

Do male researchers disregard the work of female researchers? The role of gender in citation decisions

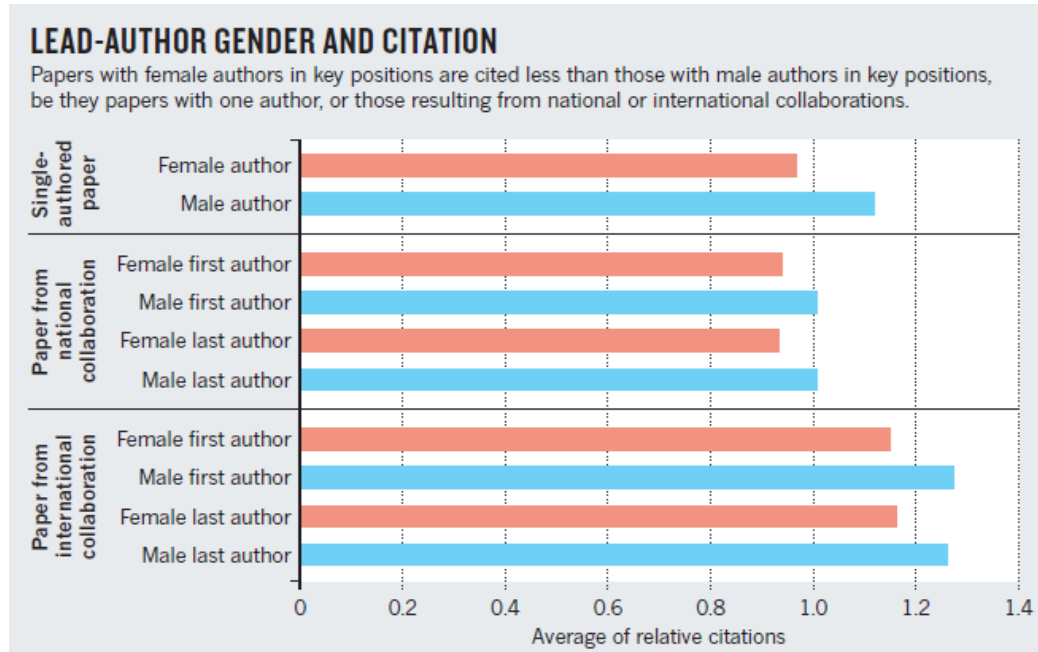
Analytical Sociology: Theory and Empirical Applications
Venice International University, Nov 2019

Alexander Tekles, Katrin Auspurg & Lutz Bornmann
LMU Munich / MPG



Motivation: The Gender Gap in Citations

- Male's work is more frequently cited than female's work

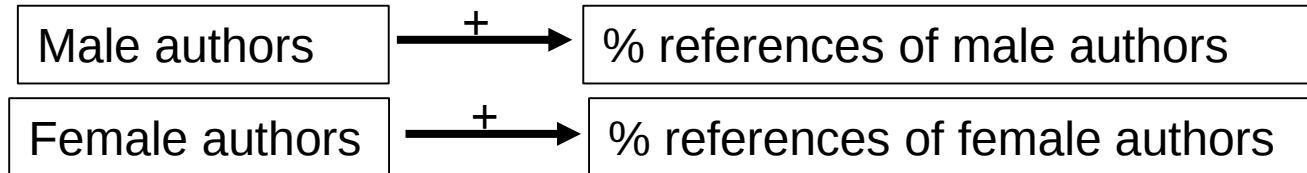


Based on 5,483,841 research papers in the Web of Science database 2008-2013 (Sugimoto 2018)

- Caused by a “gender homophily bias”?

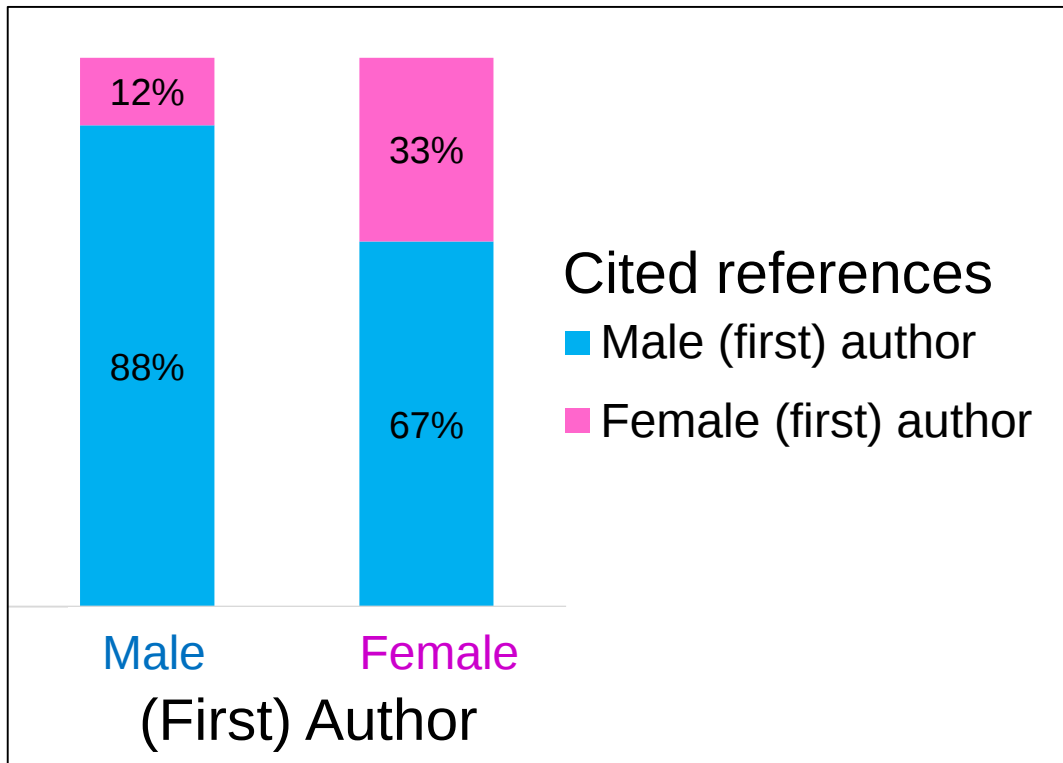
Gender Homophily Bias in Citations

- Disproportionately citing references of own gender



- Example: articles in sociology journals 1985-1994

(Davenport 1995; random sample out of SSCI-journals)

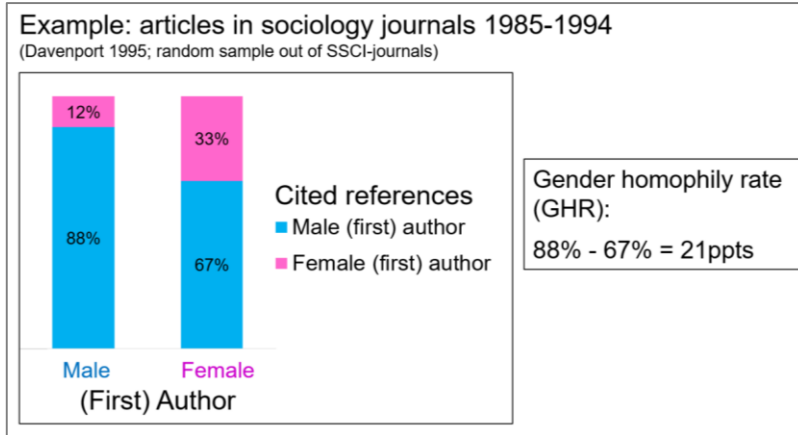


Gender homophily rate (GHR):

$$88\% - 67\% = 21\text{ppts}$$

Gender Homophily Bias in Citations

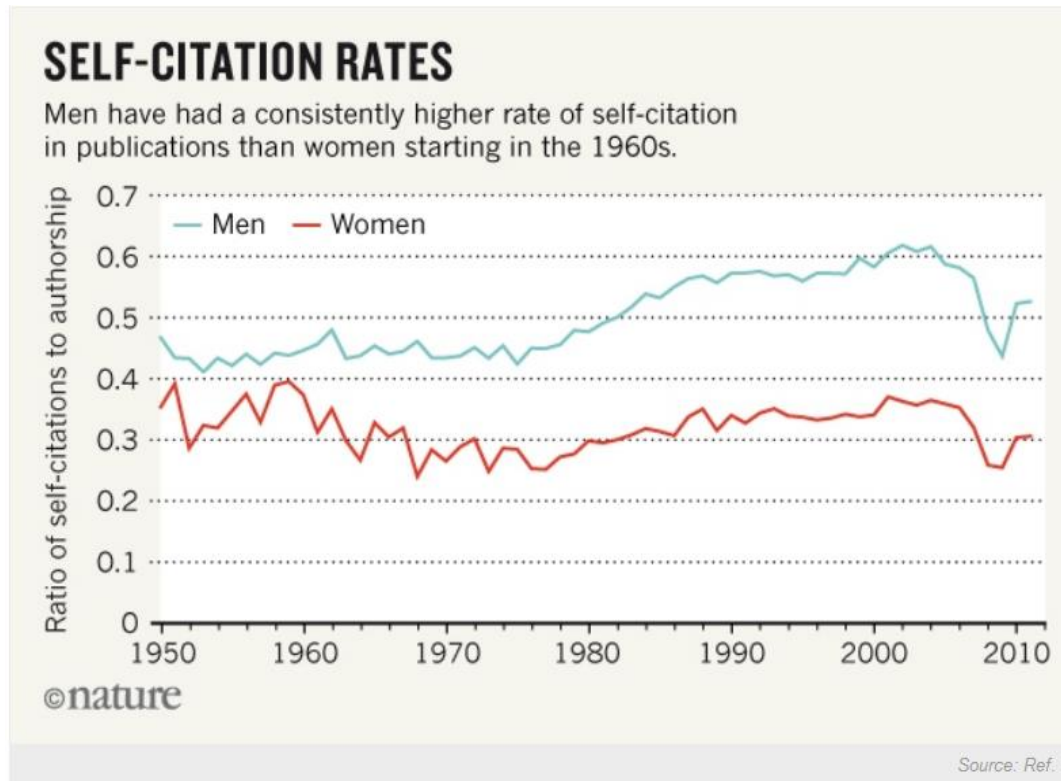
- Authors disproportionately cite references of own gender



- So far all studies report evidence for gender homophily
 - $N = 12$ studies, covering many disciplines
 - Mean GHR (also including mixed gender teams): 12ppts

Partly Caused by Self-Citations

- In particular men tend to cite themselves



- But even without self-citations studies find evidence for gender homophily bias (e.g. Ghiasi et al. 2018; Pothoff & Zimmermann 2017)

Further Explanations

- “Matilda effects”: Less recognition of the work of females (Ferber et al. 1986; Kanter 1970)
- But females might be more likely be aware of and cite females’ work (e.g. because of same-gender networks)
- “Implicit biases” (but how, and why??)
- So far there is not any clear and consistent explanation
- But there are lots of policy recommendations

Already Some Actions Taken

SWEDEN

Row over 40% gender quota for reading lists at Lund

Jan Petter Myklebust 24 November 2017



114TH 2019 ANNUAL MEETING | AUGUST 10-13, 2019 | NEW YORK, NY UNITED STATES

3156. Presidential Panel. White Supremacy and Sociology

Sheraton New York, Metropolitan Ballroom

West, Second Floor, 8:30-10:00 AM

Session Organizer: *Wendy Leo* *Northwestern University*

Presider: *Glenn Edward Bracey*, *University of Illinois at Chicago*

Sociology: An Organ of White Supremacy - *John D. Skrentny*

Racism - *Johnny E. Williams*, *University of Illinois at Chicago*

White Ignorance, Foreclosed Knowledge - *John D. Skrentny*

Damning Case of Sociology - *Jennifer C. Mueller*, *Skidmore College*

Revisiting White Sociology through the Publication Process - *David G. Embrick*, *University of Connecticut*

Revisiting White Sociology through the Publication Process - *David G. Embrick*, *University of Connecticut*

Revisiting White Sociology through the Publication Process - *David G. Embrick*, *University of Connecticut*

Revisiting White Sociology through the Publication Process - *David G. Embrick*, *University of Connecticut*

The Citation Inflation of White Masculinity - *Saida Grundy*, *Boston University*

The Citation Inflation of White Masculinity - *Saida Grundy*, *Boston University*



Ruth Pearce @NotRightRuth · 12. Aug.

Grundy: Segregated citation networks emerge because white men are not reading black feminism, queer theory etc. But also because of what we are (and are not) taught. Notes that black feminist scholars in sociology often taught only in feminist theory classes. #ASA19

1

3

14



Already Some Actions Taken



SAGE Research Methods Foundations Review Form

LEVEL

- **The content of the entry needs to be accessible - Could an upper level UG student or PG student new to the topic understand this flagship?**

☐ Yes ☐ No

Please briefly explain your reasoning, especially if your answer is no:

⋮

- **Does it appropriately cite male and female authors?**

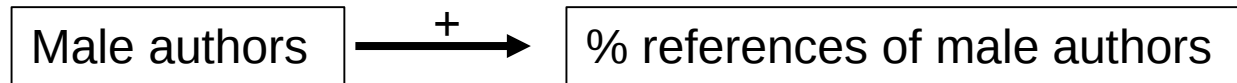
☐ Yes ☐ No

If no, please make suggestions for citations to add/delete to the author to ensure a balance:

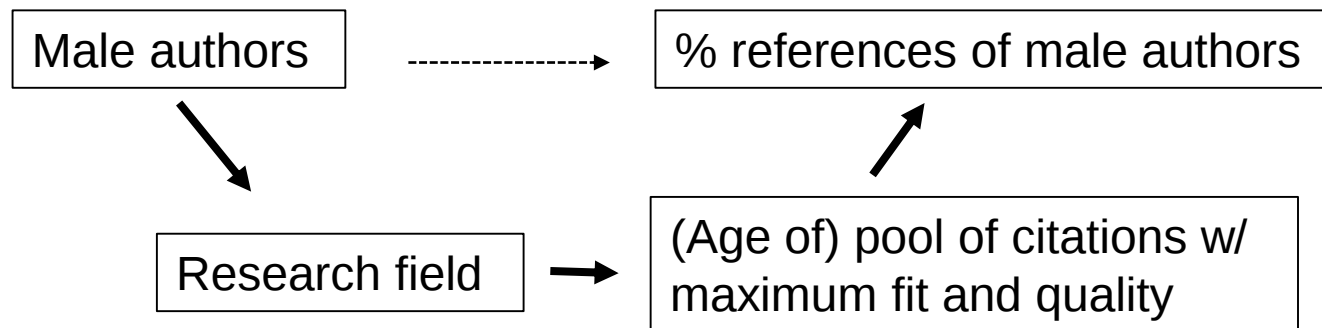
Sen 2018: p. 335

Citation Patterns Could Emerge from Unbiased Science?

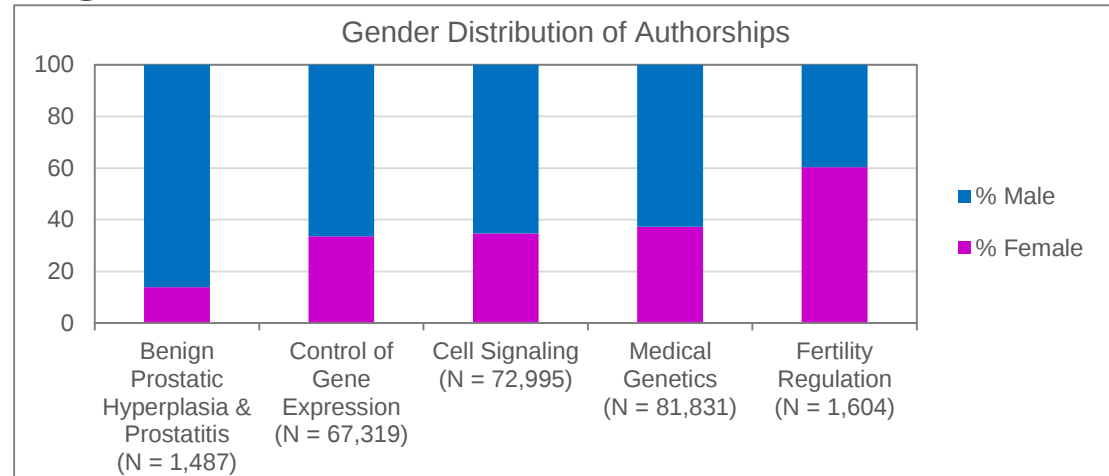
- Homophily bias would exist with a direct gender effect



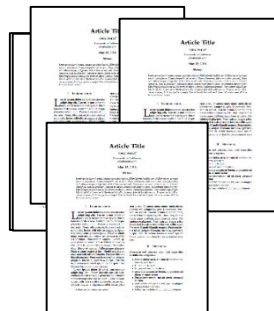
- Unbiased selection of references would exist with
 - Maximum substantive fit to research question
 - Maximum quality (rigor, impact)
 - Selection based on the whole population of existing references
- These factors could lead to indirect effects (mediators)



Horizontal Segregation: Research Field as Mediator



Authors

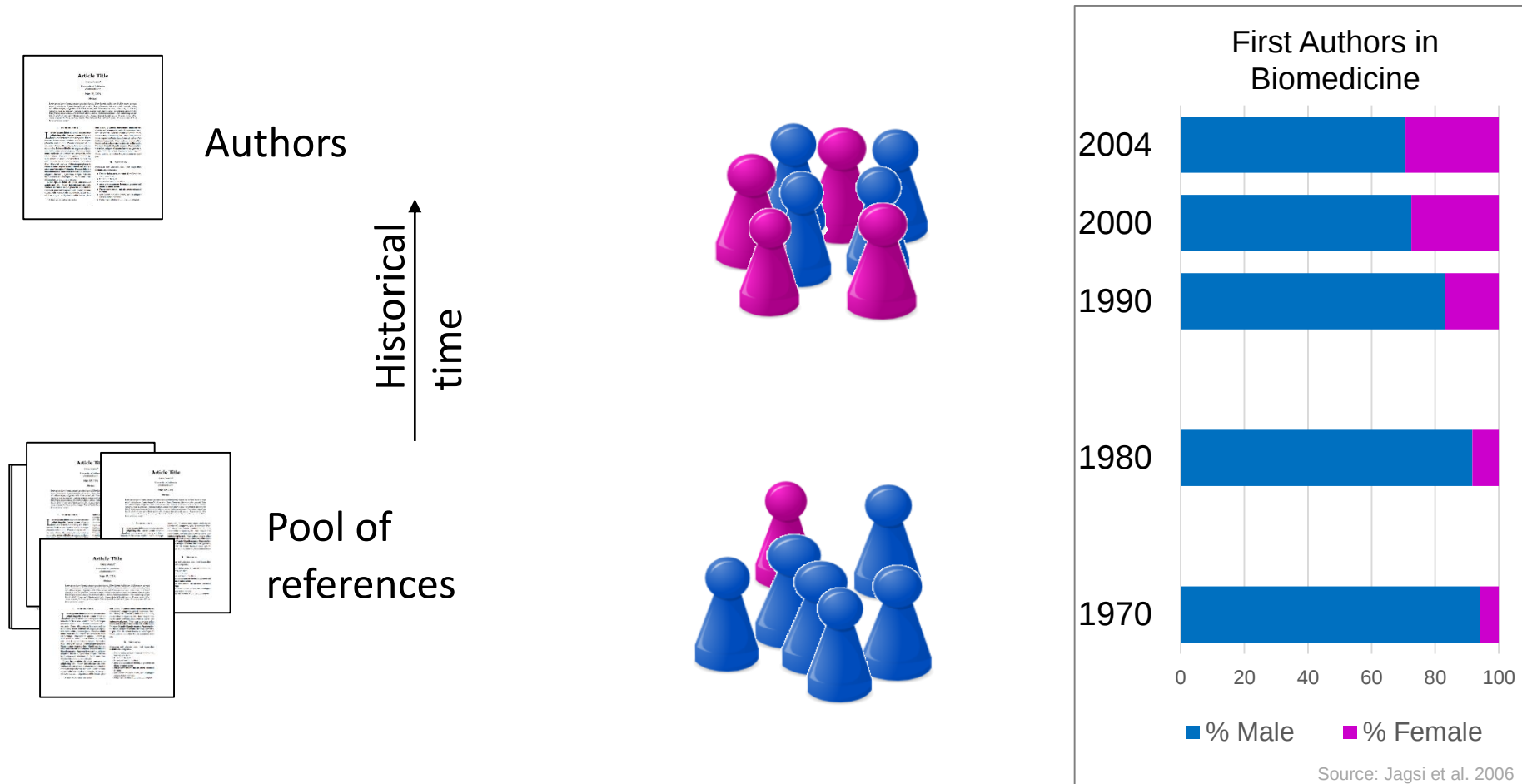


Pool of References



correlation
% (fe)male

Time Trends: Age of Research Topic as Mediator

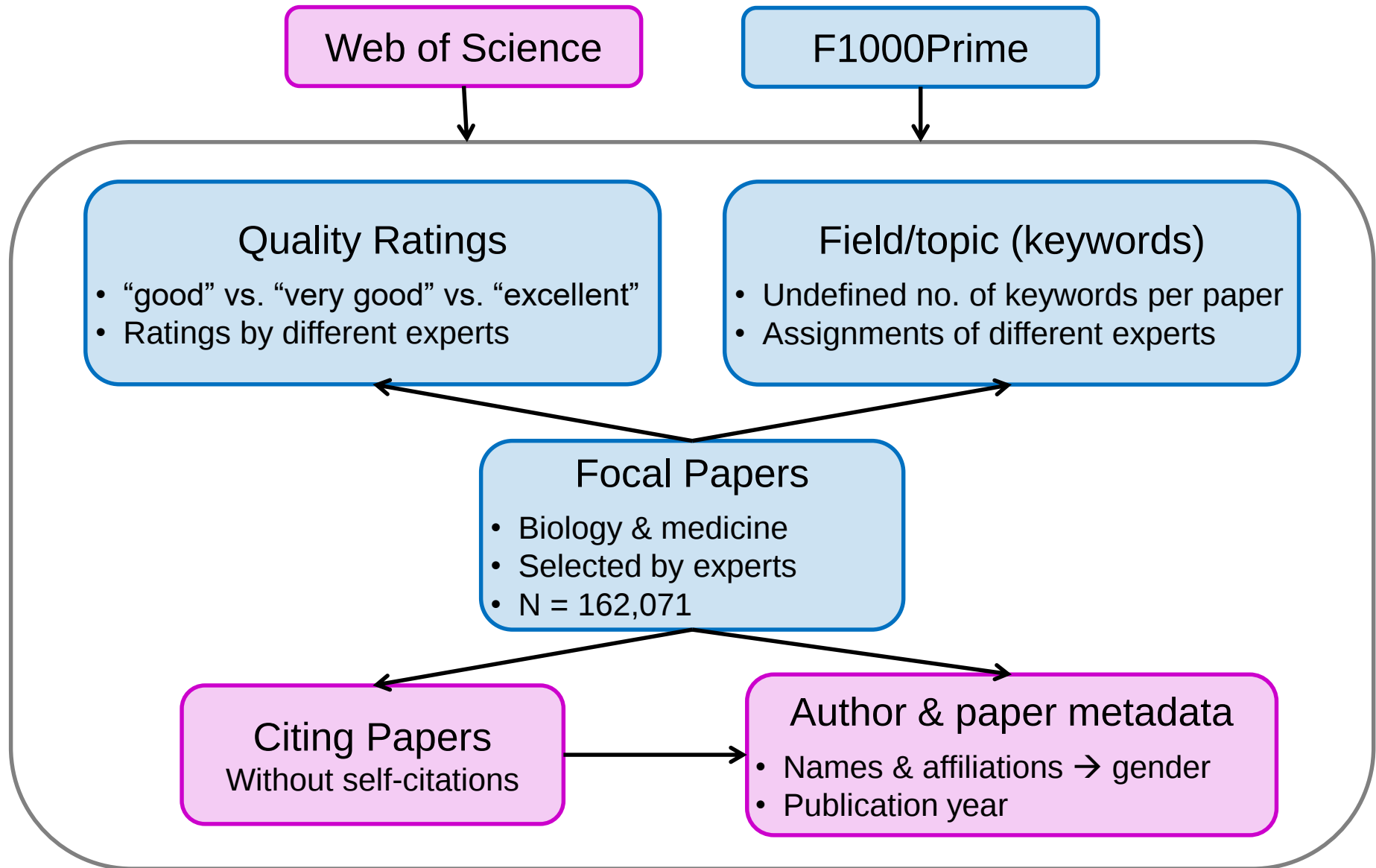


- Higher age of research topic → more references by males
- Males might more strongly focus on old, „classical“ topics (e.g. because of their higher academic age)

Our Contribution

- Does gender homophily exist when controlling mediators?
 - Sophisticated measurement of **fields/research topics**
 - Indicator for **quality** of cited papers

Data



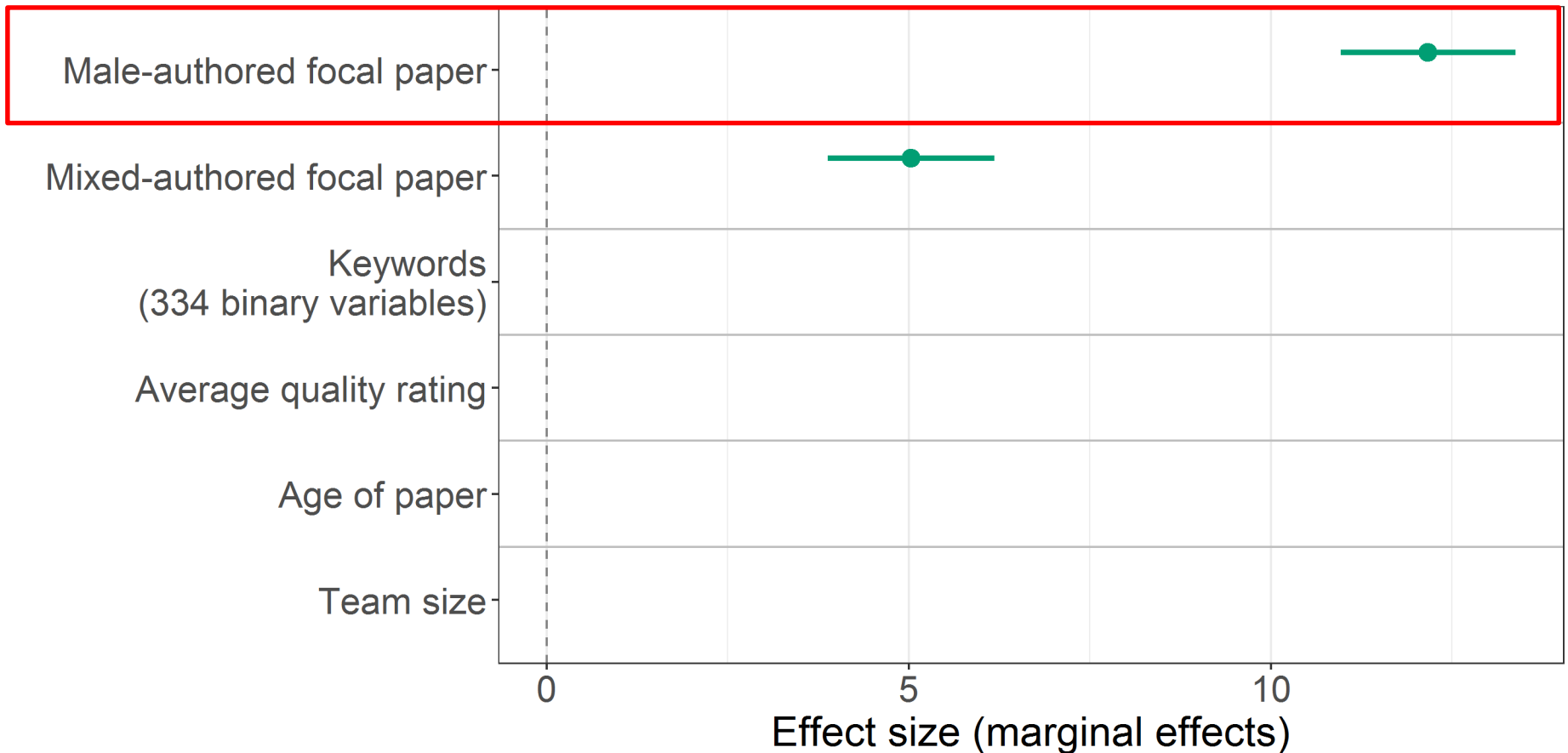
Identification Strategy

- For focal papers, estimate the effect of focal paper gender on the share of male-authored papers among citing papers
- Linear regression
 - Units: focal papers
 - Dep. variable: share of male-authored papers among citing papers
 - Homophily: effect of focal paper gender
 - Control for keywords, quality rating, age of paper, team size
- Expectation
 - Without including control variables: positive effect of focal paper being authored by males
 - After adding control variables: smaller (no) effect of focal paper being authored by males

Preliminary Results (I)

DV: Share of male-authored papers among citing papers (%)

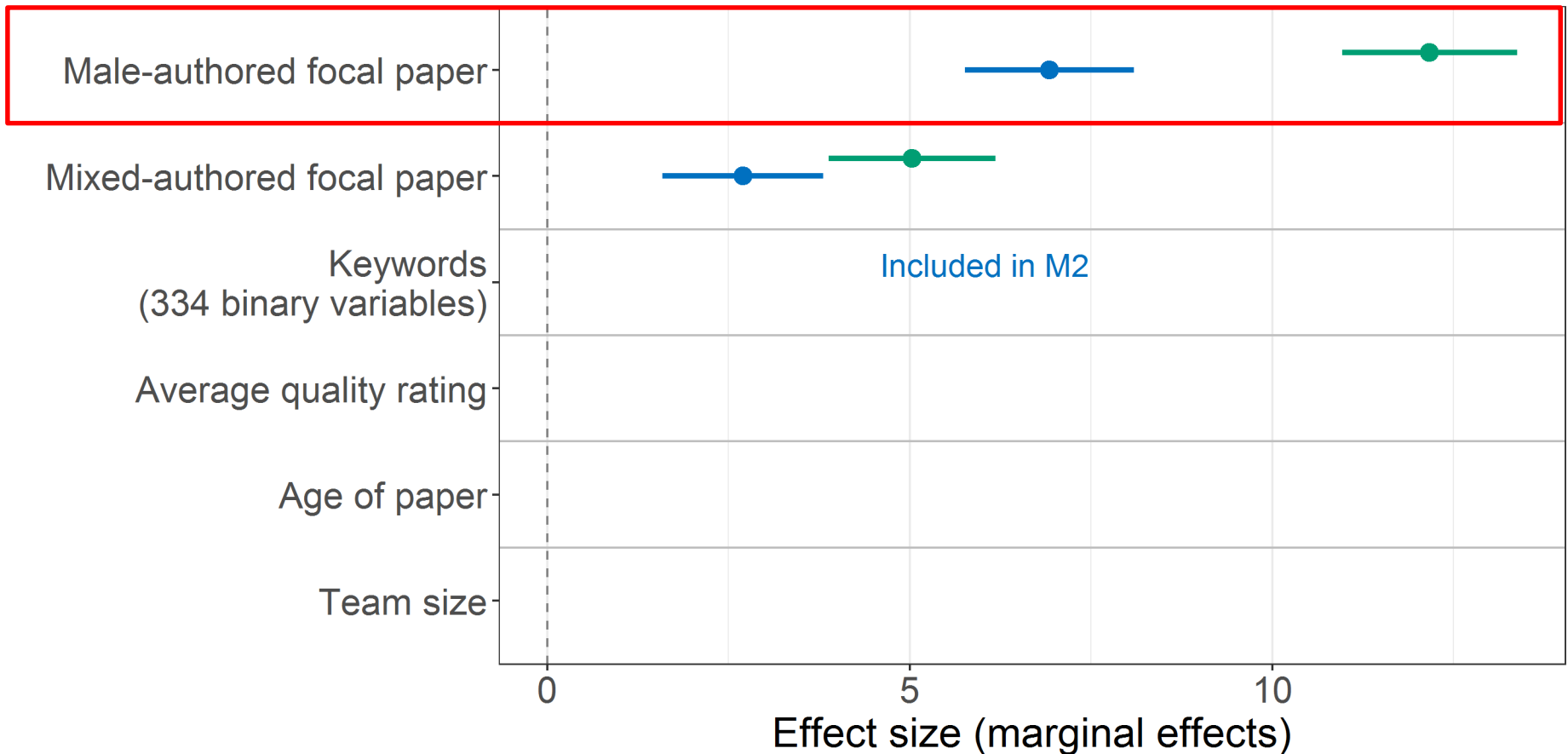
- M1: No controls (N = 42,718)



Preliminary Results (I)

DV: Share of male-authored papers among citing papers (%)

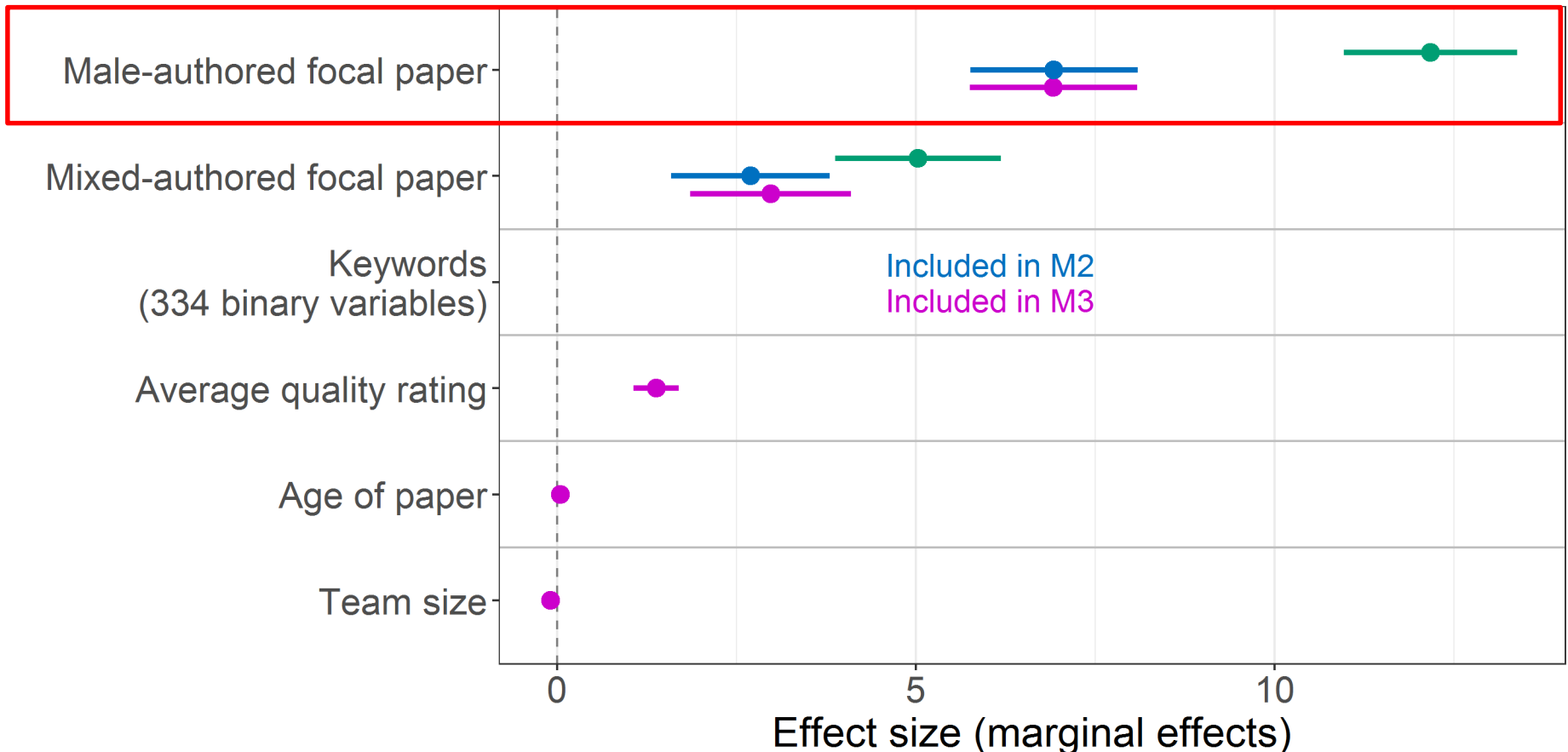
- M1: No controls (N = 42,718)
- M2: Controlling for field/topic keywords (N = 42,676)



Preliminary Results (I)

DV: Share of male-authored papers among citing papers (%)

- M1: No controls (N = 42,718)
- M2: Controlling for field/topic keywords (N = 42,676)
- M3: Full model (N = 42,676)

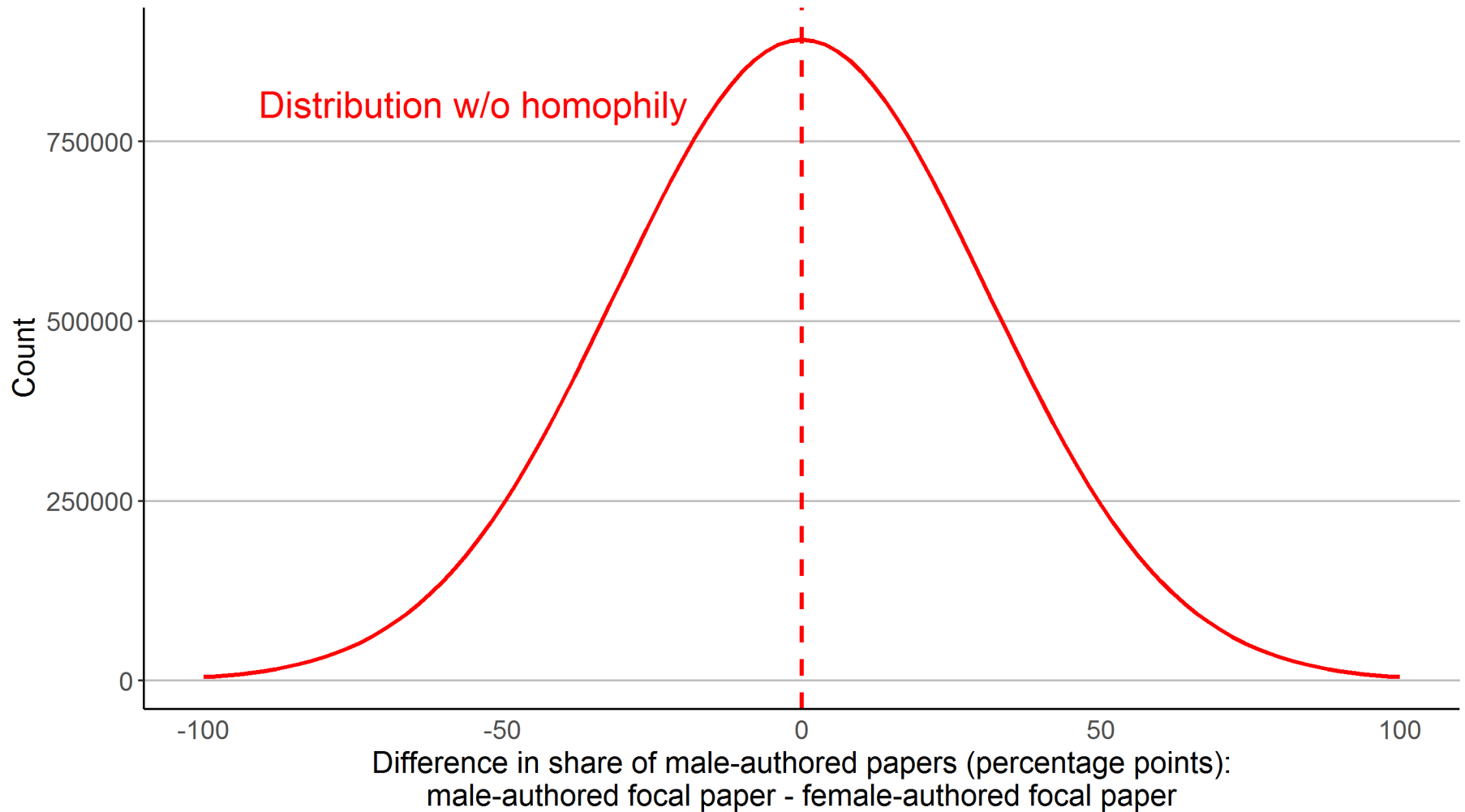


Preliminary Results (II)

- Considering binary variables for field/topic only controls effect of each keyword, independently of other keywords
- But: Fields/topics may be better represented by certain combinations of keywords
- Idea: for pairs of focal papers, use the number of shared keywords as indicator for topical similarity
- For all pairs of focal papers (one female-authored, one male-authored) with at least x shared keywords: plot histogram of the difference in the share of male-authored papers in the citing papers

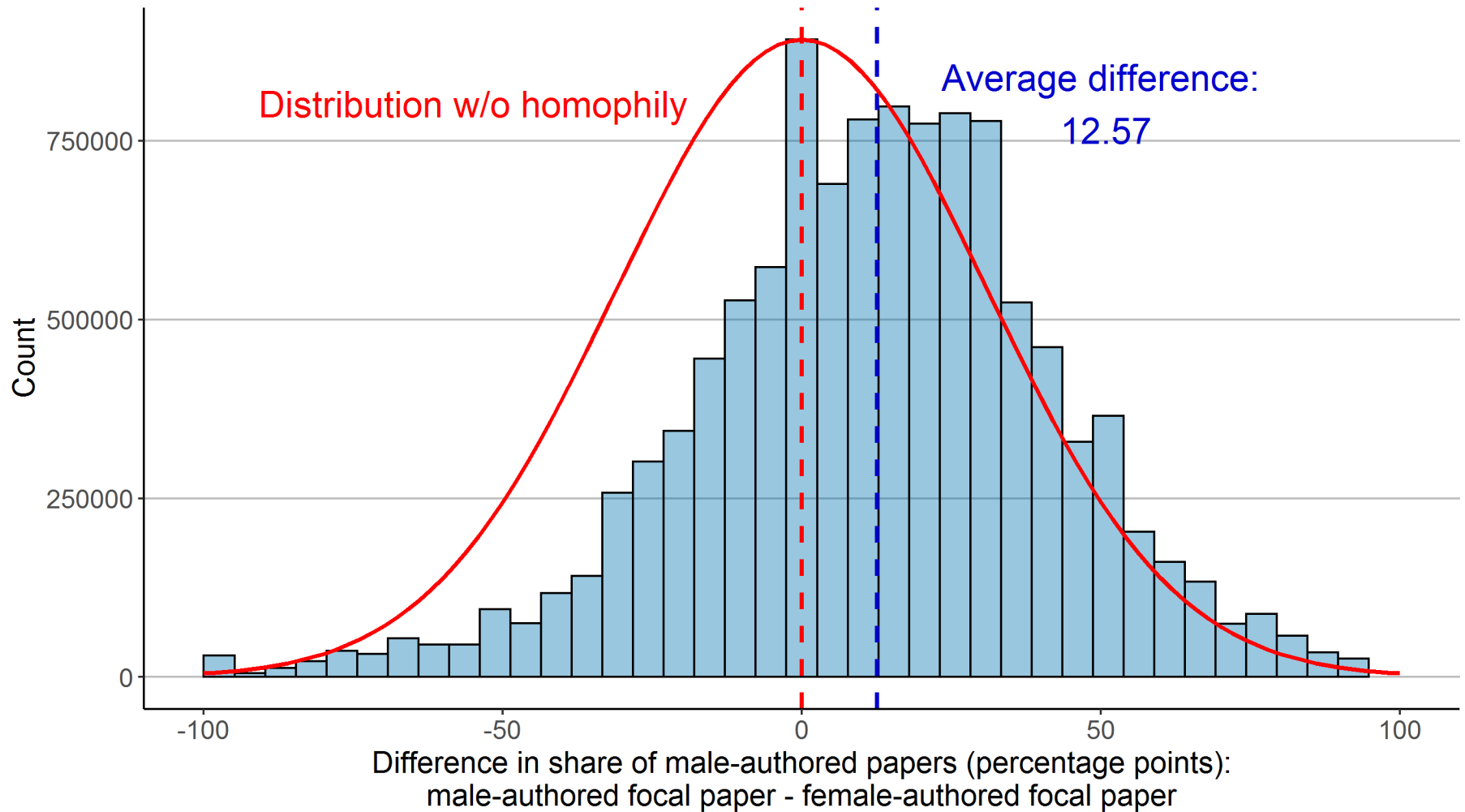
Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 0 shared sections



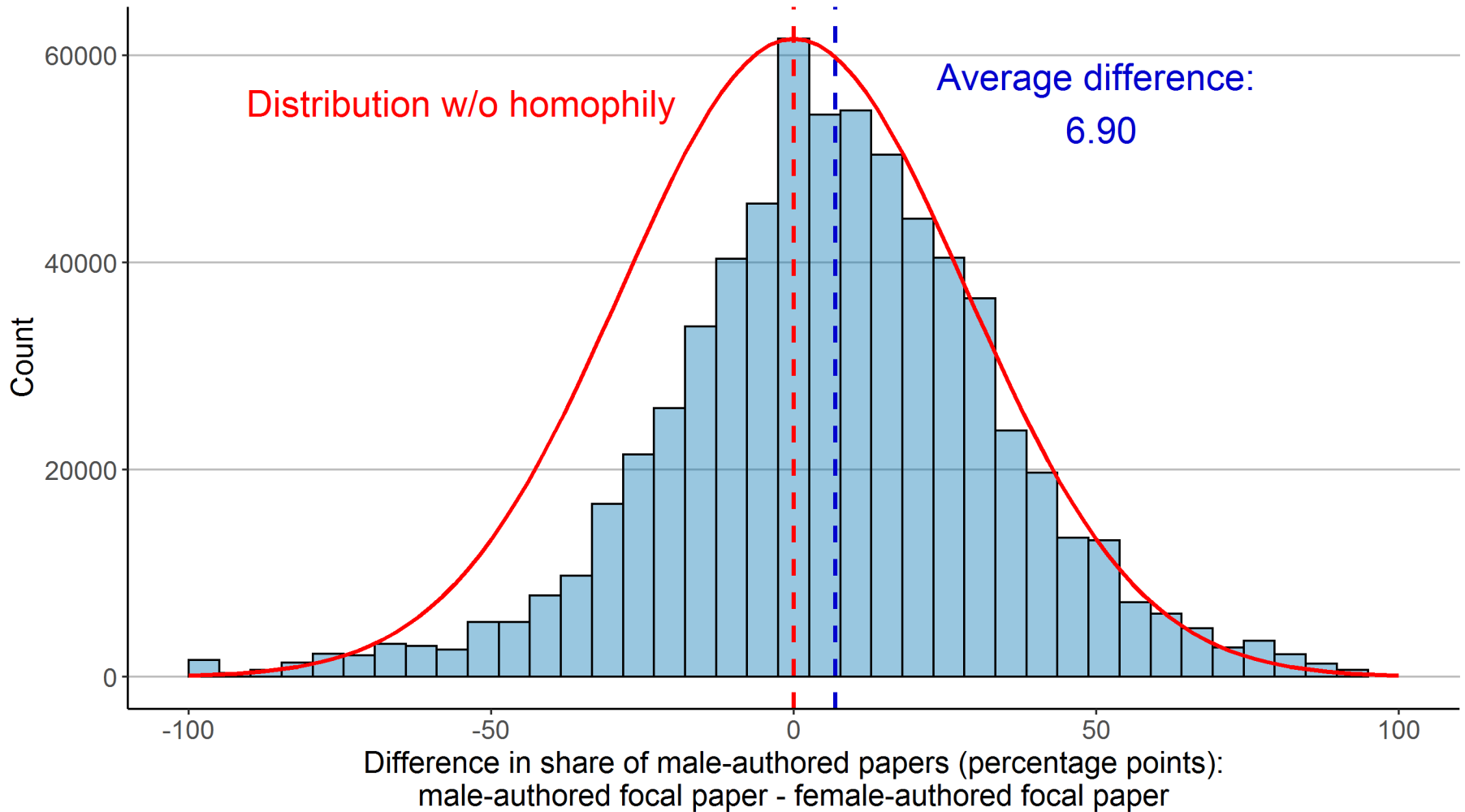
Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 0 shared sections



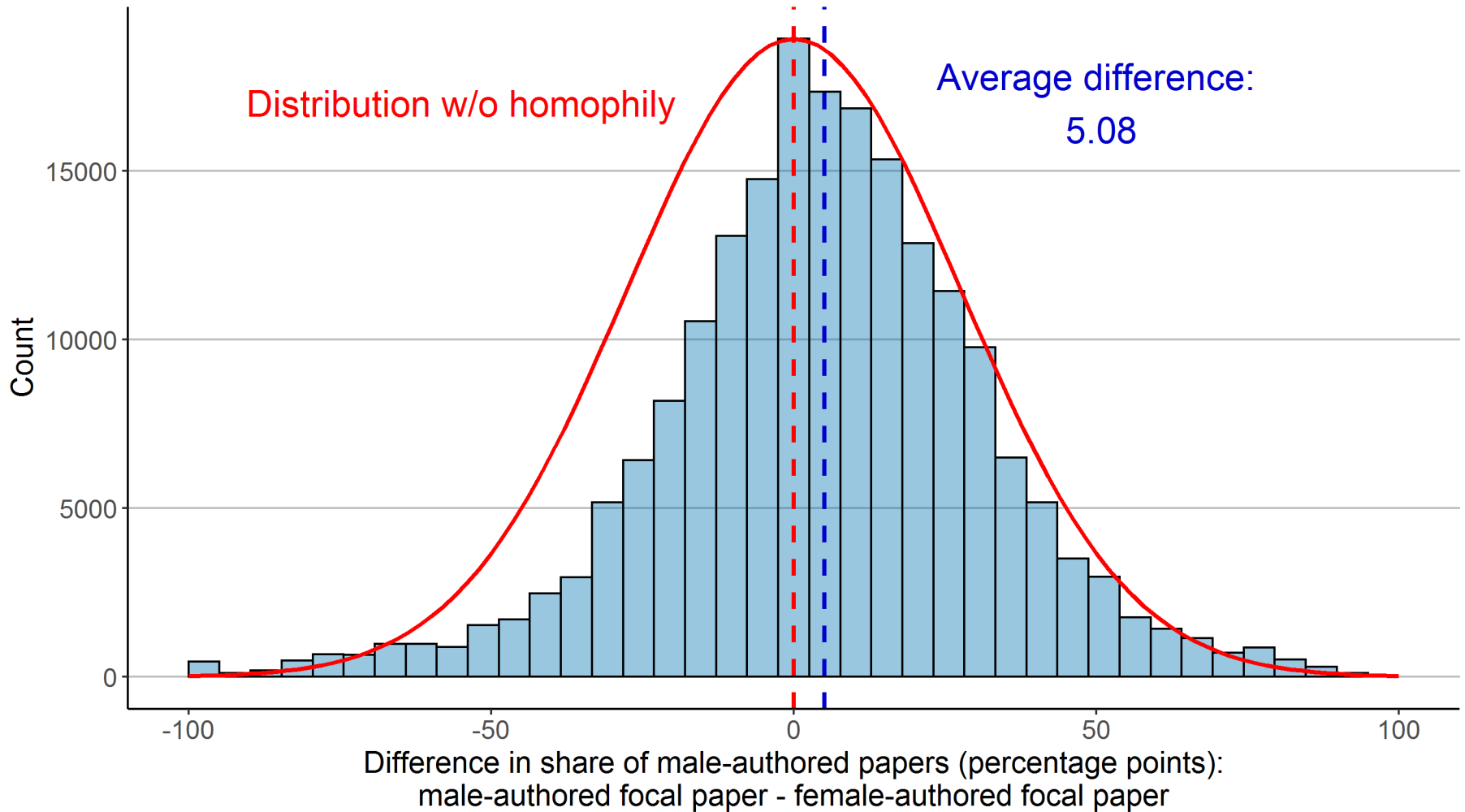
Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 1 shared sections



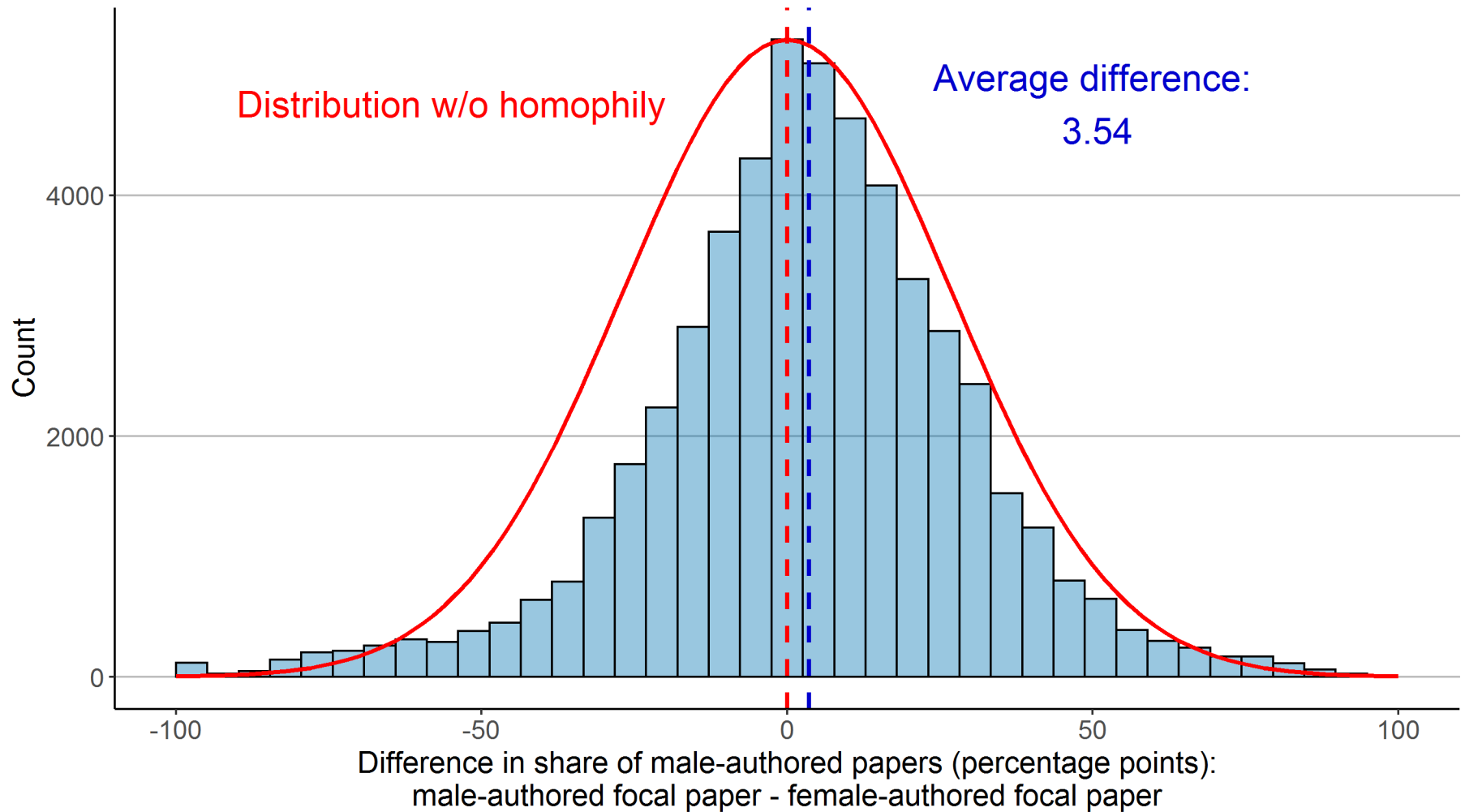
Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 2 shared sections



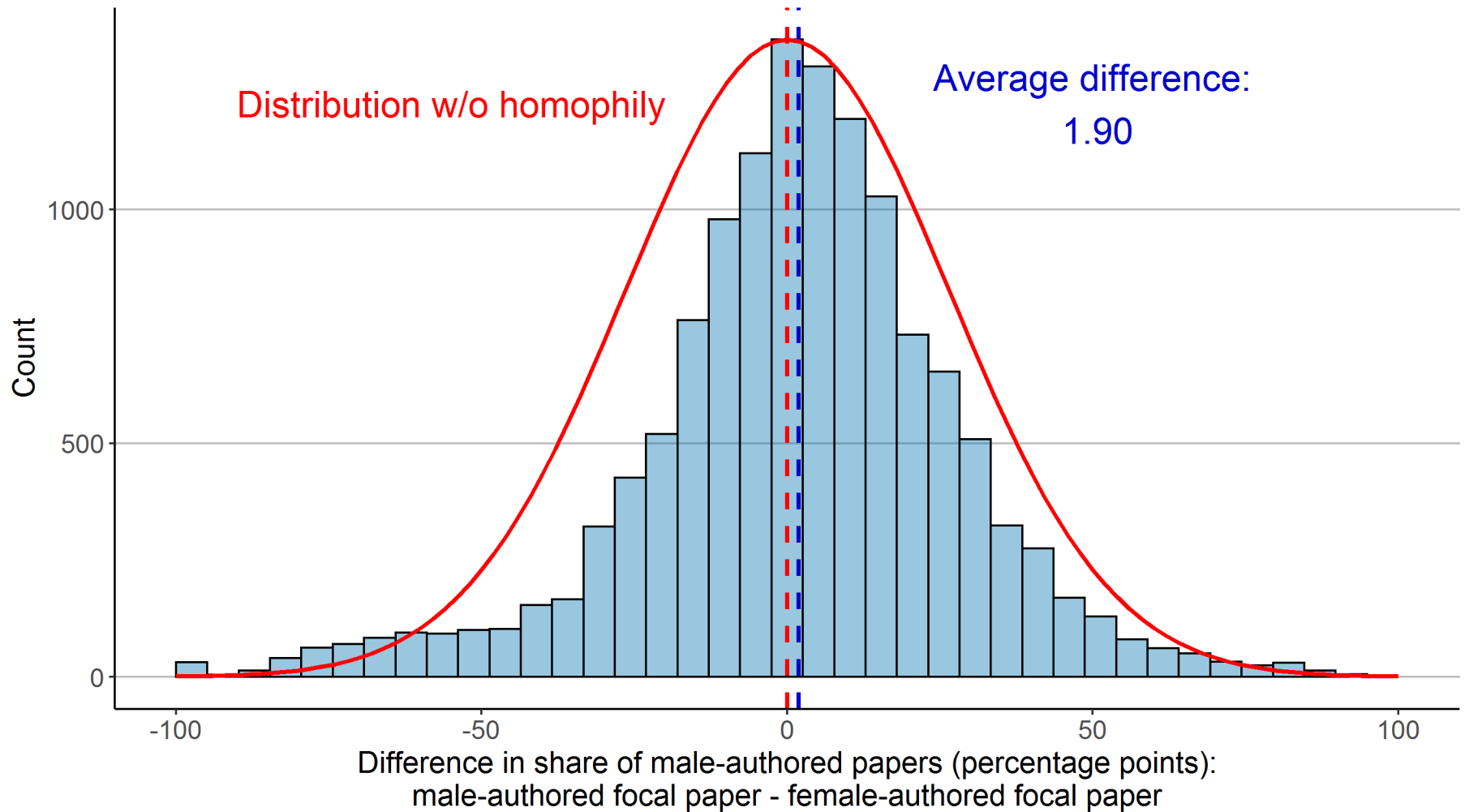
Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 3 shared sections



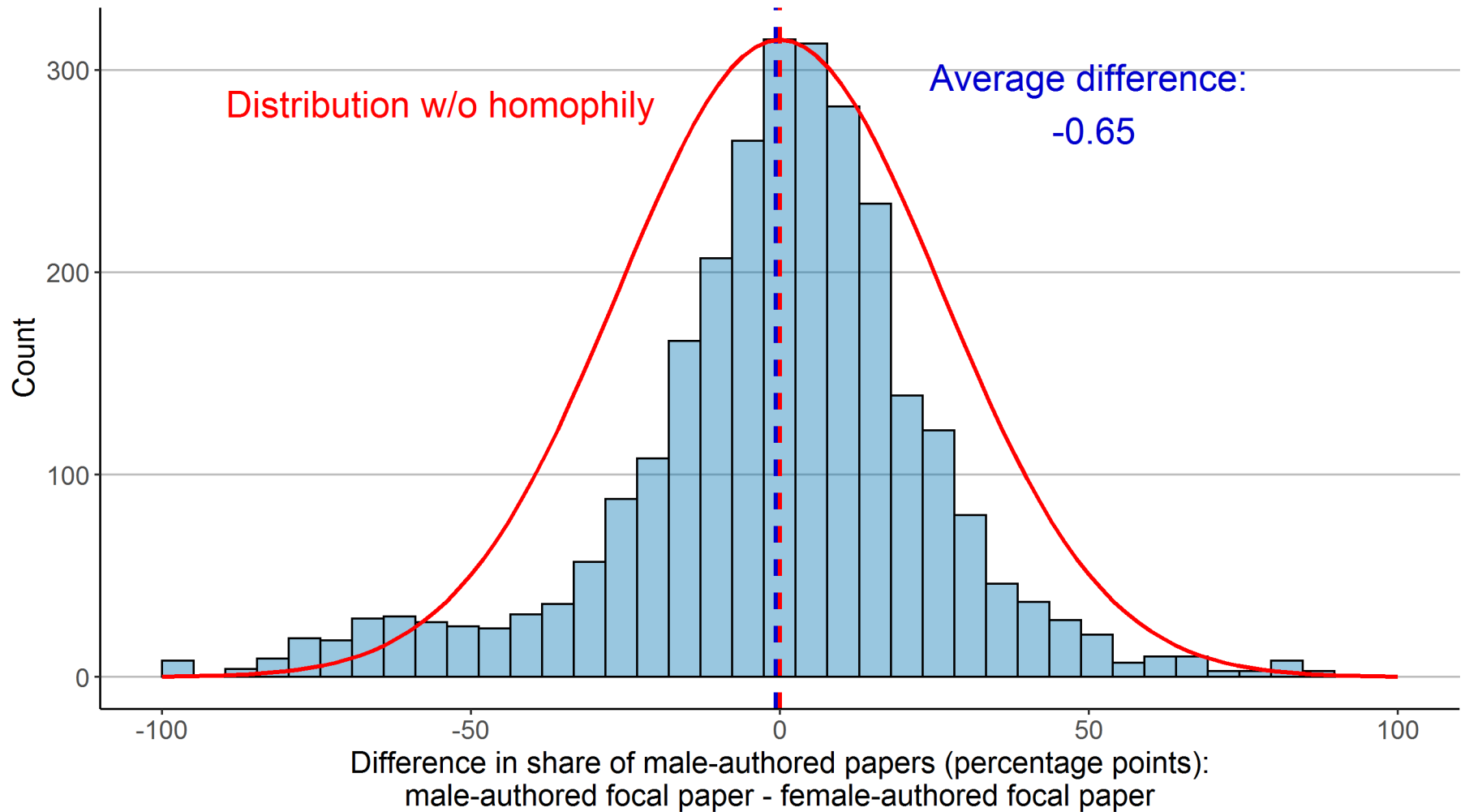
Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 4 shared sections

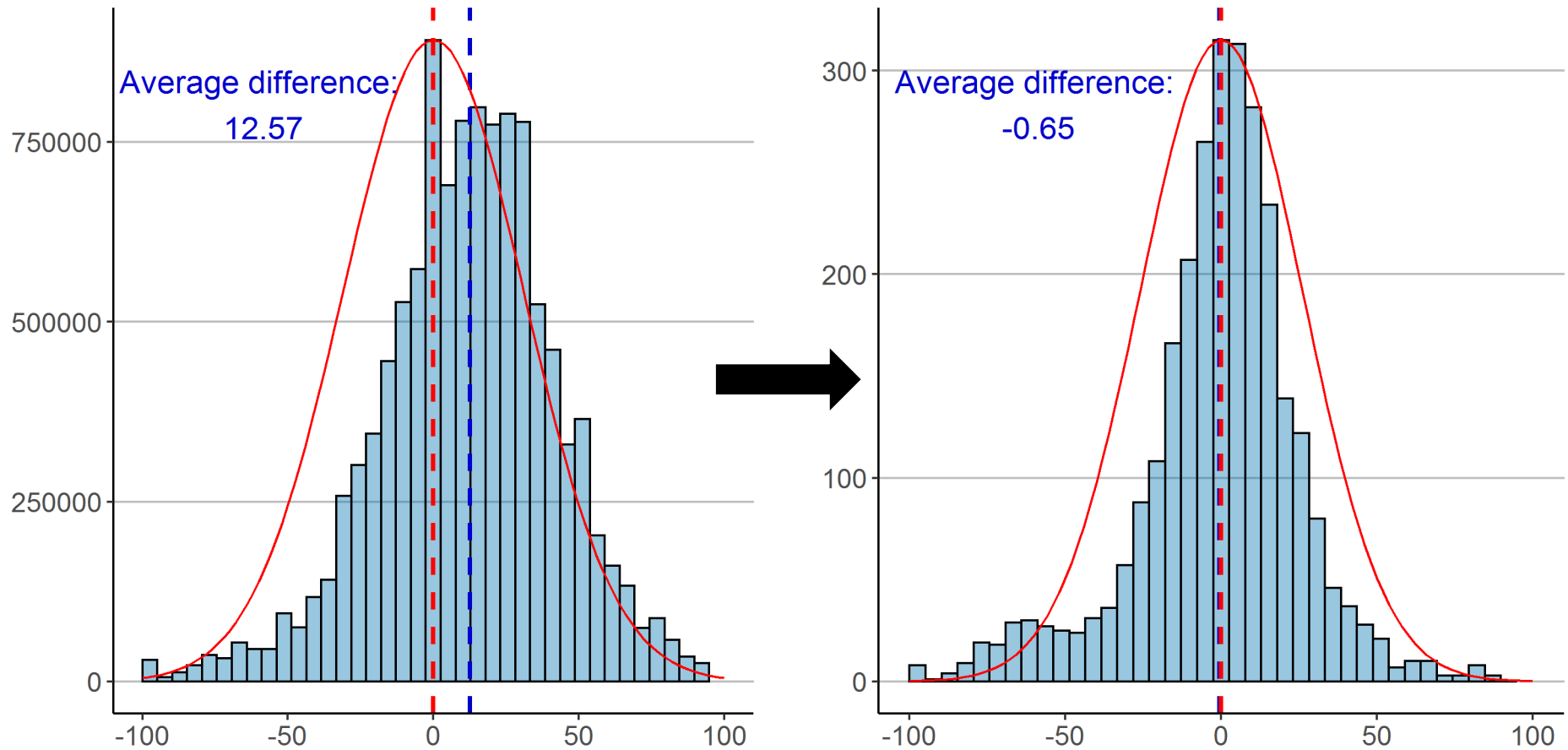


Preliminary Results (II)

Differences in share of male-authored papers among citing papers
Pairs of focal papers with at least 5 shared sections



Preliminary Results (II)



Conclusions & Outlook

- Granularity of topological classification matters
- After thoroughly controlling for field/topic, evidence for gender homophily is completely gone
- Other variables have small effects (in our selective sample)
- General take home-message: Comparing citations (e.g. for evaluations) require thorough standardizations for fields/topics
- But only first results, we still work on robustness checks
 - Different operationalizations of author team's gender
 - Analyses for non-F1000 Prime papers, e.g. in social sciences
 - Different approaches to control field/topic (e.g. similarity based on titles/abstracts)