# Publishing while female

## Are women held to higher standards? Evidence from peer review

Erin Hengel

University of Liverpool

Presentation by: Alireza Tavakoli

August 2019

# Outline

# Outline

# Journal Decision-making Process

## What is Peer Review?

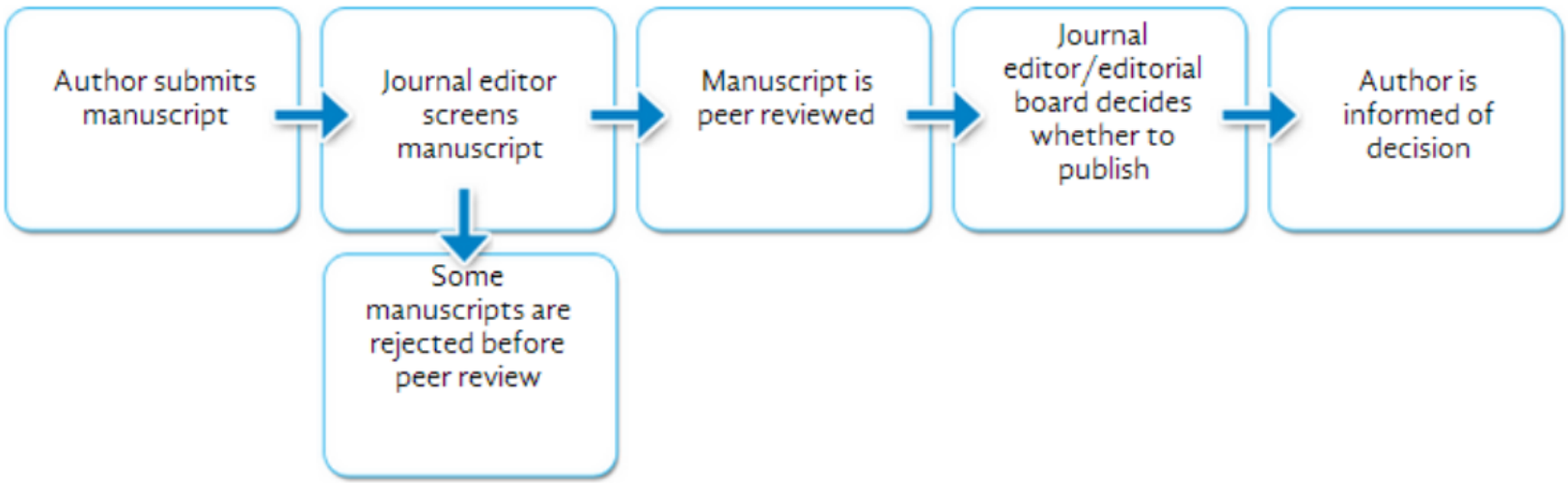



Figure 1: Peer review process and editorial decision making at journals

# Outline

# Literature Review

1. Quantity vs. quality trade-off
2. Higher standards for women
3. Gender discrimination in peer review
4. 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.

*"Whatever women do they must do twice as well as men to be thought half as good. Luckily, this is not difficult."*

-Charlotte Whitton

*"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

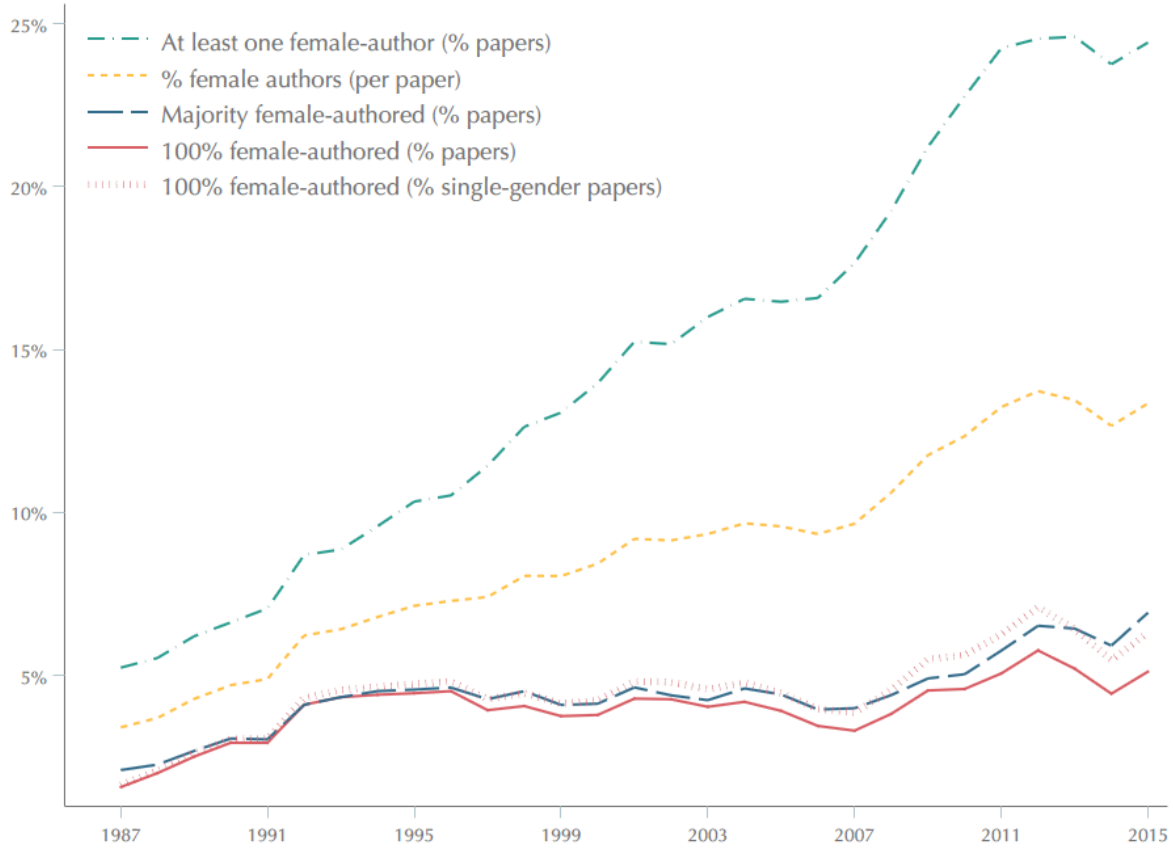
# Outline

# Paper's Contribution

Using five readability measures it finds that:

- Female-authored articles published in top economics journals are better written than equivalent papers by men.
- Peer review causes a large increase in the readability gap.
- Female-authored papers take half a year longer in peer review.
- Tougher editorial standards and/or biased referee assignment are uniquely consistent with womens observed pattern of choices.

# The Representation of Women in Top Economics Journals



# Write crisply but clearly!

## Writing clarity

1. Clear writing is valued by journals.
2. Good writing $\approx f$ (simple vocabulary, short sentences).

| Score | Formula |
|---|---|
| Flesch Reading Ease | $206.84 - 1.02 \times \frac{\text{words}}{\text{sentences}} - 84.60 \times \frac{\text{syllables}}{\text{words}}$ |
| Flesch-Kincaid | $-15.59 + 0.39 \times \frac{\text{words}}{\text{sentences}} + 11.80 \times \frac{\text{syllables}}{\text{words}}$ |
| Gunning Fog | $0.40 \times \left(\frac{\text{words}}{\text{sentences}} + 100 \times \frac{\text{polysyllabic words}}{\text{words}}\right)$ |
| SMOG | $3.13 + 5.71 \times \sqrt{\frac{\text{polysyllabic words}}{\text{sentences}}}$ |
| Dale-Chall | $3.64 + 0.05 \times \frac{\text{words}}{\text{sentences}} + 15.79 \times \frac{\text{difficult words}}{\text{words}}$ |

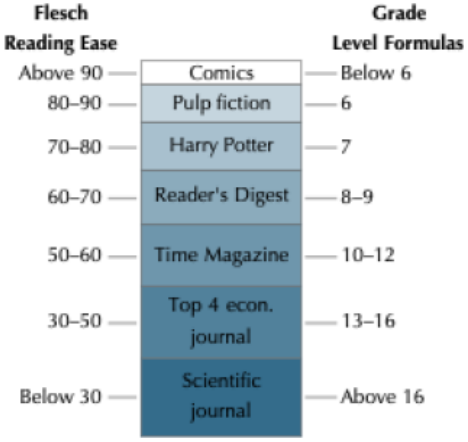



Figure 2: Calculating and interpreting readability scores

# Outline

# Data

- Every English article published in *AER*, *Econometrica*, *JPE* and *QJE* since 1950.
- NBER working papers
- Data of submit-accept times at *Econometrica* the only journal of the four to make disaggregated data on its revision process publicly available and *REStud*, a fifth highly respected economics journal.

# Outline

# 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.

# Step No. 1

# Regression

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

| | 1950–2015 | | | | | 1990–2015 | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Flesch Reading Ease | 0.89* | 0.86* | 0.82 | 0.88* | 1.13** | 0.79 | 0.79 | 0.97 |
| | (0.49) | (0.48) | (0.50) | (0.51) | (0.50) | (0.55) | (0.56) | (0.70) |
| Flesch-Kincaid | 0.17 | 0.17 | 0.17 | 0.18 | 0.21* | 0.24* | 0.26** | 0.26* |
| | (0.11) | (0.11) | (0.11) | (0.11) | (0.12) | (0.13) | (0.12) | (0.14) |
| Gunning Fog | 0.31** | 0.30** | 0.31** | 0.33** | 0.36*** | 0.39** | 0.38*** | 0.37** |
| | (0.12) | (0.12) | (0.12) | (0.13) | (0.13) | (0.15) | (0.14) | (0.16) |
| SMOG | 0.20** | 0.20** | 0.20** | 0.21** | 0.24** | 0.24** | 0.22** | 0.23* |
| | (0.09) | (0.09) | (0.09) | (0.09) | (0.09) | (0.11) | (0.10) | (0.12) |
| Dale-Chall | 0.10** | 0.10** | 0.09** | 0.10** | 0.12** | 0.13** | 0.12** | 0.15** |
| | (0.04) | (0.04) | (0.05) | (0.05) | (0.05) | (0.06) | (0.06) | (0.06) |
| Editor effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Journal effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year effects | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Journal×Year effects | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $N_j$ | | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Institution effects | | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Quality controls | | | | | ✓[1] | ✓[1] | ✓[1] | ✓[1] |
| Native speaker | | | | | ✓ | ✓ | ✓ | ✓ |
| *JEL* (primary) effects | | | | | | | ✓ | |
| *JEL* (tertiary) effects | | | | | | | | ✓ |

Figure 3: Gender differences in readability, article-level analysis

# Results

- Abstracts written only by women score about one point higher on the Flesch Reading Ease scale.
- According to the four grade-level measures, they take about 2 – 3 fewer months of schooling to understand.
- Percentage-wise, women write about 1 – 2 percent better than men.

$\Longrightarrow$ Gender difference in readability does exist.

Step No. 1 completed ✓

# Step No. 2

# Published vs. Draft

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | Working paper | Published article | Difference | Working paper | Published article | Difference |
| Flesch Reading Ease | 41.46 | 41.13 | −0.332* | 42.51 | 43.08 | 0.564 |
| | (0.26) | (0.18) | (0.185) | (0.66) | (0.43) | (0.452) |
| Flesch-Kincaid | −13.62 | −13.38 | 0.243*** | −13.53 | −13.00 | 0.531*** |
| | (0.06) | (0.05) | (0.050) | (0.15) | (0.11) | (0.122) |
| Gunning Fog | −17.28 | −17.04 | 0.242*** | −17.13 | −16.58 | 0.547*** |
| | (0.07) | (0.05) | (0.055) | (0.18) | (0.13) | (0.140) |
| SMOG | −15.14 | −15.00 | 0.135*** | −15.02 | −14.70 | 0.327*** |
| | (0.05) | (0.03) | (0.035) | (0.13) | (0.09) | (0.095) |
| Dale-Chall | −10.85 | −10.93 | −0.084*** | −10.71 | −10.70 | 0.003 |
| | (0.02) | (0.02) | (0.016) | (0.06) | (0.04) | (0.037) |

Figure 4: Textual characteristics, published papers vs. drafts

- Women's abstracts always become more readable during peer review than do those by men.

# Regressions

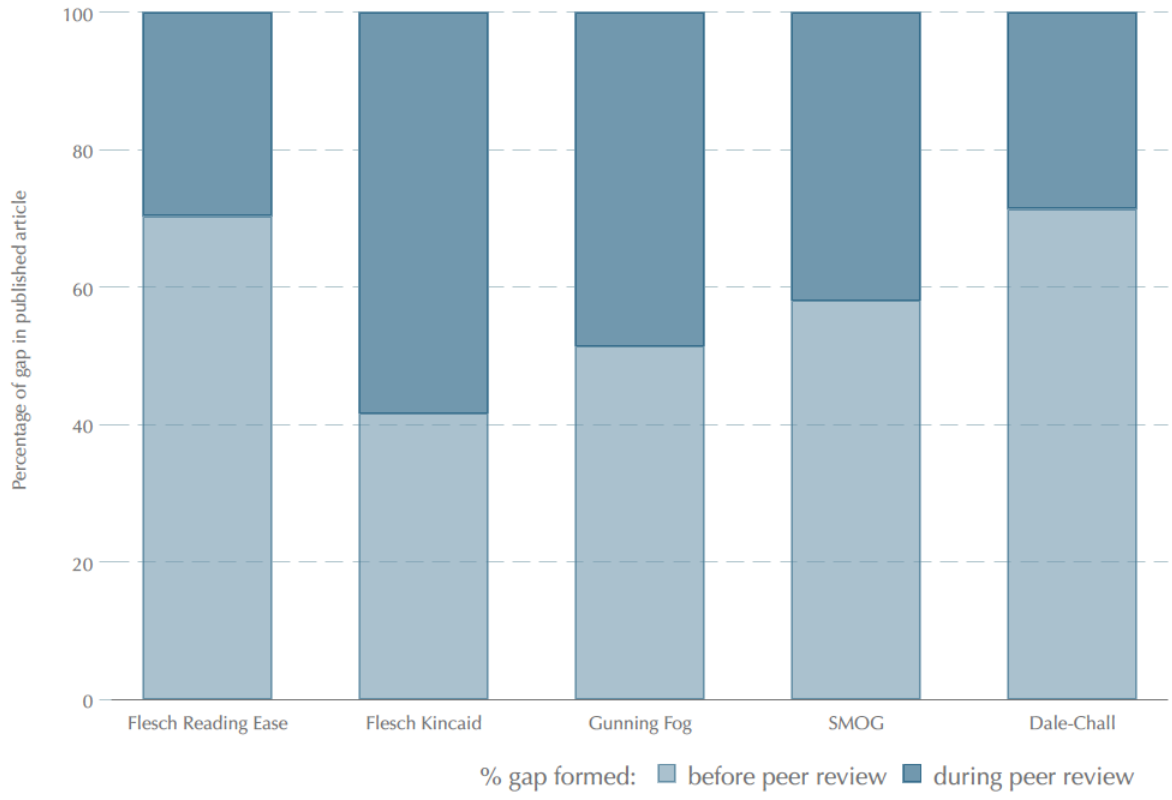
$$R_{jP} = R_{jW} + \beta_{0P} + \beta_{1P} \text{ female ratio }_j + \theta_P X_{jP} + \mu_{jP} + \varepsilon_{jP}$$

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

Figure 5: The impact of peer review on the gender readability gap

# Results

- Womens readability gains outpace mens between versions.
- Female authored working papers and published articles are both better written, but the readability gap is substantially larger in the latter.



# Timing independence

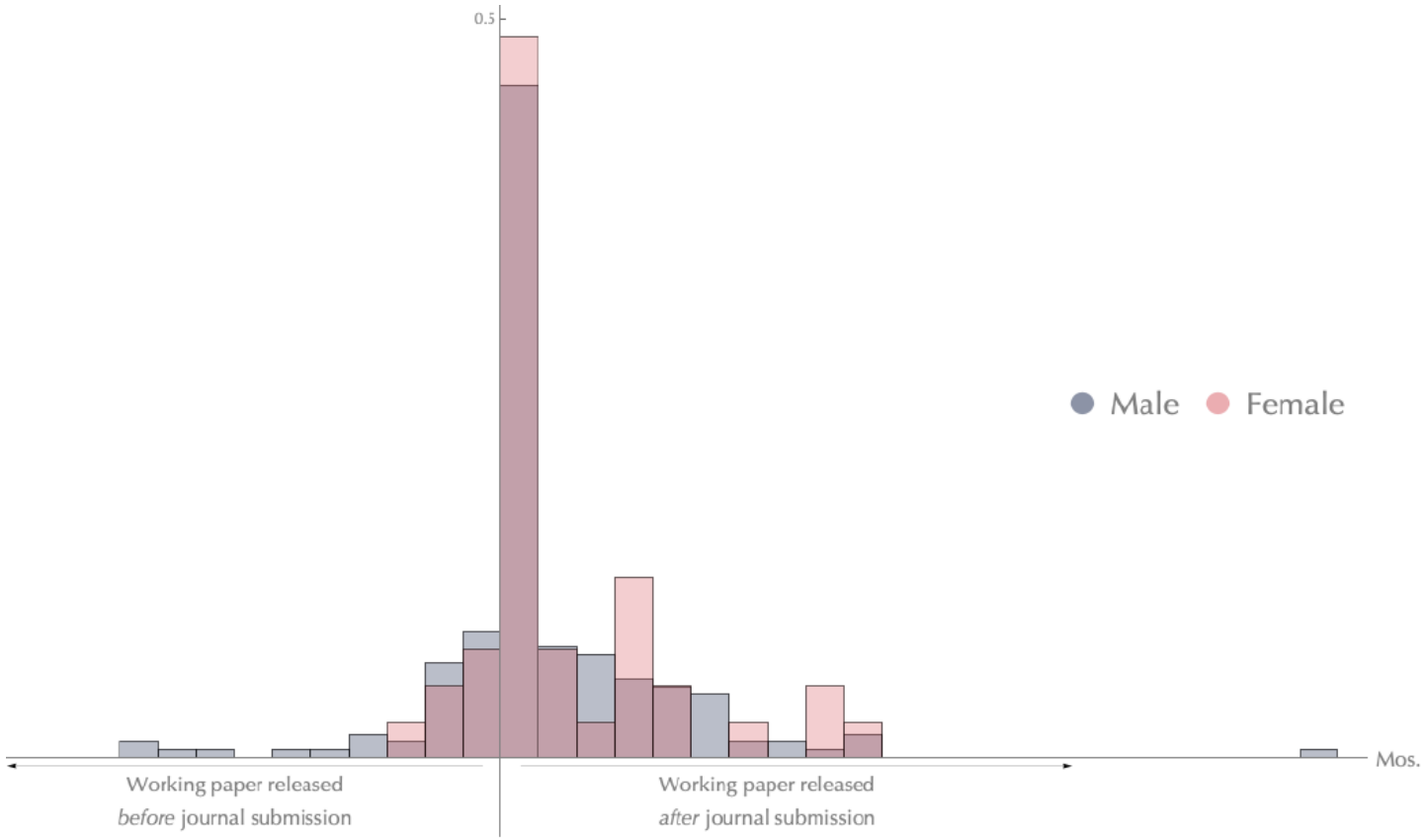



Figure 6: Distribution of months between NBER release and journal submission

# Robustness

- Timing independence is the principle assumption required to causally link the readability gap to the peer review process
- It is arguably only violated during the narrow time frame after a manuscript is released as an NBER Working Paper but before it is submitted to a top-four journal.
- Other robustness checks ✓

$\Longrightarrow$ The wider gap post-peer review confirms a causal link with peer review.

Step No. 2 completed ✓

# Step No. 3

# 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.

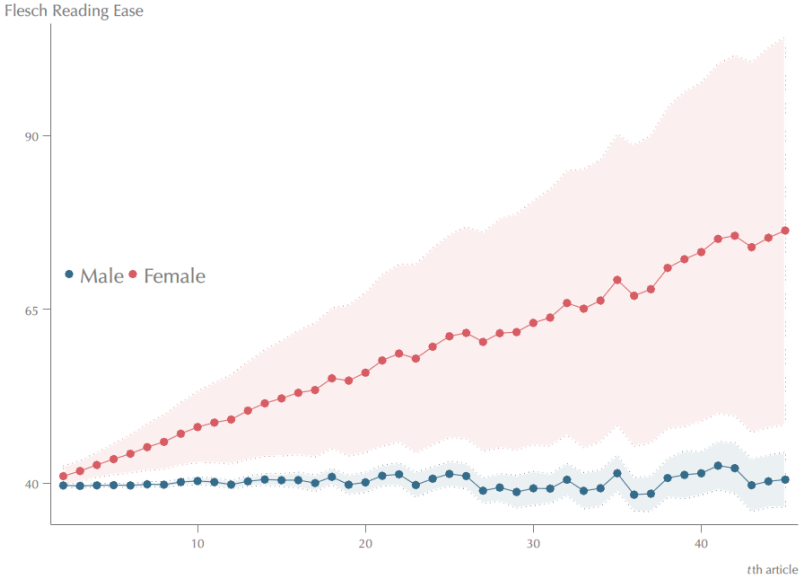
- Model an author's decision making process within a subjective expected utility framework.
- Establish 3 sufficient conditions that distinguish Possibility 1 from Possibility 2.
  1. Experienced women write better than equivalent men.
  2. Women improve their writing over time.
  3. Female-authored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

# Intuition

- Assuming preferences do not change over time.
- Authors improve readability today relative to yesterday only if they believe better writing leads to higher acceptance rates.
- Of course, over sensitivity and/or poor information may distort their beliefsand affect readabilitybut the impact declines with experience.
- Holding acceptance rates constant.

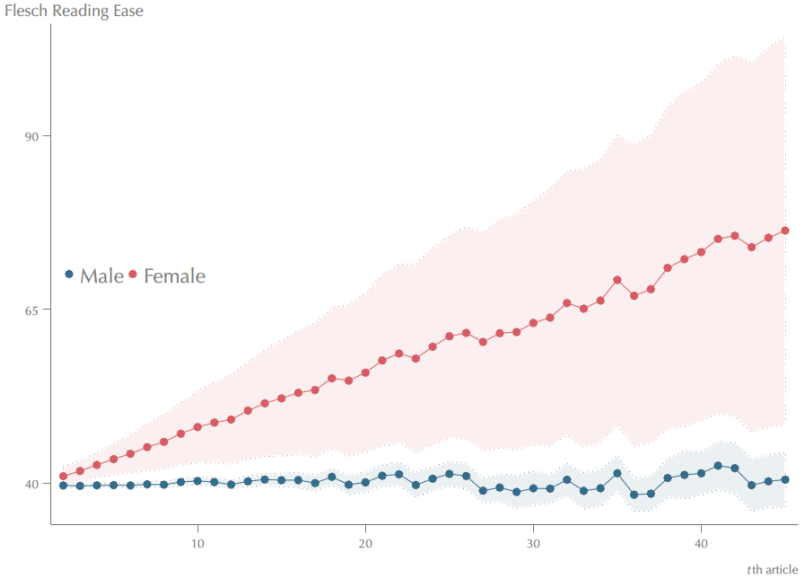
  Thus, A widening readability gap between equivalent authors is caused by discrimination i.e., asymmetric editorial standards and/or biased referee assignment beyond womens control.

# Evidence

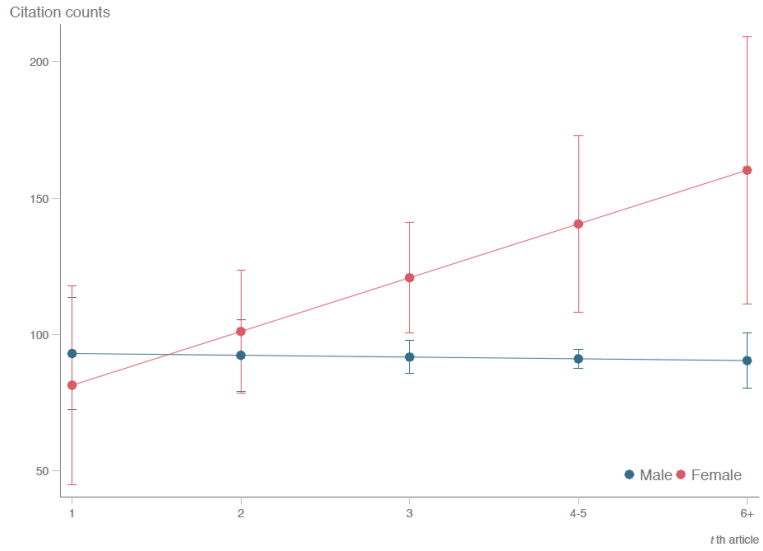



1. Experienced female economists write better than equivalent male economists.
2. Women improve their writing over time.

# Evidence

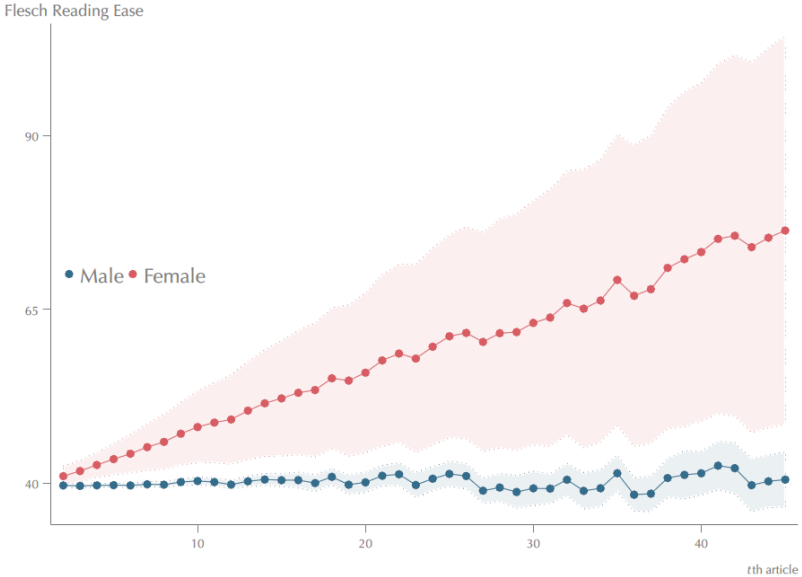


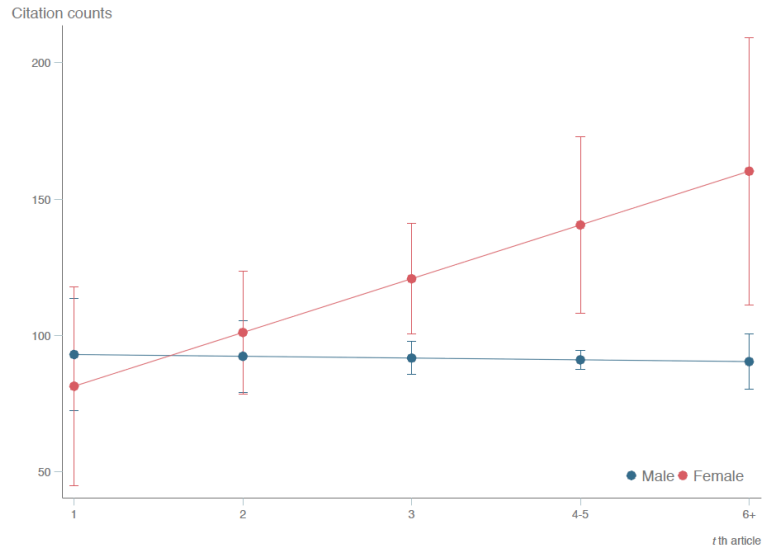

1. Experienced female economists write better than equivalent male economists.
2. Women improve their writing over time.




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

# Evidence




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

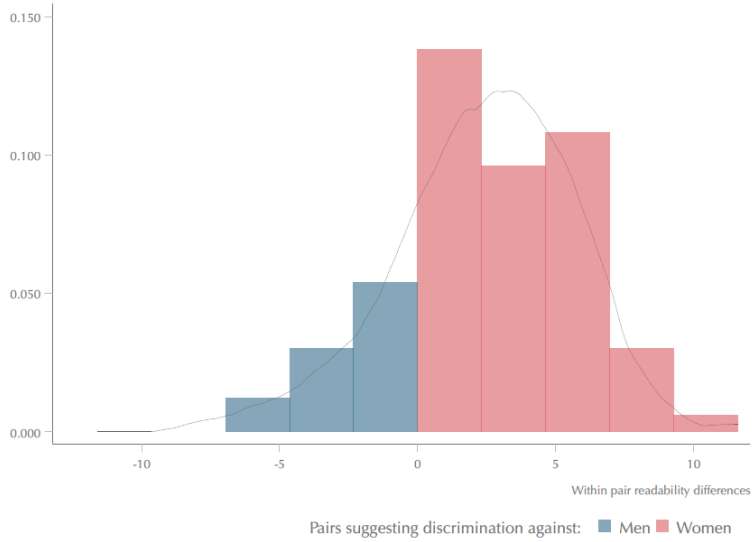
1. Experienced female economists are cited more than equivalent male economists.
2. Women increase citation counts over time..

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

# Evidence

- 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.



Step No. 3 completed ✓

# Outline

# Prolonged peer review

- Male authored papers take (on average) 18.5 months to complete all revisions; equivalent papers by women need half a year longer.

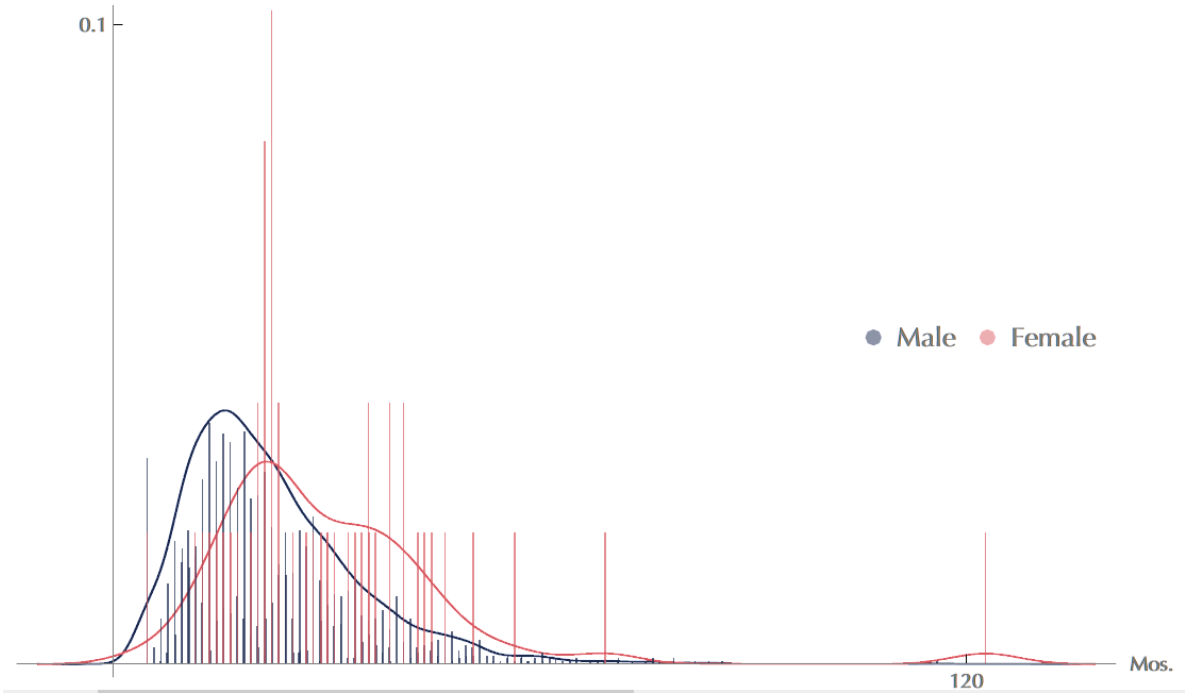



Figure 7: Distribution of review times by author sex. Blue bars represent papers written only by men; pink bars are papers written only by women.

# How do women react to higher standards?

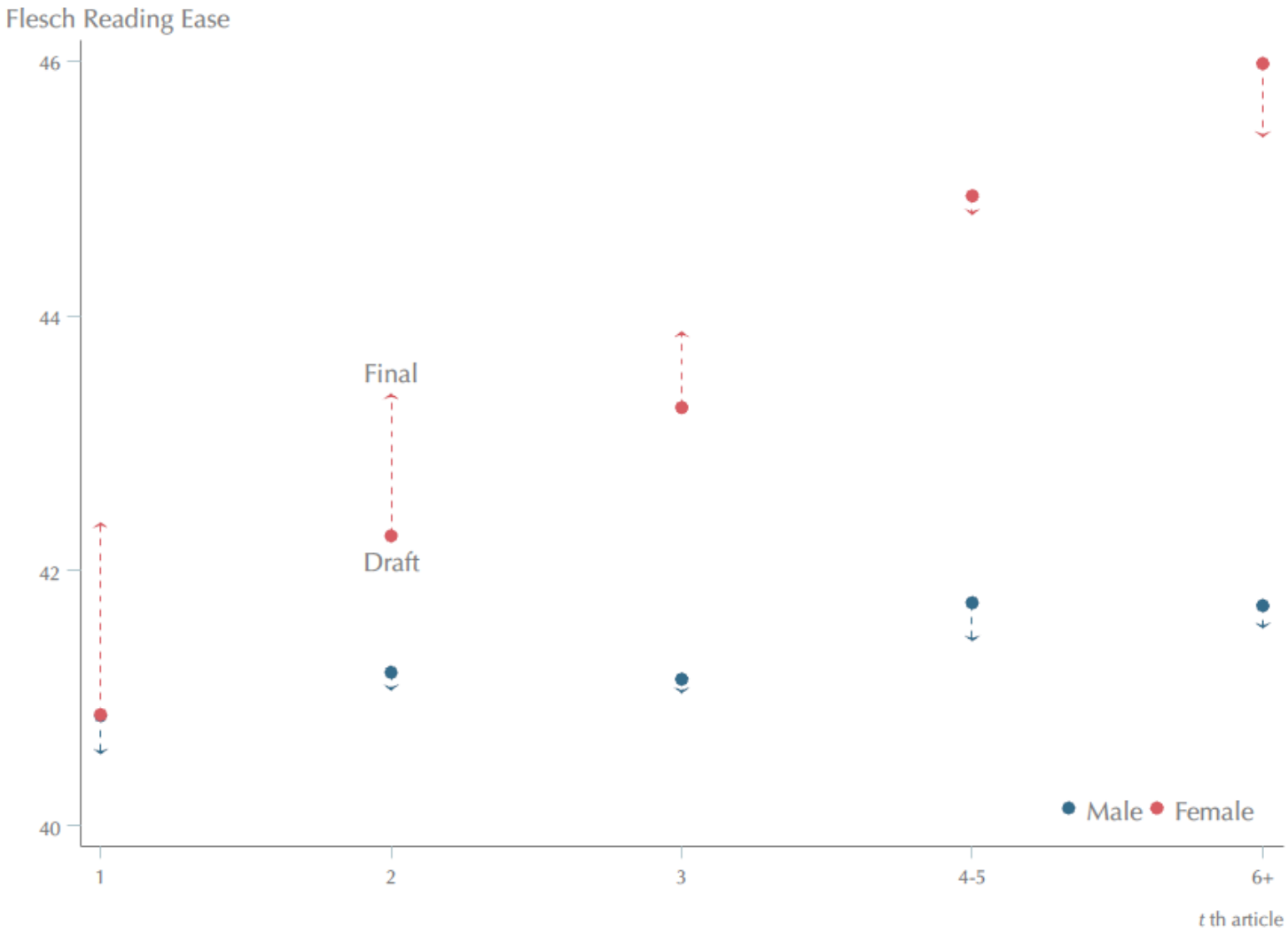



Figure 8: Readability of authors $t_{th}$ top-four publication (draft and final)

# Conclusions

## 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...

Thanks for your attention!