

# Online Appendix to Voicing disagreement in science: Missing women

David Klinowski

## A. Evidence from publications

eTOCs.- I scrape each journal's online table of contents (eTOCs) to collect information on each paper published in the period of analysis, including title, author names, publication date, volume and issue, and paper type (regular paper, comment, etc.). In JAMA, eTOCs display only the names of the first three authors of a paper and "et al." thereafter; therefore, I complement the JAMA data with data from PubMed, which include full lists of authors of JAMA papers. I also scrape names of editors in each issue of PNAS since April 2006 and each issue of Science since April 2009, which are the dates in which mastheads for these journals first become available online.

Fields within journal.- I obtain information on what field a paper is in for papers in the AER, Nature, and PNAS, which are journals that provide this information on their websites. For the AER, I collect the Journal of Economic Literature (JEL) codes assigned to each paper. JEL codes are a 20-category classification of fields within economics. I use JEL codes to assign each article one or many of 14 mutually exclusive fields, following Card and DellaVigna's (2013) classification. These fields are microeconomics, theory, macroeconomics, labor, econometrics, industrial organization, international economics, finance, public economics, health or urban economics, development, history, laboratory experiments, and other. For Nature, I collect fields from Nature's search engine, which returns papers in each of 5 major fields: biological sciences, earth sciences, health sciences, physical sciences, and social sciences. I collect all titles returned for each field search, and match titles to papers in the original dataset. As in the AER, a paper in Nature can be in multiple fields. For PNAS, I obtain fields directly from eTOCs, which sort papers into either biological sciences, physical sciences, or social sciences. Thus, a paper in PNAS is in only one field. Finally, I assign comments the same field(s) of the paper they comment.

Citation counts.- I obtain citations data from the Clarivate Analytics' Web of Science Core Collection database, accessed in January 2021 under license with Stanford University. The main variable I use from this database is the average number of citations to an article per year. For a number of comments missing in the Web of Science dataset (4 in Nature, 36 in PNAS, and 117 in Science), I compute average annual citations from Google Scholar.

Gender assignment.- I assign authors to either male, female, or missing gender using Genderize.io, a commercial service used in past research (Lerchenmueller, Sorenson, and Jena, 2019; Huang et al., 2020).<sup>1</sup> I assign a gender to an author if the gender relative frequency of the author's first name in the Genderize.io database is at least 90 percent and the name appears at least 50 times in the Genderize.io database; otherwise, I set the author's gender

---

<sup>1</sup> <https://genderize.io>. Lerchenmueller, Sorenson, and Jena (2019) report that the Genderize.io database contained 86,710 distinct names in 2017.

to missing. To examine the robustness of the results to alternative methods of gender assignment, I match author first names to names in the US Social Security Administration (SSA) birth records.<sup>2</sup> These records contain the number of occurrences and the relative gender frequency of names with 5 or more occurrences in a year in the population of US births, from 1880 through 2018. There are 98,400 distinct names in this US SSA dataset. I assign a gender to an author if the gender relative frequency of the author's first name is at least 90 percent in the US SSA dataset; otherwise, I set the author's gender to missing. Finally, for the AER only, I also assign gender manually, by looking up each author's information online. Manual assignment provides not only an additional robustness test for the results, but also a way to check the accuracy of the Genderize.io assignment in the AER sample. I am able to manually assign a gender to 99.7 percent of AER authors. Of these, 99.9 percent of authors identified as male using Genderize.io are identified as male manually, and 97.8 percent of authors identified as female using Genderize.io are identified as female manually. I assign a gender using Genderize.io to approximately 63 to 85 percent of authors in a journal. Failure to assign a gender occurs primarily when the name is Asian, as Asian names are underrepresented in Genderize.io (and in SSA birth records), or when authors sign with their initials or as a research group.

*Author seniority*.- I construct a proxy for author seniority by tracking each author's publication history in a journal. Within each journal, I match observations by author full name, and count the number of publications (regular papers and comments) each author has in the journal in the period of analysis. The cumulative number of publications an author has at a given time is then their proxy for their seniority at that time. To increase my ability to capture seniority accurately, I expand the original data as follows. For the AER, I count not only publications the author has had in the AER, but also in all other journals managed by the American Economic Association (AEA). These journals are (with the period covered in the data in parentheses), American Economic Review: Insights (2019), Journal of Economic Literature (1999-2019), Journal of Economic Perspectives (1987-2019), American Economic Journal: Applied Economics (2009-2019), American Economic Journal: Economic Policy (2009-2019), American Economic Journal: Macroeconomics (2009-2019), American Economic Journal: Microeconomics (2009-2019), and AEA Papers and Proceedings (2018-2019).<sup>3</sup> Incorporating publications from these journals adds 7,912 article-author observations to the dataset, which helps to better capture author seniority. For JAMA and PNAS, I count publications in these journals going further back in time than the original dataset. I count the publications the author has had since 2002 rather than 2013 for JAMA, and since 1997 rather than 2008 for PNAS.<sup>4</sup> This adds 26,995 article-author observations to the JAMA dataset and 184,252 article-author observations to the PNAS dataset, which, again, helps to better capture author seniority in these journals.

---

<sup>2</sup> Obtained from <https://www.ssa.gov/oact/babynames/limits.html>.

<sup>3</sup> I scraped these data from the journals eTOCs as I did for the AER.

<sup>4</sup> I scraped these data from the eTOCs, going as far back in time as JAMA and PNAS display author names on their eTOCs. I do not use these data in the main analysis because JAMA launched its comments section in 2013 and PNAS did in 2008.

**Table A1** Descriptive statistics, publications

	AER	ASR	JAMA	Nature	PNAS	Science
Period	1999-2019	2004-2019	2013-2019	2004-2019	2008-2019	1998-2019
N articles	2,240	689	2,581	13,616	47,560	17,548
Fraction comment	0.0540	0.0290	0.4661	0.0253	0.0229	0.0409
N unique authors	3,277	1,117	17,068	127,272	211,639	113,753
Gender assigned						
Genderize.io	0.8526	0.7995	0.7828	0.6274	0.7244	0.6259
SSA records	0.8374	0.8460	0.7896	0.6195	0.7178	0.6184
Manually	0.9966	-	-	-	-	-
Female						
Genderize.io	0.1316	0.3437	0.3251	0.2728	0.2833	0.2430
SSA records	0.1490	0.3737	0.3562	0.3020	0.3141	0.2686
Manually	0.1422	-	-	-	-	-
N article-author	4,140	1,158	15,541	119,947	252,034	95,679

**Notes:** *N articles* is the number of regular papers and comments combined. *Fraction comment* is the fraction of articles that are comments. *Gender assigned* is the fraction of unique authors assigned a gender. *Female* is the fraction of observations at the article-author level with a female author, among article-author observations with author assigned a gender. *N article-author* is the number of observations at the article-author level with gender assigned with Genderize.io.

**Table A2** Share of female authors in a journal

	AER	ASR	JAMA	Nature	PNAS	Science
Comment	-0.054*** (0.020)	-0.110 (0.082)	-0.126**** (0.010)	-0.056*** (0.011)	-0.080**** (0.008)	-0.078**** (0.009)
N	4,140	1,158	15,541	119,947	252,034	95,679
R <sup>2</sup>	0.0049	0.0053	0.0092	0.0021	0.0005	0.0019
Mean regular paper	0.134	0.346	0.342	0.273	0.284	0.244

**Notes:** Estimates from OLS regressions for each journal separately, regressing the author's gender on a comment indicator and year fixed effects. Observations at the article-author level. *Comment* shows the marginal effect of the comment indicator, and *Mean regular paper* shows the mean fraction of female authors for regular papers, in hundreds of a percentage point. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A3** Share of female authors in a journal, gender from SSA birth records

	AER	ASR	JAMA	Nature	PNAS	Science
Comment	-0.044** (0.022)	-0.170** (0.079)	-0.133**** (0.010)	-0.066**** (0.012)	-0.087**** (0.008)	-0.082**** (0.009)
N	4,074	1,223	15,656	119,030	251,843	94,819
R <sup>2</sup>	0.0064	0.0054	0.0096	0.0021	0.0006	0.0024
Mean regular paper	0.151	0.378	0.373	0.303	0.315	0.270

**Notes:** Replication of Table A2, assigning gender to authors using US Social Security birth records. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A4** Share of female authors in a journal, gender assigned manually

	AER	ASR	JAMA	Nature	PNAS	Science
Comment	-0.042** (0.020)	-	-	-	-	-
N	4,817	-	-	-	-	-
R <sup>2</sup>	0.0058	-	-	-	-	-
Mean regular paper	0.144	-	-	-	-	-

**Notes:** Replication of Table A2 for the AER, with gender assigned manually. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A5** Share of female authors in a journal, observations at the article level

	AER	ASR	JAMA	Nature	PNAS	Science
Comment	-0.061**** (0.015)	-0.124* (0.068)	-0.112**** (0.010)	-0.025** (0.011)	-0.056**** (0.007)	-0.046**** (0.007)
N	2,240	689	2,581	13,616	47,560	17,548
R <sup>2</sup>	0.0070	0.0172	0.0501	0.0038	0.0023	0.0056
Mean regular paper	0.118	0.286	0.275	0.155	0.193	0.144

**Notes:** Replication of Table A2, with observations at the article level. The share of female authors in an article is the fraction of female authors over all authors in the article, including authors with no gender assigned. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A6** Share of female authors in a journal, controlling for field of the article

	AER		Nature		PNAS	
	(1)	(2)	(1)	(2)	(1)	(2)
Comment	-0.054*** (0.020)	-0.044** (0.020)	-0.052**** (0.014)	-0.055**** (0.014)	-0.079**** (0.008)	-0.071**** (0.008)
Field controls	N	Y	N	Y	N	Y
N	4,140	4,140	88,471	88,471	251,881	251,881
R <sup>2</sup>	0.0049	0.0275	0.0014	0.0163	0.0005	0.0079
Mean regular paper	0.134	0.134	0.281	0.281	0.284	0.284

**Notes:** Replication of Table A2 for the AER, Nature, and PNAS, examining the effect of controlling for field within journal. For each journal, column 1 excludes field controls and column 2 includes field controls. For the AER, field controls are dummies for microeconomics, theory, macroeconomics, labor, econometrics, industrial organization, international economics, finance, public economics, health or urban economics, development, history, laboratory experiments, and other. For Nature, field controls are dummies for biological sciences, earth sciences, health sciences, physical sciences, and social sciences. For PNAS, field controls is a categorical variable that indexes biological sciences, physical sciences, and social sciences. For Nature, the period of analysis in columns 1 and 2 is restricted to years 2010 through 2019, since Nature's "Browse by Subject" search engine, which I use to observe field classification, returns results only from 2010 onward. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A7** Distribution of the cumulative total number of publications

	Cumulative total number of publications				
	1	2	3	4	5+
AER	0.6092	0.1861	0.0820	0.0429	0.0798
ASR	0.7861	0.1579	0.0381	0.0123	0.0056
JAMA	0.8438	0.1040	0.0281	0.0098	0.0143
Nature	0.7686	0.1338	0.0452	0.0206	0.0317
PNAS	0.6684	0.1890	0.0608	0.0308	0.0511
Science	0.8131	0.1194	0.0343	0.0148	0.0184

**Notes:** Fraction of authors that publish a given total number of publications in the period of analysis. Publications include regular papers and comments. Sample restricted to authors with assigned gender. Sample expanded to include publications in all AEA journals for AER authors, and publications in 2002-2019 for JAMA authors and 1997-2019 for PNAS authors.

**Table A8** Cumulative number of publications by an author

	AER	ASR	JAMA	Nature	PNAS	Science
Last author	-0.101 (0.068)	0.013 (0.043)	0.398**** (0.073)	0.835**** (0.028)	2.874**** (0.028)	0.835**** (0.021)
N	4,270	1,200	19,275	188,640	533,347	154,808
R <sup>2</sup>	0.0551	0.0799	0.0040	0.0306	0.0612	0.0439

**Notes:** Coefficient estimates from OLS regressions of the author's cumulative number of publications at the time of the observation on an indicator that the author is listed last on the paper, and year fixed effects. Observations at the article-author level. Sample is restricted to articles with at least two authors, and includes observations with no gender assigned. Cumulative number of publications is computed counting all articles. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001. If the cumulative number of publications proxies well for seniority, *Last author* is expected to be positive in JAMA, Nature, PNAS, and Science, since most fields in these journals list the most senior author last by convention (Fontanarosa, Bauchner, Flanagan, 2017; Sekara et al., 2018), and is expected to be zero or even negative in the AER and ASR, since senior authors are generally not listed last in these journals. AER papers typically list authors alphabetically or otherwise list the author who made the greatest contribution first (Ray and Robson, 2018). ASR papers typically list the author who made the greatest contribution first and list alphabetically the remaining authors.

**Table A9** Cumulative number of publications by an author

	AER	ASR	JAMA	Nature	PNAS	Science
Comment	-0.405**** (0.108)	-0.079 (0.092)	-0.597**** (0.031)	-0.497**** (0.039)	-1.239**** (0.071)	-0.226**** (0.022)
Female	-0.532**** (0.087)	-0.152**** (0.039)	-0.373**** (0.034)	-0.541**** (0.015)	-1.599**** (0.019)	-0.393**** (0.010)
N	4,140	1,158	15,541	119,947	252,034	95,679
R <sup>2</sup>	0.0617	0.0915	0.0122	0.0351	0.0184	0.0273
Mean regular paper	2.225	1.325	1.769	2.039	3.242	1.610

**Notes:** Estimates from OLS regressions for each journal separately, estimating the cumulative number of publications by the author at the time of each article published. Cumulative publications include regular papers and comments. Regressors are comment and female indicators, and year fixed effects. Observations at the article-author level. *Comment* and *Female* show the marginal effects of these indicators, while *Mean regular paper* shows the mean cumulative number of publications when the observed article is a regular paper, for males and females combined, in hundreds of a percentage point. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A10** Cumulative total number of publications by an author

	AER	ASR	JAMA	Nature	PNAS	Science
JEL author	1.789**** (0.180)	-	-	-	-	-
N	3,277	-	-	-	-	-
R <sup>2</sup>	0.1665	-	-	-	-	-

**Notes:** Coefficient estimate from an OLS regression of the author's cumulative total number of publications (the number is computed *excluding* JEL publications) on an indicator that the author has published in the JEL, and year fixed effects. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001. Results indicate that AER authors who are also JEL authors have on average 1.789 more publications than AER authors who are not JEL authors.

**Table A11** Share of female authors in a journal, controlling for author seniority

	AER	ASR	JAMA	Nature	PNAS	Science
Comment	-0.060*** (0.020)	-0.113 (0.085)	-0.140**** (0.010)	-0.067**** (0.011)	-0.103**** (0.008)	-0.082**** (0.009)
N	4,140	1,158	15,541	119,947	252,034	95,679
R <sup>2</sup>	0.0185	0.0189	0.0176	0.0155	0.0335	0.0174
Mean regular paper	0.134	0.347	0.344	0.274	0.284	0.245

**Notes:** Replication of Table A2, including as additional control the author's cumulative number of publications at the time of the observation. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A12** Probability of solo-authorship

<b>Panel a Main-Effects Model</b>						
	AER	ASR	JAMA	Nature	PNAS	Science
Female	0.030*	0.039*	-0.008****	-0.001****	-0.002****	-0.001****
	(0.016)	(0.023)	(0.002)	(0.000)	(0.000)	(0.000)
Comment	0.067**	0.134	0.164****	0.025****	0.090****	0.065****
	(0.029)	(0.086)	(0.008)	(0.004)	(0.006)	(0.006)
N	4,140	1,158	15,541	119,947	252,034	95,679
R <sup>2</sup>	0.0168	0.0205	0.1454	0.0093	0.0331	0.0338
<b>Panel b Interacted Model</b>						
	AER	ASR	JAMA	Nature	PNAS	Science
Female	0.038**	0.044*	0.000	-0.001****	-0.001****	-0.001****
	(0.016)	(0.023)	(0.000)	(0.000)	(0.000)	(0.000)
Comment	0.087***	0.193*	0.181****	0.032****	0.100****	0.073****
	(0.031)	(0.102)	(0.010)	(0.005)	(0.007)	(0.007)
Female x Comment	-0.259****	-0.244	-0.077****	-0.028****	-0.050****	-0.047****
	(0.036)	(0.174)	(0.017)	(0.006)	(0.012)	(0.011)
N	4,140	1,158	15,541	119,947	252,034	95,679
R <sup>2</sup>	0.0189	0.0225	0.1509	0.0109	0.0347	0.0361

**Notes:** Coefficient estimates from OLS regressions of the probability that the article-author observation is a solo-authored, on female and comment indicators and year fixed effects in Panel a, and including also the *female x comment* interaction in Panel b. Observations at the article-author level. Robust std errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A13** Distribution of annual citations

		Pct 1	Pct 5	Pct 25	Pct 50	Pct 75	Pct 95	Pct 99	Max
<b>AER</b>	Regular paper	0.1	0.8	2.9	5.7	11.4	27.9	61.2	143.9
	Comment	0.0	0.1	0.5	1.1	2.7	8.5	10.5	16.7
<b>ASR</b>	Regular paper	0.5	1.2	3.2	5.8	9.1	19.6	29.8	70.2
	Comment	0.0	0.1	0.4	0.9	1.6	6.3	8.8	8.8
<b>JAMA</b>	Regular paper	0.3	1.2	5.5	11.2	22.5	61.9	116.8	595.9
	Comment	0.0	0.0	0.0	0.1	0.3	1.1	1.8	21.2
<b>Nature</b>	Regular paper	3.1	5.3	11.2	19.5	34.1	85.9	191.3	1087.4
	Comment	0.1	0.5	1.5	2.9	5.6	16.9	29.6	37.7
<b>PNAS</b>	Regular paper	0.5	1.3	3.2	5.6	9.7	22.3	42.6	332.5
	Comment	0.0	0.1	0.4	0.7	1.6	4.3	9.4	42.1
<b>Science</b>	Regular paper	1.8	3.6	8.4	15.0	28.1	73.4	162.3	2057.8
	Comment	0.0	0.2	0.6	1.2	2.7	7.4	15.1	32.0

**Notes:** Average annual citations to the article in the given percentile. Values are adjusted for differences in citations across fields and years since publication, following past work (Radicchi, Fortunato, and Castellano, 2008; Perry and Reny, 2016; Heckman and Moktan, 2020). The average annual citations to an article are rescaled by the average annual citations of all articles in the same field(s) and publication year as the article in question. For journals for which I do not observe fields (ASR, JAMA, and Science), I adjust only for year of publication.

**Table A14** Seniority of authors on the paper

	a. All authors on the paper			b. Last author on the paper		
	Paper not commented on	Paper commented on	p-val diff	Paper not commented on	Paper commented on	p-val diff
AER	2.226	2.459	0.203	2.132	1.950	0.323
ASR	1.327	1.248	0.392	1.314	1.160	0.101
Nature	2.044	1.699	0.000	2.724	2.226	0.001
PNAS	3.245	2.710	0.000	6.482	5.101	0.000
Science	1.615	1.460	0.000	2.419	1.896	0.000

**Notes:** Cumulative number of publications of authors on papers commented and not commented on. Estimates from journal-specific OLS regressions of the outcome on an indicator that the paper is commented on and year fixed effects, with the sample restricted to regular papers. Panel b further restricts the sample to the last author on the paper. P-value of the difference from robust standard errors.

**Table A15** Likelihood that paper is commented on

	AER	ASR	JAMA	Nature	PNAS	Science
Last author is/was editor	-	-	-	-	-0.001 (0.002)	-0.014*** (0.004)
Seniority of last author	-	-	-	-	-0.001*** (0.000)	-0.006*** (0.001)
N	-	-	-	-	36,335	5,551
R <sup>2</sup>	-	-	-	-	0.0088	0.0076
Mean likelihood	-	-	-	-	0.0089	0.0267

**Notes:** Estimates from OLS regressions for each journal separately of the likelihood that the paper receives a comment. Sample restricted to the last author on the paper, for regular papers. Indicator of editor is 1 if author is currently, or has previously been, an editor in the journal. Regressions control for year fixed effects. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table A16** Share of female authors on the paper

	a. All authors on the paper			b. Last author on the paper		
	Paper not commented on	Paper commented on	p-val diff	Paper not commented on	Paper commented on	p-val diff
AER	0.134	0.150	0.598	0.133	0.146	0.791
ASR	0.346	0.343	0.967	0.383	0.402	0.881
Nature	0.274	0.255	0.127	0.116	0.095	0.368
PNAS	0.284	0.288	0.701	0.178	0.193	0.489
Science	0.245	0.233	0.138	0.132	0.158	0.161

**Notes:** Share of female authors on papers commented and not commented on. Estimates from journal-specific OLS regressions of the outcome on an indicator that the paper is commented on and year fixed effects, with the sample restricted to regular papers. Panel b further restricts the sample to the last author on the paper. P-value of the difference from robust standard errors.

**Table A17** Share of female authors on the paper commented on

	a. All authors on the paper			b. Last author on the paper		
	Male commenter	Female commenter	p-val diff	Male commenter	Female commenter	p-val diff
AER	0.129	0.103	0.799	0.134	0.286	0.391
ASR	0.354	0.321	0.822	0.350	0.500	0.537
Nature	0.214	0.225	0.452	0.202	0.197	0.863
PNAS	0.251	0.291	0.003	0.209	0.214	0.811
Science	0.215	0.243	0.035	0.171	0.172	0.994

**Notes:** Share of female authors on papers commented on by male and female commenters. Estimates from journal-specific OLS regressions of the outcome on an indicator that the author of the comment is a female and year fixed effects (year fixed effects excluded for AER and ASR due to small sample). Panel b further restricts the sample to the last author on the paper. P-value of the difference from robust standard errors.

**Table A18** Seniority of the authors on the paper commented on

	a. All authors on the paper			b. Last author on the paper		
	Male commenter	Female commenter	p-val diff	Male commenter	Female commenter	p-val diff
AER	2.083	1.066	0.048	1.754	1.429	0.296
ASR	1.209	1.138	0.468	1.150	1.000	0.084
Nature	1.547	1.525	0.727	1.831	2.040	0.135
PNAS	2.611	2.470	0.128	4.282	4.373	0.743
Science	1.419	1.449	0.389	1.835	1.851	0.859

**Notes:** Cumulative number of publications of authors on papers commented on by male and female commenters. Estimates from journal-specific OLS regressions of the outcome on an indicator that the author of the comment is a female and year fixed effects (year fixed effects excluded for AER and ASR panel b due to small sample). Panel b further restricts the sample to the last author on the paper. P-value of the difference from robust standard errors.



**John B. Holbein** @JohnHolbein

A junior scholar comes to you & asks your advice on whether it's worth it professionally to pursue a replication/reproduction of a prominent scholar's work that has a flaw.

What risks from doing a replication would you highlight, if any?

**John B. Holbein** @JohnHolbein · Jan 21, 2021

Replying to @JohnHolbein

[If possible, I'd like to keep this conversation as focused as possible on the personal/professional risks/rewards from doing a replication. Assume the hypothetical junior scholar knows that replication is valuable, often undervalued, and that the record should be corrected.]

**Chris Chambers** @chrisc77 · Jan 21, 2021

Replying to @JohnHolbein

It really boils down to risk: potentially make enemies depending on the personality of the researcher(s) you're attempting to replicate

reward: make new friends, likely get a well cited pub out of it

**Chris Chambers** @chrisc77 · Jan 21, 2021

Replying to @chrisc77 and @JohnHolbein

If they go ahead with it, would strongly recommend they use the @RegReports format to do it. High risk of file drawer effect otherwise.

**Chris Hilton** @Chris749 · Jan 21, 2021

Replying to @chrisc77 @JohnHolbein and @RegReports

Also, if the reason for potential repercussions is disagreement on the 'flaw', then a stage 1 review before data collection/analysis will help to mitigate this risk - do other experts agree on the problem?

**Chris Chambers** @chrisc77 · Jan 21, 2021

Replying to @Chris749 @JohnHolbein and @RegReports

Yes. I've got a slide on this from a talk...back in the heady days of going places and giving talks...

Also here is a link to a results-blind replication policy we offer at Royal Society Open Science [royalsocietypublishing.org/page/replicati...](https://royalsocietypublishing.org/page/replicati...)

**11. "How do Registered Reports support replication studies?"**

- Conspiracy of circumstances tells us not to bother doing direct (close) replications
  - Method sections are often too vague to allow precise replication
  - Chronic lack of power in novel research means that replications often require very large samples sizes
  - Attempting to exactly repeat a previous experiment can be seen in some fields (e.g. psychology) as an act of aggression (cf. physics)
  - Motivated reasoning by reviewers can impede publication
  - Many journals prioritise novelty and see replications as unpublishable
- RR's: have proposed replication experiment reviewed and provisionally accepted before you invest substantial resources into doing it; potentially involve original authors in peer review of the protocol; **motivated reasoning is prevented**

**Aaron Martin** @Aaron\_R\_Martin · Jan 21, 2021

Replying to @chrisc77 and @JohnHolbein

The idea of making enemies because a junior scholar designs to identify and correct an analytical problem in a published article is absurd.

**Chris Chambers** @chrisc77 · Jan 21, 2021

Replying to @Aaron\_R\_Martin and @JohnHolbein

Sure is.

**Austin Nichols** @AustnNichols · Jan 21, 2021

Replying to @JohnHolbein

Strongly depends on the nature of the flaw and its correction, but here are two good examples that I include in my Public Econ courses:

1. @rothstein\_jesse (2007) [eml.berkeley.edu/~jrothst/replic...](https://eml.berkeley.edu/~jrothst/replic...) led to changes in @stata merge commands
2. Deardorff (1976) led to [jstor.org/stable/2525720](https://jstor.org/stable/2525720)

**Jonathan Wai** @JonathanLWai · Jan 21, 2021

Replying to @JohnHolbein

@jsgoldstein, has done this

**Jessica Goldstein Holmes** @jgholmes · Jan 21, 2021

Replying to @JonathanLWai and @JohnHolbein

And wouldn't change it! Such a fantastic learning experience.

**Nicole Boyson** @nikir1

Replying to @JohnHolbein

Here's a con. A group of researchers replicated one of my papers with a prominent coauthor. They found an error in a Figure. They rewrote a whole paper stating that the error in the Figure rendered our paper null and void. 1/

**Nicole Boyson** @nikir1 · Jan 21, 2021

Replying to @nikir1 and @JohnHolbein

The journal called us in. We agreed with the error, but not with the paper being null and void since the figure was mostly motivational, and none of the underlying data in the regressions was incorrect. The regressions told the story much more than the figure. 2/

**Nicole Boyson** @nikir1 · Jan 21, 2021

Replying to @nikir1 and @JohnHolbein

We immediately wrote and published an erratum. But the whole process went on for like 3 years, with the journal deciding whether or not to publish this author's paper. The journal finally decided against it. 3/

**Nicole Boyson** @nikir1 · Jan 21, 2021

Replying to @nikir1 and @JohnHolbein

Two of the authors on the rebuttal paper were pretty junior. This process did not help their careers -- the waiting time alone was a nightmare for them. 4/

**Nicole Boyson** @nikir1 · Jan 21, 2021

Replying to @nikir1 and @JohnHolbein

If the flaw you see is material and you can write a paper that addresses it, I am all for it. If it's a bit of a "gotcha, your figure is wrong." I think it's a bad call. I have no bias against these authors and neither does my prominent coauthor, but it was a waste of time. 5/5

**Cameron Piercy, PhD** @cameronpiercy · Jan 21, 2021

Replying to @JohnHolbein

Short answer: Take the path of least resistance to establish your scholarship.

**Cameron Piercy, PhD** @cameronpiercy · Jan 21, 2021

Replying to @cameronpiercy and @JohnHolbein

Long answer: This is a question of both good science and getting published. Sounds like a wonderful idea for someone who does not have the pressure of a job-hunt/tenure clock. Save that idea and keep working on contributions which you expect will have stronger publication appeal.

**David Brady** @DaveBrady72 · Jan 21, 2021

Replying to @JohnHolbein

The most important criteria are whether the replication will improve collective understanding of an important social issue. If the topic is not actually important or if the replication doesn't really change understanding, life is too short. And keep it respectful.

**Graeme Blair** @graemedblair · Jan 21, 2021

Replying to @JohnHolbein

I've replicated and been replicated. I'd say do it!

To reduce bad feelings, be human and generous, avoid making assumptions about how error happened (you might be wrong!), & especially, depersonalize language—this is about paper's findings, not the authors you are replicating.

**Alexander Wuttke** @alex.wuttke@scienc... @Kunk... · Jan 21, 2021

Replying to @JohnHolbein

If they decide to replicate, maybe this is helpful

[cambridge.org](https://cambridge.org)

Replicate Others as You Would Like to Be Replic...  
Replicate Others as You Would Like to Be  
Replicated Yourself - Volume 54 Issue 2

**Macartan Humphreys** @maqartan · Jan 21, 2021

Replying to @JohnHolbein

I'd suggest they share the flaw with the authors and suggest they post an erratum. The junior scholar will get less credit but it seems to me to be a more constructive and collegial approach to science. Consider a standalone piece only if the original authors stonewall.

**Aaron Martin** @Aaron\_R\_Martin · Jan 21, 2021

Replying to @JohnHolbein

I tried this. The authors refused to provide the original data. Journal said it was not their responsibility to provide the data. The 1st version got an R&R, rejected after revisions, and it was very clear that Reviewer 2 was the author of the original text. tldr; nothing to gain

**Nicole Boyson** @nikir1 · Jan 21, 2021

Replying to @Aaron\_R\_Martin and @JohnHolbein

Sorry the original authors stonewalled. More journals are requiring original data be provided when you submit (or at least strongly encouraging it), and I applaud this approach.

**Colleen's eating popcorn** @ColleenByron · Jan 21, 2021

Replying to @JohnHolbein

From a practical standpoint it depends on who that prominent scholar is. Some embrace theory refinement, most see it as a personal attack. Not hard to tell who is who. If prominent means "towering" & petty run, don't walk, away.

**Colleen's eating popcorn** @ColleenByron · Jan 21, 2021

Replying to @ColleenByron and @JohnHolbein

If prominent means elder statesperson with an abiding affection for jrs, then replicate & even talk to said scholar about the research design & roll out. Those scholars are flattered & might boost you into the preeminent journal of your field by copublishing.

**Elliott Green** @ElliottDGreen · Jan 21, 2021

Replying to @JohnHolbein

IRRC the Journal of Applied Econometrics has a special section for replications from top economics journals, and thus replication can lead to publication without any bad feelings (cf. the JAE replication of the Nunn/Wantchekon AER slavery/trust paper).

**Chris Buddenhagen PhD** @CEBuddenhagen · Jan 21, 2021

Replying to @JohnHolbein

Addressing misconceptions should be foundational to science, learning and evidence based inquiry. Go for it.

**Elizabeth Wrigley-Field** @wrigleyfield@f... @ewrigL... · Jan 21, 2021

Replying to @JohnHolbein

I have a few unpublished drafts along these lines from when I was a grad student and postdoc. Advice I got was, "If you can make a positive point, instead of framing it as a critique, then make the positive point. If you can't... maybe do something else."

**Sebastian Karcher** @adam42smith · Jan 21, 2021

Replying to @JohnHolbein

N=1, but the experience Dave & I had with the reproducibility & extension of a number of papers for [doi.org/10.1111/squ.1...](https://doi.org/10.1111/squ.1...), while in grad school was very positive. This is pretty well cited & authors were incredibly nice about it, even opened some connections/doors 1/2

**Sebastian Karcher** @adam42smith · Jan 21, 2021

Replying to @adam42smith and @JohnHolbein

2/2 Notes 1) the reproducibility part was trivial: all tables reproduced push-button 2) None of the papers came apart on extension/correction; it just tweaked results 3) Our paper proposed and demonstrated a specific and constructive alternative, i.e. made an active contribution

**Paul Whaley** @dimplejug@mastodon... @paul... · Jan 21, 2021

Replying to @JohnHolbein

The risk of a bad reaction from the "rival" group can be reduced by making them aware of the work. So long as they aren't awful people, they won't mind - generally, people just don't like being ambushed. I have healthy somewhat adversarial relationships with lots of researchers.

**Paul Whaley** @dimplejug@mastodon... @paul... · Jan 21, 2021

Replying to @paul\_msg and @JohnHolbein

I kind of have to be OK at this because my job is, basically, to criticise the hell out of my peer-group's research. 😊

**D. Brian Blank, PhD** @FinProfBlank · Jan 21, 2021

Replying to @JohnHolbein

Warning: N=1 - We wrote a paper on a weakness (changed primary result and conclusion) of a JFQA. We published it in JCF, though the researchers were probably only slightly more prominent than us, so that may differ too much. [sciencedirect.com/science/article...](https://sciencedirect.com/science/article...)

The following media includes potentially sensitive content. [View](#)

**Fernando Martel García** @fmg\_twtr · Jan 21, 2021

Replying to @JohnHolbein

If their objective function is to achieve tenure etc then invest time on activities likely to result in publication in a high impact journal ASAP.

I am doubtful replication has the highest ROI in this respect. But everyone's mileage varies.

**Regina Bateson** @regina\_bateson · Jan 21, 2021

Replying to @JohnHolbein

Is the jr scholar a man or woman? I think the potential downsides are greater for a woman, who is more at risk of being labeled difficult, not nice, etc. Other dimensions of identity are relevant too (ex: race). Costs/benefits depend, in part, on power & status in the discipline.

**Jason Reifer** @JasonReifer · Jan 22, 2021

Replying to @JohnHolbein

What kind of flaw and how consequential is it?

**Will Lowe** @conjugaltepor · Jan 22, 2021

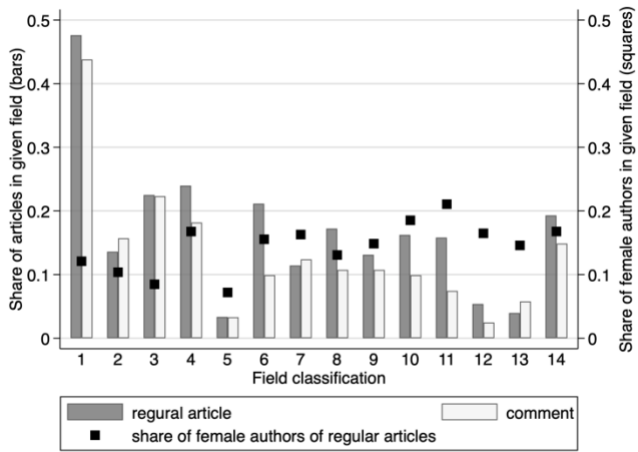
Replying to @JohnHolbein

The Powell Doctrine seems sensible: the flaw can be made obvious to the subfield, who also appreciate its importance, and can be eliminated with overwhelmingly plausible argument. And there's a plan for dealing with +0 publications, an uncorrected subfield, and +n enemies.

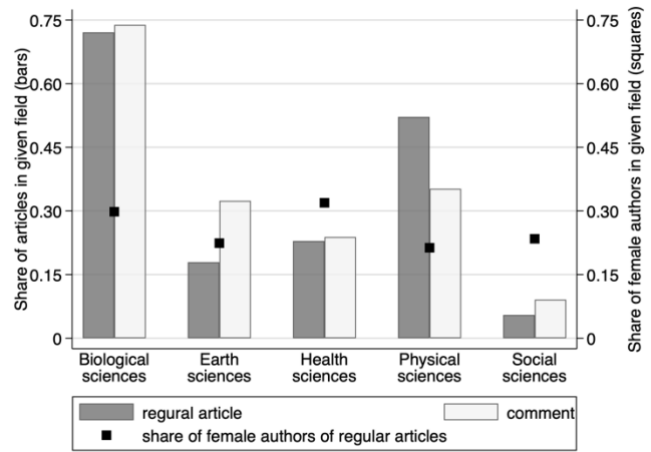
Figure A1 [Twitter thread](#)



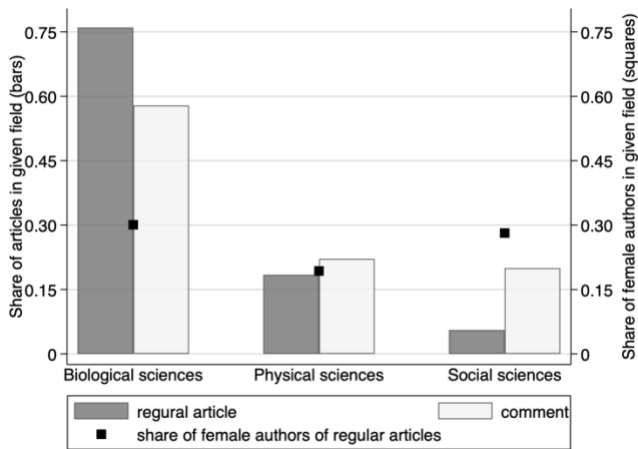
### AER



### Nature

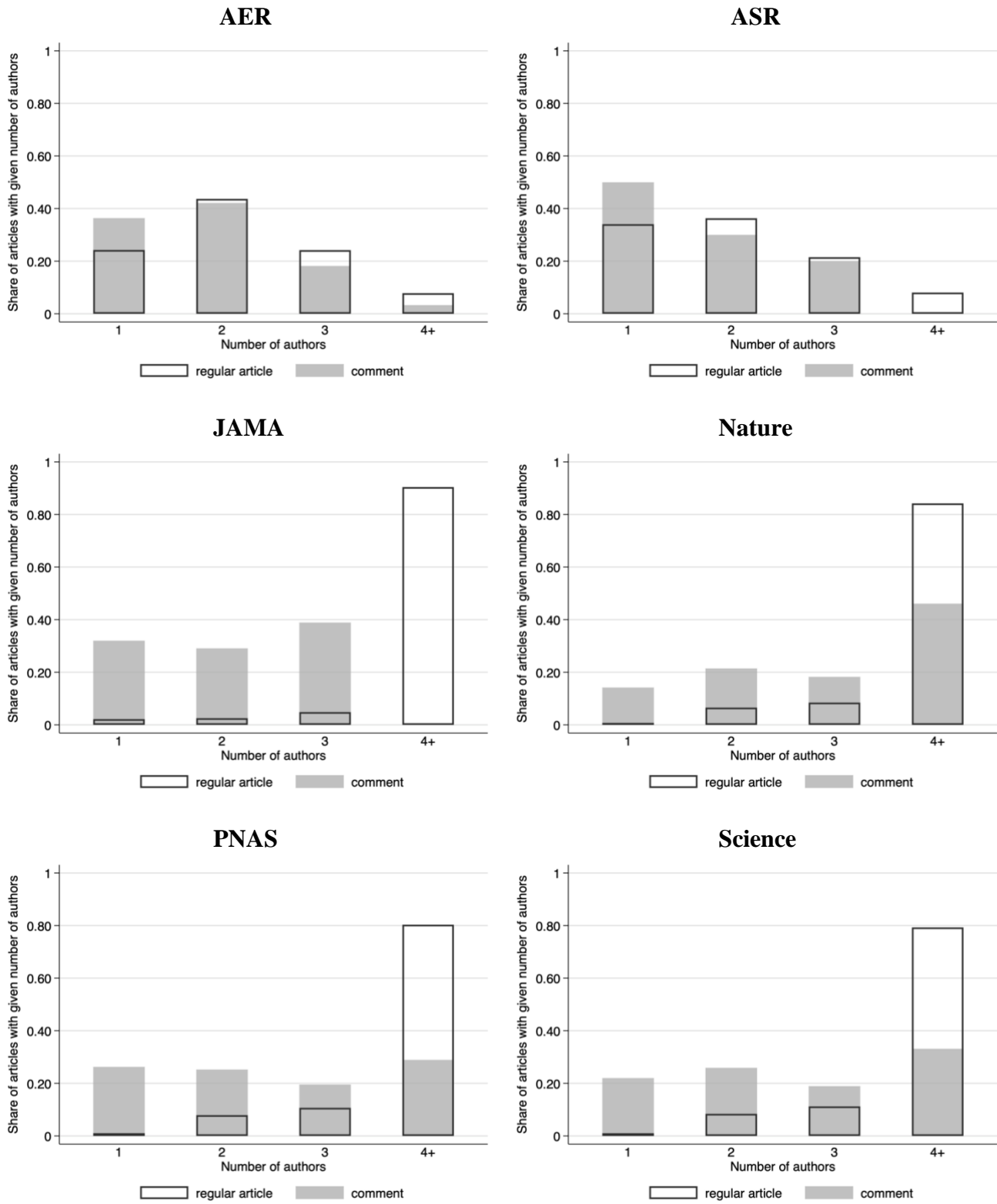


### PNAS

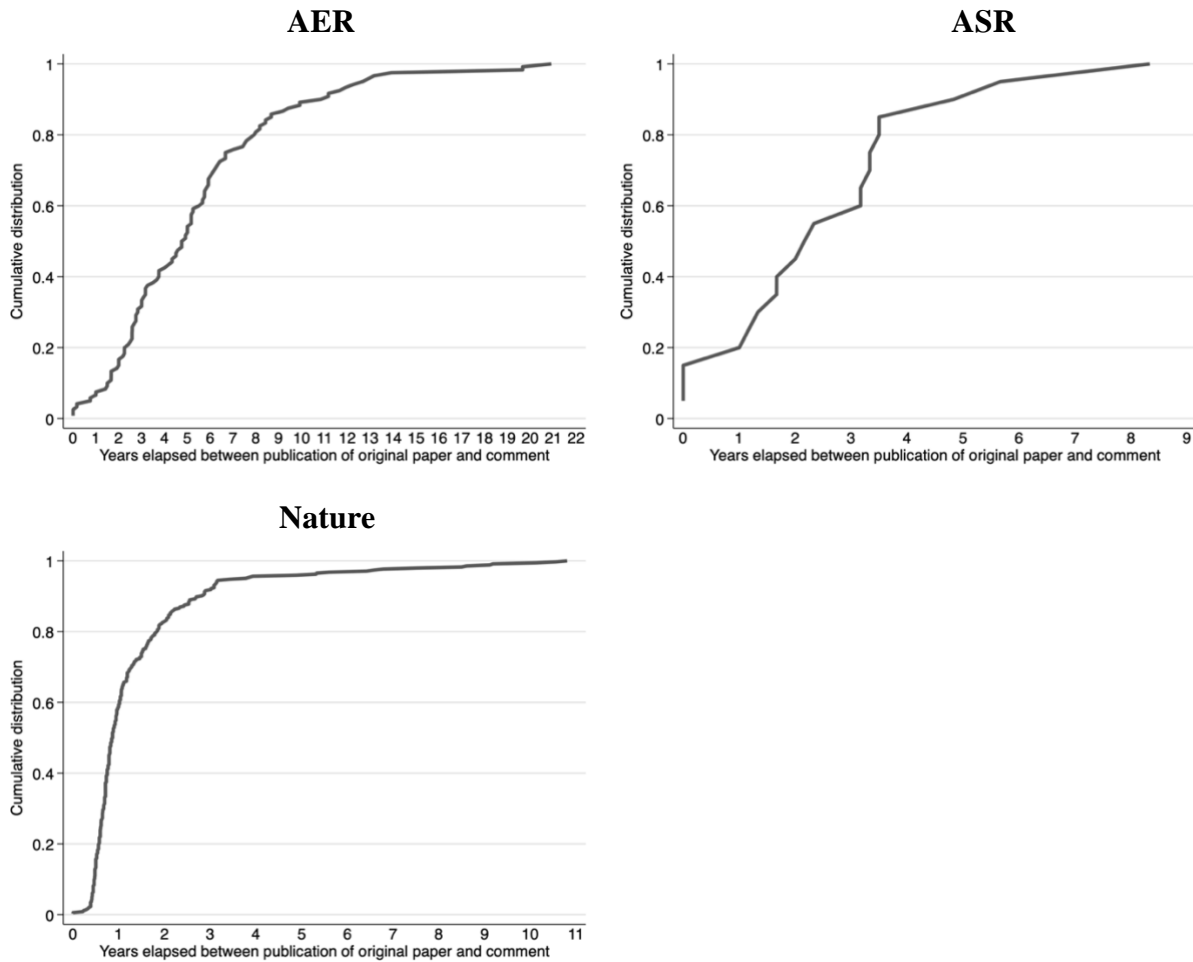


**Figure A2** Distribution of articles by field within journal

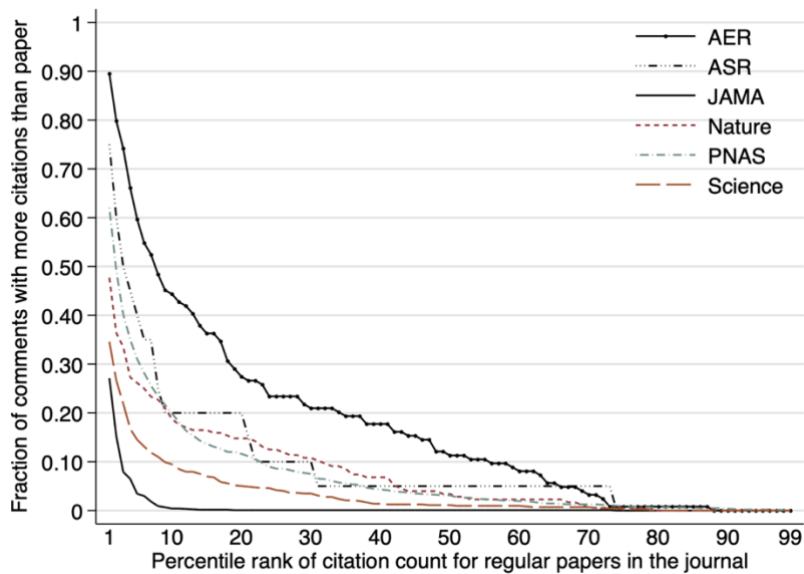
**Notes:** The bars show the share of regular papers (dark bars) and comments (light bars) in a given field in the journal. The squares indicate the share of female authors of regular papers in the given field. Fields in the ARE are: 1-microeconomics, 2-theory, 3-macroeconomics, 4-labor, 5-econometrics, 6-industrial organization, 7-international economics, 8-finance, 9-public economics, 10-health or urban economics, 11-development, 12-history, 13-laboratory experiments, and 14-other.



**Figure A3** Distribution of the number of authors in an article

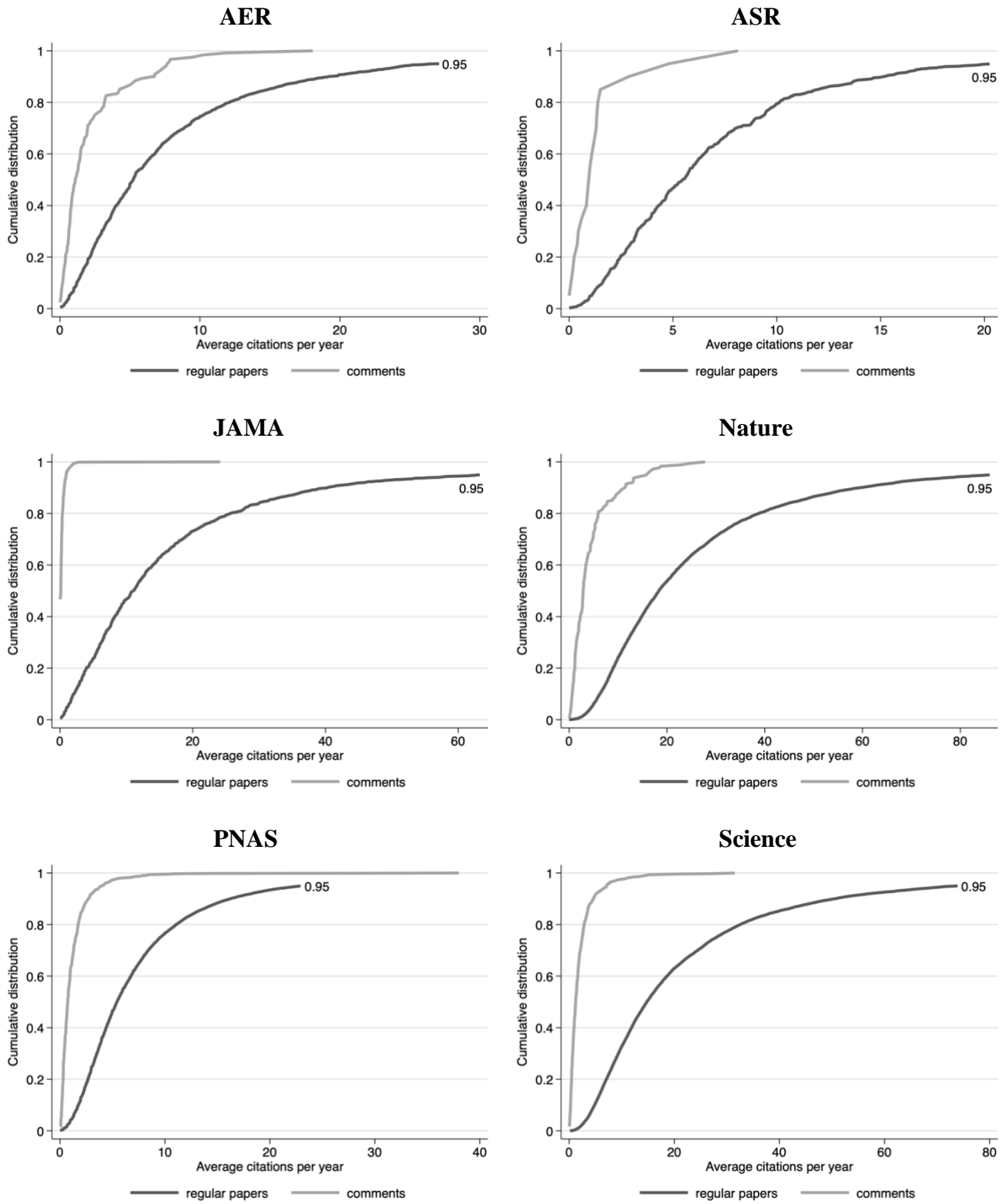


**Figure A4** Time lag between the publication of the original paper and the publication of the comment



**Figure A5** Citations to comments relative to citations to regular papers in the same journal

**Notes:** This figure plots the fraction of comments in a journal (y-axis) that receive more citations than those received by a regular paper in a given percentile of the distribution of all regular papers in the same journal (x-axis). Citations are average citations per year. Observations at the article level.



**Figure A6** Cumulative distribution of the average annual citations

**Notes:** The CDF for regular papers is truncated at 0.95 to help visualization.

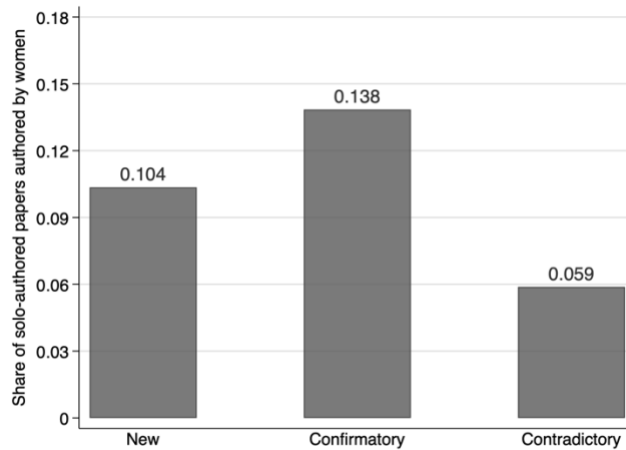
## B. Evidence from preprints

I construct the dataset by scraping the search results from bioRxiv's search engine. I search for all papers in each field-paper type combination. For each paper, I collect bibliometrics including title, names of authors, doi (which reveals posting date), field, and type of paper.

*Fields.*- The field of the paper is assigned by the author(s) at the time they post the paper on bioRxiv. Only one field can be assigned. They fields are animal behavior and cognition, biochemistry, bioengineering, bioinformatics, biophysics, cancer biology, cell biology, clinical trials, developmental biology, ecology, epidemiology, evolutionary biology, genetics, genomics, immunology, microbiology, molecular biology, neuroscience, paleontology, pathology, pharmacology and toxicology, physiology, plant biology, scientific communication and education, synthetic biology, systems biology, and zoology.

I then assign gender to authors using Genderize.io, as described above.

*Gender assignment.*- I assign gender to authors using genderize.io as detailed above.



**Figure B1** Share of female authors in solo-authored bioRxiv preprints

**Notes:** these values are not significantly different from each other, likely due the small number of solo-authored replications with gender assigned (65 *Confirmatory Results* and 51 *Contradictory Results*, see Table B4)

**Table B1** Descriptive statistics, bioRxiv preprints

	All Results		<i>New Results</i>		<i>Confirmatory Results</i>		<i>Contradictory Results</i>	
	N	Gender assigned	Share	Female	Share	Female	Share	Female
Anim Beh & Cog	3,646	0.715	0.959	0.315	0.018	0.431	0.023	0.345
Biochemistry	10,274	0.705	0.988	0.310	0.007	0.425	0.005	0.184
Bioengineering	5,904	0.644	0.987	0.269	0.011	0.215	0.002	0.417
Bioinformatics	25,638	0.683	0.977	0.234	0.020	0.237	0.003	0.122
Biophysics	10,357	0.635	0.984	0.239	0.012	0.217	0.004	0.093
Cancer Biology	16,673	0.686	0.992	0.356	0.007	0.315	0.001	0.111
Cell Biology	17,895	0.697	0.991	0.371	0.006	0.419	0.003	0.210
Clinical Trials	777	0.698	0.986	0.410	0.008	0.333	0.006	0.000
Dev Biology	9,690	0.676	0.991	0.390	0.008	0.370	0.002	0.278
Ecology	11,844	0.723	0.964	0.300	0.023	0.331	0.013	0.260
Epidemiology	7,910	0.646	0.954	0.361	0.038	0.367	0.008	0.246
Evo Biology	18,193	0.742	0.983	0.285	0.009	0.274	0.008	0.240
Genetics	34,551	0.702	0.984	0.356	0.008	0.336	0.009	0.273
Genomics	34,437	0.701	0.989	0.315	0.008	0.275	0.003	0.212
Immunology	12,329	0.695	0.992	0.380	0.006	0.312	0.002	0.375
Microbiology	28,607	0.687	0.987	0.369	0.010	0.414	0.003	0.321
Molec Biology	11,336	0.670	0.986	0.347	0.011	0.358	0.004	0.275
Neuroscience	57,859	0.697	0.979	0.312	0.015	0.332	0.006	0.297
Paleontology	378	0.761	0.984	0.228	0.000	.	0.016	0.167
Pathology	2,093	0.670	0.957	0.379	0.040	0.429	0.003	0.429
Pharma and Toxic	2,682	0.648	0.972	0.332	0.025	0.299	0.003	0.500
Physiology	4,412	0.683	0.978	0.346	0.015	0.269	0.007	0.300
Plant Biology	9,533	0.645	0.986	0.336	0.010	0.283	0.004	0.421
Sci Comm & Edu	1,980	0.694	0.943	0.392	0.048	0.302	0.008	0.125
Synthetic Biology	2,828	0.697	0.991	0.251	0.008	0.435	0.001	0.667
Systems Biology	8,650	0.715	0.986	0.267	0.012	0.269	0.002	0.278
Zoology	1,186	0.654	0.958	0.300	0.034	0.425	0.008	0.200

**Notes:** *N* is the number of paper-author observations with gender assigned, and *Gender assigned* is the fraction of unique authors assigned a gender. For each type of paper in bioRxiv (*New Results*, *Confirmatory Results*, and *Contradictory Results*), *Share* is the fraction of paper-author observations of that paper type, and *Female* is the fraction of paper-author observations that are female among paper-author observations with gender assigned.



**Table B2** Share of female authors in bioRxiv

	a. All authors			b. First author		
	1	2	3	1	2	3
<i>Confirmatory</i>	-0.001 (0.007)	0.005 (0.007)	0.008 (0.007)	-0.020 (0.017)	-0.011 (0.017)	-0.001 (0.017)
<i>Contradictory</i>	-0.061**** (0.011)	-0.060**** (0.011)	-0.053**** (0.011)	-0.142**** (0.022)	-0.139**** (0.022)	-0.126**** (0.022)
N	351,662	351,662	351,662	46,446	46,446	46,446
R <sup>2</sup>	0.0001	0.0087	0.0103	0.0007	0.0221	0.0275
Field FE	N	Y	Y	N	Y	Y
Year FE	N	N	Y	N	N	Y
Mean <i>New Results</i>	0.323	0.323	0.323	0.343	0.343	0.342

**Notes:** Coefficient estimates from OLS regressions of the author's gender on the type of paper, with *New Results* as the omitted category. Column 2 includes field fixed effects, and column 3 includes field and year fixed effects. Observations at the paper-author level. Estimates in hundreds of a percentage point. The sample includes all authors of all papers in Panel a, and the first author of each paper in Panel B. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table B3** Share of female authors in bioRxiv, observations at the paper level

	1	2	3
<i>Confirmatory</i>	-0.035**** (0.007)	-0.033**** (0.007)	-0.031**** (0.007)
<i>Contradictory</i>	-0.053**** (0.010)	-0.055**** (0.010)	-0.050**** (0.010)
N	68,975	68,975	68,975
R <sup>2</sup>	0.0009	0.0296	0.0330
Field FE	N	Y	Y
Year FE	N	N	Y
Mean <i>New Results</i>	0.208	0.208	0.208
<i>Confirmatory</i> vs. <i>Contradictory</i> p-value	0.131	0.065	0.108

**Notes:** Replication of Table B2 Panel a, with observations at the paper level. The share of female authors in a paper is the fraction of female authors over all authors in the paper, including authors with no gender assigned. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

**Table B4** Share of female authors in bioRxiv, solo-authored papers

	1	2	3
<i>Confirmatory</i>	0.035 (0.044)	0.032 (0.043)	0.033 (0.044)
<i>Contradictory</i>	-0.045 (0.034)	-0.045 (0.037)	-0.044 (0.037)
N	1,420	1,420	1,420
R <sup>2</sup>	0.0014	0.0402	0.0553
Field FE	N	Y	Y
Year FE	N	N	Y
Mean <i>New Results</i>	0.104	0.104	0.104

**Notes:** Replication of Table B2 on the sample restricted to solo-authored papers. Robust standard errors in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01, \*\*\*\*p<0.001.

### C. Experimental evidence

Beliefs elicitation.- After choosing whether to contest, P1 answers two belief elicitation questions. First, P1 states his belief that he will solve the task correctly, with a number from 0 to 100 percent. This question is unincentivized in order to avoid hedging. The unincentivized question is also easy for subjects to understand and fast to implement. Then, P1 indicates how many P2 participants he thinks solved the task correctly out of 20 randomly-chosen P2 participants other than the participant he is matched to. P1 receives \$1 for guessing correctly and \$0 otherwise. Incentivizing this belief is less problematic, since P1 does not control others' performance.

Deduct-a-\$1 treatment.- In this treatment, P1 is asked to choose one of two payoff tables. The tables describe the payoffs to P1 and P2 in each contingency, where a contingency is a combination between state *A* or *B* for P1 and states *C*, *D*, or *E* for P2. *A* or *B* are analogous to P1 solving the task correctly or incorrectly, respectively, in the *contest treatment*. *C*, *D*, or *E* are analogous to P2 solving the task correctly, solving it incorrectly and receiving \$0, or solving it incorrectly and receiving \$1 in the *contest treatment*, respectively. Payoffs in each contingency match the payoffs under the analogous contingency in the *contest treatment*. The two payoff tables differ from each other only in the payoffs in contingency (*A, E*), which is the contingency analogous to conditions (i) and (ii) holding in the *contest treatment*. P1 makes his choice over payoff tables without knowing the realized contingency. P1 never receives feedback on P2's realized state. The state probabilities are exogenous and shown to P1. I selected probabilities for *A*, *B*, *C*, *D*, and *E* to approximately match typical beliefs P1 held in the first wave of the *contest treatment*. (I collected data in two waves, as described below). Subjects were randomized into either the top or bottom set of state probabilities:

$$(p(A), p(B), p(C), p(D), p(E)) = \begin{cases} (0.50, 0.50, 0.50, 0.25, 0.25) \\ (0.80, 0.20, 0.76, 0.12, 0.12) \end{cases}$$

Implementation of the experiment.- I conducted the experiment on Prolific in August 2021. A total of 301 subjects participated in the *Contest treatment* and 188 in the *Deduct-a-\$1 treatment*. I collected the data in two waves. In the first wave, 209 subjects participated in the *Contest treatment*. I then used the performance beliefs elicited from these subjects to calibrate the contingency probabilities in the *Deduct-a-\$1 treatment*. In the second wave, conducted two weeks after the first, 188 subjects participated in the *Deduct-a-\$1 treatment*, and 92 subjects participated in a replication of the *Contest treatment*. The analysis in the main text pools observations across the two waves. The results replicate in each wave separately.

Incentives to contest treatments.- I conducted two additional treatments that introduce an incentive for P1 to contest. These treatments are identical to the *Contest treatment* described in the main text, except for the payoffs P1 and P2 receive if P1's choice to contest is implemented. In the *50-50 treatment*, if P1's choice to contest is implemented, the false positive \$1 is split equally between P1 and P2. That is, P1's earnings increase by 50¢ and P2's earnings are reduced by only 50¢. In the *Take \$1 treatment*, if P1's choice to contest is implemented, the false

positive \$1 is transferred entirely to P1. That is, P1's earnings increase by \$1 and P2's earnings are reduced by \$1. These treatments were administered in an across-subject design as a part of the same study involving the *Contest treatment* described in the main text. A total of 504 subjects participated in the *50-50 treatment* (293 in the contest condition and 211 in the matrices condition), and 497 subjects participated in the *Take \$1 treatment* (298 in the contest condition and 199 in the matrices condition). In these treatments, men continue to choose to contest at a significantly higher rate than women, although now this gender gap is entirely explained by gender differences in distributional preferences (Figure C1).

**Table C1** Payoff tables in the *Deduct-a-\$1 treatment*

<i>Payoff Matrix 1</i>				<i>Payoff Matrix 2</i>					
		P2					P2		
		<i>C</i>	<i>D</i>	<i>E</i>			<i>C</i>	<i>D</i>	<i>E</i>
Subject	<i>A</i>	( 1, 1 )	( 1, 0 )	( 1, 1 )	Subject	<i>A</i>	( 1, 1 )	( 1, 0 )	( 1, 0 )
	<i>B</i>	( 0, 1 )	( 0, 0 )	( 0, 1 )		<i>B</i>	( 0, 1 )	( 0, 0 )	( 0, 1 )

**Notes:** Subjects choose between Payoff Matrix 1 and Payoff Matrix 2. Values in parentheses are dollar payoffs to the self (left) and the passive participant P2 (right) in each contingency. Matrices vary only in the payoff to P2 in contingency *AE*.

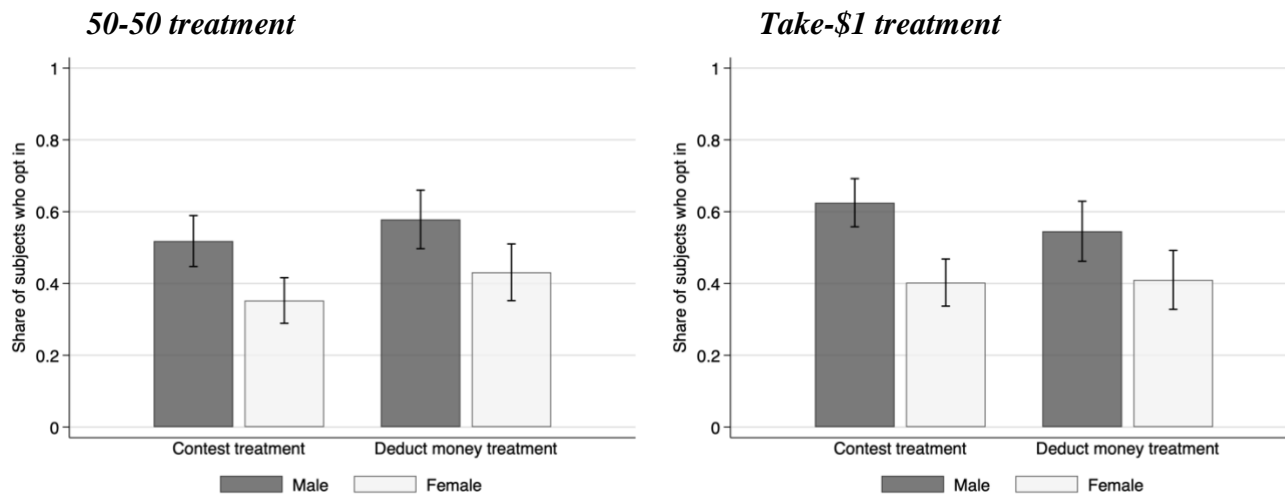
**Table C2** Descriptive statistics, experiment

	Contest treatment			Deduct-a-\$1 treatment		
	Men	Women	p-val diff	Men	Women	p-val diff
Age	34.2	24.4	0.000	31.0	24.2	0.000
Race.Asian	0.05	0.08	0.234	0.10	0.16	0.272
Race.Black	0.08	0.03	0.044	0.09	0.07	0.525
Race.Latino	0.03	0.06	0.189	0.04	0.10	0.110
Race.White	0.74	0.70	0.541	0.63	0.59	0.539
Race.Other/Mix	0.10	0.12	0.597	0.13	0.09	0.341
Residence.Northeast	0.16	0.29	0.006	0.20	0.19	0.794
Residence.South	0.38	0.30	0.141	0.33	0.28	0.468
Residence.Midwest	0.19	0.19	0.974	0.26	0.24	0.866
Residence.West	0.23	0.18	0.225	0.18	0.27	0.172
Educ.HS or less	0.08	0.13	0.144	0.17	0.19	0.784
Educ.Some college	0.14	0.37	0.000	0.20	0.32	0.065
Educ.College or more	0.77	0.50	0.000	0.62	0.49	0.065
High school in US	0.96	0.96	0.940	0.95	0.97	0.548
Correct first round	0.62	0.54	0.193	-	-	-
Correct additional round	0.64	0.58	0.255	-	-	-
Belief self	81.75	76.28	0.011	-	-	-
Belief others	14.46	13.18	0.011	-	-	-
N subjects	159	142	-	98	90	-

**Table C3** Choice to opt in in the *Contest treatment* and the *Deduct-a-\$1 treatment*

	Contest treatment		Deduct-a-\$1 treatment		Both treatments	
	(1)	(2)	(1)	(2)	(1)	(2)
Female	-0.241**** (0.050)	-0.191**** (0.053)	-0.119*** (0.043)	-0.108** (0.044)	-0.119*** (0.043)	-0.088* (0.046)
Contest					0.239**** (0.054)	0.213**** (0.054)
Female x Contest					-0.122* (0.066)	-0.115* (0.067)
Belief self		0.001 (0.002)				
Belief others		0.002 (0.007)				
Correct answer		-0.004 (0.054)				
Indifferent		0.060 (0.100)				
Prob. treatment				-0.012 (0.046)		
Controls	N	Y	N	Y	N	Y
R <sup>2</sup>	0.0702	0.1670	0.0371	0.1191	0.1060	0.1553
N	301	301	188	188	489	489

**Notes:** Coefficient estimates from OLS regressions of the probability of opting in; that is, choosing to contest in the *Contest treatment* or to deduct the \$1 in the *Deduct-a-\$1 treatment*. Panel *Both treatments* pools the data across both treatments. Column (1) includes only a female indicator. In panel *Contest treatment*, column (2) adds beliefs of own performance, beliefs of others' performance, an indicator that the subject's answer is correct in the additional incentivized round, an indicator that the subject anticipates zero probability of answering the task correctly or anticipates that 20 out of 20 other subjects answered the task correctly. In panel *Deduct-a-\$1 treatment*, column (2) adds *Prob. treatment*, which is an indicator of the treatment condition in which the contingency probabilities are  $p(A) = 0.5$ ,  $p(B) = 0.5$ ,  $p(C) = 0.5$ ,  $p(D) = 0.25$ ,  $p(E) = 0.25$ . Controls are the subject's region of residence, an indicator for white race, educational attainment categories, age categories, and experience with Prolific categories. Robust standard errors in parentheses. \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ , \*\*\*\* $p < 0.001$ .

**Figure C1** Take-up of contest option and equivalent payoff matrix

**Notes:** Replication of Figure 4 for treatments with incentives for P1 to contest.

## References

- Card, D., and DellaVigna, S., 2013. Nine facts about top journals in economics. *Journal of Economic Literature*, 51.1: 144-161.
- Chen, D.L., Schonger, M., and Wickens, C., 2016. oTree—An open-source platform for laboratory, online, and field experiment. *Journal of Behavioral and Experimental Finance*, 9.1: 88-97.
- Fontanarosa, P., Bauchner, H., and Flanagin, A., 2017. Authorship and team science. *Journal of the American Medical Association*, 318.24: 2433-2437.
- Heckman, J.J., and Moktan, S., 2020. Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, 58.2: 419-470.
- Huang, J., Gates, A.J., Sinatra, R., and Barabási, A.L., 2020. Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences of the United States of America*, 117.9: 4609-4616.
- Lerchenmueller, M.J., Sorenson, O., and Jena, A.B., 2019. Gender differences in how scientists present the importance of their research: Observational study. *bmj*, 367.
- Perry, M., and Reny, P.J., 2016. How to count citations if you must. *American Economic Review*, 106.9: 2722-2741.
- Radicchi, F., Fortunato, S., and Castellano, C., 2008. Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, 105.45: 17268-17272.
- Ray, D., and Robson, A., 2018. Certified random: A new order for coauthorship. *American Economic Review*, 108.2: 489-520.
- Sekara, V., Deville, P., Ahnert, S.E., Barabasi, A-L, Sinatra, R., and Lehmann, S., 2018. The chaperone effect in scientific publishing. *Proceedings of the National Academy of Sciences of the United States of America*, 115.50: 12603-12607.