing the anonymity of associate editors and senior referees, disclosing review histories, and fully open peer review. Active reviewers are less favorable towards open peer review and having a formal appeal process. Researchers from US/Canada are less likely to support open peer review. Female respondents are more likely to support double-blind review. These results are robust to using an ordered logit specification.[13]

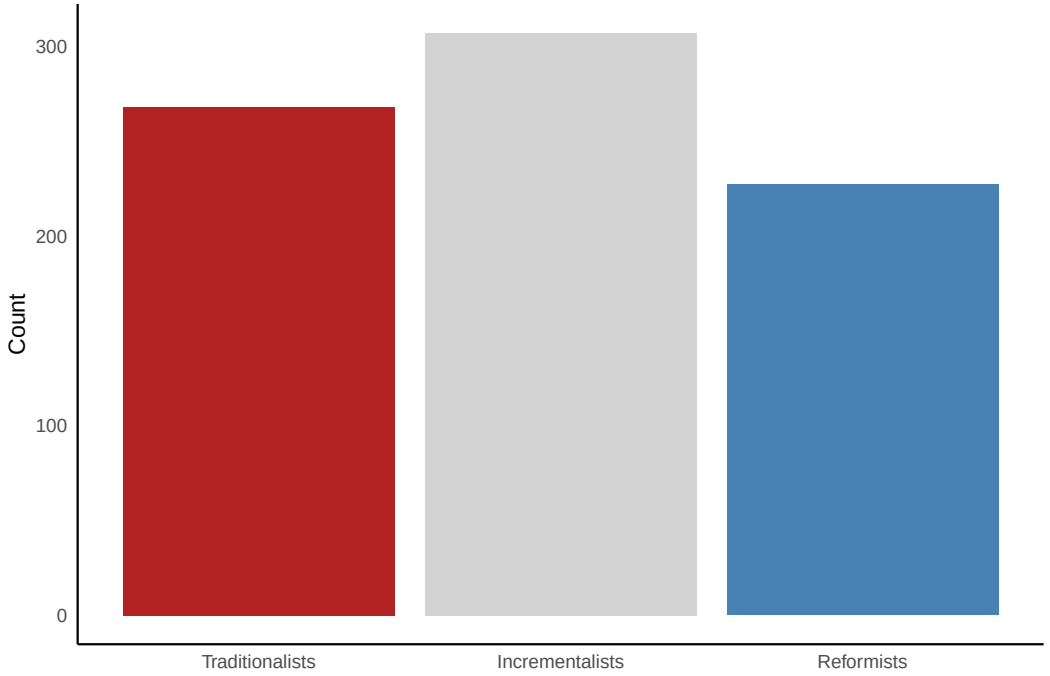
[13]The ordered logit specification results in magnitudes that are somewhat larger and slight differences in significance levels. However, the signs of most of the estimates remain the same, except for a few estimates close to zero. The main exception to this pattern of similarity is for the "share report" policy, where "Female" is statistically significant at the 5% level in the ordered logit specification but is marginally significant in the OLS regression.

Figure B.2: Correlation matrix showing pairwise relationships between survey responses



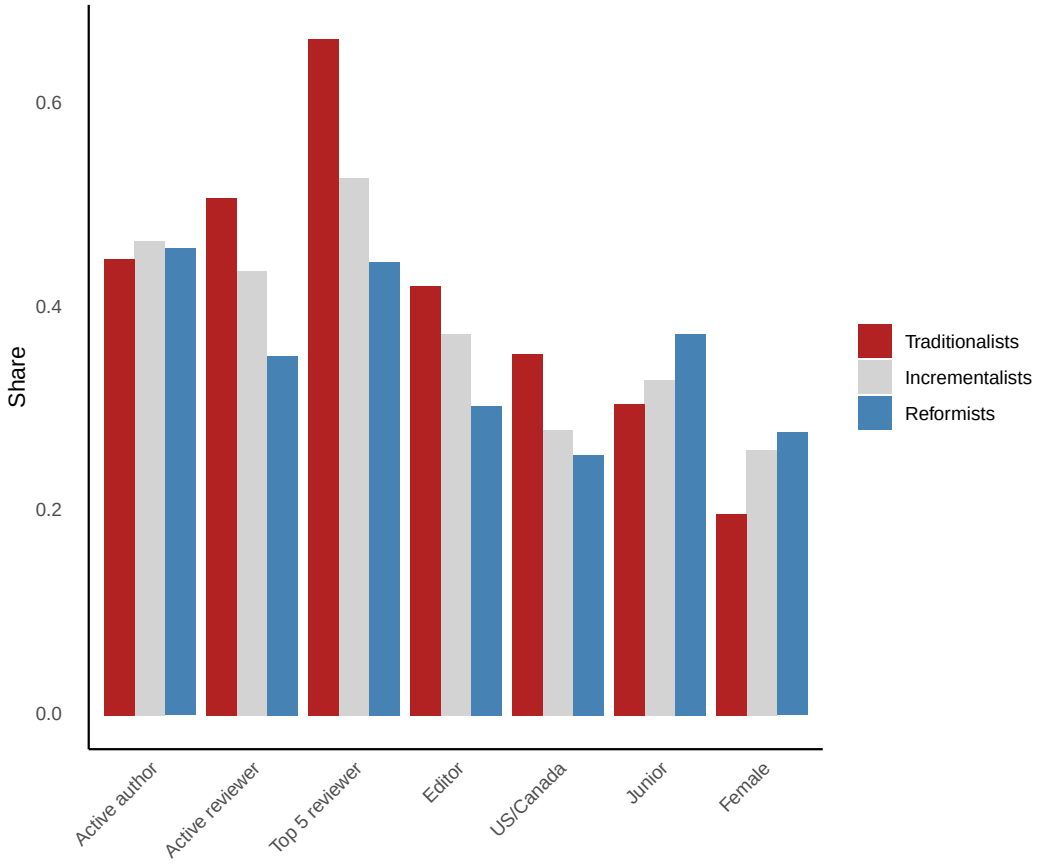
Notes: This figure shows pairwise polychoric correlations between attitudes toward each pair of the main policy questions discussed in the paper. Each correlation is based on the maximum number of observations with non-missing values for both questions in the pair. The sample size is $N = 802$ for most correlations, except for those involving the "disqualifying referees" and "double-blind review" questions, where $N = 112$.

Figure B.3: Number of respondents per LCA class



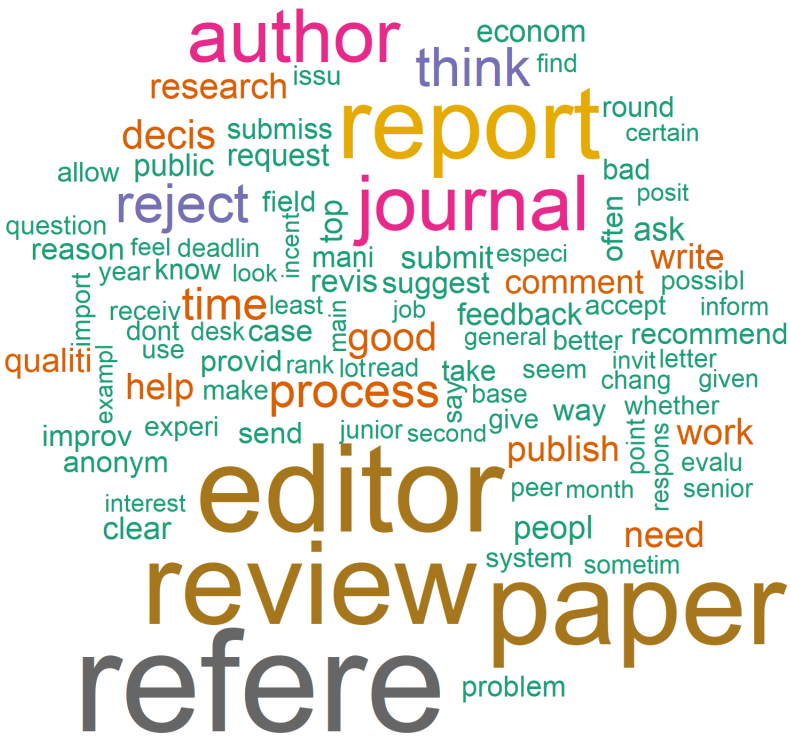
Notes: This figure shows the number of respondents assigned to each of the three classes identified by the latent class analysis (LCA). The shares of respondents in each class are 33%, 38%, and 28% for Traditionalists, Incrementalists, and Reformists, respectively. The classification is based on responses to all of the main policy questions discussed in the paper, except those on "disqualifying referees" and "double-blind review". The sample size is $N = 802$.

Figure B.4: Share of LCA class belonging to each demographic group



Notes: This figure shows the share of respondents in each LCA class who belong to a given demographic group. Each bar represents the share of respondents who belong to a demographic group within a specific class, not the share of the overall sample. The sample size is $N = 802$.

Figure B.5: Frequency of words mentioned in comments



Notes: This figure shows the 100 most frequent words (condensed to their common roots) mentioned in open-text responses to our survey after removing common stop words. The size of each word is proportionate to its frequency. This word cloud combines responses to the questions "Are there other proposals you would like to make to improve the quality of peer reviews or the peer-review process more generally" **[Q21]** and "Please enter below any additional suggestion(s) to improve your experience as a referee" **[Q38]**. This figure is based on 506 comments left by 268 unique respondents.

Figure B.6: Frequency of unique words mentioned in transparency-related comments

Notes: This figure shows the 9 most frequent and unique words (condensed to their common roots) mentioned in transparency-related open-text responses to our survey after removing common stop words. As a first step, all responses were manually categorized into one of 119 categories. Of these, 23 categories corresponding to 113 comments were classified by GPT-4o as transparency-related. Word frequencies were then calculated for transparency- and non-transparency-related comments separately. We then identify uniquely transparency-related words by keeping the most frequent words in the transparency-related comments except for those that also appear in the non-transparency-related comments. The size of each word is proportionate to its frequency. This word cloud combines responses to the questions "Are there other proposals you would like to make to improve the quality of peer reviews or the peer-review process more generally" **[Q21]** and "Please enter below any additional suggestion(s) to improve your experience as a referee" **[Q38]**. For the purpose of clarity, we do not include words that appear less than three times in the transparency comments, but our overall conclusions are not sensitive to this choice.

C Survey questions

Consent Form

Principal Investigators: Gary Charness (UCSB), Anna Dreber (Stockholm School of Economics), and Séverine Toussaert (Oxford)

Description: This is a survey on peer review, which should take about 15-20 minutes of your time. We are interested in your view of the current peer-review process and how it can be improved.

Eligibility Criteria: You are eligible to participate in this survey if, **over the last two years**, (i) you **completed at least one peer review**; and (ii) you **received referee reports** on a paper you submitted for publication in a peer-reviewed journal.

Risks and benefits: There are no physical or emotional risks associated with this study that would go beyond the risks of daily life. Your participation in this study may improve the peer-review process and, therefore, benefit the scientific community. In addition, we will give $100 (cash or gift certificate) to two people drawn randomly from the respondents; you will be asked to leave your email address in a separate survey link if you wish to be entered in the lottery.

Confidentiality: The information collected in this survey may be published in a report or a journal article and presented to interested parties, including possibly, but not exclusively, members of editorial boards or scientific committees. In no circumstances will your identity or personal involvement in this study be disclosed. No personal data (e.g., your IP address) will be collected, except for your email address if you wish to be emailed the report and/or participate in the prize draw (this information will not be connected to your survey responses and will be destroyed after the prize draw). Other information (e.g., survey responses, time of the survey) will be kept by the researchers and may be used for future studies.

Your rights as a participant: Participation is entirely voluntary. You may leave the survey at any time without any penalty or prejudice.

Ethics approval: This research has been reviewed according to the ethics procedures for research involving human subjects of the University of Oxford (approval # ECONCIA-21-21-20). If you wish to raise any concerns about this study to the ethics committee, please email ethics@economics.ox.ac.uk.

Please indicate below that you have read the above, that you meet the eligibility criteria, and that you are willing to participate in this online survey.

Yes, proceed to the survey YES/NO

Your experience of the peer-review process as an author

[Q1]: Over the last two years, how many times did you submit a paper to an economics journal? Please include only first-time submissions (not revisions), with submissions of the same paper to different journals counted separately. [Dropdown with numbers]

[Q2]: How would you rate the overall quality of the referee reports you received over this period? Please indicate what approximate percentage of reports were of the following quality (total should sum to 100):

Very low	[ ]
Fairly low	[ ]
Average	[ ]
Fairly high	[ ]
Very high	[ ]
Total	[100]

[Q3]: What were the characteristics of the low-quality reports? Please tick all that apply:

☐ Inaccurate statements about what the paper does or does not do

☐ Overly short report

☐ Very vague and unconstructive comments

☐ Written with an aggressive tone

☐ Personal insults

☐ Unrealistic demands

☐ Inconsistent demands

☐ Other - please specify: [TEXT BOX]

[Q4]: A referee report can achieve multiple objectives. How important do you consider each of the following objectives? Please rank 1-4 in order of importance (with 1 being most important) by dragging and dropping the various items: [1= most important, 2, 3; 4 = least important]

- Help editor reach an informed decision on the paper
- Give general comments that improve the paper
- Provide detailed feedback on the paper
- Make precise suggestions that improve the paper

[Q5]: As an author, what do you expect from the peer-review process? Please rank 1-3 in order of importance (with 1 being most important): [1 = most important, 2; 3 = least important]

- Getting useful feedback on my work
- A timely decision (whether good or bad)
- Getting a reasonable and well-substantiated decision

Improving the quality of peer reviews

[Q6]: Below is a list of proposals to improve peer reviews. On a scale from 1 to 5, how useful do you find each of them? [1 = not useful at all; 2, 3, 4; 5 = extremely useful]

i. Providing a set of guidelines for writing referee reports.

ii. Providing doctoral training on how to write peer reviews.

iii. Making the history of (anonymous) reviews and authors' responses publicly available.

iv. Removing the anonymity of senior referees.

v. Removing the anonymity of associate editors.

vi. Somehow grading reports and rewarding referees for high-quality reports.

vii. Encouraging the use of a platform that tracks referee activity in a centralized way.

viii. Making all reports available to all of the reviewers and making sure reviewers know this is being done.

Guidelines for writing a report

[Q7]: What type of comments do you find most useful or would you like to see more of? Please make 3 selections from the following list:

☐ Comments about the presentation of the results

☐ Suggestions to improve the existing analysis

☐ Suggestions about possible extensions

☐ Comments that help me clarify the contribution of the paper relative to the literature

☐ Comments about shortening/restructuring the paper

☐ Comments that put in perspective the assumptions made in the paper

☐ Comments about missing previous work and references

☐ Robustness checks

[Q8]: Do you think journals or associations should provide a template for referee reports? [YES/NO]

Information disclosure

[Q9]: In other disciplines, such as public health/medicine, many journals have an open peer-review process: referees sign their reports and the entire review history (including responses to referees) is disclosed. On a scale from 1 to 5, how favorable would you be to an open review policy? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q10]: What if this only applied to senior reviewers? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q11]: Another recent trend is to make the history of reports/responses to referees publicly available in an anonymized way unless the reviewers choose to disclose their identity; see e.g., Nature Communcations. On a scale from 1 to 5, how favorable would you be to such a policy? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

Tracking referee activity

[Q12]: At the moment, there is no centralized system that would allow journal editors to:

- check how many peer-review requests a researcher has recently received across all journals.
- find suitable referees who might be currently available to provide a peer review.

One platform called Publons allows researchers to document their (verified) peer-review activity and to register their interest in doing peer reviews for journals. However, it is not widely used at the moment in economics.

On a scale from 1 to 5, how favorable would you be to the more widespread use of Publons or a similar type of platform? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

Recognition

[Q13]: Do you think that referees would do a better job if they were better rewarded for their work? [YES/NO]

[Q14]: How should referees be rewarded? Please tick all that apply:

☐ Excellence in refereeing awards based on specific criteria
☐ Payment for timely completion e.g., as at the *American Economic Review*
☐ Discount on submissions to the publisher
☐ Other - please specify: [TEXT BOX]

Improving the peer-review process more generally

[Q15]: What do you think is an appropriate time length to give to reviewers to submit their reports (in weeks)? [Dropdown: From 1 to 16+ weeks]

[Q16]: How do you feel about the policy of having desk rejections? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q17]: The American Economic Association started a new journal in 2017 called *AER: Insights*. This journal follows a model close to the one of medicine, with the endeavor to accept or reject papers without having to go through a lengthy revision process. Like the papers that *AER: Insights* is looking to publish, reports are supposed to be short and to the point. The whole process is supposed to be fast.

How favorable are you to this type of model? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q18]: In the case of a rejection, the norm is not to challenge the decision made by the Editor or the views of the referees. This norm is not always followed in practice.

How favorable would you be to a policy allowing the authors to submit a (single) response to the referees and the Editor? The referees would be under no obligation to provide additional comments; a "cooling period" could be required before the authors can send their response. There would be no guarantee of the referees taking this rebuttal into account, and the decision would be final after the comment period. [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q19] (Only included in later survey versions): At journals such as *Management Science*, the review process is double-blind i.e., the identity of both the authors and the referees is kept anonymous. How favorable are you to double-blind reviewing? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q20] (Only included in later survey versions): In some fields, authors are allowed to suggest that certain reviewers should be disqualified from reviewing their work. How favorable are you to this possibility? [1 = not favorable at all; 2, 3, 4; 5 = very favorable]

[Q21]: Are there other proposals you would like to make to improve the quality of peer reviews or the peer-review process more generally? [TEXT BOX]

Your experience of the peer-review process as a referee

[Q22]: On average, approximately how many referee reports do you write per year? [Dropdown with numbers]

[Q23]: What percentage of the time do you write referee reports for the following types of journals? (total should sum to 100):

Top-5 journal	[ ]
Top field journal	[ ]
Other journal in Economics	[ ]
Journals in other disciplines	[ ]

Total [100]

[Q24]: Have you occupied or are you currently occupying an editorial position? [YES/NO]

[Q25]: Usually, how much time do you spend on a referee report, including reading the paper and writing the report? [Dropdown: Less than one hour, 1 or 2 hours, Half a working day, 1 day, 2 days, More than 2 days]

[Q26]: Over the past two years, what percentage of the time were you late submitting a referee report? [0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]

[Q27] [*If Q26 answer* > 0]: On average, what was your delay? [Dropdown: 1 day, More than 1 day & less than 1 week, 1-2 weeks, 3-4 weeks, More than a month]

[Q28]: What do you think is a reasonable number of reports to be assigned per year? [Dropdown with numbers]

[Q29]: Did you reject a request to referee over these past two years? [YES/NO]

[Q30]: How many times did you reject a request to referee? [Dropdown with numbers]

[Q31]: What were the main reasons? Please tick all that apply:

☐ Conflict of interest
☐ Inability to judge the paper
☐ Too remote from your research field
☐ Lack of time
☐ Low quality paper
☐ Lower-ranked journal
☐ Other - please specify: [TEXT BOX]

[Q32]: How many times did you feel tempted to decline a report even if you ended up fulfilling the request? [Dropdown with numbers]

[Q33]: When you were tempted to decline a report, what were the main reasons? Please tick all that apply:

☐ Conflict of interest

☐ Inability to judge the paper

☐ Too remote from your research field

☐ Lack of time

☐ Low quality paper

☐ Lower-ranked journal

☐ Other - please specify: [TEXT BOX]

[Q34]: How do you feel about people refereeing papers by co-authors or friends?

- This should never happen.
- This should happen as little as possible but cannot be avoided sometimes.
- This is not a problem as long as the editor is aware of the potential conflict of interest.
- This is not a problem and there is no reason to inform the editor.

[Q35]: What do you see as the biggest benefits of being a referee? Please rank 1-5 in order of importance (with 1 being most important) by dragging and dropping the various items: [1 = most important; 2, 3, 4; 5 = least important]

i. I can help to ensure the right papers are published or rejected

ii. I can get to know the editors and make myself known.

iii. I can learn from the opinion of the other referees and the editor.

iv. I can attentively read papers I would never read otherwise.

v. Being a referee makes me a better writer.

[Q36]: How important do you consider your role as a referee? [1 = most important; 2, 3, 4; 5 = least important]

[Q37]: How could your experience as a referee be improved? Please rank 1-4 in order of importance [1 = most important; 2, 3; 4 = least important]

i. There is a global annual limit on how many papers I am requested to review.

ii. The editors give clear guidance of what they would like to learn from my report.

iii. The editors systematically share their decision and the other reports.

iv. The editors assign me only papers that are related to my research.

[Q38]: Please enter below any additional suggestion(s) to improve your experience as a referee: [TEXT BOX]

A little more about you

[Q39]: How many papers have you published in your career up to now? Please indicate a ballpark estimate. [TEXT BOX]

[Q40]: What are your key areas of research? Please select all that apply:

- Applied econometrics
- Applied microeconomics
- Behavioral economics
- Decision theory
- Development economics
- Economic history
- Econometric theory
- Experimental economics
- Financial economics
- Game theory
- Industrial organization
- International trade
- Labor economics
- Macroeconomics
- Microeconomic theory
- Political economy
- Public economics
- Structural econometrics

- Urban economics
- Other - indicate: [TEXT BOX]

[Q41]: What is your gender? [Dropdown: Male, Female, Other, Prefer not to say]

[Q42]: What is your age? [Dropdown: Under 30, 30-39, 40-49, 50-59, 60-69, 70+, Prefer not to say]

[Q43]: What is your position? [Dropdown: PhD candidate, Post-doctoral researcher, Assistant professor, Associate professor, Full professor, Prefer not to say]

[Q44]: In what country is your job located? [Dropdown]

[Q45]: Finally, if you have any comments about the survey itself, feel free to add in the text box below: [TEXT BOX]