Are economics conferences gender-neutral? Evidence from Ireland*

Margaret Samahita† Kevin Devereux‡

July 6, 2022

Abstract

We study gender inequality in conference acceptance using data from the Irish Economic Association annual conference from 2016 to 2022. While no gender gap is observed in the organisers' acceptance decisions, there is indication of in-group gender bias at the reviewer stage. In particular, male reviewers persistently give higher scores to male-authored papers. Evidence suggests that the bias stems from unconscious stereotyping against lesser-known female authors. Anonymisation eliminates the in-group bias of male reviewers but introduces a bias in favour of male authors for female reviewers.

JEL codes: A1, D83, J16, J71

Keywords: academic conference, discrimination, gender

*We thank Míde Griffin for her excellent research assistance and the Irish Economic Association conference organisers 2016-2022 for generously sharing their submissions data. We also thank seminar participants at University College Dublin, NUI Galway, the Central Bank of Ireland and the Irish Economic Association annual conference for helpful comments. This work was supported by the Irish Economic Association and the Irish Society for Women in Economics.

†School of Economics and Geary Institute for Public Policy, University College Dublin. E-mail: margaret.samahita@ucd.ie

‡School of Economics, Peking University, China. E-mail: kevin.devereux@pku.edu.cn

1 Introduction

Economists are increasingly interested in studying the gender disparity in their profession, in an effort to improve gender equality. Papers have documented a gender gap in research output, which is partly explained by differences in collaboration network (Ductor et al., forth.), invitation to present at seminars (Doleac et al., 2021) and treatment during these seminars (Dupas et al., 2021), teaching evaluation (Mengel et al., 2019; Boring, 2017; Boring and Philippe, 2021), recognition for group work (Sarsons et al., 2021), citations (Koffi, 2021), the publication process (Card et al., 2020; Hengel and Moon, 2020; Hengel, forth.), as well as evaluation by other economists in an online professional forum (Wu, 2020) and reference letters (Eberhardt et al., 2022)—all of which may contribute to the underrepresentation of females in economics, especially at the top of the profession (Bayer and Rouse, 2016; Ginther and Kahn, 2021; Lundberg and Stearns, 2019).

In this paper, we focus on acceptance into an academic conference to study whether and at which stage of the review process gender bias is detected. Acceptance to conferences is an important input in the researcher production function as it allows researchers to disseminate their work to a larger audience. In previous findings, Chari and Goldsmith-Pinkham (2017) find no gender disparity in NBER conferences while Hospido and Sanz (2021) find bias against female authors in three European economic conferences. We contribute new evidence from Ireland using a novel dataset of all submissions into the Irish Economic Association (IEA) annual conference from 2016 to 2022. We complement this dataset with additional data on all authors' PhD year to proxy for experience. Across the six years studied (excluding 2020 when the conference was cancelled), around 61% of papers are accepted.

Each submission is sent to a reviewer who assigns an acceptance score (1-5) and rates the paper in four components of quality (technical merit, readability, originality, and, from 2017 onward, relevance—collectively referred to as *component scores*). While reviewers do not make the acceptance decisions, these scores are communicated to the conference organizers who take the information into account but may exercise discretion in their final

decision, for example to achieve a balance in gender, institutions represented or other dimension.

Prior to 2021, submissions are not blind; organisers and reviewers can infer the gender of the author(s), thus enabling us to study whether organisers have different likelihoods of accepting male- and female-authored papers, and if reviewers assign acceptance and component scores differently. Our data reveals that conference organisers are gender-neutral: the proportion of male authors on a paper does not significantly affect the likelihood of acceptance. Similarly, the acceptance scores given by reviewers, which may influence organisers' decisions, are not significantly different for an increasing share of male co-authors on the paper. Hence, overall, organisers are successful in achieving gender balance at the conferences.

Since the above finding is based on analyses aggregating all reviewers, we next explore whether reviewers' gender matters for their decisions given that one of the suggested explanations for gender bias in conference acceptance is male reviewers' stereotype against female authors (Hospido and Sanz, 2021). We find consistently sizeable effects indicating that reviewers indeed display an in-group gender bias: female(male) reviewers give higher acceptance scores to papers submitted by a higher share of female(male) authors. However, a large part of this bias is explained by the quality of the paper, as captured by its component scores. We therefore proceed by using component scores as an outcome variable. While we do not observe a gender gap in the scores assigned by female reviewers, we find that male reviewers consistently give lower scores to female-authored papers.

To explore the possibility that the gender bias is due to an information asymmetry driven by less well-known authors, we study several indicators for being well-known: having a higher share of authors at an Irish institution (or one with an active working paper series which proxies for a larger Irish institution), who have a PhD, with a Professor title, and the total number of years post-PhD for all authors. The results are consistent with the information asymmetry explanation: where the paper authors are less well-known, male reviewers display a bias in favour of an increasing share of male authors—this bias is reduced with an increase in familiarity. No such bias is observed for female reviewers.

Finally, we check whether reviewers' *perception* of information asymmetry can help explain the gender gap by male reviewers. We do not observe that reviewers are more confident when evaluating male-authored papers. Neither is reviewer confidence associated with a smaller gender gap for male reviewers. Hence, the observed gender gap appears to stem from unconsciously implementing a stereotype against female-authored papers who are less well-known.

In 2021, the organisers moved to a blind submission process, thus allowing us to study whether the above gender gap persists despite anonymisation. While, as expected, male reviewers' gender bias disappears, for female reviewers the blinding surprisingly *introduces* a male author bias that is greater than their initial in-group bias. As a result, blinded papers are scored higher by female reviewers when there is an increasing share of male authors. While we hesitate to speculate given the relatively small number of observations driving the bias, we note that this additional bias is not sufficient to affect acceptance rates.

Our finding of no *aggregate* gender gap in acceptance contrasts with Hospido and Sanz (2021), the closest paper to ours, which finds a bias against female-authored papers despite a similar share of male reviewers (76% vs our 79%). This may reflect the organisers' conscious decision to ensure gender neutrality. Our mixed findings on the impact of blinding on gender discrimination contrast with other studies. Some have found that blinding reduces discrimination (Goldin and Rouse, 2000; Ross et al., 2006; Åslund and Skans, 2012; Tomkins et al., 2017) or that it has no effect in the context of an environmental economics conference (Carlsson et al., 2012). Other findings suggest that when submissions are blinded, those reviewers who would like to favour females no longer can (Krause et al., 2012a,b; Behaghel et al., 2015)—or in other words, when submissions are *not* blinded reviewers make an effort to reduce their bias.

The remainder of the paper proceeds as follows. Section 2 provides some background about the conferences and describes the data. We present results in Section 3 and Section 4 concludes.

2 Background and Data

The IEA Annual Conference is the main event of the Irish Economic Association and takes place in May at different institutions around Ireland.[1] In addition to contributed sessions, the conference features two international plenary lectures. Both national and international economists can submit papers for presentation, a subset of which are selected to be included in the conference's program.[2] Selection is made by a team of reviewers which consists of around fifty economists from Irish institutions, the board members of the IEA and the conference organisers. Each paper is allocated to one reviewer according to field of research, and each reviewer reviews on average three papers per conference—although the local organisers typically review more. Between 2016-2022, the average yearly number of submissions is 150 and around 61% are accepted.

We were able to collect each year's submissions data from the local organisers. Combined, there are 902 observations, each of which represents a single submission of a paper. The dataset includes the submission details input by the submitter: paper title and abstract, primary field,[3] secondary field (optional), whether the submission is a student paper, and the submitter details: title, full name, email, institution, country, for the submitter and all other co-authors. The dataset also includes the reviewer assigned to each paper, their score between 1 (unacceptable) to 5 (excellent) for the paper's technical merit, readability, originality and relevance to the conference, and their confidence within the knowledge area (1-5).[4] The reviewer also gives the paper an overall acceptance score (1-5) with a

[1] The 2020 conference was cancelled due to COVID-19. The local organisers thus moved the event to be held virtually in 2021. In 2022 the conference was back to an in-person format.

[2] In 2016-2018, there were a number of special sessions. Dropping the 24 papers accepted as part of these sessions yields similar results.

[3] Agricultural and natural resource economics; Economic development and growth; Economic history; Economic systems; Energy economics; Financial economics; Health, education, and welfare economics; History of economic thought; Household finance and consumption; Industrial organisation; International economics; Labour/demographic economics; Macroeconomics; Microeconomics; Public economics; Regional/real estate/transport economics and Teaching economics.

[4] Reviewer score for the paper's relevance and their confidence are only available from 2017 onwards. Reviewer scores are missing for 47 submissions in 2022 where scores are not recorded, these are last minute reviews done offline by the organisers and colleagues due to assigned reviewers not submitting their reviews on time.

short comment to the organisers. A higher score increases the likelihood of acceptance, a decision that is made by the organisers and which we also observe in the data. The gender of each submitting author and reviewer is coded manually by a research assistant and double-checked by us manually.[5] We also cross-check the submitter gender with the title stated when submitting (Mr, Mrs or Ms) if available. Additionally, the gender of all other co-authors is coded manually and cross-checked with their title if available.

The final decision whether to accept or reject the paper is made by the conference organisers, however reviewers are highly influential with the likelihood of acceptance increasing with reviewers' scores from 0% (reviewer score of 1), 7% (2), 28% (3), 68% (4), and 90% (5). Organisers exercise some discretion in order to get an even balance among the different institutions and may reject a paper with a high score when the submitter has submitted multiple papers.

We complement the above conference dataset with additional data on researcher experience. We do a manual search of the submitter and all co-authors using personal and institutional webpages, CV and LinkedIn to collect data on whether they have a PhD and the PhD graduation year. We cross-check this information with the stated title at submission (Dr or Prof), which we use to assign the PhD dummy in case of missing data.[6]

Up until 2019, submissions are not blind: reviewers are able to see the name, and thus infer the gender, of each submitter and any co-author(s). In 2021 and 2022 the IEA implemented a blind submission system whereby the submission platform concealed the author names from the reviewer. However, reviewers who are also the conference organisers could still see the names of all authors—such submissions are thus coded as unblinded, along with all submissions from 2016-2019.[7]

[5]Gender data is missing for five reviewers who wish to remain anonymous.

[6]We also collect data on paper quality: number of citations up to the year before the conference, total number of citations, publication and the journal ranking as at 31 December 2021. Controling for the citation up to the year before the conference produces qualitatively the same results. For data up to 2021, where sufficient time has passed to control for total citation and publication, including these controls also does not change our results.

[7]We also code as unblinded the 17% of submissions in 2021 where the authors did not remove their names from the pdf submitted. Results are qualitatively similar when dropping all unblinded observations from 2021 and 2022.

3 Results

3.1 Is paper acceptance into the conference gender-balanced?

We first present summary statistics on the characteristics of authors, papers and reviewers in Table 1. Around 70% of submitters are male and the same proportion have a PhD. Around 10% are Professors, 60% are affiliated with an Irish institution and 40% at an Irish institution with an active working paper (WP) series, which proxy for relatively larger institutions in Ireland.[8] When looking at the share of authors on a paper with the above characteristics, the proportions are similar for males and PhDs but a higher share of authors are Professors (over 15%) and fewer are affiliated with an Irish Institution (just over 50%, or around 35% at an Irish institution with a WP).

The sample of authors who submit in the blinded and unblinded years are different. Those submitting in the blinded years appear to be more experienced with the submitter having almost 9 years post-PhD vs 6 in the unblinded case. The total experience of all authors on the average paper, obtained by summing years post-PhD for all authors, is also higher for the blinded case (16 vs 12) and there are more authors on the average paper (2.2 vs 2.0).

Papers tend to score just under 4 in all dimensions. The proportion of male reviewers is significantly higher for unblinded papers than blinded papers (81% vs 69%) and overall acceptance rate is also significantly higher for blinded papers (72% vs 68%). This is driven by the 2021 conference which was held online due to COVID-19, meaning that it was able to accommodate more presentations than offline (78% acceptance rate). At the same time, possibly due to academics' difficulty in juggling commitments during the pandemic, the total number of submissions (132, 29% by females) was lower than in previous years.[9]

Figure 1 displays the trend in paper submission and acceptance by submitter gender.

[8]The five institutions are the Central Bank of Ireland, the Economic and Social Research Institute, Maynooth University, Trinity College Dublin and University College Dublin.

[9]Consistent with the narrative that the pandemic had a disproportionately negative effect on junior female researchers, the female submitters in 2021 were more experienced than in previous years, as measured by the number of years post PhD (8.12 vs 4.14, t-test, $p = 0.0094$) and the proportion who are Professors (10.5% vs 4.9%, two-sample test of proportion, $p = 0.1780$).

Table 1: Summary statistics

	Unblinded		Blinded		Difference		
Variable	Mean	SD	Mean	SD	Diff	SE	N
Submitter male	0.713	(0.453)	0.675	(0.469)	-0.038	(0.037)	902
Submitter PhD	0.723	(0.448)	0.716	(0.452)	-0.007	(0.036)	902
Submitter years post-PhD	5.605	(8.147)	8.968	(8.094)	3.363***	(0.806)	679
Submitter Professor	0.100	(0.301)	0.108	(0.311)	0.008	(0.025)	902
Submitter at Irish inst	0.593	(0.492)	0.619	(0.487)	0.025	(0.040)	902
Submitter at WP inst	0.410	(0.492)	0.392	(0.489)	-0.018	(0.040)	902
Number of authors on paper	1.963	(1.068)	2.227	(1.230)	0.264***	(0.090)	902
Share of authors male	0.731	(0.361)	0.690	(0.377)	-0.041	(0.030)	902
Share of authors PhD	0.738	(0.386)	0.759	(0.358)	0.022	(0.031)	902
All authors years post-PhD	12.391	(15.543)	15.763	(15.791)	3.372***	(1.264)	902
Share of authors Professor	0.150	(0.277)	0.186	(0.282)	0.036	(0.023)	902
Share of authors at Irish inst	0.528	(0.455)	0.542	(0.446)	0.013	(0.037)	902
Share of authors at WP inst	0.363	(0.435)	0.346	(0.438)	-0.017	(0.035)	902
Student paper	0.219	(0.414)	0.237	(0.426)	0.018	(0.034)	902
Score: Technical merit	3.717	(0.919)	3.743	(0.934)	0.026	(0.077)	855
Score: Readability	3.772	(0.920)	3.902	(0.915)	0.129*	(0.077)	855
Score: Originality	3.534	(0.905)	3.541	(0.912)	0.007	(0.076)	855
Score: Relevance	3.811	(0.997)	3.956	(0.895)	0.145*	(0.084)	701
Score: Acceptance	3.859	(1.151)	3.918	(1.153)	0.059	(0.096)	855
Reviewer confidence	3.770	(0.848)	3.776	(0.811)	0.006	(0.072)	701
Reviewer male	0.812	(0.391)	0.691	(0.463)	-0.122***	(0.033)	897
Paper accepted	0.578	(0.494)	0.722	(0.449)	0.144***	(0.039)	902
Observations	708		194		902		

Notes: Summary statistics of submissions.

While the majority of papers, between 66-76%, are submitted by males each year, the acceptance rates of male and female submitters are similar (60% vs 60%, two-sample test of proportion, $p = 0.5396$) and in no year is the difference statistically significant. While the final decision to accept or reject a paper by the conference organisers appears to be gender-neutral, this decision may involve some discretion to achieve gender balance in the conference. Hence, we next look at the acceptance scores assigned by reviewers. As shown in Figure 2, male and female submitters receive similar acceptance scores from reviewers (3.871 vs 3.873, t-test, $p = 0.9834$) and in no year is the difference statistically significant (two-sample t-tests).

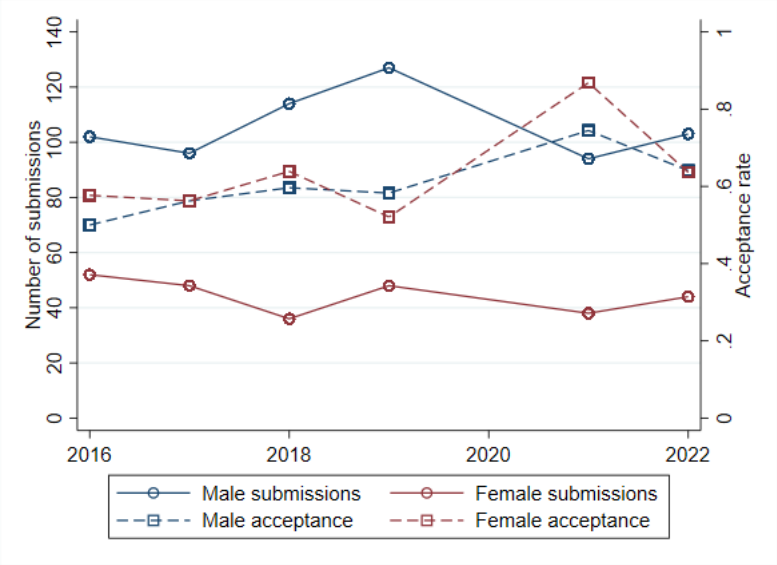

Figure 1: Trends in paper submission and acceptance by submitter gender

We explore other dimensions of diversity such as the field each submission falls into. Figure 3 shows the acceptance rate for various fields and the corresponding number of submissions. Excluding extremely small fields with fewer than 10 submissions, the acceptance rate varies from 39% (Agriculture and natural resource economics) to 75% (Economic history). Figure 4 plots the average acceptance scores for male and female submitters by their paper's primary field, sorted in order of increasing share of male reviewers. Although all fields have a majority of male reviewers, there is no clear trend showing that the gender share of reviewers in a particular field affects the acceptance

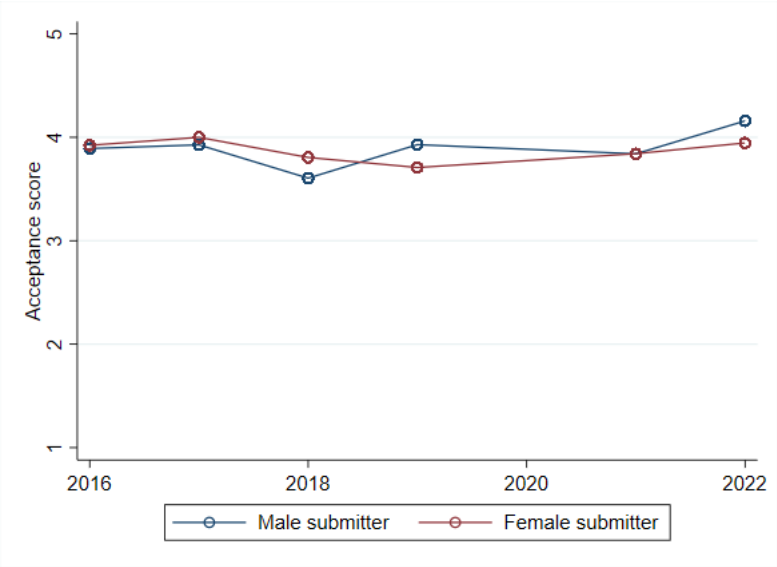

Figure 2: Paper acceptance score by submitter gender

scores given to male and female submitters differently.

Acceptance rate is higher for a non-student paper relative to a student paper (64% vs 48%, $p < 0.0001$), submitters with a PhD than without (63% vs 54%, $p = 0.0107$), submitters with a Professor title than without (74% vs 59%, $p = 0.0068$) and submitters from an Irish vs non-Irish institution (75% vs 40%, $p < 0.0001$). For submitters from an Irish institution, acceptance rate is higher for institutions with an active WP series than without (81% vs 63%, $p < 0.0001$).

To support our finding of no gender gap in acceptance, we proceed by considering the following linear probability model:

$$Accept_{iy} = \beta_1 ProportionMale_{iy} + \alpha_y + \varepsilon_{iy} \tag{1}$$

where Accept is a dummy variable which equals 1 if submission i in year y is accepted into the conference. We also use each paper's acceptance (1-5) as determined by the reviewer as an alternative outcome variable.[10] ProportionMale is a continuous variable between 0 and 1 indicating the share of male authors on the paper. α_y is year fixed effects,

[10]Results are robust to logit and ordered logit specifications, respectively.

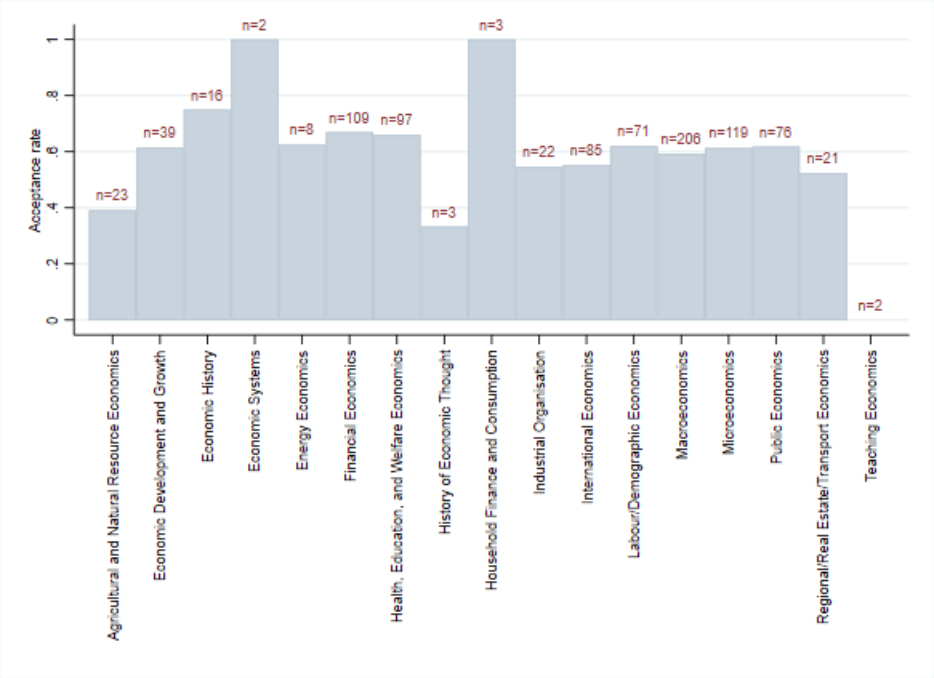

Figure 3: Paper acceptance by primary field

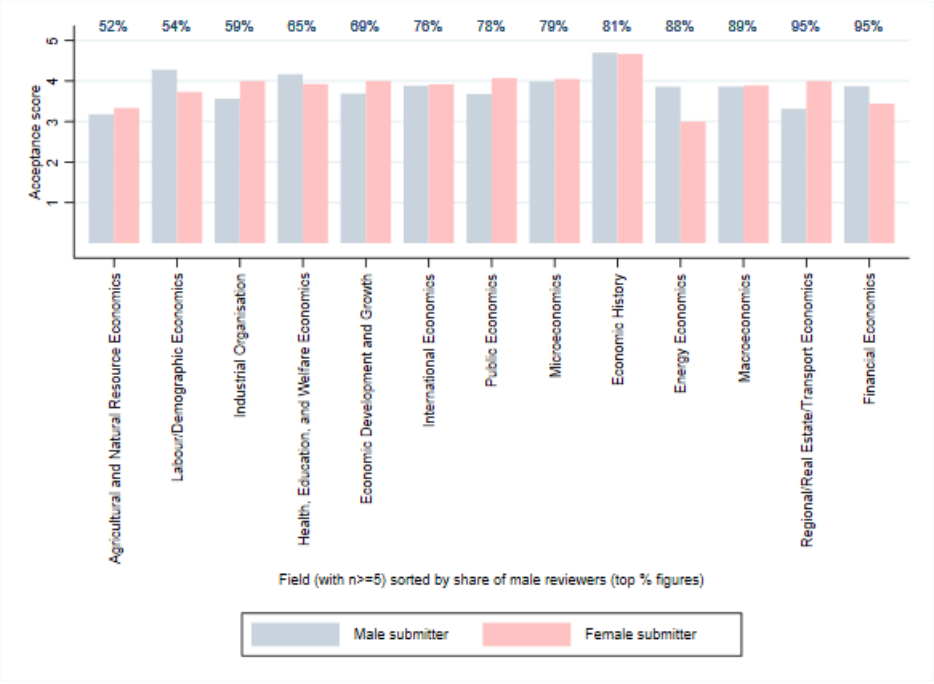

Figure 4: Paper acceptance score by gender and primary field in increasing share of male reviewers

ε_{iy} is an error term, and β_1 is the parameter of interest. The unit of observation is an individual submission.

We also investigate whether blinding has any effect on paper acceptance by testing the following model:

$$Accept_{iy} = \beta_1 ProportionMale_{iy} + \beta_2 Blinded_{iy} + \beta_3 ProportionMale \times Blinded_{iy} + \alpha_y + \varepsilon_{iy} \quad (2)$$

Blinded is a dummy variable which equals 1 if the author names are blinded on the paper (this applies to all submissions in 2021-2022 reviewed by someone other than the conference organisers). The parameter of interest is $\beta_1 + \beta_3$ which indicate the extent of any gender bias despite blinding. In additional specifications we also include field fixed effects; a set of controls—including the number of authors on the paper, the proportion of authors with a PhD, the proportion of authors affiliated with an Irish institution and the proportion of authors affiliated with an Irish institution with WP—; reviewer scores on the paper's technical merit, readability, originality, and relevance (2017 onward); and reviewer fixed effects.

The results of model (1) are in columns (1) and (7) of Table 2, showing no significant effect of gender on the likelihood of acceptance nor on the reviewers' acceptance scores. These confirm the above descriptive analyses whereby organisers and reviewers do not appear to discriminate between male and female submitting authors.

Given the null effect of gender on organisers' acceptance decision, it is unsurprising that blinding also has no effect as shown in columns (2-6). However, when considering reviewers' acceptance scores, blinding is associated with around 0.5 higher scores (on a scale from 1-5) given to male authors, though the effect disappears once we control for the paper component scores (technical merit, readability, originality and relevance).

To sum up, we find no gender difference in the likelihood of paper acceptance into the conference or in reviewers' acceptance scores.

Result 1. *The likelihood of paper acceptance does not vary in the share of male authors.*

Table 2: Paper acceptance

	Likelihood of acceptance						Reviewers' acceptance score					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Proportion male	0.004	-0.002	0.010	0.051	-0.005	-0.000	0.129	0.037	0.098	0.148	-0.087	-0.071
	(0.045)	(0.053)	(0.054)	(0.046)	(0.044)	(0.049)	(0.105)	(0.120)	(0.120)	(0.113)	(0.065)	(0.065)
Blinded		0.036	0.053	0.069	0.051	-0.001		-0.529**	-0.482*	-0.329	-0.117	-0.168
		(0.093)	(0.093)	(0.086)	(0.079)	(0.098)		(0.256)	(0.262)	(0.252)	(0.136)	(0.163)
Prop male x Blinded		0.035	0.031	0.030	-0.022	0.003		0.398	0.376	0.349	0.161	0.187
		(0.100)	(0.100)	(0.093)	(0.084)	(0.095)		(0.247)	(0.247)	(0.236)	(0.130)	(0.140)
TE: $\beta_1 + \beta_3$		0.033	0.042	0.081	-0.027	0.002		0.435**	0.474**	0.497**	0.074	0.116
SE		(0.085)	(0.085)	(0.081)	(0.073)	(0.083)		(0.216)	(0.214)	(0.207)	(0.115)	(0.126)
N	902	902	902	902	701	701	855	855	855	855	701	701
Adj. R-sq	0.020	0.019	0.021	0.194	0.483	0.480	0.006	0.008	0.036	0.102	0.770	0.806
Year FE	X	X	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X	X			X	X	X	X
Controls				X	X	X				X	X	X
Scores					X	X					X	X
Reviewer FE						X						X

Notes: OLS regressions of paper acceptance (columns 1-6) and reviewers' acceptance scores (columns 7-12). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. Scores include the reviewer's assessment of the paper's technical merit, readability, originality, and relevance (2017 onward). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Our finding contrasts with Hospido and Sanz (2021) who find that, even after including a range of controls, a 1 percentage point (pp) increase in the share of male authors leads to a 0.054 pp increase in the probability of paper acceptance in three European economic conferences. Their suggested explanation is a stereotype against female authors driven by male referees against lesser-known authors. Therefore, in the following, we examine whether such dynamics are present in our setting by focusing on the gender effect for male and female reviewers separately, which may be masked in the pooled analyses above.

3.2 Does the reviewer gender matter?

We first present evidence in Figure 5 showing that, when separated by gender, reviewers give similar acceptance scores to male and female submitters. No difference is detected for male reviewers' scores for male vs female submitters (3.86 vs 3.83, t-test, $p = 0.7366$), neither do we detect a difference for female reviewers' scores for male vs female submitters (3.87 vs 4.00, t-test, $p = 0.4310$).

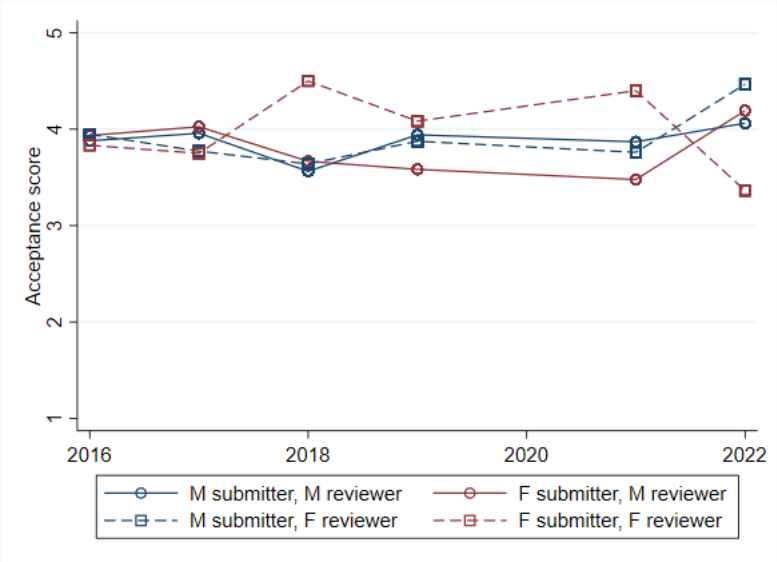

Figure 5: Paper acceptance score by submitter and reviewer gender

We next plot reviewers' scores for each paper's technical merit, readability, originality and relevance in Figure 6. Aside from paper relevance, in which female reviewers give

higher scores to female submitters (4.02 vs 3.75, t-test, $p = 0.0916$), no significant difference is detected for either male or female reviewers' individual component scores, when scoring male vs female submitters.

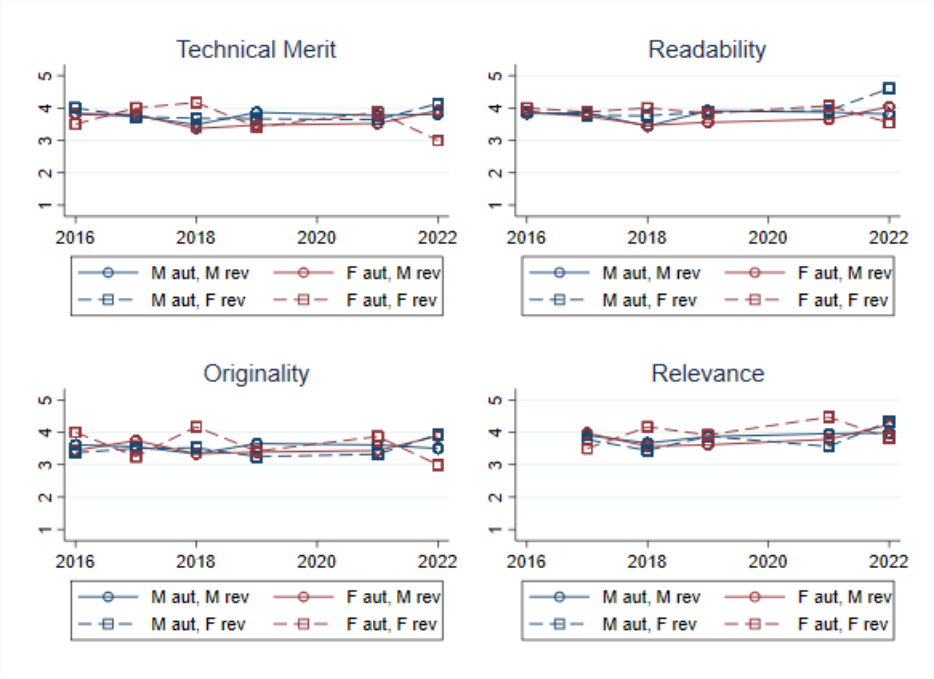

Figure 6: Paper score components by submitter and reviewer gender

To further explore the role of the reviewer's gender in acceptance decisions controling for other characteristics, and the effect of blinding, we test the previous specifications for male and female reviewers separately. The results are shown in Table 3. Male reviewers do not display significant gender bias when author names are observed—while all-male-authored papers get around 0.25 extra points, the effect disappears when controling for paper component scores. Blinding has no significant effect on the gender gap. On the contrary, female reviewers appear to consistently favour female authors when names are observed. Surprisingly, blinding *introduces* a male-author bias greater than the initial in-group bias—in column (10), going from a paper with all female authors to one with all male authors, scores are higher by almost a whole number—though the effect again disappears when controling for paper component scores.

Since it appears that paper component scores explain a significant part of the gender

Table 3: Paper acceptance scores

	Male reviewers						Female reviewers					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Proportion male	0.212*	0.169	0.210	0.246*	-0.034	-0.025	-0.147	-0.476*	-0.394*	-0.428*	-0.301**	-0.284**
	(0.124)	(0.138)	(0.140)	(0.132)	(0.078)	(0.081)	(0.192)	(0.243)	(0.225)	(0.225)	(0.131)	(0.134)
Blinded		-0.677**	-0.643**	-0.394	-0.193	-0.318*		-0.118	0.109	-0.116	0.130	0.041
		(0.293)	(0.309)	(0.300)	(0.152)	(0.185)		(0.545)	(0.510)	(0.499)	(0.300)	(0.354)
Prop male x Blinded		0.192	0.189	0.111	0.148	0.181		1.120***	1.166***	1.353***	0.313	0.419
		(0.306)	(0.306)	(0.289)	(0.148)	(0.166)		(0.394)	(0.373)	(0.424)	(0.274)	(0.279)
TE: $\beta_1 + \beta_3$		0.361	0.399	0.357	0.113	0.156		0.644**	0.772***	0.925***	0.012	0.135
SE		(0.274)	(0.273)	(0.261)	(0.132)	(0.151)		(0.311)	(0.284)	(0.335)	(0.223)	(0.237)
N	662	662	662	662	533	533	188	188	188	188	163	163
Adj. R-sq	0.012	0.018	0.047	0.109	0.777	0.810	-0.023	0.017	0.083	0.126	0.762	0.803
Year FE	X	X	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X	X			X	X	X	X
Controls				X	X	X				X	X	X
Scores					X	X					X	X
Reviewer FE						X						X

Notes: OLS regressions of paper acceptance scores. Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. Scores include the reviewer's assessment of the paper's technical merit, readability, originality, and relevance (2017 onward). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

gap exhibited by reviewers, we repeat the above analyses using the average scores out of the paper's technical merit, readability, originality, and relevance as an outcome variable. The results are shown in Table 4. Without blinding, only male reviewers appear to display significant in-group gender bias, giving an all-male-authored paper around 0.2 extra points relative to an all-female-authored paper. While the coefficients for female reviewers also indicate in-group gender bias, with higher scores given to female-authored papers, these coefficients are statistically insignificant.

Does blinding remove these in-group gender biases? When names are unobserved, only paper quality should matter. For male reviewers, blinding appears to temper reviewers' male bias and the overall effect of the proportion of authors who are male is no longer significant. Surprisingly, as for acceptance scores above, for female reviewers blinding introduces a male-author bias. The effect size is substantial: going from an all-female-authored paper to an all-male-authored paper, the average component score increases by 0.75 (on a scale of 1-5). This result is driven by the 30 papers reviewed by female reviewers in 2022, their scores correlate strongly with the share of male authors on a paper ($\rho = 0.5691$, $p = 0.0024$). This correlation is not significant at the 5% level for male or female reviewers in any other year.

In Tables 5-8 we explore whether the above bias is driven by any one of the four paper components. Technical merit and readability appear to be the main components driving the results. Without blinding, male reviewers judge male-authored papers to be more readable and have more technical merit, however blinding removes the significance of readability while the gender effect on technical merit stands. For female reviewers, neither technical merit nor readability are increasing in male authors prior to blinding. However, when papers are blinded, a higher share of male authors is associated with a higher score in technical merit. The effect size is twice as big for papers reviewed by females than males. Female reviewers also give higher readability scores to male-authored papers subsequent to blinding.

We therefore summarise our second result:

Table 4: Paper component scores

	Male reviewers					Female reviewers				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Proportion male	0.180**	0.187*	0.209**	0.232**	0.210**	0.019	-0.198	-0.139	-0.128	-0.276
	(0.085)	(0.096)	(0.098)	(0.093)	(0.096)	(0.139)	(0.171)	(0.169)	(0.166)	(0.171)
Blinded		-0.261	-0.243	-0.086	-0.224		-0.168	-0.067	-0.307	-0.218
		(0.194)	(0.206)	(0.200)	(0.223)		(0.315)	(0.295)	(0.305)	(0.361)
Prop male x Blinded		-0.042	-0.031	-0.079	-0.093		0.735**	0.789***	0.891***	1.029***
		(0.201)	(0.202)	(0.196)	(0.186)		(0.289)	(0.279)	(0.305)	(0.278)
TE: $\beta_1 + \beta_3$		0.145	0.178	0.153	0.117		0.537**	0.650***	0.763***	0.754***
SE		(0.177)	(0.177)	(0.174)	(0.159)		(0.233)	(0.213)	(0.241)	(0.232)
N	662	662	662	662	662	188	188	188	188	188
Adj. R-sq	0.021	0.023	0.070	0.118	0.215	-0.024	0.006	0.057	0.129	0.239
Year FE	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X			X	X	X
Controls				X	X				X	X
Reviewer FE					X					X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Paper technical merit scores

	Male reviewers					Female reviewers				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Proportion male	0.240**	0.191*	0.200*	0.216**	0.182	0.259	-0.050	-0.046	-0.032	-0.221
	(0.103)	(0.114)	(0.115)	(0.109)	(0.115)	(0.169)	(0.203)	(0.216)	(0.221)	(0.223)
Blinded		-0.379	-0.368	-0.219	-0.661**		-0.283	-0.276	-0.270	-0.393
		(0.259)	(0.264)	(0.258)	(0.311)		(0.378)	(0.376)	(0.378)	(0.476)
Prop male x Blinded		0.232	0.251	0.205	0.299		1.047***	1.188***	1.091***	1.319***
		(0.263)	(0.263)	(0.256)	(0.257)		(0.360)	(0.357)	(0.377)	(0.387)
TE: $\beta_1 + \beta_3$		0.423*	0.451*	0.421*	0.480**		0.997***	1.142***	1.059***	1.098***
SE		(0.237)	(0.238)	(0.234)	(0.237)		(0.297)	(0.271)	(0.285)	(0.308)
N	662	662	662	662	662	188	188	188	188	188
Adj. R-sq	0.019	0.019	0.065	0.092	0.152	-0.006	0.042	0.098	0.152	0.168
Year FE	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X			X	X	X
Controls				X	X				X	X
Reviewer FE					X					X

Notes: OLS regressions of technical merit scores. Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Paper readability scores

	Male reviewers					Female reviewers				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Proportion male	0.207**	0.246**	0.250**	0.278**	0.246**	0.095	-0.132	-0.121	-0.088	-0.252
	(0.101)	(0.115)	(0.116)	(0.111)	(0.116)	(0.172)	(0.207)	(0.209)	(0.189)	(0.228)
Blinded		-0.042	0.020	0.184	0.100		-0.294	-0.196	-0.560	-0.588
		(0.224)	(0.236)	(0.234)	(0.315)		(0.377)	(0.384)	(0.404)	(0.485)
Prop male x Blinded		-0.194	-0.148	-0.196	-0.176		0.764**	0.841**	0.990**	1.188***
		(0.242)	(0.243)	(0.236)	(0.240)		(0.363)	(0.354)	(0.380)	(0.425)
TE: $\beta_1 + \beta_3$		0.052	0.102	0.081	0.071		0.632**	0.720***	0.902***	0.936***
SE		(0.213)	(0.214)	(0.211)	(0.212)		(0.298)	(0.276)	(0.313)	(0.358)
N	662	662	662	662	662	188	188	188	188	188
Adj. R-sq	0.024	0.023	0.058	0.097	0.174	-0.013	0.001	-0.006	0.050	0.148
Year FE	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X			X	X	X
Controls				X	X				X	X
Reviewer FE					X					X

Notes: OLS regressions of readability scores. Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Paper originality scores

	Male reviewers					Female reviewers				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Proportion male	0.101	0.128	0.169	0.179*	0.156	-0.121	-0.393*	-0.261	-0.262	-0.334
	(0.096)	(0.107)	(0.109)	(0.105)	(0.107)	(0.179)	(0.215)	(0.208)	(0.217)	(0.211)
Blinded		-0.225	-0.242	-0.102	-0.222		-0.471	-0.300	-0.520	0.104
		(0.237)	(0.246)	(0.240)	(0.233)		(0.406)	(0.343)	(0.386)	(0.462)
Prop male x Blinded		-0.144	-0.127	-0.173	-0.243		0.910**	0.946**	1.061**	0.807**
		(0.235)	(0.231)	(0.226)	(0.199)		(0.384)	(0.370)	(0.408)	(0.384)
TE: $\beta_1 + \beta_3$		-0.016	0.042	0.006	-0.087		0.518	0.685**	0.799**	0.473
SE		(0.210)	(0.205)	(0.202)	(0.166)		(0.318)	(0.301)	(0.322)	(0.334)
N	662	662	662	662	662	188	188	188	188	188
Adj. R-sq	0.005	0.008	0.059	0.084	0.237	-0.017	0.001	0.071	0.095	0.255
Year FE	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X			X	X	X
Controls				X	X				X	X
Reviewer FE					X					X

Notes: OLS regressions of originality scores. Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Paper relevance scores

	Male reviewers					Female reviewers				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Proportion male	0.195*	0.216	0.234*	0.271**	0.263*	-0.116	-0.165	-0.061	-0.080	-0.163
	(0.114)	(0.136)	(0.139)	(0.137)	(0.142)	(0.193)	(0.253)	(0.266)	(0.239)	(0.251)
Blinded		-0.374	-0.367	-0.209	-0.099		0.413	0.576	0.111	0.227
		(0.231)	(0.255)	(0.251)	(0.285)		(0.409)	(0.398)	(0.406)	(0.501)
Prop male x Blinded		-0.094	-0.124	-0.160	-0.220		0.166	0.105	0.402	0.810**
		(0.244)	(0.252)	(0.251)	(0.237)		(0.380)	(0.380)	(0.410)	(0.387)
TE: $\beta_1 + \beta_3$		0.122	0.110	0.110	0.043		0.001	0.044	0.321	0.646*
SE		(0.204)	(0.210)	(0.214)	(0.193)		(0.284)	(0.272)	(0.326)	(0.333)
N	533	533	533	533	533	163	163	163	163	163
Adj. R-sq	0.015	0.022	0.036	0.105	0.206	0.004	0.007	0.041	0.179	0.207
Year FE	X	X	X	X	X	X	X	X	X	X
Field FE			X	X	X			X	X	X
Controls				X	X				X	X
Reviewer FE					X					X

Notes: OLS regressions of relevance scores (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Result 2. *There are indications of reviewers exhibiting in-group gender bias when determining a paper's acceptance and component scores with author names observable. While blinding removes the bias of male reviewers, it overturns the bias of female reviewers, causing them to favour male-authored papers.*

3.3 What drives gender bias?

Evidence in Hospido and Sanz (2021) suggest that bias favoring male authors may be due to a stereotype against female authors driven by a lack of information about the paper quality as judged from its content—as a result reviewers resort to other (potentially uninformative) signals when assigning scores, such as gender when author names are observable.

To explore whether information asymmetry/unfamiliarity associated with female authors drives reviewers' scoring bias in favour of males, we interact the gender dependent variable with other variables which may proxy for familiarity: the proportion of authors affiliated with an Irish institution, affiliated with an Irish institution with WP, with a PhD, who are Professors and all authors' total number of years post-PhD.

Consistent with the information asymmetry explanation in Hospido and Sanz (2021), male reviewers' bias is concentrated among papers with all authors from outside Ireland (see the first row of Table 9). The bias is reduced with an increasing share of Irish-affiliated authors, shown by the negative coefficients in the second row. Similar patterns are observed for the proportion of authors affiliated with an Irish institution with WP (Table 10), with a PhD (Table 11), with a Professor title (Table 12), and authors' total number of years post-PhD (Table 13). We do not observe these patterns for female reviewers, unsurprising given female reviewers do not display a bias when assigning component scores in the unblinded case.

In the case of a blinded submission, it is not possible to rely on gender-based stereotypes. As expected, male reviewers' bias against lesser-known female authors is no longer observed. However, it is surprising that the coefficient of Proportion male is

consistently positive and sizeable for female reviewers, though not always significant, in Tables 9-13—indicating that for truly unfamiliar papers, an increasing share of male authors is associated with higher quality. Some indicators of familiarity, such as the proportion of authors at an Irish institution or an Irish institution with a WP, reduce the gender bias—the interaction coefficients, while statistically insignificant, are large and negative. While these papers are blinded, having authors affiliated with an Irish institution increases the likelihood that the reviewer may recognise the paper or have seen it presented—perhaps in such cases female reviewers attempt to favour more female authors.[11] However, we hesitate to over-interpret these results since these are based on a much smaller number of observations (56).

Finally, we check whether reviewers' *perception* of information asymmetry can help explain the gender bias using data on reviewer confidence during the paper evaluation. Figure 7 plots the reviewer's confidence (1-5) in the knowledge area. Male reviewers in fact appear to be more confident when evaluating an (unblinded) female submitter's paper than a male submitter's paper, but the difference is not significant (3.88 vs 3.73, t-test, $p = 0.1007$). A similar pattern is observed for female reviewers (3.93 vs 3.71, t-test, $p = 0.2190$). However, higher confidence is not associated with lower gender gap, as shown in Table 14. In fact, when reviewers have zero confidence, the bias tends to favour the opposite gender. Hence, *perceived* lack of confidence is not the driver of the gender gap exhibited by male reviewers above, which instead appears to be driven by unconscious stereotyping.

Result 3. *Male reviewers exhibit a bias against lesser-known female authors when author names are observable. This bias appears to stem from unconscious stereotyping, rather than driven by a perceived lack of confidence during evaluation, and is removed by blinding.*

[11] In April 2022 we conducted a survey of the 2021 reviewers (38% response rate). The responses reveal that around 29% of reviewers recognised, or could identify an author on, at least one of the blinded papers they reviewed. An additional 18% looked up the title of at least one of the papers they reviewed. While the sample is small, these results are in line with Blank (1991) who finds that, when reviewing blinded submissions, around 50% of referees believe they could identify the author(s) and 45% could indeed correctly identify the author(s) of the paper.

Table 9: Paper scores by share of male authors

	Male reviewers				Female reviewers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unblinded				
Proportion male	0.241	0.289*	0.272*	0.312**	-0.298	-0.262	-0.288	-0.243
	(0.154)	(0.152)	(0.145)	(0.148)	(0.254)	(0.250)	(0.234)	(0.242)
Prop male x Prop Irish	-0.044	-0.096	-0.036	-0.174	0.226	0.191	0.312	-0.031
	(0.207)	(0.206)	(0.200)	(0.205)	(0.351)	(0.347)	(0.347)	(0.376)
Prop Irish	0.311*	0.359**	0.044	0.117	0.113	0.082	-0.220	-0.109
	(0.166)	(0.166)	(0.196)	(0.203)	(0.298)	(0.305)	(0.320)	(0.349)
N	535	535	535	535	132	132	132	132
Adj. R-sq	0.051	0.101	0.137	0.215	0.001	0.004	0.106	0.160
				Blinded				
Proportion male	-0.041	-0.043	-0.088	0.443	0.803	1.295**	1.034	0.911
	(0.260)	(0.282)	(0.288)	(0.469)	(0.557)	(0.579)	(0.631)	(1.079)
Prop male x Prop Irish	0.334	0.362	0.402	-0.153	-0.271	-0.736	-0.620	-0.858
	(0.411)	(0.422)	(0.434)	(0.718)	(0.619)	(0.736)	(0.744)	(1.184)
Prop Irish	0.036	0.081	0.054	0.479	0.434	0.511	0.205	0.286
	(0.319)	(0.309)	(0.376)	(0.622)	(0.525)	(0.585)	(0.626)	(1.017)
N	127	127	127	127	56	56	56	56
Adj. R-sq	0.015	0.032	0.012	-0.033	0.036	0.220	0.255	-0.005
Year FE	X	X	X	X	X	X	X	X
Field FE		X	X	X		X	X	X
Controls			X	X			X	X
Reviewer FE				X				X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Paper scores by share of male authors

	Male reviewers				Female reviewers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unblinded				
Proportion male	0.283**	0.347**	0.336**	0.368***	-0.236	-0.235	-0.199	-0.284
	(0.136)	(0.136)	(0.132)	(0.135)	(0.202)	(0.196)	(0.196)	(0.209)
Prop male x Prop WP inst	-0.082	-0.184	-0.194	-0.352*	0.074	0.134	0.205	0.119
	(0.199)	(0.193)	(0.187)	(0.193)	(0.401)	(0.422)	(0.435)	(0.487)
Prop WP inst	0.454***	0.513***	0.504***	0.586***	0.395	0.239	0.198	0.484
	(0.158)	(0.156)	(0.177)	(0.181)	(0.350)	(0.394)	(0.417)	(0.456)
N	535	535	535	535	132	132	132	132
Adj. R-sq	0.074	0.121	0.139	0.218	0.026	0.016	0.101	0.161
				Blinded				
Proportion male	0.324	0.315	0.287	0.518	0.737*	0.982***	0.726*	0.544
	(0.240)	(0.246)	(0.256)	(0.378)	(0.369)	(0.334)	(0.422)	(0.849)
Prop male x Prop WP inst	-0.448	-0.354	-0.385	-0.374	-0.406	-0.470	-0.294	-0.468
	(0.373)	(0.412)	(0.434)	(0.691)	(0.456)	(0.452)	(0.509)	(1.088)
Prop WP inst	0.491*	0.475	0.290	0.114	0.632*	0.457	0.598	0.754
	(0.274)	(0.290)	(0.368)	(0.496)	(0.369)	(0.355)	(0.375)	(0.723)
N	127	127	127	127	56	56	56	56
Adj. R-sq	0.003	0.009	0.012	-0.028	0.076	0.220	0.243	-0.036
Year FE	X	X	X	X	X	X	X	X
Field FE		X	X	X		X	X	X
Controls			X	X			X	X
Reviewer FE				X				X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Paper scores by share of male authors

	Male reviewers				Female reviewers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unblinded				
Proportion male	0.336*	0.352*	0.387**	0.327*	-0.104	0.079	-0.071	-0.218
	(0.190)	(0.181)	(0.177)	(0.183)	(0.269)	(0.257)	(0.258)	(0.268)
Prop male x Prop PhD	-0.227	-0.229	-0.198	-0.164	-0.090	-0.328	-0.130	-0.063
	(0.228)	(0.220)	(0.212)	(0.219)	(0.365)	(0.415)	(0.403)	(0.406)
Prop PhD	0.466**	0.473***	0.423**	0.416**	0.450	0.555*	0.258	0.252
	(0.180)	(0.176)	(0.168)	(0.183)	(0.303)	(0.310)	(0.318)	(0.334)
N	535	535	535	535	132	132	132	132
Adj. R-sq	0.048	0.097	0.139	0.214	0.012	0.020	0.101	0.160
				Blinded				
Proportion male	0.268	0.284	0.162	0.462	0.512	0.675*	0.680*	0.114
	(0.383)	(0.403)	(0.440)	(0.559)	(0.326)	(0.349)	(0.349)	(0.823)
Prop male x Prop PhD	-0.213	-0.186	-0.029	-0.183	-0.003	0.069	-0.188	0.554
	(0.472)	(0.495)	(0.520)	(0.777)	(0.444)	(0.387)	(0.465)	(1.424)
Prop PhD	0.248	0.236	0.156	0.099	0.235	0.352	0.734*	-0.113
	(0.370)	(0.347)	(0.375)	(0.518)	(0.325)	(0.268)	(0.380)	(1.089)
N	127	127	127	127	56	56	56	56
Adj. R-sq	-0.012	-0.009	0.004	-0.033	0.023	0.238	0.239	-0.043
Year FE	X	X	X	X	X	X	X	X
Field FE		X	X	X		X	X	X
Controls			X	X			X	X
Reviewer FE				X				X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Paper scores by share of male authors

	Male reviewers				Female reviewers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unblinded				
Proportion male	0.204**	0.261**	0.285***	0.254**	-0.250	-0.244	-0.220	-0.323
	(0.103)	(0.103)	(0.098)	(0.101)	(0.181)	(0.181)	(0.182)	(0.198)
Prop male x Prop Prof	-0.383	-0.865*	-0.669	-0.657	0.653	0.673	1.199*	1.176
	(0.422)	(0.494)	(0.470)	(0.476)	(0.615)	(0.684)	(0.686)	(0.794)
Prop Prof	0.644**	1.011**	0.763*	0.692*	0.179	-0.023	-0.855	-0.852
	(0.322)	(0.400)	(0.390)	(0.396)	(0.498)	(0.580)	(0.621)	(0.785)
N	535	535	535	535	132	132	132	132
Adj. R-sq	0.041	0.093	0.145	0.218	0.026	0.014	0.106	0.165
				Blinded				
Proportion male	0.092	0.127	0.060	0.132	0.367	0.590**	0.404	0.447
	(0.212)	(0.206)	(0.217)	(0.282)	(0.234)	(0.253)	(0.256)	(0.968)
Prop male x Prop Prof	0.750	0.924	1.276	2.554	1.454	1.464	1.881	0.088
	(0.841)	(0.957)	(0.998)	(1.943)	(0.905)	(0.938)	(1.264)	(4.701)
Prop Prof	-0.752	-1.074	-1.371	-2.954	-1.431**	-1.411*	-1.882*	-0.967
	(0.723)	(0.846)	(0.886)	(1.778)	(0.709)	(0.815)	(1.076)	(3.647)
N	127	127	127	127	56	56	56	56
Adj. R-sq	-0.011	-0.000	0.011	0.022	0.056	0.234	0.274	0.044
Year FE	X	X	X	X	X	X	X	X
Field FE		X	X	X		X	X	X
Controls			X	X			X	X
Reviewer FE				X				X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Paper scores by share of male authors

	Male reviewers				Female reviewers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unblinded				
Proportion male	0.291***	0.322***	0.333***	0.281***	-0.233	-0.192	-0.230	-0.399**
	(0.111)	(0.108)	(0.105)	(0.106)	(0.190)	(0.185)	(0.188)	(0.190)
Prop male x Total years post-PhD	-0.014**	-0.016**	-0.012*	-0.010	0.001	-0.001	0.008	0.021
	(0.007)	(0.007)	(0.007)	(0.007)	(0.013)	(0.013)	(0.015)	(0.016)
Total years post-PhD	0.018***	0.019***	0.014**	0.013**	0.012	0.013	-0.002	-0.017
	(0.005)	(0.006)	(0.006)	(0.006)	(0.010)	(0.010)	(0.014)	(0.015)
N	535	535	535	535	132	132	132	132
Adj. R-sq	0.052	0.101	0.144	0.220	0.041	0.058	0.101	0.167
				Blinded				
Proportion male	0.198	0.246	0.217	0.784**	0.514*	0.726**	0.528*	0.227
	(0.245)	(0.265)	(0.272)	(0.359)	(0.281)	(0.290)	(0.293)	(0.773)
Prop male x Total years post-PhD	-0.006	-0.008	-0.008	-0.036**	0.002	0.007	0.005	0.012
	(0.011)	(0.012)	(0.012)	(0.017)	(0.010)	(0.009)	(0.011)	(0.023)
Total years post-PhD	0.008	0.010	0.012	0.024	-0.000	0.001	0.004	-0.011
	(0.007)	(0.007)	(0.008)	(0.015)	(0.006)	(0.005)	(0.010)	(0.017)
N	127	127	127	127	56	56	56	56
Adj. R-sq	-0.006	-0.001	0.006	0.019	0.009	0.216	0.236	-0.097
Year FE	X	X	X	X	X	X	X	X
Field FE		X	X	X		X	X	X
Controls			X	X			X	X
Reviewer FE				X				X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

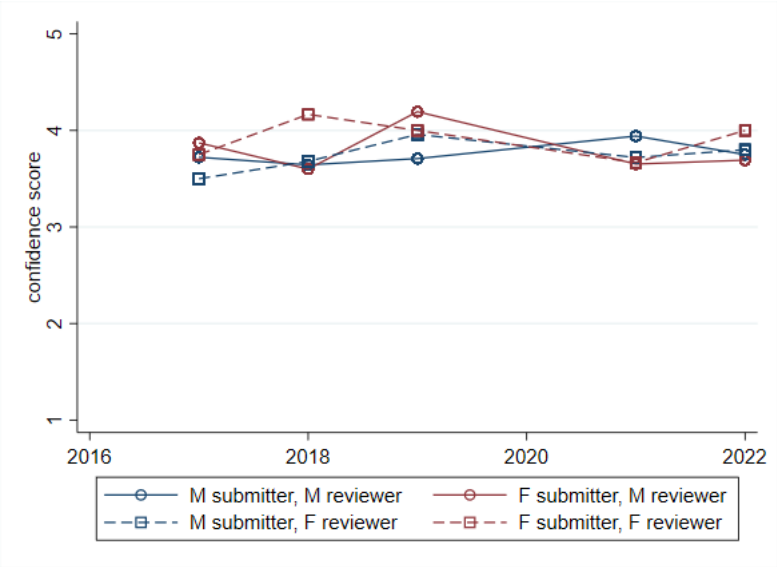

Figure 7: Reviewer confidence by submitter and reviewer gender

Table 14: Paper scores by share of male authors

	Male reviewers				Female reviewers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unblinded				
Proportion male	-0.162	-0.152	-0.293	-0.682	0.008	0.230	0.690	0.971
	(0.533)	(0.520)	(0.488)	(0.511)	(0.817)	(0.807)	(0.733)	(0.952)
Prop male x Reviewer confidence	0.110	0.104	0.146	0.228*	-0.018	-0.050	-0.173	-0.261
	(0.146)	(0.140)	(0.132)	(0.135)	(0.212)	(0.206)	(0.186)	(0.224)
Reviewer confidence	0.150	0.149	0.084	0.144	0.281	0.414**	0.527***	0.747***
	(0.115)	(0.107)	(0.103)	(0.112)	(0.178)	(0.190)	(0.161)	(0.228)
N	406	406	406	406	107	107	107	107
Adj. R-sq	0.098	0.140	0.183	0.291	0.045	0.090	0.262	0.316
				Blinded				
Proportion male	-0.501	-0.459	-0.333	-0.819	-0.115	0.496	0.734	-0.059
	(0.801)	(0.897)	(0.896)	(1.436)	(1.163)	(1.328)	(1.261)	(3.466)
Prop male x Reviewer confidence	0.159	0.153	0.115	0.298	0.173	0.074	-0.044	0.152
	(0.198)	(0.228)	(0.231)	(0.362)	(0.330)	(0.346)	(0.328)	(0.810)
Reviewer confidence	0.131	0.121	0.130	0.044	-0.126	-0.055	0.022	0.090
	(0.161)	(0.194)	(0.196)	(0.264)	(0.237)	(0.242)	(0.228)	(0.706)
N	127	127	127	127	56	56	56	56
Adj. R-sq	0.055	0.050	0.047	0.013	0.014	0.196	0.217	-0.095
Year FE	X	X	X	X	X	X	X	X
Field FE		X	X	X		X	X	X
Controls			X	X			X	X
Reviewer FE				X				X

Notes: OLS regressions of paper component scores, the average of technical merit, readability, originality and relevance (2017 onward). Controls include number of authors on paper, proportion of authors with PhD, proportion of authors affiliated with an Irish institution and proportion of authors affiliated with an Irish institution with WP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4 Discussion and concluding remarks

We use conference submission data to examine gender diversity in the Irish Economic Association's annual conference from 2016-2022. Acceptance into the conference is measured using i) the organisers' acceptance decision and ii) the reviewer's acceptance score, which influences i). We do not observe gender difference in the former, suggesting that organisers are successful in achieving a gender balance in conference presenters. Neither do we observe gender difference in reviewers' acceptance scores, pooling reviewers of both genders.

However, the above result masks important heterogeneities as, when author names are observable, we find indications that male and female reviewers judge each paper's acceptance differently depending on the share of male authors on a paper. In particular, reviewers exhibit an in-group gender bias. As this bias is somewhat reduced when controling for the paper's component scores (technical merit, readability, originality and relevance), the bias may simply reflect differing paper quality which is correlated with the gender of the author. For example, male reviewers give higher acceptance scores to male authors because their papers are of better quality—which could be due to lower quality of female-authored papers in male-dominated fields.

We next turn to the component scores themselves to see if these also exhibit gender gaps. While component scores assigned by female reviewers do not differ much for male and female authors, male reviewers give higher component scores to male-authored papers. As suggested by Hospido and Sanz (2021), male reviewers' discrimination against female authors may reflect a stereotype used to address the problem of information asymmetry. We find suggestive evidence along these lines: male reviewers' bias exists only for less well-known authors, as proxied by affiliation, the proportion of PhD or Professor authors and the overall experience of all authors on a paper (measured by total years post-PhD). However, higher reviewer confidence does not eliminate the gender bias, which suggests that the bias is not driven by reviewers being conscious of the information asymmetry.

Finally, we exploit a change in 2021 to anonymised submission to study whether the gender bias persists. While blinding successfully removes male reviewers' gender bias, it surprisingly overturns female reviewers bias such that they now favour male-authored papers. This bias is somewhat reduced for papers with Irish authors, which may increase the reviewer's familiarity with it despite blinding, thus suggesting that (where possible) reviewers attempt to favour female authors. On the other hand, this finding also suggests that when a paper is completely unfamiliar and reviewers have to rely purely on its content to judge quality, a paper with a higher share of male authors tends to do better especially in technical merit and readability. However, we hesitate to speculate further given the small number of observations driving the result. Rather, more data and further research possibly in other contexts are needed to draw strong conclusions, for example by allocating the same paper to both male and female reviewers.

References

Åslund, O. and Skans, O. N. (2012). Do anonymous job application procedures level the playing field? *Industrial and Labor Relations Review*, 65(1):82–107.

Bayer, A. and Rouse, C. E. (2016). Diversity in the economics profession: A new attack on an old problem. *Journal of Economic Perspectives*, 30(4):221–42.

Behaghel, L., Crépon, B., and Le Barbanchon, T. (2015). Unintended effects of anonymous resumes. *American Economic Journal: Applied Economics*, 7(3):1–27.

Blank, R. M. (1991). The effects of double-blind versus single-blind reviewing: Experimental evidence from the American Economic Review. *The American Economic Review*, pages 1041–1067.

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145:27–41.

Boring, A. and Philippe, A. (2021). Reducing discrimination in the field: Evidence from an awareness raising intervention targeting gender biases in student evaluations of teaching. *Journal of Public Economics*, 193:104323.

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2020). Are referees and editors in economics gender neutral? *The Quarterly Journal of Economics*, 135(1):269–327.

Carlsson, F., Löfgren, Å., and Sterner, T. (2012). Discrimination in scientific review: A natural field experiment on blind versus non-blind reviews. *The Scandinavian Journal of Economics*, 114(2):500–519.

Chari, A. and Goldsmith-Pinkham, P. (2017). Gender representation in economics across topics and time: Evidence from the NBER Summer Institute. National Bureau of Economic Research Working Paper No. 23953.

Doleac, J. L., Hengel, E., and Pancotti, E. (2021). Diversity in economics seminars: Who gives invited talks? *AEA Papers and Proceedings*, 111:55–59.

Ductor, L., Goyal, S., and Prummer, A. (forth.). Gender and collaboration. *The Review of Economics and Statistics*.

Dupas, P., Modestino, A. S., Niederle, M., Wolfers, J., et al. (2021). Gender and the dynamics of economics seminars. National Bureau of Economic Research Working Paper No. 28494.

Eberhardt, M., Facchini, G., and Rueda, V. (2022). Gender differences in reference letters: Evidence from the economics job market. Centre for Economic Policy Research Discussion Paper No. 16960.

Ginther, D. K. and Kahn, S. (2021). Women in academic economics: Have we made progress? *AEA Papers and Proceedings*, 111:138–42.

Goldin, C. and Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4):715–741.

Hengel, E. (forth.). Publishing while female. *Economic Journal*.

Hengel, E. and Moon, E. (2020). Gender and equality at top economics journals. Working Paper.

Hospido, L. and Sanz, C. (2021). Gender gaps in the evaluation of research: Evidence from submissions to economics conferences. *Oxford Bulletin of Economics and Statistics*, 83(3):590–618.

Koffi, M. (2021). Gendered citations at top economic journals. *AEA Papers and Proceedings*, 111:60–64.

Krause, A., Rinne, U., and Zimmermann, K. F. (2012a). Anonymous job applications in Europe. *IZA Journal of European Labor Studies*, 1(1):1–20.

Krause, A., Rinne, U., and Zimmermann, K. F. (2012b). Anonymous job applications of fresh Ph.D. economists. *Economics Letters*, 117(2):441–444.

Lundberg, S. and Stearns, J. (2019). Women in economics: Stalled progress. *Journal of Economic Perspectives*, 33(1):3–22.

Mengel, F., Sauermann, J., and Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, 17(2):535–566.

Ross, J. S., Gross, C. P., Desai, M. M., Hong, Y., Grant, A. O., Daniels, S. R., Hachinski, V. C., Gibbons, R. J., Gardner, T. J., and Krumholz, H. M. (2006). Effect of blinded peer review on abstract acceptance. *JAMA*, 295(14):1675–1680.

Sarsons, H., Gërxhani, K., Reuben, E., and Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political Economy*, 129(1):101–147.

Tomkins, A., Zhang, M., and Heavlin, W. D. (2017). Reviewer bias in single-versus double-blind peer review. *Proceedings of the National Academy of Sciences*, 114(48):12708–12713.

Wu, A. H. (2020). Gender bias among professionals: An identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880.