

Publishing while female

Are women held to higher standards?

Evidence from peer review.

Erin Hengel

University of Liverpool

Gender and Career Progression Conference

14 May 2018

Background

- ▶ Women are underrepresented in economics (2016):
 - ▶ Roughly 30 percent of new PhDs.
 - ▶ Just under 30 percent of assistant professors.
 - ▶ 25 percent of associate professors.
 - ▶ Almost 15 percent of full professors.

Background

- ▶ Women are underrepresented in economics (2016):
 - ▶ Roughly 30 percent of new PhDs.
 - ▶ Just under 30 percent of assistant professors.
 - ▶ 25 percent of associate professors.
 - ▶ Almost 15 percent of full professors.
- ▶ Women are **really** underrepresented in publications at top economics journals (2015).
 - ▶ The average ratio of female authors barely broke 15 percent.
 - ▶ Only 7.5 percent of papers were majority female-authored.
 - ▶ Just 4 percent were written entirely by women.
 - ▶ *QJE* did not publish a single exclusively female-authored paper in 2015... or 2016.... or 2017...
 - ▶ ...in four of the last fifteen years covered by the data (2001–2015), *Econometrica* and *JPE* didn't either.

Background

- ▶ Women are underrepresented in economics (2016):
 - ▶ Roughly 30 percent of new PhDs.
 - ▶ Just under 30 percent of assistant professors.
 - ▶ 25 percent of associate professors.
 - ▶ Almost 15 percent of full professors.
- ▶ Women are **really** underrepresented in publications at top economics journals (2015).
 - ▶ The average ratio of female authors barely broke 15 percent.
 - ▶ Only 7.5 percent of papers were majority female-authored.
 - ▶ Just 4 percent were written entirely by women.
 - ▶ *QJE* did not publish a single exclusively female-authored paper in 2015... or 2016.... or 2017...
 - ▶ ...in four of the last fifteen years covered by the data (2001–2015), *Econometrica* and *JPE* didn't either.

Is peer review Affirmative Action for men?

Background

Women are held to higher standards

- ▶ Men are rated more competent when compared to otherwise equally competent women (Foschi, 1996).
- ▶ Male undergraduate biology students underestimated female classmates' ability (Grunspan et al., 2016).
- ▶ Female graduate students are rated less qualified for laboratory management positions (Moss-Racusin et al., 2012).
- ▶ When collaborating with men, women are given less credit for their mutual work (Heilman and Haynes, 2005; Sarsons, 2017).
- ▶ Manuscripts by female authors are rated lower quality (Goldberg, 1968; Paludi and Bauer, 1983; Krawczyk and Smyk, 2016).

Background

Women are held to higher standards

- ▶ Men are rated more competent when compared to otherwise equally competent women (Foschi, 1996).
- ▶ Male undergraduate biology students underestimated female classmates' ability (Grunspan et al., 2016).
- ▶ Female graduate students are rated less qualified for laboratory management positions (Moss-Racusin et al., 2012).
- ▶ When collaborating with men, women are given less credit for their mutual work (Heilman and Haynes, 2005; Sarsons, 2017).
- ▶ Manuscripts by female authors are rated lower quality (Goldberg, 1968; Paludi and Bauer, 1983; Krawczyk and Smyk, 2016).

“Women must do twice as well to be thought half as good.”

–Charlotte Whitton

Gender discrimination in peer review

Are women's papers held to higher standards in peer review?

- ▶ No evidence gender impacts acceptance rates (see Blank, 1991; Gilbert et al., 1994; Ceci et al., 2014).

Gender discrimination in peer review

Are women's papers held to higher standards in peer review?

- ▶ No evidence gender impacts acceptance rates (see Blank, 1991; Gilbert et al., 1994; Ceci et al., 2014).
- ▶ Most papers undergo major referee-requested revisions (Abrevaya and Hamermesh, 2012).

Gender discrimination in peer review

Are women's papers held to higher standards in peer review?

- ▶ No evidence gender impacts acceptance rates (see Blank, 1991; Gilbert et al., 1994; Ceci et al., 2014).
- ▶ Most papers undergo major referee-requested revisions (Abrevaya and Hamermesh, 2012).
- ▶ Are referees, *e.g.*, more likely to double-check technical details, demand robustness checks or require clearer exposition in a female-authored paper?
 - ▶ If so, then female-authored papers should be better quality on the dimension in which they are held to higher standards.

Gender discrimination in peer review

Are women's papers held to higher standards in peer review?

- ▶ No evidence gender impacts acceptance rates (see Blank, 1991; Gilbert et al., 1994; Ceci et al., 2014).
- ▶ Most papers undergo major referee-requested revisions (Abrevaya and Hamermesh, 2012).
- ▶ Are referees, *e.g.*, more likely to double-check technical details, demand robustness checks or require clearer exposition in a female-authored paper?
 - ▶ If so, then female-authored papers should be better quality on the dimension in which they are held to higher standards.

“I have no doubt that one of [discrimination's] results has been that those women who do manage to make their mark are much abler than their male colleagues.”

–Milton Friedman

Writing clarity

1. Clear writing is valued by journals.
 - ▶ Stated explicitly in submission guidelines.
 - ▶ “Evaluate adequacy of the language” is one of the most frequent tasks editors make of referees (Chauvin et al, 2016).

Writing clarity

1. Clear writing is valued by journals.
 - ▶ Stated explicitly in submission guidelines.
 - ▶ “Evaluate adequacy of the language” is one of the most frequent tasks editors make of referees (Chauvin et al, 2016).
2. Every article abstract published in the *AER*, *Econometrica*, *JPE* and *QJE* since 1950.

Writing clarity

1. Clear writing is valued by journals.
 - ▶ Stated explicitly in submission guidelines.
 - ▶ “Evaluate adequacy of the language” is one of the most frequent tasks editors make of referees (Chauvin et al, 2016).
2. Every article abstract published in the *AER*, *Econometrica*, *JPE* and *QJE* since 1950.
 - ▶ Readability scores highly correlated across abstract, introduction and discussion sections of a paper (Hartley et al., 2003; Plavén-Sigra et al., 2017).

Correlation with other measures of reading comprehension

1. Good writing $\approx f$ (simple vocabulary, short sentences).

Correlation with other measures of reading comprehension

1. Good writing $\approx f$ (simple vocabulary, short sentences).
 - ▶ Flesch Reading Ease, Flesch-Kincaid, Gunning Fog, SMOG and Dale-Chall.

Correlation with other measures of reading comprehension

1. Good writing $\approx f$ (simple vocabulary, short sentences).
 - ▶ Flesch Reading Ease, Flesch-Kincaid, Gunning Fog, SMOG and Dale-Chall.
 - ▶ Developed primarily for adults.

Correlation with other measures of reading comprehension

1. Good writing $\approx f$ (simple vocabulary, short sentences).
 - ▶ Flesch Reading Ease, Flesch-Kincaid, Gunning Fog, SMOG and Dale-Chall.
 - ▶ Developed primarily for adults.
 - ▶ Tested on technical documents (especially military training/regulation manuals) and consistently correlate with reading comprehension (see DuBay, 2004).

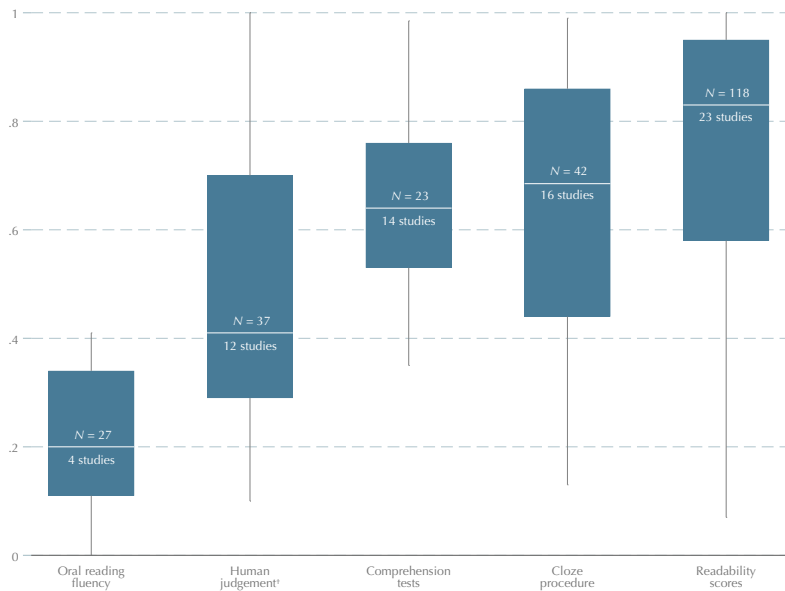
Correlation with other measures of reading comprehension

1. Good writing $\approx f$ (simple vocabulary, short sentences).
 - ▶ Flesch Reading Ease, Flesch-Kincaid, Gunning Fog, SMOG and Dale-Chall.
 - ▶ Developed primarily for adults.
 - ▶ Tested on technical documents (especially military training/regulation manuals) and consistently correlate with reading comprehension (see DuBay, 2004).
 - ▶ Used in research, particularly in finance and political science (see Benoit et al., 2017, and Loughran and McDonald, 2016).

Correlation with other measures of reading comprehension

1. Good writing $\approx f$ (simple vocabulary, short sentences).
 - ▶ Flesch Reading Ease, Flesch-Kincaid, Gunning Fog, SMOG and Dale-Chall.
 - ▶ Developed primarily for adults.
 - ▶ Tested on technical documents (especially military training/regulation manuals) and consistently correlate with reading comprehension (see DuBay, 2004).
 - ▶ Used in research, particularly in finance and political science (see Benoit et al., 2017, and Loughran and McDonald, 2016).
 - ▶ Linked to trustworthiness, believability, intelligence (Oppenheimer, 2016).

Writing clarity



Strategy

Identification

1. Establish that there is a gender difference in readability.
2. Causally link this difference to the peer review process.
3. Establish sufficient conditions to verify discrimination is present in academic publishing.
 - ▶ Show evidence that these conditions are satisfied on average for two different measures of research quality: readability *and* citation counts.
 - ▶ Use a matching estimator to estimate the causal impact of higher readability standards in peer review.

Strategy

Identification

1. Establish that there is a gender difference in readability.
2. Causally link this difference to the peer review process.
3. Establish sufficient conditions to verify discrimination is present in academic publishing.
 - ▶ Show evidence that these conditions are satisfied on average for two different measures of research quality: readability *and* citation counts.
 - ▶ Use a matching estimator to estimate the causal impact of higher readability standards in peer review.

Consequences

- ▶ **Behaviourial change.** As women update beliefs about referees' standards, they increasingly meet those standards before peer review.
- ▶ **Time tax.** Female-authored papers take longer in peer review.

Article-level analysis

$$R_j^s = \beta_0 + \beta_1 \text{female ratio}_j + \theta \mathbf{X}_j + \varepsilon_j.$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Flesch Reading Ease	0.90* (0.48)	0.87* (0.48)	0.83* (0.50)	0.81 (0.48)	0.97* (0.50)	0.52 (0.53)	0.92 (0.71)
Flesch-Kincaid	0.19* (0.11)	0.18 (0.11)	0.18 (0.11)	0.19* (0.11)	0.22* (0.12)	0.23* (0.12)	0.25* (0.14)
Gunning Fog	0.33*** (0.12)	0.33*** (0.12)	0.33*** (0.12)	0.33*** (0.13)	0.37*** (0.14)	0.34** (0.14)	0.36** (0.16)
SMOG	0.21** (0.09)	0.21** (0.09)	0.22** (0.09)	0.21** (0.09)	0.23** (0.10)	0.19* (0.10)	0.23* (0.12)
Dale-Chall	0.10** (0.04)	0.10** (0.04)	0.10** (0.05)	0.09** (0.04)	0.11** (0.05)	0.09* (0.05)	0.13** (0.06)
Editor effects	✓	✓	✓	✓	✓	✓	✓
Journal effects	✓	✓	✓	✓	✓	✓	✓
Year effects		✓	✓	✓	✓	✓	✓
Journal×Year effects			✓	✓	✓	✓	✓
Institution effects				✓	✓	✓	✓
Quality controls					✓ ¹	✓ ¹	✓ ¹
Native speaker					✓	✓	✓
JEL (primary) effects						✓	
JEL (tertiary) effects							✓

Notes. 9,122 articles in (1)–(5); 5,216 articles in (6); 5,777 articles—including 561 from *AER Papers & Proceedings*—in (7). Figures represent the coefficient on female ratio from an OLS regression on the relevant readability score. Quality controls denoted by ✓¹ include citation count and max. T_j fixed effects. Standard errors clustered on editor in parentheses. ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

Article-level analysis

$$R_j^s = \beta_0 + \beta_1 \text{female ratio}_j + \theta \mathbf{X}_j + \varepsilon_j.$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Flesch Reading Ease	0.90* (0.48)	0.87* (0.48)	0.83* (0.50)	0.81 (0.48)	0.97* (0.50)	0.52 (0.53)	0.92 (0.71)
Flesch-Kincaid	0.19* (0.11)	0.18 (0.11)	0.18 (0.11)	0.19* (0.11)	0.22* (0.12)	0.23* (0.12)	0.25* (0.14)
Gunning Fog	0.33*** (0.12)	0.33*** (0.12)	0.33*** (0.12)	0.33*** (0.13)	0.37*** (0.14)	0.34** (0.14)	0.36** (0.16)
SMOG	0.21** (0.09)	0.21** (0.09)	0.22** (0.09)	0.21** (0.09)	0.23** (0.10)	0.19* (0.10)	0.23* (0.12)
Dale-Chall	0.10** (0.04)	0.10** (0.04)	0.10** (0.05)	0.09** (0.04)	0.11** (0.05)	0.09* (0.05)	0.13** (0.06)
Editor effects	✓	✓	✓	✓	✓	✓	✓
Journal effects	✓	✓	✓	✓	✓	✓	✓
Year effects		✓	✓	✓	✓	✓	✓
Journal×Year effects			✓	✓	✓	✓	✓
Institution effects				✓	✓	✓	✓
Quality controls					✓ ¹	✓ ¹	✓ ¹
Native speaker					✓	✓	✓
JEL (primary) effects						✓	
JEL (tertiary) effects							✓

Notes. 9,122 articles in (1)–(5); 5,216 articles in (6); 5,777 articles—including 561 from *AER Papers & Proceedings*—in (7). Figures represent the coefficient on female ratio from an OLS regression on the relevant readability score. Quality controls denoted by ✓¹ include citation count and max. T_j fixed effects. Standard errors clustered on editor in parentheses. ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

Female-authored abstracts are 1–2 % more clearly written.

Author-level analysis

$$R_{jit}^s = \beta_0 R_{it-1}^s + \beta_1 \text{female ratio}_j + \beta_2 \text{female ratio}_j \times \text{male}_i + \theta \mathbf{X}_j + \alpha_i + \varepsilon_{it}.$$

	Flesch Reading Ease	Flesch- Kincaid	Gunning Fog	SMOG	Dale- Chall
Female ratio (women)	2.37** (1.00)	0.35* (0.20)	0.66*** (0.24)	0.47** (0.19)	0.23** (0.10)
Female ratio (men)	0.57 (1.31)	0.10 (0.25)	0.15 (0.29)	0.09 (0.21)	0.10 (0.11)
N_j	✓	✓	✓	✓	✓
Editor effects	✓	✓	✓	✓	✓
Journal effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Journal × Year effects	✓	✓	✓	✓	✓
Institution effects	✓	✓	✓	✓	✓
Quality controls	✓ ¹	✓ ¹	✓ ¹	✓ ¹	✓ ¹
Native speaker	✓	✓	✓	✓	✓

Notes. Sample 9,186 observations (2,827 authors). Figures from first-differenced, IV estimation of the regression equation (Arellano and Bover, 1995; Blundell and Bond, 1998). Quality controls denoted by ✓¹ include citation count and max. T_j fixed effects. Regressions weighted by $1/N_j$; standard errors adjusted for two-way clustering on editor and author (in parentheses). ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

Author-level analysis

$$R_{jit}^s = \beta_0 R_{it-1}^s + \beta_1 \text{female ratio}_j + \beta_2 \text{female ratio}_j \times \text{male}_i + \theta \mathbf{X}_j + \alpha_j + \varepsilon_{it}.$$

	Flesch Reading Ease	Flesch- Kincaid	Gunning Fog	SMOG	Dale- Chall
Female ratio (women)	2.37** (1.00)	0.35* (0.20)	0.66*** (0.24)	0.47** (0.19)	0.23** (0.10)
Female ratio (men)	0.57 (1.31)	0.10 (0.25)	0.15 (0.29)	0.09 (0.21)	0.10 (0.11)
N_j	✓	✓	✓	✓	✓
Editor effects	✓	✓	✓	✓	✓
Journal effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Journal×Year effects	✓	✓	✓	✓	✓
Institution effects	✓	✓	✓	✓	✓
Quality controls	✓ ¹	✓ ¹	✓ ¹	✓ ¹	✓ ¹
Native speaker	✓	✓	✓	✓	✓

Notes. Sample 9,186 observations (2,827 authors). Figures from first-differenced, IV estimation of the regression equation (Arellano and Bover, 1995; Blundell and Bond, 1998). Quality controls denoted by ✓¹ include citation count and max. T_j fixed effects. Regressions weighted by $1/N_j$; standard errors adjusted for two-way clustering on editor and author (in parentheses). ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

Everyone writes better when co-authoring with women!

- ▶ Female-authored abstracts are 2–6 % more clearly written.
- ▶ Convex relationship between readability and female ratio.

Causal impact of peer review

	FGLS			OLS
	Working paper	Published article	Difference	Change in score
Flesch Reading Ease	2.26** (1.00)	3.21*** (1.21)	0.95* (0.57)	0.94 (0.60)
Flesch-Kincaid	0.31 (0.23)	0.75*** (0.28)	0.44** (0.18)	0.44** (0.19)
Gunning Fog	0.44* (0.24)	0.86*** (0.29)	0.42** (0.19)	0.42** (0.20)
SMOG	0.33** (0.15)	0.56*** (0.19)	0.24** (0.12)	0.24* (0.12)
Dale-Chall	0.32*** (0.10)	0.45*** (0.11)	0.13** (0.05)	0.13** (0.05)
Editor effects	✓	✓		✓
Journal effects	✓	✓		✓
Year effects	✓	✓		
Journal × Year effects	✓	✓		✓
Quality controls	✓ ²	✓ ²		✓ ³
Native speaker	✓	✓		✓

Notes. Sample 1,709 NBER working papers; 1,707 published articles. Estimates exclude 279 pre-internet double-blind reviewed articles. Column one standard errors clustered by editor in parentheses. Columns two and three standard errors clustered by year and robust to cross-model correlation in parentheses. Column five standard errors clustered by year in parentheses. Quality controls denoted by ✓² include citation count, max. T_j and max. t_j ; ✓³ includes max. t_j , only. ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

Causal impact of peer review

	FGLS			OLS
	Working paper	Published article	Difference	Change in score
Flesch Reading Ease	2.26** (1.00)	3.21*** (1.21)	0.95* (0.57)	0.94 (0.60)
Flesch-Kincaid	0.31 (0.23)	0.75*** (0.28)	0.44** (0.18)	0.44** (0.19)
Gunning Fog	0.44* (0.24)	0.86*** (0.29)	0.42** (0.19)	0.42** (0.20)
SMOG	0.33** (0.15)	0.56*** (0.19)	0.24** (0.12)	0.24* (0.12)
Dale-Chall	0.32*** (0.10)	0.45*** (0.11)	0.13** (0.05)	0.13** (0.05)
Editor effects	✓	✓		✓
Journal effects	✓	✓		✓
Year effects	✓	✓		
Journal × Year effects	✓	✓		✓
Quality controls	✓ ²	✓ ²		✓ ³
Native speaker	✓	✓		✓

Notes. Sample 1,709 NBER working papers; 1,707 published articles. Estimates exclude 279 pre-internet double-blind reviewed articles. Column one standard errors clustered by editor in parentheses. Columns two and three standard errors clustered by year and robust to cross-model correlation in parentheses. Column five standard errors clustered by year in parentheses. Quality controls denoted by ✓² include citation count, max. T_j and max. t_j ; ✓³ includes max. t_j , only. ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

Peer review causes a large increase in the readability gap

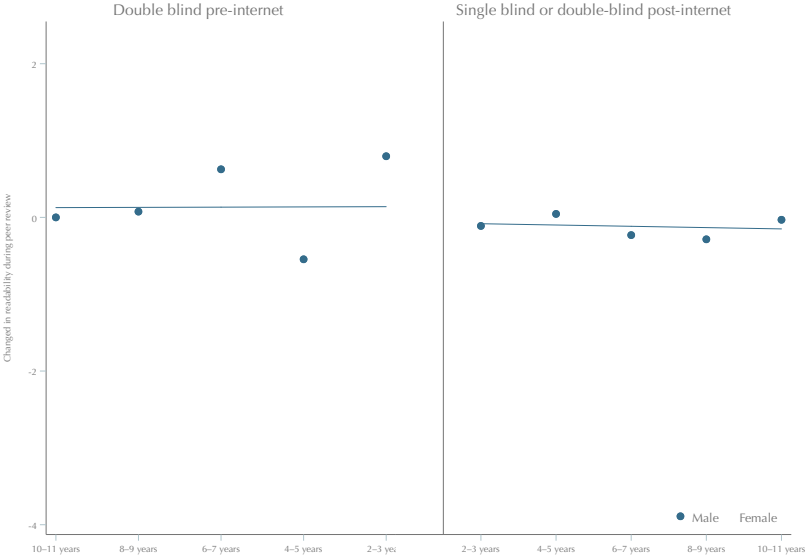
Causal impact of peer review

	Flesch Reading Ease	Flesch-Kincaid	Gunning Fog	SMOG	Dale-Chall
Non-blind	0.93 (0.60)	0.43** (0.19)	0.41** (0.20)	0.23* (0.12)	0.12** (0.05)
Blind	-1.51 (3.05)	-0.56 (0.70)	-0.54 (0.82)	-0.36 (0.59)	-0.13 (0.18)
Difference	2.44 (3.14)	1.00 (0.75)	0.95 (0.87)	0.59 (0.61)	0.25 (0.18)
Editor effects	✓	✓	✓	✓	✓
Journal effects	✓	✓	✓	✓	✓
Journal × Year effects	✓	✓	✓	✓	✓
Quality controls	✓ ³	✓ ³	✓ ³	✓ ³	✓ ³
Native speaker	✓	✓	✓	✓	✓

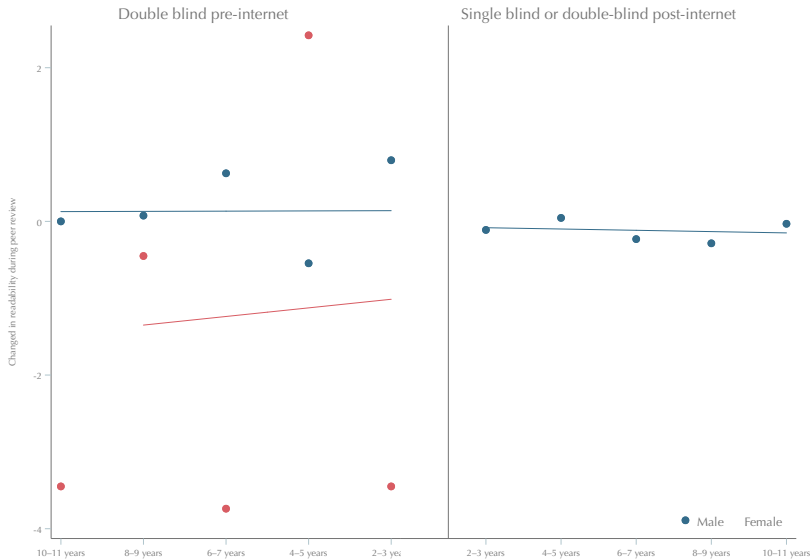
Notes. Sample 1,988 NBER working papers; 1,986 published articles. Standard errors clustered by year in parentheses. Quality controls denoted by ✓³ includes max. t_j , only. ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

No significant gap under double-blind review.

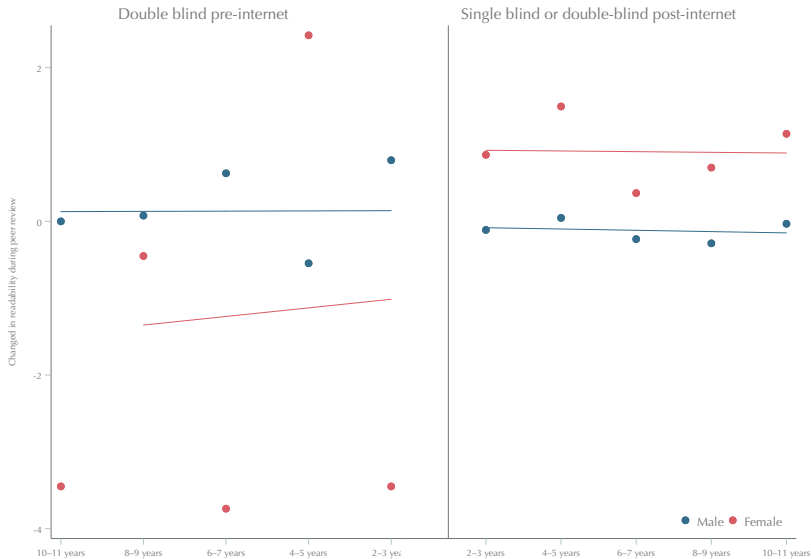
Causal impact of peer review



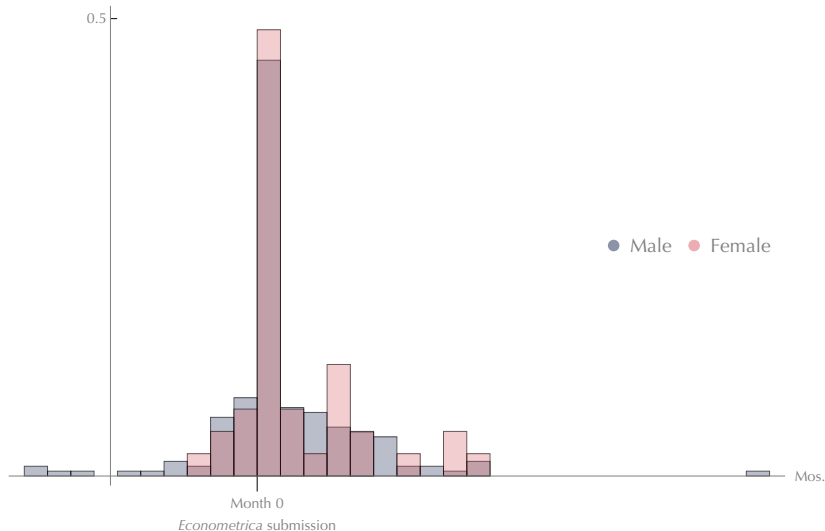
Causal impact of peer review



Causal impact of peer review



NBER Working Papers



- ▶ Female-authored manuscripts are submitted to journals *first* and released as NBER Working Papers only *afterwards*.

Causal impact of discrimination: theory

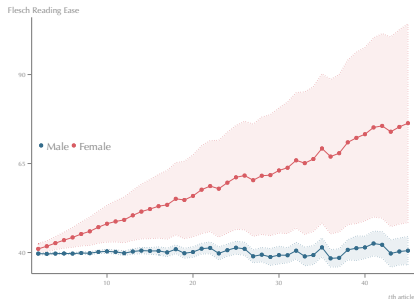
Why does peer review cause women to write more clearly?

Possibility 1 Women voluntarily write better papers—e.g., they're more sensitive to referee criticism.

Possibility 2 Better written papers are women's response to higher standards imposed by referees and/or editors.

- ▶ Model an author's decision making process within a subjective expected utility framework.
- ▶ Establish 3 sufficient conditions that distinguish Possibility 1 from Possibility 2.
 1. Experienced women write better than equivalent men.
 2. Women improve their writing over time.
 3. Female-authored papers are accepted no more often than equivalent male-authored papers.

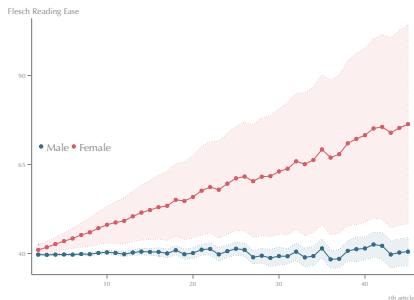
Causal impact of discrimination: evidence (I)



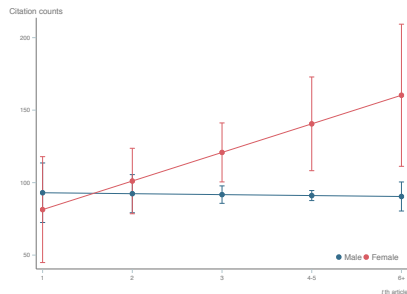
1. Experienced female economists write better than equivalent male economists
2. Women improve their writing over time.

No female advantage in acceptance rates (Ceci et al., 2014).

Causal impact of discrimination: evidence (I)



1. Experienced female economists write better than equivalent male economists
2. Women improve their writing over time.



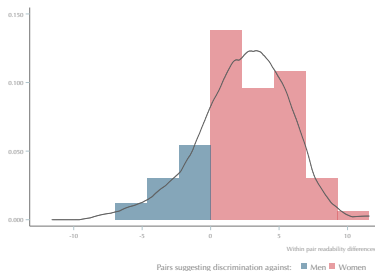
1. Experienced female economists are cited more than equivalent male economists.
2. Women increase citation counts over time.

No female advantage in acceptance rates (Ceci et al., 2014).

Causal impact of discrimination: evidence (II)

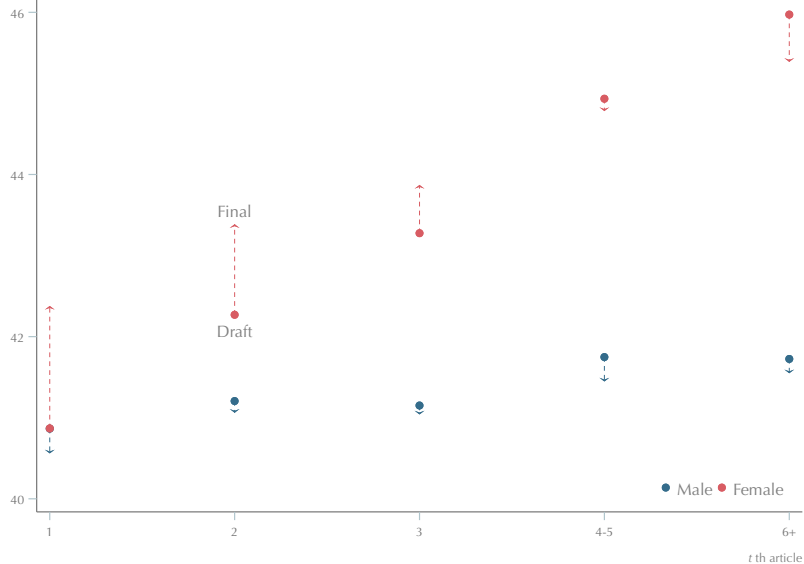
- ▶ Use a matching estimator to account for the fact that each condition must hold for the same author in two different situations:
 - ▶ Before and after gaining experience.
 - ▶ When compared to an equivalent, experienced author of the opposite gender.
- ▶ Matches based on ten observable characteristics: primary *JEL* category, citation counts, decade, institution, *etc.*

- ▶ Evidence of discrimination in 60–70 percent of matched pairs.
 - ▶ Subtracted experienced male scores from experienced female scores within each of these matched pairs.



Behavioural changes

Flesch Reading Ease



Prolonged peer review

	(1)	(2) ^a	(3)	(4)	(5)	(6)
Female ratio	5.29** (2.01)	6.63*** (2.16)	6.64*** (2.14)	5.54*** (2.05)	6.65*** (2.15)	8.80*** (2.72)
Max. t_j	-0.16** (0.07)	-0.17** (0.07)	-0.17** (0.07)	-0.16** (0.07)	-0.16** (0.07)	-0.17* (0.09)
No. pages	0.18*** (0.03)	0.18*** (0.03)	0.18*** (0.03)	0.18*** (0.03)	0.18*** (0.03)	0.21*** (0.04)
N	1.02** (0.44)	0.97** (0.44)	0.96** (0.44)	1.01** (0.44)	0.97** (0.44)	1.149 (0.70)
Order	0.22** (0.09)	0.22** (0.09)	0.22** (0.09)	0.22** (0.09)	0.22** (0.09)	0.50** (0.22)
No. citations	0.00 (0.000)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00*** (0.00)
Mother			-6.66** (2.68)		-10.93*** (3.21)	-17.67*** (3.29)
Birth				-2.25 (3.36)	7.58* (4.17)	12.34** (5.59)
Constant	37.71*** (2.04)	37.60*** (2.08)	37.79*** (2.05)	37.69*** (2.05)	37.89*** (2.06)	14.85*** (2.79)
Editor effects	✓	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓	✓
Institution effects	✓	✓	✓	✓	✓	✓
<i>JEL</i> (primary) effects						✓
No. observations	2,626	2,610	2,626	2,626	2,626	1,281

Notes. Sample 2,626 articles. Standard errors clustered by year in parentheses. ***, ** and * statistically significant at 1%, 5% and 10%.

^a Excludes papers authored only by women who gave birth (9 articles) and/or had a child younger than five (16 articles) during peer review.

Conclusions for academia

Implications for measuring productivity

- ▶ Women may produce better quality output. . .
- ▶ But quality costs time, so women produce less.
- ▶ Women appear less productive than they actually are.

“Publishing Paradox” may not be so paradoxical. . .

References I

- Abrevaya, J. and D. S. Hamermesh (2012). “Charity and Favoritism in the Field: Are Female Economists Nicer (to Each Other)?” *Review of Economics and Statistics* 94(1), pp. 202–207.
- Arellano, M. and O. Bover (1995). “Another Look at the Instrumental Variable Estimation of Error-components Models”. *Journal of Econometrics* 68(1), pp. 29–51.
- Blank, R. M. (1991). “The Effects of Double-blind versus Single-blind Reviewing: Experimental Evidence from the American Economic Review”. *American Economic Review* 81(5), pp. 1041–1067.
- Blundell, R. and S. Bond (1998). “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”. *Journal of Econometrics* 87(1), pp. 115–143.

References II

- Ceci, S. J. et al. (2014). “Women in Academic Science: A Changing Landscape”. *Psychological Science in the Public Interest* 15(3), pp. 75–141.
- DuBay, W. H. (2004). *The Principles of Readability*. Costa Mesa, California: Impact Information.
- Foschi, M. (1996). “Double Standards in the Evaluation of Men and Women”. *Social Psychology Quarterly* 59(3), pp. 237–254.
- Gilbert, J. R., E. S. Williams, and G. D. Lundberg (1994). “Is There Gender Bias in JAMA’s Peer Review Process?” *Journal of the American Medical Association* 272(2), pp. 139–142.
- Goldberg, P. (1968). “Are Women Prejudiced against Women?” *Trans-action* 5(5), pp. 28–30.
- Grunspan, D. Z. et al. (2016). “Males Under-estimate Academic Performance of Their Female Peers in Undergraduate Biology Classrooms”. *PLOS ONE* 11(2), pp. 1–16.

References III

- Hartley, J., J. W. Pennebaker, and C. Fox (2003). “Abstracts, Introductions and Discussions: How Far Do They Differ in Style?” *Scientometrics* 57(3), pp. 389–398.
- Heilman, M. E. and M. C. Haynes (2005). “No Credit Where Credit Is Due: Attributional Rationalization of Women’s Success in Male-female Teams”. *Journal of Applied Psychology* 90(5), pp. 905–916.
- Krawczyk, M. and M. Smyk (2016). “Author’s Gender Affects Rating of Academic Articles: Evidence from an Incentivized, Deception-free Laboratory Experiment”. *European Economic Review* 90, pp. 326–335.
- Moss-Racusin, C. A. et al. (2012). “Science Faculty’s Subtle Gender Biases Favor Male Students”. *Proceedings of the National Academy of Sciences* 109(41), pp. 16474–16479.
- Paludi, M. A. and W. D. Bauer (1983). “Goldberg Revisited: What’s in an Author’s Name”. *Sex Roles* 9(3), pp. 387–390.

References IV

Plavén-Sigray, P. et al. (2017). “The Readability Of Scientific Texts Is Decreasing Over Time”. *bioRxiv*, p. 119370.

Sarsons, H. (2017). “Recognition for Group Work: Gender Differences in Academia”. *American Economic Review* 107(5), pp. 141–145.