Online Appendix for "Gendered Citations at Top Economic Journals"

Marlène Koffi
University of Toronto

January 2021

A Citing Journals' Discipline

There are around 20,000 scientific journals whose papers cite at least one article originally published in the Top Five journals. At this stage, I further identify citing articles by fields to separate citations from economic journals from non-economic journals. I use the average ranking of economics journals over the last 10 years provided by RePEc/IDEAS.[1] The RePEc ranking is one of the most comprehensive rankings of economic journals since it includes a high number of journals (2465). A potential problem with this ranking is that it also includes non-economic journals; for example, the *American Political Science Review*, one of the most prominent journals in political science, is also present in this ranking. To extract these non-economic journals, I employ the field-based journal classification used by Angrist et al. (2020). Thus, journals that are not included in my classification either relate to another field (sociology, political science, mathematics, etc.) or are very low ranked. For simplicity, I will categorize all the remaining not listed in RePEc (after correction) as non-economic journals.

For each of the following disciplines: Accounting, Anthropology, Computer Science, Management, Marketing, Mathematics, Medicine, Operations research, Physics, Political science, Psychology, Public health, Sociology, Statistics, Law, and Interdisciplinary (Angrist et al. (2020)), I compile a list of related journals (more than 300 journals in each of the above disciplines) based on the journal ranking by field provided by Clarivate. The selected disciplines correspond to the one identified by Angrist et al. (2020) as having important cross-field citations with economics.

[1] RePEc/IDEAS: https://ideas.repec.org/top/top.journals.simple.html.
Clarivate Analytics: https://jcr.clarivate.com/JCRHomePageAction.action?.

B Determination of the style

I construct a machine learning classifier to define the methodology's style ("theory", "empirics", or "econometrics"). In doing so, I randomly select and download 2000 papers from the Top Five journals. For each paper, I estimate the methodology by counting instances of certain keywords within the paper. For example, a word like "theorem" is more likely to refer to a theoretical paper, while a word like "randomized experiment" is more likely to refer to an empirical paper. Using this training sample as an input in a statistical learning algorithm, I classify the articles in the Top Five baseline database with their titles, their abstracts, their keywords, and their JEL code used as predictors.

Papers have been classified into three styles, depending on the research method: theoretical, empirical, and econometric. A paper will be said to be empirical if it uses the data to estimate parameters and contains words such as "standard error", "table"; it will be said to be theoretical when it uses words like "theorem", "proposition", "equilibrium". Note the difference between theoretical economics and theoretical econometrics which will be classified as "econometric").[2]

To categorize articles based on style, I use a machine-learning algorithm. To do this, I construct my training sample as follows. I randomly select a set of 2000 articles proportionally to the fraction of articles belonging to each newspaper. Those articles are downloaded in their entirety. I count the number of occurrences of words referring to each category mentioned above. Articles for which the fraction of the words in a category represents more than 90% of the total words recorded are classified as belonging to said category. For example, an article in which the words "instrumental variables", "data", "standard error" come up most often and constitute more than 90% of the targeted words will be of "empirical" style. Because many articles use both an empirical and a theoretical style, the training sample consists of articles that offer the most certainty about the method. To distinguish theoretical economics from theoretical econometrics, I identify articles with the main JEL code "C" and manually classify those related to theoretical econometrics. In sum, the training sample comprises 838 observations (15% of the total Top Five sample). Further, the classification uses as features: the 3-digit JEL codes, abstracts, titles, and keywords. The best classifier is the random forest (this classifier is chosen from a set of other classifiers). It maximizes the accuracy of the test sample with an 80-20 split. The accuracy rate is almost 85%. The algorithm allows us to estimate the probability that a given article is written in one style or another. Due to the increasingly loose line between styles, relative probabilities seem to be a better way to characterize papers' style.

[2]See Koffi (2020) for more details on the current procedure, Card et al. (2020) for the selection of the words, Angrist et al. (2020) for similar machine learning applications.

C Race recognition based on Names

The race of the authors is not available in the database. I build a race classification algorithm based on the last name of the authors. To do this, I use data from the US census of 2000 and 2010. The Census data gives 167,000 surnames with at least 100 occurrences in the United States, the fraction of people relating to the different races: White, Black, Asian, Hispanic, and Native American. Because I am using the same data as Hofstra et al. (2020), I use their thresholds to determine races. Thus, individuals will be classified as "white" if the relative use of their first name for the "white" category is greater than 0.83, "Asian" 0.93, "Hispanic / Latino" 0.75, and 0.5 for "Black" will be considered as white. The classifier will make it possible to determine the race of those who are not present in the census base, which concerns more than half of the observations. Using census data as a training sample, predictors are automatically constructed by the algorithm and relate to the characters included to form the names. The best classifier is the Logistic Classifier with an L2 penalty with a maximal Recall rate of 70%. The recall rate is the number of true positives divided by the sum of true positives and false negatives. I consider "Hispanic / Latino" and "Black" as minorities. The minorities group represents around 0.20 of the authors (number quite close to the racial minority mean of 0.246 in a discipline-year found by Hofstra et al. (2020)), and the fraction of articles with at least one author from a minority group is roughly 30%.

D Additional Tables

D.1 Summary Statistics

Table 1: Summary Statistics

Journal	Gender				Overall
	Male	Female	Mixed	Unknown	
AER	1,334	98	353	43	1,828
ECA	1,190	36	175	16	1,417
JPE	818	45	154	16	1,033
QJE	764	41	190	11	1,006
RES	905	48	170	24	1,147
Total	5,011	268	1,042	110	6,431

The table shows the distribution of articles in the database by journals and gender. Those articles were published between 1990 and 2019.

D.2 Journals that cite the most articles in Top Five

Table 2: Most Citing Journals

	Journals	Number
1	*American Economic Review*	*16,650*
2	*Journal of Economic Behavior And Orga..*	*11,593*
3	*Journal of Economic Theory*	*9,970*
4	*Journal of Econometrics*	*8,958*
5	*European Economic Review*	*8,904*
6	*Econometrica*	*8,755*
7	*Review of Economic Studies*	*8,432*
8	*Journal of International Economics*	*8,301*
9	*Journal of Public Economics*	*8,245*
10	*Economic Journal*	*8,037*
11	*Applied Economics*	*7,944*
12	*Economics Letters*	*7,676*
13	*Games And Economic Behavior*	*7,613*
14	*Journal of Economic Dynamics And Cont..*	*7,266*
15	*Quarterly Journal of Economics*	*7,068*

The table shows the 15 journals that cite the most an article published in the Top Five journals.

D.3 Most Citing Journals by Gender

Table 3: Most Citing Journals by Gender

Male	Female
J ECON GROWTH	ECON EDUC REV
PHYSICS (DISC)	PUBLIC HEALTH (DISC)
THEORETICAL ECONOMICS	J DEV STUD
J ECON THEORY	J HEALTH ECON
J MATH ECON	J HUM RES
ECON TH	AEJ APPLIED
GAMES AND ECON BEH	J POP ECON
MACRO DYN	WORLD DEV
ECONOMETRIC TH	MEDECINE (DISC)
J MONEY CREDIT BANK	J DEV ECON
J ECON MAN SC	SOCIOLOGY (DISC)
COMPUTER SCIENCE (DISC)	AEJ ECON POL
REV ECON DYN	MULTIDISCIPLINE (DISC)
ECA	REG SCI AND URB ECON
RAND J ECON	J ECON PERSP

This table shows the journals (if in economics) or disciplines (if not economics) that cite the most female or male papers based on the relative share of citations allocated to each gender. "DISC." stands for "discipline" , indicating that this is not a single journal, but a group of journals from a discipline different from economics.

D.4 Top 25 percentiles journals of each Ranking

RePEc/IDEAS	Clarivate Analytics	Kodrzycki and Yu (2006)
American Economic Journal Applied Economics	*American Economic Journal Applied Economics*	*American Economic Review*
American Economic Journal Macroeconomics	*American Economic Journal Macroeconomics*	*Econometrica*
American Economic Review	*American Economic Review*	*Journal of Econometrics*
Annual Review of Economics	*Annual Review of Economics*	*Journal of Economic Literature*
Brookings Papers On Economic Activity	*Brookings Papers On Economic Activity*	*Journal of Economic Perspectives*
Econometrica	*Cambridge Journal of Regions Economy and Society*	*Journal of Economic Theory*
Economic Journal	*Ecological Economics*	*Journal of Finance*
Economic Policy	*Econometrica*	*Journal of Financial Economics*
Journal of Applied Econometrics	*Economic Geography*	*Journal of Monetary Economics*
Journal of Business And Economic Statistics	*Economic Journal*	*Journal of Political Economy*
Journal of Econometrics	*Economic Policy*	*Quarterly Journal of Economics*
Journal of Economic Growth	*Energy Economics*	*Review of Economic Studies*
Journal of Economic Literature	*Energy Policy*	*Review of Economics And Statistics*
Journal of Economic Perspectives	*Food Policy*	*Review of Financial Studies*
Journal of Finance	*Journal of Accounting And Economics*	
Journal of Financial Economics	*Journal of Economic Geography*	
Journal of Financial Intermediation	*Journal of Economic Growth*	Combes and Linnemer (2010)
Journal of International Economics	*Journal of Economic Literature*	*American Economic Review*
Journal of Labor Economics	*Journal of Economic Perspectives*	*Econometrica*
Journal of Monetary Economics	*Journal of Environmental Economics and Management*	*Economic Journal*
Journal of Political Economy	*Journal of Finance*	*Journal of Econometrics*
Journal of Public Economics	*Journal of Financial Economics*	*Journal of Economic Theory*
Journal of the European Economic Association	*Journal of Human Resources*	*Journal of Finance*
Quarterly Journal of Economics	*Journal of Labor Economics*	*Journal of Financial Economics*
Rand Journal of Economics	*Journal of Policy Analysis And Manage..*	*Journal of Monetary Economics*
Review of Economic Studies	*Journal of Political Economy*	*Journal of Political Economy*
Review of Economics And Statistics	*Journal of The Association of Environ..*	*Journal of Public Economics*
Review of Financial Studies	*Journal of the European Economic Association*	*Quarterly Journal of Economics*
World Bank Economic Review	*Journal of Transport Geography*	*Rand Journal of Economics*
	Pharmacoeconomics	*Review of Economic Studies*
	Quarterly Journal of Economics	*Review of Economics And Statistics*
	Regional Studies	
	Review of Economic Studies	
	Review of Economics And Statistics	Kalaitzidakis, Mamuneas and Stengos (2011)
	Review of Environmental Economics and Policy	*American Economic Review*
	Review of Financial Studies	*Econometrica*
	Small Business Economics	*Economic Journal*
	Socioeconomic Review	*Journal of Econometrics*
	Transportation Research Part Apolicy ..	*Journal of Economic Literature*
	Transportation Research Part Bmethodo..	*Journal of Economic Perspectives*
	Transportation Research Part Elogisti..	*Journal of Economic Theory*
	Value In Health	*Journal of Finance*
	World Bank Research Observer	*Journal of Monetary Economics*
	World Development	*Journal of Political Economy*
		Quarterly Journal of Economics
		Review of Economic Studies
		Review of Economics And Statistics
		Review of Financial Studies

D.5 Citation and Prominence of Citing Papers Journals

Table 4: Citation and Top Journals

	Grouping 1 (1)	Grouping 2 (2)	Grouping 3 (3)
Female	-0.139	-0.108	-0.075
	(0.073)	(0.065)	(0.066)
N	5279	5279	5279
R-sqr	0.384	0.401	0.407

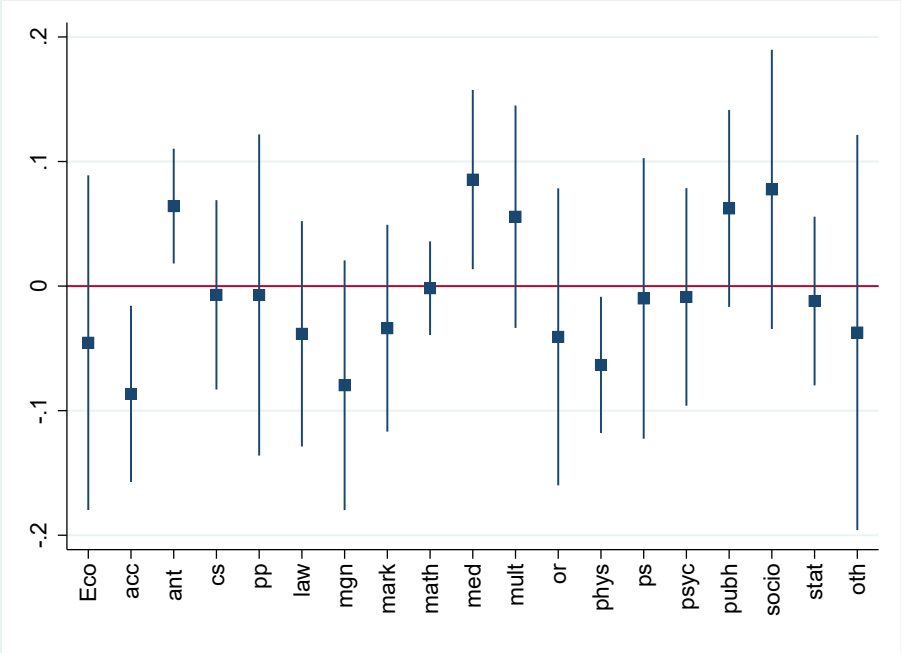
This table shows the estimates of the main regression in the paper, restricting the cited papers to the Top Five publications by all-male or all-female teams. Each estimate corresponds to the female citation premium, where citation comes from a specified subset of top economic journals. Grouping 1 includes: *American Economic Review, Econometrica, Journal of Econometrics, Journal of Economic Growth, Journal of Economic Literature, Journal of Economic Perspectives, Journal of Economic Theory, Journal of Finance, Journal of Financial Economics, Journal of International Economics, Journal of Labor Economics, Journal of Monetary Economics, Journal of Political Economy, Quarterly Journal of Economics, Review of Economic Studies, and Review of Financial Studies.* Grouping 2 includes grouping 1 and: *American Economic Journal Applied Economics, American Economic Journal Economic Policy, American Economic Journal Macroeconomics, American Economic Journal Microeconomics, Brookings Papers On Economic Activity, Economic Journal, Experimental Economics, Games And Economic Behavior, International Economic Review, Journal of Accounting And Economics, Journal of Applied Econometrics, Journal of Business And Economic Statistics, Journal of Economic Behavior And Organization, Journal of Financial And Quantitative Analysis, Journal of Human Resources, Journal of Law And Economics, Journal of Money Credit And Banking, Journal of Public Economics, Journal of The European Economic Association, Journal of Urban Economics, Rand Journal of Economics, Review of Economic Dynamics, Review of Economics And Statistics.* Grouping 3 includes grouping 1 and: *Journal of Health Economics, Journal of Development Economics.* The *Journal of Health Economics* and the *Journal of Development Economics* tend to increase the premium in favor of women.

E Additional Graphs

E.1 Citation Premium by Disciplines

Outside of economics, women have a positive citation premium in anthropology, medicine, public health and sociology (consistent with the descriptive analysis). Those are fields in which there are a substantial fraction of female authors.

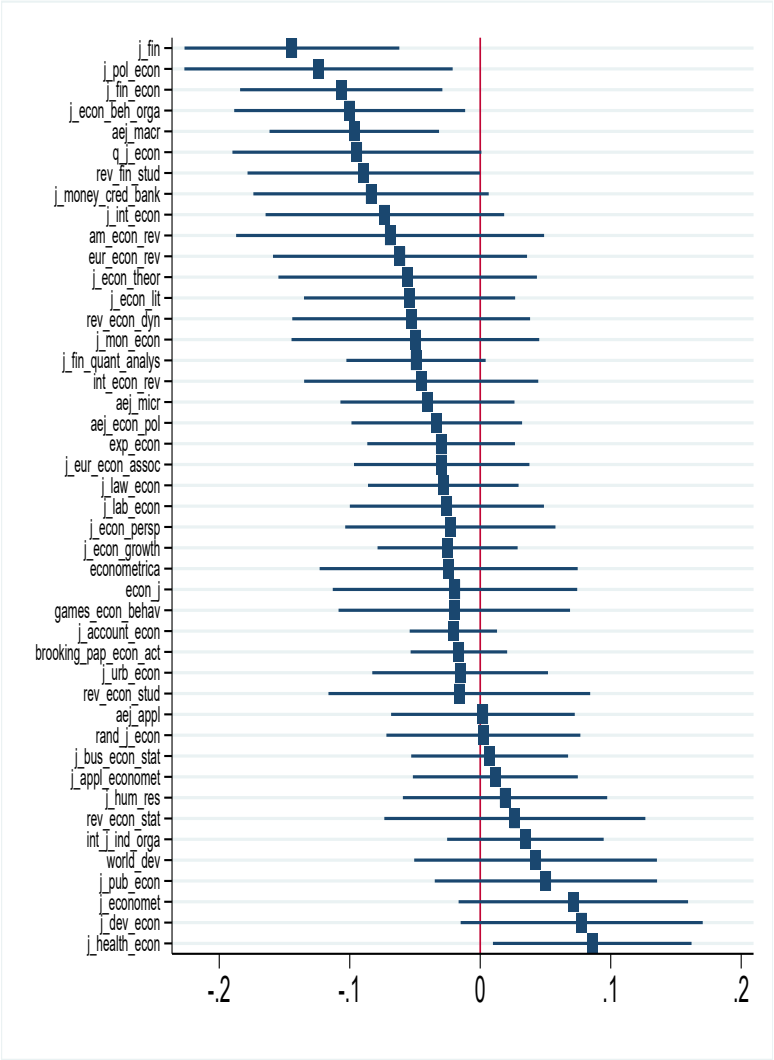
Figure 1: Citation premium by discipline of the citing article



This figure shows the gender citation premium by discipline, restricting the cited papers to the Top Five publications by all-male or all-female teams. Each full square corresponds to the estimate of β_F in the main equation for a subset of journals belonging to the specified discipline. acc: Accounting, ant: Anthropology, cs: Computer Science, mgn: Management, mark: Marketing, math: Mathematics, med: Medicine, or: Operations research, phys: Physics, ps: Political science, psyc: Psychology, pubh: Public health, socio: Sociology, stat: Statistics, oth: Other, law: Law, and mult: Interdisciplinary. For example, the full square for accounting gives the citation premium for female versus male focusing on a subset of citing papers published in journals in accounting. Controls include fields, style, journal· year dummies, affiliation, publication, and number of authors.

E.2 Citation premium by Journals

Figure 2: Citation premium by journal of the citing article



This figure shows the gender citation premium by journals, restricting the cited papers to the Top Five publications by all-male or all-female teams. The sample of journals is composed of journals appearing in most top 50 rankings. Each full square corresponds to the estimate of β_F in the main equation for the specified journal. Controls include fields, style, journal· year dummies, affiliation, publication, and number of authors.

References

Angrist, Josh, Pierre Azoulay, Glenn Ellison, Ryan Hill, and Susan Feng Lu. 2020. "Inside job or deep impact? Extramural citations and the influence of economic scholarship." *Journal of Economic Literature*, 58(1): 3–52.

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri. 2020. "Are Referees and Editors in Economics Gender Neutral?" *The Quarterly Journal of Economics*, 135(1): 269–327.

Combes, Pierre-Philippe, and Laurent Linnemer. 2010. "Inferring missing citations: A quantitative multi-criteria ranking of all journals in economics."

Hofstra, Bas, Vivek V Kulkarni, Sebastian Munoz-Najar Galvez, Bryan He, Dan Jurafsky, and Daniel A McFarland. 2020. "The Diversity–Innovation Paradox in Science." *Proceedings of the National Academy of Sciences*, 117(17): 9284–9291.

Kalaitzidakis, Pantelis, Theofanis P Mamuneas, and Thanasis Stengos. 2011. "An updated ranking of academic journals in economics." *Canadian Journal of Economics/Revue canadienne d'économique*, 44(4): 1525–1538.

Kodrzycki, Yolanda K, and Pingkang Yu. 2006. "New approaches to ranking economics journals." *The BE Journal of Economic Analysis & Policy*, 5(1).

Koffi, Marlène. 2020. "Innovative Ideas and Gender Inequality." Unpublished.