

Online Appendix

Racial Inequality and Publication in Economics

Marlène Koffi* Roland Pongou† Leonard Wantchekon‡

May 2024

1 Descriptive statistics; White and Non-White

	White	Non-White	p-value
200 journals	3.165	2.202	0.000
35 high-impact	0.700	0.486	0.000
Top 5 journals	0.087	0.039	0.000

This table displays the average number of publications in each set of journals for White authors compared to non-White authors.

*University of Toronto, 150 St. George St., Toronto, ON M5S 3G7, Canada, marlene.koffi@utoronto.ca.

†University of Ottawa, 120 University, ON K1N 6N5, Canada, rpongou@uottawa.ca

‡Princeton University, 321 Bendheim Hall, Princeton, NJ 08544, USA, lwantche@princeton.edu

2 Descriptive statistics by racial categories

	Race category					
	White	Black	Hispanic	Asian	Unclear	Total
N	85,694	786	8,513	37,805	30,207	163,005
Average Publications (200 journals)	3.165	2.963	2.741	2.064	2.592	2.781
Average Publications (35 high-impact journals)	0.7	0.543	0.497	0.482	0.466	0.595
Average Publications (Top 5 journals)	0.087	0.07	0.066	0.032	0.052	0.067

This table presents the detailed descriptive statistics of authors in the database, categorized according to their racial classification. The set of 35 high-impact journals corresponds to the set of journals in Card et al. [2020].

3 Publication in Top 5 journals by racial categories

	Outcome: Number of Top 5 publications			
	(1)	(2)	(3)	(4)
Black	-0.0005 (0.0018)	-0.0000 (0.0017)	-0.0001 (0.0011)	-0.0000 (0.0011)
Hispanic	-0.0004 (0.0006)	-0.0003 (0.0006)	-0.0007 (0.0004)	-0.0007 (0.0004)
Asian	-0.0033 (0.0003)	-0.0028 (0.0003)	-0.0018 (0.0002)	-0.0017 (0.0002)
Unclear	-0.0022 (0.0003)	-0.0016 (0.0003)	-0.0010 (0.0002)	-0.0009 (0.0002)
Number of past publications			0.1111 (0.0040)	0.1107 (0.0039)
Number of past coauthors				0.0001 (0.0000)
Observations	1867183	1867183	1867183	1867183
R-sqr	0.0003	0.0173	0.0678	0.0679
Mean DV	0.0058	0.0058	0.0058	0.0058

This table illustrates the relation between race and the number of publications in the top five journals. The sample comprises 163,005 individuals, tracked from the date of their initial appearance in the database (or starting from 1990) until 2019. Column (1) provides estimates of the model without incorporating any control variables. Column (2) introduces controls for differences in research fields, experience, the scope of affiliations, and time-fixed effects. Column (3) includes the history of publications in the top five journals, and Column (4) includes the count of individual coauthors prior to a specified time t . Standard errors are in parentheses and clustered at the author level.

4 Publication in a set of 35 high-impact journals

	Outcome: Number of Top 35 publications			
	(1)	(2)	(3)	(4)
Non-White	-0.0032 (0.0011)	-0.0048 (0.0010)	-0.0061 (0.0007)	-0.0055 (0.0007)
Unclear	-0.0137 (0.0010)	-0.0100 (0.0009)	-0.0071 (0.0006)	-0.0067 (0.0006)
Number of past publications			0.1552 (0.0024)	0.1499 (0.0021)
Number of past coauthors				0.0023 (0.0002)
Observations	1867183	1867183	1867183	1867183
R-sqr	0.0004	0.0699	0.1276	0.1287
Mean DV	0.0516	0.0516	0.0516	0.0516

This table illustrates the relation between race and the number of publications in the top 35 journals. The sample comprises 163,005 individuals, tracked from the date of their initial appearance in the database (or starting from 1990) until 2019. Column (1) provides estimates of the model without incorporating any control variables. Column (2) introduces controls for differences in research fields, experience, the scope of affiliations, and time-fixed effects. Column (3) includes the history of publications in the top 35 journals, and Column (4) includes the count of individual coauthors prior to a specified time t . Standard errors are in parentheses and clustered at the author level.

5 Publication in the database of 200 journals

	Outcome: Number of publications			
	(1)	(2)	(3)	(4)
Non-White	-0.0103 (0.0022)	-0.0402 (0.0021)	-0.0181 (0.0012)	-0.0181 (0.0012)
Unclear	-0.0198 (0.0022)	-0.0284 (0.0020)	-0.0138 (0.0012)	-0.0137 (0.0012)
Number of past publications			0.4184 (0.0038)	0.4099 (0.0031)
Number of past coauthors				0.0021 (0.0004)
Observations	1867183	1867183	1867183	1867183
R-sqr	0.0002	0.2507	0.3608	0.3609
Mean DV	0.2353	0.2353	0.2353	0.2353

This table illustrates the relation between race and the number of publications in the top 200 journals. The sample comprises 163,005 individuals, tracked from the date of their initial appearance in the database (or starting from 1990) until 2019. Column (1) provides estimates of the model without incorporating any control variables. Column (2) introduces controls for differences in research fields, experience, the scope of affiliations, and time-fixed effects. Column (3) includes the history of publications in the top 200 journals, and Column (4) includes the count of individual coauthors prior to a specified time t . Standard errors are in parentheses and clustered at the author level.

6 Alternative dependant variable specifications

	Outcome: Number of publications					
	Dummy		Negative binomial		Random Effect	
	(1)	(2)	(3)	(4)	(5)	(6)
Non-White	-0.0014 (0.0001)	-0.0014 (0.0001)	-0.3074 (0.0325)	-0.2969 (0.0324)	-0.0015 (0.0002)	-0.0015 (0.0002)
Unclear	-0.0009 (0.0001)	-0.0009 (0.0001)	-0.1728 (0.0324)	-0.1622 (0.0324)	-0.0010 (0.0002)	-0.0009 (0.0002)
Number of past publications	0.0931 (0.0024)	0.0929 (0.0024)	1.8712 (0.0232)	1.7293 (0.0255)	0.1111 (0.0040)	0.1109 (0.0040)
Number of past coauthors		0.0006 (0.0001)		0.3479 (0.0235)		0.0007 (0.0001)
Observations	1867183	1867183	1867183	1867183	1867183	1867183
R-sqr	0.0641	0.0641				
Mean DV	0.0053	0.0053	0.0058	0.0058	0.0058	0.0058

This table illustrates the relation between race and the number of publications in the top five journals. The sample comprises 163,005 individuals, tracked from the date of their initial appearance in the database (or starting from 1990) until 2019. The odd-numbered columns incorporate all control variables except for the one pertaining to the number of coauthors, whereas the even-numbered columns include this variable. Columns (1) and (2) are associated with a model wherein the dependent variable is a binary indicator reflecting whether the author has any publications in the top five journals at a given time. Columns (3) and (4) are linked to a model that utilizes a negative binomial estimation procedure. Columns (5) and (6) are related to a model employing a random effect estimation approach. Standard errors, presented in parentheses, are clustered at the author level.

7 Pooling dates

	Outcome: Number of Top 5 publications			
	(1)	(2)	(3)	(4)
Non-White	-0.0126 (0.0011)	-0.0093 (0.0011)	-0.0073 (0.0008)	-0.0067 (0.0008)
Unclear	-0.0100 (0.0012)	-0.0070 (0.0011)	-0.0048 (0.0008)	-0.0043 (0.0008)
Past publications			0.5451 (0.0219)	0.5355 (0.0211)
Past coauthors				0.0025 (0.0003)
Observations	442629	442629	442629	442629
R-sqr	0.0007	0.0437	0.1620	0.1638
Mean DV	0.0246	0.0246	0.0246	0.0246

This table illustrates the relation between race and the number of publications in the top five journals. The sample comprises 163,005 individuals, tracked from the date of their initial appearance in the database (or starting from 1990) until 2019. In this model, the dates and the observations are aggregated into five-year intervals. Column (1) provides estimates of the model without incorporating any control variables. Column (2) introduces controls for differences in research fields, experience, the scope of affiliations, and time-fixed effects. Column (3) includes the history of publications in the top five journals, and Column (4) includes the count of individual coauthors prior to a specified time t . Standard errors are in parentheses and clustered at the author level.

References

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