Overview

This replication package available here contains 6 folders: 'data', 'graph_stata', 'prog_SAS_data', 'prog_Stata_analysis', 'results' and 'texstata'. The 'data' folder contains 3 subfolders: 'export_txt', 'SAS' and 'source'. 2 main files run all the code to generate data for the figure and the 30 tables in the paper. First, SAS is used to create the data necessary for the statistical analysis. All corresponding programs, including the master program (0.master_prog.sas) are in the 'prog_SAS_data' folder. They create all the datasets necessary for the statistical analysis in the 'export_txt' subfolder from the data in the 'source' subfolder (for this step, some intermediary SAS dataset are created in the 'SAS' subfolder). Then, Stata is used to create the figure and tables from the data of the 'export_txt' subfolder. All corresponding programs, including the master program (0.master.do) are in the 'prog_Stata_analysis' folder. The figure is saved in the 'graph_stata' folder and the tables are saved in the 'texstata' folder in .tex format. All these filed are then called from the paper and the specific file compiling all results in the 'results' folder.

Data Availability and Provenance Statements

The data used in this paper have been manually collected by the authors. The information on the population of French academics between 1990-2007 was collected from departments' websites and contacts with the French Ministry of Higher Education and Research and the CNRS. The information on publications was collected from EconLit. The data was anonymized to protect individuals' privacy. Similar information was also used in Combes et al. (2008), Bosquet and Combes (2017) and Bosquet et al. (2019).

Statement about Rights

- ☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

- ☒ All data are publicly available.

Details

The data used to support the findings of this study have been deposited in the Zenedo repository (https://doi.org/10.5281/zenodo.6381554). The data were collected by the authors.

Dataset list

All the source datasets are provided in the 'data\source' folder of this replication package and are available both in .sas7dbat (SAS) and in .txt (open source) formats.

- bscaufrpanelt is the transposed panel of authors (variable aumain2): for each year and affiliation (variable in), it indicates the age of the author, the weight of the author in the corresponding affiliation (variable wp), a dummy variable equal to 1 for women (variable indf). Variable indfw is the product of indf and wp. Variable resul indicates the result of the agregation contest if the author was candidate that year, variable clagreg indicates his/her ranking in case of success and nbpost the number of jobs opened through the contest between 1992 and 2004. The other variables are time-varying publication scores used as dependent variables for regressions in the paper. Scores starting with ma3, t2 and t1 assume a 3-year, 2-year and 1-year delay of publication, respectively. Then, r1 indicates solo-authored equivalent total score while s1, c1, l1 and f1 indicate solo-authored equivalent publication scores of solo-authored, co-authored, co-authored with peers and co-authored with at least one female peer publications. r2, s2, c2, l2 and f2 indicate

similar scores that take the quality of publications into account (used for robustness Tables 13 and 14) and j1 to j18 indicate JEL codes at the first letter level: 1=A, 2=B, etc.

- bscau1stpub includes various measures of lifetime average outputs/scores for each author (variable aumain2). Scores starting with ma3, t2 and t1 assume a 3-year, 2-year and 1-year delay of publication, respectively. a1 then indicates that the score is the number of solo-authored equivalent publications. Finally, j1 to j18 indicate JEL codes at the first letter level: 1=A, 2=B, etc.

- bscau1stpubjelpy includes various measures of time-varying lifetime average outputs/scores for each author (variable aumain2) at the first letter JEL code level (j1=A, j2=B; etc.), starting from the first publication in that particular JEL code. Scores starting with ma3, t2 and t1 assume a 3-year, 2-year and 1-year delay of publication, respectively. b1 then indicates that the score is the number of solo-authored equivalent publications whereas b2 indicates an additional weighting for the quality of publications (used for robustness Tables 13 and 14) and h1 indicates that the score is the number of solo-authored equivalent publications until date t, i.e. using past production only (used for robustness Tables A11 and A12).

- bscdegau includes time-varying degressive publication stocks' scores at the aggregate (variable d1) and JEL levels (variable d1j + figure corresponding to first letter: j1=A, j2=B, etc.) used to compute 'Average academics' output' for Tables 1 and 2 of the paper.

- distin gives the geographic distances (variable dist), in kilometers, between all pairs of French universities/academic institutions (variables in and in2).

- nomination_agreg_19922008 gives the information about the biannual agregation contests from 1992 to 2008: for each successful candidate (variable aumain2), the ranking (variable clagreg), the status and city before the contest (variables function_1 and ville_1), the university of nomination (variable in) and the affiliations (variables univ1 to univ3 for individuals with several affiliations). The variable notcand_1 is a dummy variable equal to 1 if the successful candidate was not candidate to the previous contest.

Computational requirements

- Operating System: Windows 7 or later

Software Requirements

- SAS 9.4
- Stata (code was last run with version 15)
  - estout
  - ftools
  - reghdfe
  - ranktest
  - ivreg2
  - ivreghdfe
  - _gwtmean
  - the program "0.setup.do" will install all dependencies locally, and should be run once.

Memory and Runtime Requirements

Summary

Approximate time needed to reproduce the analyses on a standard (CURRENT YEAR) desktop machine:

- The SAS code to create the data runs in approximately 5 minutes
- 2 hours are needed to general the figure and all the tables of the paper with Stata

Description of programs/code

- Programs in prog_SAS_data extract and reformat all datasets referenced above to create the datasets used for the statistical/econometric analyses. The file prog_SAS_data/0.master_prog.sas will run them all.
- Programs in prog_Stata_analysis generate all tables and figures in the main body and appendices of the article. The program prog_Stata_analysis/0.master.do will run them all. Each program called from 0.master.do identifies the table or figure it creates (e.g., table5.do). Output files are called appropriate names (table5.tex, figureA1.png) and should be easy to correlate with the manuscript.

Instructions to Replicators

- Download the data files referenced above. Each should be stored in the prepared subdirectories of data/, in the format that you download them in.
- Edit prog_SAS_data/0.master_prog.sas to adjust the paths to directories of source datasets, intermediary SAS datasets and output .txt datasets needed for the statistical/econometric analysis and then run it.
- Edit prog_Stata_analysis/0.master.do to adjust the paths to directories of Stata programs, needed datasets and the 'texstata' and 'graph_Stata' folders for outputs, and then run it.

Details

- prog_SAS_data:
  - If running programs individually, note that ORDER IS IMPORTANT.
  - prog_SAS_data/0.master.sas will run them all in sequence, which should take about 5 minutes.
- Prog_Stata_analysis:
  - Order does not matter, all programs can be run in parallel, if needed.
  - prog_Stata_analysis/0.master.do will run them all in sequence, which should take about 2 hours.

List of tables and programs

Figure/Table #	Program (in folder prog_Stata_analysis)	Output file	Necessary data
Table 1	table1.do	texstata/table1.tex	breglocchoice.txt
Table 2	table2.do	texstata/table2.tex	bregagregchoice.txt
Table 3	table3_A2_A3.do	texstata/table3.tex texstata/table3se2.tex texstata/table3se3.tex	bregauinma3.txt bregautjinma3.txt
Table 4	table4_5_A8_A11_A12_A17_A18.do	texstata/table4.tex	bregautjinma3_last10.txt
Table 5	table4_5_A8_A11_A12_A17_A18.do	texstata/table5.tex	bregautjinma3_last10.txt
Table 6	table6_A19.do	texstata/table6.tex	bregautjinma3.txt
Table 7	table7_A20.do	texstata/table7.tex	bregautjinma3.txt
Table A1	tableA1panelA.do tableA1panelB.do	desc_nbobs_ag.tex desc_nbobs_agrege.tex desc_ttest_ag.tex desc_ttest_ag_agrege.tex	bregauinma3.txt
Table A2	table3_A2_A3.do	texstata/tableA2.tex	bregauinma3.txt bregautjinma3.txt
Table A3	table3_A2_A3.do	texstata/tableA3.tex	bregautjinma3.txt
Table A4	tableA4_A5.do	texstata/tableA4.tex	bregautjint1_last10.txt
Table A5	tableA4_A5.do	texstata/tableA5.tex	bregautjint1_last10.txt
Table A6	tableA6_A7.do	texstata/tableA6.tex	bregautjinm12_last10.txt
Table A7	tableA6_A7.do	texstata/tableA7.tex	bregautjinm12_last10.txt
Table A8	table4_5_A8_A11_A12_A17_A18.do	texstata/tableA8.tex	bregautjinma3_last10.txt
Table A9	tableA9_A10.do	texstata/tableA9.tex	bregautjinma3_no1st10.txt
Table A10	tableA9_A10.do	texstata/tableA10.tex	bregautjinma3_no1st10.txt
Table A11	table4_5_A8_A11_A12_A17_A18.do	texstata/tableA11.tex	bregautjinma3_last10.txt
Table A12	table4_5_A8_A11_A12_A17_A18.do	texstata/tableA12.tex	bregautjinma3_last10.txt
Table A13	tableA13_A14.do	texstata/tableA13.tex	bregautjinma3_last10.txt

Table A14	tableA13_A14.do	texstata/tableA14.tex	bregautjinma3_last10.txt
Table A15	tableA15_A16.do	texstata/tableA15.tex	bregautjinma3_ivR3.txt
Table A16	tableA15_A16.do	texstata/tableA16.tex	bregautjinma3_ivR3.txt
Table A17	table4_5_A8_A11_A12_A17_A18.do	texstata/tableA17.tex texstata/tableA17se2.tex texstata/tableA17se3.tex	bregautjinma3_last10.txt
Table A18	table4_5_A8_A11_A12_A17_A18.do	texstata/tableA18.tex texstata/tableA18se2.tex texstata/tableA18se3.tex	bregautjinma3_last10.txt
Table A19	table6_A19.do	texstata/tableA19.tex texstata/tableA19se2.tex texstata/tableA19se3.tex	bregautjinma3.txt
Table A20	table7_A20.do	texstata/tableA20.tex texstata/tableA20se2.tex texstata/tableA20se3.tex	bregautjinma3.txt
Table A21	tableA21.do	texstata/tableA21.tex	bregautjinma3_last10.txt
Table A22	tableA22.do	texstata/tableA22.tex	bregautjinma3.txt
Table A23	tableA23.do	texstata/tableA23.tex	bregautjinma3_last10.txt
Figure A1	figureA1.do	graph_stata/Trends_control_all.png	bregautjinma3_last10.txt

References

Bosquet, C. and Combes, P.P (2017), 'Sorting and agglomeration economies in French economics departements', *Journal of Urban Economics*, vol. 101, pp. 27-44.

Bosquet, C., Combes, P.P. and Garcia-Peñalosa, C. (2019), 'Gender and promotions: Evidence from academic economists inFrance', *Scandinavian Journal of Economics*, vol. 121, pp. 1020-1053.

Bosquet, C., Combes, P.P., Henry, E. and Mayer, T. (2022), 'Peer effects in academic research: Senders and receivers - Replication package', *Economic Journal*, data deposited at Zenodo: https://doi.org/10.5281/zenodo.6381554.

Combes, P.P., Linnemer, L. and Visser, M. (2008), 'Publish or peer-rich? The role of skills and networks in hiring economics professors', *Labour Economics*, vol. 15(3), pp. 423-441.