Online Appendix

Measuring Science: Performance Metrics and the Allocation of Talent

Sebastian Hager* Carlo Schwarz† Fabian Waldinger‡

July 9, 2024

The Appendix presents details on data collection and additional results:

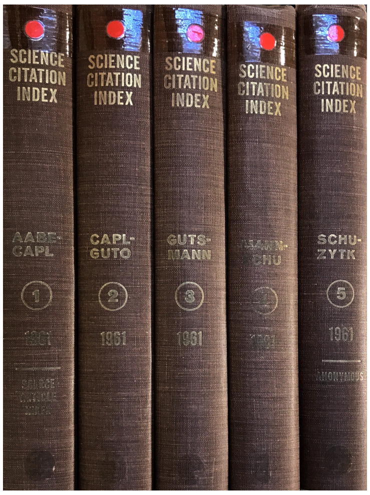
- Appendix A provides further background on the SCI.
- Appendix B provides details on data collection.
- Appendix C reports robustness checks and additional findings on the analysis of assortative matching in Section II.
- Appendix D reports additional findings on the heterogeneity analysis in Section III.
- Appendix E reports additional findings on the analysis of career outcomes in Section IV.

*Ludwig-Maximilians-Universität München, sebastian.hager@econ.lmu.de.
†Università Bocconi, Department of Economics, IGIER, PERICLES, CEPR, CAGE, carlo.schwarz@unibocconi.it.
‡Ludwig-Maximilians-Universität München, CEPR, CESifo, fabian.waldinger@econ.lmu.de.

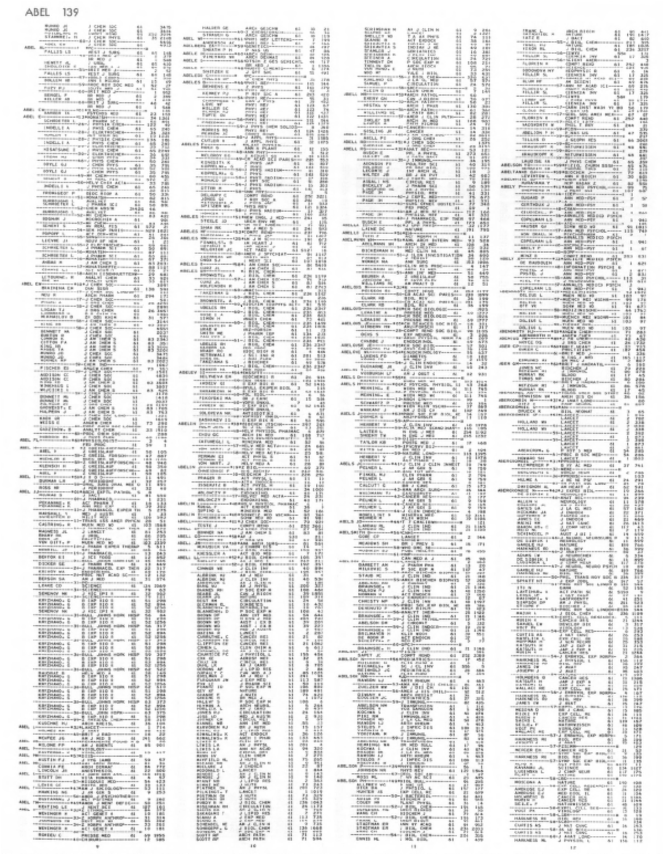
A Background on the SCI

Figure A.1: Entry in Science Citation Index

(a) The 1961 SCI volume



(b) A page in the 1961 SCI



Notes: Panel (a) shows the five books of the 1961 SCI. Panel (b) shows a sample page in the 1961 volume of the SCI.

Figure A.2: Example of Citing Journal List

Science Citation Index - 1961
Source Journals
Arranged by Full Title

ACTA ALLERGOLOGICA	ACT ALLERG
ACTA ANAESTHESIOLOGICA SCANDINAVICA	ACT ANAE SC
ACTA ANATOMICA	ACT ANATOM
ACTA BIOCHIMICA POLONICA	ACT BIOCH P
ACTA BIOLOGICA ACADEMIAE SCIENTIARUM HUNGARICAE	ACT BIOL H
ACTA BIOLOGICA ET MEDICA GERMANICA	ACT BIO MED
ACTA CHEMICA SCANDINAVICA	ACT CHEM SC
ACTA CHIMICA ACADEMIAE SCIENTIARUM HUNGARICAE	ACT CHIM H
ACTA CHIRURGICA ACADEMIAE SCIENTIARUM HUNGARICAE	ACT CHIR H
ACTA CIENTIFICA VENEZOLANA	ACT CIENT V
ACTA CRYSTALLOGRAPHICA	ACT CRYST
ACTA CYTOLOGICA	ACT CYTOL
ACTA DERMATO-VENEREOLOGICA	ACT DER-VEN
ACTA ENDOCRINOLOGICA	ACT ENDOCR
ACTA ENDOCRINOLOGICA SUPPLEMENTUM	ACT ENDOCR
ACTA GENETICA ET STATISTICA MEDICA	ACT GENET S
ACTA GENETICAE MEDICAE ET GEMELLOLOGIAE	ACT GENET M
ACTA HAEMATOLOGICA	ACT HAEMAT
ACTA HEPATO-SPLENOLOGICA	ACT HEP-SPL
ACTA HISTOCHEMICA	ACT HISTOCH
ACTA MEDICA ACADEMIAE SCIENTIARUM HUNGARICAE	ACT MED H
AGRICULTURAL AND BIOLOGICAL CHEMISTRY	AGR BIOL CH
AGRONOMY JOURNAL	AGRON J
AMERICAN DOCUMENTATION	AM DOCUMENT
AMERICAN HEART JOURNAL	AM HEART J
AMERICAN JOURNAL OF ANATOMY	AM J ANAT
AMERICAN JOURNAL OF BOTANY	AM J BOTANY
AMERICAN JOURNAL OF CARDIOLOGY	AM J CARD
AMERICAN JOURNAL OF CLINICAL NUTRITION	AM J CLIN N
AMERICAN JOURNAL OF CLINICAL PATHOLOGY	AM J CLIN P
AMERICAN JOURNAL OF DIGESTIVE DISEASES	AM J DIG DI
AMERICAN JOURNAL OF DISEASES OF CHILDREN	AM J DIS CH
AMERICAN JOURNAL OF GASTROENTEROLOGY	AM J GASTRO
AMERICAN JOURNAL OF HUMAN GENETICS	AM J HU GEN
AMERICAN JOURNAL OF HYGIENE	AM J HYG
AMERICAN JOURNAL OF MATHEMATICS	AM J MATH
AMERICAN JOURNAL OF MEDICINE	AM J MED
AMERICAN JOURNAL OF OBSTETRICS AND GYNECOLOGY	AM J OBST G
AMERICAN JOURNAL OF OPHTHALMOLOGY	AM J OPHTH
AMERICAN JOURNAL OF ORTHODONTICS	AM J ORTHOD
AMERICAN JOURNAL OF PATHOLOGY	AM J PATH
AMERICAN JOURNAL OF PHARMACEUTICAL EDUCATION	AM J PHA ED

Notes: This figure shows the first page of the "Source Journal List" of the 1961 SCI (Garfield, 1963). This is a complete list of all 613 citing journals, from which citations were indexed for the 1961 SCI. We construct visible citations based on this list and the analogous lists from the 1964 to 1969 SCIs (see Section I.B).

Figure A.3: Internal Correspondence at Ohio State University

THE OHIO STATE UNIVERSITY LIBRARIES
Chemistry Library
COLUMBUS 10, OHIO

LEWIS C. BRANSCOMB
Director of Libraries

September 21, 1965

Professor Hyman W. Kritzer
Assistant Director, Public Services
Main Library, 1858 Neil Avenue
Campus

Dear Professor Kritzer:

I would like to place a subscription for the Science Citation Index for Chemistry Library. In view of the scarcity of funds for periodical subscriptions, I am sending the order to you for your approval and assistance.

A poll of some of the faculty in the Chemistry Department revealed that, without exception, all the younger men would like to have SCI available for use in Chemistry Library. The senior faculty are not as emphatic but would like to have it available also.

Since SCI definitely complements the Chemical Abstracts approach to the literature and since placing it in Chemistry Library would mean fairly close access for other science departments in the area, I think special consideration should be given to our request for a subscription.

Sincerely,

Virginia E. Yagello
Chemistry Librarian

VEY:kk

Enclosures: Correspondence to Chemistry faculty with comments attached.

Notes: In this letter, the chemistry librarian at Ohio State University requested a second copy of the SCI to be placed in the library of the chemistry department, in addition to the existing copy at the medical library. It shows that as early as 1965 there was large demand by chemists at Ohio State University to use the SCI. We thank archivists at Ohio State University Library for sharing this document.

B Further Details on Data

B.1 Data on Scientists

B.1.1 Linking Faculty Rosters with Publication and Citation Data

As described in the main text, we link scientists with their publications and citations using the linking algorithm developed in Iaria et al. (2022). The links are based on the academic's surname, first name or initials (depending on whether first names are available), country, city, and subject. The matching is based on the primary subject of each academic (e.g., physics) to reduce the number of false positives. To harmonize affiliations across the faculty rosters and the *Web of Science*, we rely on *Google Maps API*.

B.1.2 Coding Minority Status

In Section III, we report results on the heterogeneous effect of citation metrics. In particular, in Section III.C, we report differential results for women and for people with Asian, Hispanic, and Jewish names.

We use information in the faculty rosters to tag scientists as members of one of these groups. Gender coding relies on information on gender that can be directly observed in the faculty rosters (e.g., Miss in front of the first name) and the first names of scientists (see Iaria et al. (2022)).

We code Jewish names based on the approach in Benetti et al. (2023). Using their classification of Jewish names results in an overly conservative classification of Jewish scientists. We therefore lower the cut-off for classifying names as distinctively Jewish to 5 (instead of 10). However, results remain very similar when using the cut-off used in Benetti et al. (2023).

The coding of Hispanic names is based on data from the U.S. Census. We draw a list of Hispanic names from Name Census (2023b). From this list, we select all surnames with a conditional probability of self-identifying as Hispanic of more than 25%. We then tag all academics who have one of these names as Hispanic.

Similarly, we use data from the U.S. Census to code Asian names. We draw a list of the most common Asian names from Name Census (2023a). From this list, we select all surnames with a conditional probability of self-identifying as Asian or Pacific Islander of more than 50%.[1] We then tag all academics who have one of these names as Asian.

B.1.3 Data on NSF Grants

For the analysis in Section IV.B, we match scientists in our faculty rosters with historical records on grants by the National Science Foundation (NSF). We digitize entries on all

[1] The different cutoffs for Asian and Hispanic names reflect different assimilation patterns of the various immigrant groups. Results are very similar if we impose the same cutoffs for both groups.

grants listed in the 1969 Annual Report of the NSF.[2] We then match principal investigators from these grants to the scientists in our data based on first names, last name, and subject.

B.2 Department Rankings

The following six tables list the top 20 departments according to our self-constructed rankings (by average citations and by average publications in a department) and according to survey-based rankings from the 1960s and 1970s. Across all rankings similar departments are ranked among the top 20 departments.

Table B.1: Top 20 Departments: Biochemistry

Rank	Citations Ranking	Publications Ranking	Cartter Ranking	Roose-Andersen Ranking
1	Stanford	Washington	Harvard	Harvard
2	Rockefeller	Harvard	U.C. Berkeley	Stanford [2]
3	Johns Hopkins	Stanford	Stanford	U.C. Berkeley [2]
4	Washington	U.C. Berkeley	Rockefeller	Rockefeller
5	Harvard	Dartmouth	Wisconsin	Wisconsin
6	Kentucky	Wisconsin	M.I.T.	Cal. Tech.
7	U.C. Berkeley	Michigan	Cal. Tech.	M.I.T.
8	Dartmouth	Kentucky	Johns Hopkins	Brandeis [8]
9	Wisconsin	Johns Hopkins	Brandeis	Cornell [8]
10	Michigan	Virginia Polytechnic Institute	Illinois	Johns Hopkins [8]
11	U.C. Davis	U.C. Davis	Columbia	Duke [11]
12	Brandeis	Kansas [12]	Case Western Reserve	U.C.L.A. [11]
13	Case Western Reserve	Saint Louis [12]	N.Y.U.	U.C. San Diego [13]
14	Utah	Rockefeller	Washington	Washington [13]
15	Duke	Duke	Duke	Yeshiva University [13]
16	U.C.L.A.	U.C.L.A.	Michigan	Chicago [16]
17	Columbia	Columbia	Pennsylvania [17]	Illinois [16]
18	Pennsylvania	Case Western Reserve	Yeshiva University [17]	Princeton [16]
19	Chicago	Rice	Chicago	Case Western Reserve [19]
20	Rochester	Brandeis	U.C.L.A.	N.Y.U. [19]

Notes: This table lists the top 20 biochemistry departments based on four different department rankings. The first column reports our self-constructed ranking based on the average number of citations (between 1956 and 1969, to publications between 1956 and 1969) of all scientists employed at the department in 1969. The second column reports our self-constructed ranking based on the average number of publications (between 1956 and 1969) of all scientists employed at the department in 1969. The third column reports the ranking from Cartter (1966). The fourth column reports the ranking from Roose and Andersen (1970). Where departments are ranked equally (in any of the four rankings), a superscript indicates their rank. In the analysis, they are given the same rank.

[2] These data were generously shared by Dan Gross.

Table B.2: Top 20 Departments: Biology

Rank	Citations Ranking	Publications Ranking	Cartter Ranking	Roose-Andersen Ranking
1	Rockefeller	Albion College	U.C. Berkeley	Harvard
2	Albion College	Millikin	Harvard	U.C. Berkeley
3	Harvard	Texas	Cal. Tech.	M.I.T.
4	Princeton	Georgetown College	Johns Hopkins	Cal. Tech.
5	U.C. San Diego	Rockefeller [5]	Rockefeller	Rockefeller
6	Stanford	U.C. San Diego [5]	Wisconsin	Wisconsin
7	Cal. Tech.	U.C. Riverside	Illinois	Stanford
8	Texas	Wisconsin	Michigan	Washington
9	U.C. Berkeley	U.C. Berkeley	Stanford	U.C. San Diego [9]
10	Syracuse	Stanford	Minnesota	Yale [9]
11	Brandeis	U.C. Davis	Indiana [11]	Chicago
12	Yale	Brandeis	Princeton [11]	Illinois
13	Chicago	Princeton	Cornell	Cornell
14	M.I.T.	Notre Dame	Yale	U.C. Davis
15	U.C. Santa Barbara	Whitman College	Purdue [15]	Michigan
16	Notre Dame	Mount Holyoke College	U.C.L.A. [15]	Duke
17	Johns Hopkins	Alma College	Case Western Reserve	U.C.L.A.
18	Whitman College	U.C. Santa Barbara	Washington	Johns Hopkins
19	Washington	Central College Pella [19]	Chicago	Brandeis
20	U.C. Davis	Harvard [19]	Pennsylvania	Indiana

Notes: This table lists the top 20 biology departments based on four different department rankings. The first column reports our self-constructed ranking based on the average number of citations (between 1956 and 1969, to publications between 1956 and 1969) of all scientists employed at the department in 1969. The second column reports our self-constructed ranking based on the average number of publications (between 1956 and 1969) of all scientists employed at the department in 1969. The third column reports the ranking from Cartter (1966). While the Cartter ranking does not report rankings for biology overall, it does report rankings for five subfields of biology (Bacteriology/Microbiology, Botany, Entomology, Physiology, and Zoology). Based on these rankings, we construct an overall score for biology by taking the average rank of a department in the five reported subfields of biology. The fourth column reports the ranking from Roose and Andersen (1970). While the Roose-Andersen ranking does not report results for biology overall, it does report rankings for eight subfields of biology (Botany, Developmental Biology, Entomology, Microbiology, Molecular Biology, Physiology, Population Biology, and Zoology). Based on these rankings, we construct an overall score for biology by taking the average rank of a department in the eight reported subfields of biology. Where departments are ranked equally (in any of the four rankings), a superscript indicates their rank. In the analysis, they are given the same rank.

Table B.3: Top 20 Departments: Chemistry

Rank	Citations Ranking	Publications Ranking	Cartter Ranking	Roose-Andersen Ranking
1	U.C. Irvine	U.C. Santa Barbara	Harvard	Harvard
2	Stanford	Thiel College	Cal. Tech.	Cal. Tech.
3	Harvard	Stanford	U.C. Berkeley	Stanford [3]
4	U.C. Santa Barbara	U.C. Riverside	M.I.T.	U.C. Berkeley [3]
5	U.C.L.A.	U.C. Irvine	Stanford	M.I.T.
6	U.C. Riverside	Southern California	Illinois	Illinois
7	Cal. Tech.	College of Forestry at Syracuse	Columbia [7]	U.C.L.A.
8	Northwestern	Iowa State	Wisconsin [7]	Chicago [8]
9	Southern California	Utah	U.C.L.A.	Columbia [8]
10	College of Forestry at Syracuse	U.C. Davis	Chicago	Cornell [8]
11	Thiel College	Northwestern	Cornell	Wisconsin [8]
12	U.C. Berkeley	Texas	Yale	Yale
13	Iowa State	U.C.L.A.	Princeton	Princeton
14	Rice	Case Western Reserve	Northwestern	Northwestern
15	Illinois	Pennsylvania	Minnesota	Iowa State [15]
16	Utah	Illinois	Iowa State	Purdue [15]
17	Notre Dame	Johns Hopkins	Ohio State [17]	Ohio State [17]
18	U.C. Santa Cruz	Iowa State	Purdue [17]	Texas [17]
19	Columbia	Michigan	Michigan	U.C. San Diego [17]
20	Texas	Harvard	Indiana	Indiana

Notes: This table lists the top 20 chemistry departments based on four different department rankings. The first column reports our self-constructed ranking based on the average number of citations (between 1956 and 1969, to publications between 1956 and 1969) of all scientists employed at the department in 1969. The second column reports our self-constructed ranking based on the average number of publications (between 1956 and 1969) of all scientists employed at the department in 1969. The third column reports the ranking from Cartter (1966). The fourth column reports the ranking from Roose and Andersen (1970). Where departments are ranked equally (in any of the four rankings), a superscript indicates their rank. In the analysis, they are given the same rank.

Table B.4: Top 20 Departments: Mathematics

Rank	Citations Ranking	Publications Ranking	Cartter Ranking	Roose-Andersen Ranking
1	Princeton	U.C. Santa Barbara	Harvard	Harvard [1]
2	Chicago	U.C. Riverside	U.C. Berkeley	U.C. Berkeley [1]
3	Stanford	Harvard	Princeton	Princeton
4	Institute for Advanced Study	Princeton	Chicago	Chicago
5	Harvard	Carnegie-Mellon	M.I.T.	M.I.T.
6	Columbia	Washington	Stanford	Stanford
7	Johns Hopkins	Chicago	Yale	Yale
8	Brandeis	Johns Hopkins	N.Y.U.	N.Y.U.
9	U.C. Berkeley	Rockefeller	Columbia	Wisconsin
10	Virginia Polytechnic Institute	Stanford	Wisconsin	Columbia [10]
11	Rockefeller	Washington Saint Louis	Michigan	Michigan [10]
12	U.C. San Diego	Columbia	Illinois	Cornell [12]
13	Washington	Virginia	Cornell	Illinois [12]
14	Carnegie-Mellon	U.C. San Diego	Cal. Tech.	U.C.L.A.
15	Wisconsin	Wisconsin	Minnesota	Brandeis [15]
16	Yale	Brandeis	U.C.L.A.	Brown [15]
17	Washington Saint Louis	Yale	Washington	Cal. Tech. [15]
18	Case Institute of Technology	Institute for Advanced Study	Brown	Minnesota [18]
19	Brown	Minnesota	Brandeis	Pennsylvania [18]
20	Cornell	Michigan	Johns Hopkins	Washington [18]

Notes: This table lists the top 20 mathematics departments based on four different department rankings. The first column reports our self-constructed ranking based on the average number of citations (between 1956 and 1969, to publications between 1956 and 1969) of all scientists employed at the department in 1969. The second column reports our self-constructed ranking based on the average number of publications (between 1956 and 1969) of all scientists employed at the department in 1969. The third column reports the ranking from Cartter (1966). The fourth column reports the ranking from Roose and Andersen (1970). Where departments are ranked equally (in any of the four rankings), a superscript indicates their rank. In the analysis, they are given the same rank.

Table B.5: Top 20 Departments: Medicine

Rank	Citations Ranking	Publications Ranking	Cole-Lipton Ranking
1	Rockefeller	New Mexico	Harvard
2	Harvard	Minnesota Rochester	Johns Hopkins [2]
3	Utah	Rutgers	Stanford [2]
4	U.C. San Diego	U.C. San Diego	U.C. San Francisco
5	Minnesota Rochester	Harvard	Yale
6	Rutgers	Amherst College	Columbia
7	Washington	Loretto Heights College	Duke
8	M.I.T.	Medical College of Virginia	Michigan
9	Texas	M.I.T.	Cornell
10	U.C. San Francisco	Washington	Washington Saint Luis
11	Johns Hopkins	U.C.L.A.	Pennsylvania
12	Minnesota	Johns Hopkins	Minnesota
13	U.C.L.A.	Utah	U.C.L.A.
14	Florida	Minnesota	Albert Einstein College
15	New Mexico	Florida [15]	Chicago Pritzker [15]
16	Kansas	Rockefeller [15]	Washington [15]
17	Medical College of Virginia	U.C. San Francisco	Case Western Reserve
18	Washington Saint Louis	Southern California	Rochester
19	Stanford	Mississippi	Colorado
20	Columbia	Wagner College	U.C. San Diego

Notes: This table lists the top 20 biochemistry departments based on four different department rankings. The first column reports our self-constructed ranking based on the average number of citations (between 1956 and 1969, to publications between 1956 and 1969) of all scientists employed at the department in 1969. The second column reports our self-constructed ranking based on the average number of publications (between 1956 and 1969) of all scientists employed at the department in 1969. The third column reports the ranking from Cole and Lipton (1977). Since Cartter (1966) and Roose and Andersen (1970) do not report rankings for medical schools, we use the ranking by Cole and Lipton (1977) for medicine. Where departments are ranked equally (in any of the three rankings), a superscript indicates their rank. In the analysis, they are given the same rank.

Table B.6: Top 20 Departments: Physics

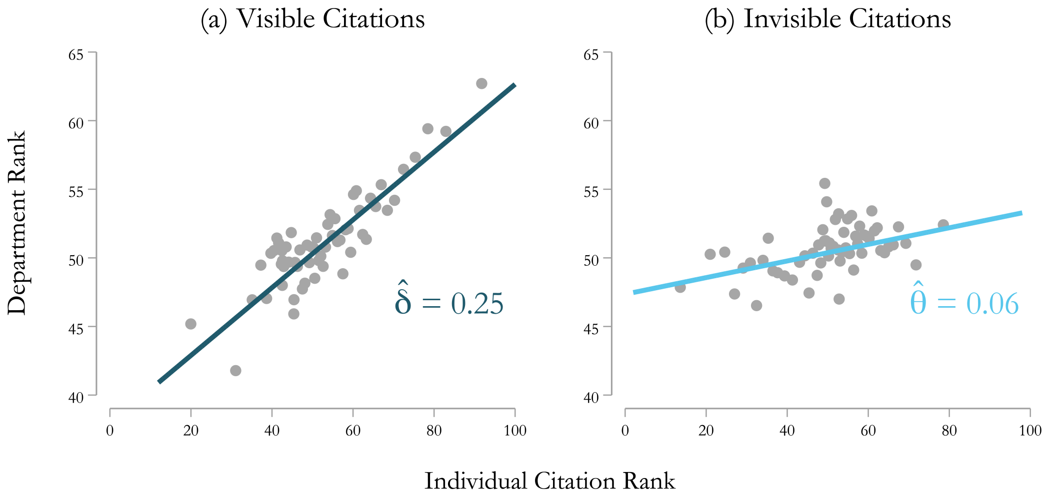
Rank	Citations Ranking	Publications Ranking	Cartter Ranking	Roose-Andersen Ranking
1	U.C. San Diego	U.C. Riverside	U.C. Berkeley	Cal. Tech. [1]
2	U.C. Riverside	U.C. San Diego	Cal. Tech.	Harvard [1]
3	U.C. Berkeley	Lycoming College	Harvard	U.C. Berkeley [1]
4	Chicago	U.C. Santa Barbara	Princeton	Princeton
5	Rockefeller	Kentucky Wesleyan College	Stanford	M.I.T. [5]
6	Stanford	Goshen College	M.I.T.	Stanford [5]
7	Princeton	Chicago	Columbia	Columbia [7]
8	Columbia	Harvard	Illinois	Illinois [7]
9	U.C. Santa Barbara	Rockefeller	Cornell	Chicago [9]
10	Harvard	U.C. Irvine	Chicago	Cornell [9]
11	Pennsylvania	Columbia	Yale	U.C. San Diego [11]
12	U.C. Irvine	Stanford	Wisconsin	Yale [11]
13	Brown	Princeton	Michigan [13]	Wisconsin
14	Carnegie-Mellon	Pennsylvania	Rochester [13]	Michigan [14]
15	Cal. Tech.	Pittsburgh	Pennsylvania	Pennsylvania [14]
16	Pittsburgh	Brown	Maryland	Maryland [16]
17	State University of New York	U.C. Berkeley	Minnesota	Rockefeller [16]
18	Washington	Iowa State	Washington	Rochester
19	Illinois	Washington	Johns Hopkins [19]	U.C.L.A.
20	Johns Hopkins	Notre Dame	U.C.L.A. [19]	Minnesota

Notes: This table lists the top 20 physics departments based on four different department rankings. The first column reports our self-constructed ranking based on the average number of citations (between 1956 and 1969, to publications between 1956 and 1969) of all scientists employed at the department in 1969. The second column reports our self-constructed ranking based on the average number of publications (between 1956 and 1969) of all scientists employed at the department in 1969. The third column reports the ranking from Cartter (1966). The fourth column reports the ranking from Roose and Andersen (1970). Where departments are ranked equally (in any of the four rankings), a superscript indicates their rank. In the analysis, they are given the same rank.

C Assortative Matching: Additional Results and Robustness

C.1 Graphical Representation of Specification 1

Figure C.1: Specification 1: Illustration of Results



Notes: The figure illustrates the results from Equation (1), see Table 3, Specification 1. Panel (a) shows a bin-scatter plot with the visible citation percentile rank on the horizontal axis and the department rank on the vertical axis, conditional on invisible citations and publication controls. Panel (b) shows a binned scatter plot with the invisible citation percentile rank on the horizontal axis and the department rank on the vertical axis, conditional on visible citations and publication controls. The slopes are significantly different from each other; the p-value from a t-test of no difference is < 0.001.

C.2 Robustness Checks

In this section, we show that the main results are robust to various changes to the analysis. First, in Appendix C.2.1, we show that results are similar for alternative measures of the department rank. Second, in Appendix C.2.2, we show results are similar for alternative performance measures of individual scientists. Third, in Appendix C.2.3, we show that the results are robust to different ways of assigning percentile ranks to scientists and departments. Last, in Appendix C.2.4, we show that the results hold in different subsamples. To reduce the number of tables, we report all robustness checks using the specification equivalent to column (3) in Table 3, Specification 1. The results are very similar across specifications using alternative control variables, corresponding to columns (1), (2), (4), and (5) in Table 3.

C.2.1 Alternative Department Rankings

First, we consider alternative department rankings. The main results (Table 3) are estimated with department ranks based on the leave-out mean of citations as the dependent variable. The results are robust to using rankings based on the mean of citations, i.e., including citations of the focal scientist (Table C.1, Panel A, column (2)). Instead of using department rankings based on citations, we can use scientists' publication counts to construct department rankings. This leaves the results almost unchanged (Table C.1, Panel A, columns (3) and (4)).

Our results also hold if we construct department rankings based on the scientific output of departments in the 1956 cross-section (Table C.1, Panel B). While 1956 rankings have the advantage that they are determined before the introduction of the SCI, they are not available for universities that only enter the data after 1956. Moreover, the 1956 rankings may suffer from higher measurement error, because we measure department composition before hiring and moving decisions were actually made. Ranking departments on the basis of 1956 rankings results in a 25 percent smaller sample. Nevertheless, the results remain qualitatively unchanged.

Our results are also robust to using external department rankings, which do not rely on citation or publication data. We draw on subject-specific reputational rankings from Roose and Andersen (1970) and Cartter (1966) to construct analogous department percentile ranks. To avoid unnecessary sample selection for this robustness check, departments that are not listed in these rankings are assigned the percentile rank between 1 and the lowest-ranked department.[3] As these rankings do not cover medical schools, we supplement

[3] This is necessary because these external rankings cover fewer departments than our data. Furthermore, Roose and Andersen (1970) and Cartter (1966) do not contain rankings for biology as a whole but for specific subfields of biology (Botany, Developmental Biology, Entomology, Microbiology, Molecular Biology, Physiology, Population Biology, and Zoology in the Roose-Andersen ranking; Botany, Entomology, Microbiology, Physiology, and Zoology in the Cartter ranking). For both the Roose-Andersen ranking and the Cartter ranking, we construct an overall ranking for biology by calculating the average rank of a department in the subfields of biology.

Table C.1: Robustness Check: Alternative Measures of Department Quality

	Dependent Variable: Department Rank			
	(1)	(2)	(3)	(4)
Department Ranking Based on:	Leave-Out Mean of Citations	Mean of Citations	Leave-Out Mean of Publications	Mean of Publications
Panel A: Department Rankings From 1969				
Visible Citations	0.280	0.320	0.286	0.318
	(0.035)	(0.030)	(0.034)	(0.028)
Invisible Citations	0.062	0.078	0.047	0.053
	(0.021)	(0.020)	(0.020)	(0.019)
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001
Observations	27,315	27,315	27,315	27,315
R^2	0.153	0.207	0.150	0.210
Dependent Variable Mean	50.40	50.20	50.37	50.16
Panel B: Department Rankings From 1956				
Visible Citations	0.169	0.178	0.158	0.175
	(0.038)	(0.039)	(0.037)	(0.039)
Invisible Citations	0.027	0.028	0.006	0.009
	(0.026)	(0.027)	(0.026)	(0.027)
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001
Observations	21,269	21,269	21,269	21,269
R^2	0.066	0.066	0.061	0.063
Dependent Variable Mean	50.29	55.59	50.26	56.27
Subject Fixed Effects	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes

Notes: The table reports the estimates of Equation (1) with alternative department rankings as dependent variables. In Panel A, department rankings are based on the 1969 cross-section of scientists; in Panel B, they are based on the 1956 cross-section. For departments that did not exist in 1956, the 1956 ranking cannot be computed. This results in a smaller sample size in Panel B. In column (1), the dependent variable is the department rank, based on the leave-out mean of citations in the department of scientist i (as in Table 3). In column (2), the department rank is based on the mean of citations in the department. In column (3), the department rank is based on the leave-out mean of publications in the department. In column (4), the department rank is based on the mean of publications in the department. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

these rankings with the first comprehensive ranking of medical schools by Cole and Lipton (1977). We report the results of these tests in Table C.2, column (4). The estimates show that our results are very similar if we use independently compiled reputation-based rankings.

Instead of percentile ranks, we can also use the reputational rankings from Cartter (1966) and Roose and Andersen (1970) to construct indicators for being in a top-ranked department. According to both rankings, we assign each scientist an indicator for whether they worked in a top-five, top-ten, or top-twenty department. In line with our main results, a scientist with a higher visible citation rank was more likely to work in a top department in 1969. For example, a ten-percentile increase in visible citations increased the probability of being affiliated with a top-twenty department by 2.94 percentage points (i.e., a 13.5 percent increase). In contrast, invisible citations had a much smaller effect on the assortativeness of the match to a top-department (Table C.2, columns (1)-(3)).

Table C.2: Robustness Check: External Department Ranking

	Dependent Variable: Indicator			*Dep. Rank*
	(1) Top 5	(2) Top 10	(3) Top 20	(4)
Panel A: Cartter Ranking				
Visible Citations	0.00077	0.00156	0.00294	0.224
	(0.00037)	(0.00039)	(0.00044)	(0.031)
Invisible Citations	0.00023	0.00059	0.00083	0.046
	(0.00018)	(0.00025)	(0.00032)	(0.022)
P-value (Visible = Invisible)	0.282	0.066	0.001	< 0.001
Observations	27,315	27,315	27,315	27,315
R^2	0.050	0.061	0.097	0.104
Dependent Variable Mean	0.04	0.12	0.22	50.15
Panel B: Roose-Andersen Ranking				
Visible Citations	0.00084	0.00166	0.00282	0.249
	(0.00037)	(0.00040)	(0.00043)	(0.032)
Invisible Citations	0.00025	0.00067	0.00096	0.039
	(0.00019)	(0.00025)	(0.00032)	(0.022)
P-value (Visible = Invisible)	0.234	0.061	0.004	< 0.001
Observations	27,315	27,315	27,315	27,315
R^2	0.053	0.065	0.099	0.116
Dependent Variable Mean	0.05	0.12	0.22	50.15
Subject Fixed Effects	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes

Notes: The table reports the estimates of Equation (1), where the dependent variable is based on the reputation-based department rankings by Cartter (1966) and Roose and Andersen (1970). Since these rankings do not cover medical schools, for medicine we supplement them with the ranking of medical schools by Cole and Lipton (1977). In columns (1)-(3), the dependent variable is an indicator for whether scientist i was employed at a top-5, top-10, or top-20 department. In column (4), the dependent variable is the rank of scientist i's department. To avoid unnecessary sample selection, we assign departments that are not listed in these rankings to the average rank between 1 and the lowest-ranked department. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

C.2.2 Alternative Transformations of Individual Citation Counts

We also show that results are robust to using alternative ways of measuring the performance of individual scientists.

For the main results, we count citations independently of the number of co-authors on the cited papers. In Table C.3, column (2), we report results of Specification 1, where citations to each paper are divided by the number of authors of the paper. The results are very similar.

Another concern could be that the results are driven by differences in the distributions of visible and invisible citations. Larger measurement error for invisible citations could potentially explain the smaller and insignificant coefficient for invisible citations. We address this concern with a robustness check in which we only use citations from 1956 to 1965 to construct visible and invisible citation ranks. This leads to similar distributions of visible and invisible citations.[4] For these alternative variables, measurement error concerns would, if anything, disproportionately downward bias the coefficient on visible citations. Using these alternative individual citation ranks leaves our results qualitatively unchanged (Table C.3, column (3)).

A further concern is that one "superstar" paper may place a scientist at the top of the citation distribution. However, having many moderately cited papers might be a better signal of quality than having very few highly cited papers. To account for both the number of cited papers and for the citations they receive, we use the h-index (e.g., Hirsch, 2005; Ellison, 2013) as an alternative performance metric. A scientist has an h-index of h if h of their papers have at least h citations each. We calculate the h-index of visible and invisible citations for each scientist. We then transform the h-index into the percentile rank for two reasons: first, this makes the coefficient directly comparable to the main results. Second, different scientific subjects have different publication and citation patterns. An h-index of three (i.e., having at least three publications with at least three citations) therefore indicates very different quality percentiles in each subject. For example, in medicine, a subject where scientists publish many papers and receive many citations, an h-index of three indicates poorer performance than in mathematics, a subject where scientists publish relatively few papers and receive a lot fewer citations. When we use percentiles of the visible and invisible h-indices as the explanatory variable, we confirm our main results (Table C.3, column (4)).

We also show that the results are similar if we standardize visible and invisible citations at the subject-level (Table C.3, column (5)). As standardized citations contain large outliers, we show that the results are also robust to winsorizing citation counts at the 99th percentile and then standardizing citation counts (Table C.3, column (6)). Further, the results are also similar if we use the inverse hyperbolic sine transformation of citations (Table C.3, column (7)).

[4] For citations measured in 1956-1965 the summary statistics are as follows. Visible citations: mean 14.3, standard deviation 41.4; invisible citations: mean 17.3, standard deviation 52.1.

Table C.3: Robustness Check: Alternative Transformations of Citation Counts

	Dependent Variable: Department Rank						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable Transformation:	Main Specification	Co-Author Weighted Citations	Only 1956-65 Citations	H-Index	Standard-ized	Winsorized & Std.	Inverse Hyperbolic Sine
Visible Citations	0.280	0.288	0.209	0.267	2.484	4.631	3.294
	(0.035)	(0.034)	(0.029)	(0.033)	(0.693)	(0.543)	(0.567)
Invisible Citations	0.062	0.062	0.119	0.081	0.367	1.461	1.268
	(0.021)	(0.022)	(0.025)	(0.021)	(0.545)	(0.416)	(0.309)
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P-value (Visible = Invisible)	< 0.001	< 0.001	0.009	< 0.001	0.063	< 0.001	0.002
Observations	27,315	27,315	27,315	27,315	27,315	27,315	27,315
R^2	0.153	0.157	0.139	0.152	0.105	0.116	0.149
Dependent Variable Mean	50.40	50.40	50.40	50.40	50.40	50.40	50.40

Notes: The table reports the estimates of Equation (1) for alternative transformations of visible and invisible citations. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist i. In column (1), the explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. In column (2), citation counts are divided by the number of authors of a paper and then transformed as in column (1). In column (3), citation counts are based only on citations from 1956-1965 (instead of 1956-1969), and then transformed as in column (1). In column (4), the explanatory variables are scientist i's h-index values based on visible and invisible citations, which are then transformed into the percentile rank. In column (5), we standardize citations by subject. In column (6), we standardize citations by subject, but to reduce the weight of outliers, we winsorize citation counts at the 99th percentile before standardizing them. In column (7), we transform citations using the inverse hyperbolic sine. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

C.2.3 Scientists and Departments with Zero Citations

When more than one percent of scientists have zero citations, a unique percentile rank cannot be assigned to these scientists. For example, in physics, 30.37% of observations have zero citations. For these scientists, there is no unique percentile in the distribution of citations. In our main analysis, we assign the mid-point between the 1st and the 31st percentile, i.e., a percentile rank of 15.5, to each of these observations. Alternatively, we can assign all of these observations to the 1st percentile (Min.-Point in Table C.4) or to the 31st percentile (Max.-Point). Reassuringly, the exact construction of percentile ranks of scientists with zero citations has no qualitative impact on the findings (Table C.4, columns (2) and (3)). A similar issue can occur for scientists with very low citation counts, e.g., one citation. We treat them accordingly.

Another way of assigning the percentile rank to scientists with zero citations is to spread the specific percentile rank randomly within the group of scientists with zero citations. In the above example of physicists with zero citations, this means that each of these scientists' percentile rank is independently drawn from a uniform distribution from 1 to 31. The results using this alternative transformation are similar to the main results column (4).

Table C.4: Robustness Check: Alternative Percentile Rank Definititions

	Dependent Variable: Department Rank			
	(1)	(2)	(3)	(4)
Variable Transformation:	Mid-Point (Main Spec.)	Min.-Point	Max.-Point	Random For 0 Cit.
Visible Citations	0.280	0.211	0.361	0.238
	(0.035)	(0.022)	(0.059)	(0.029)
Invisible Citations	0.062	0.048	0.107	0.068
	(0.021)	(0.014)	(0.036)	(0.015)
Subject Fixed Effects	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001
Observations	27,315	27,315	27,315	27,315
R^2	0.153	0.155	0.148	0.147
Dependent Variable Mean	50.40	50.03	50.76	50.40

Notes: The table reports the estimates of Equation (1) for alternative constructions of the percentile rank transformation. In all columns, the dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist i. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. The columns differ in how percentile ranks are assigned to brackets that comprise multiple percentiles. In column (1), departments and individuals without citations are assigned a percentile according to the midpoint between 1 and the lowest percentile with positive citations. In column (2), departments and individuals without citations are assigned to the first percentile. In column (3), departments and individuals without citations are assigned to the lowest percentile with positive citations. In column (4), individuals without citations are randomly assigned to a percentile rank within the bracket of zero citations. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

C.2.4 Alternative Sample Restrictions

We also show that the results are robust to restricting the sample in various ways. In particular, the findings are robust to excluding scientists with zero citations (Table C.5, column (2)). This test shows that our findings are not driven by scientists without citations. We also show that the results are robust to excluding scientists in small departments because department ranks may be less precisely calculated in small departments. For this test, we restrict the sample to all scientists in departments with more than 10 scientists (Table C.5, column (3)).

Table C.5: Robustness Check: Alternative Sample Restrictions

	Dep. Var.: Department Rank		
Sample Restriction:	(1) Full Sample	(2) Num. of Cit. > 0	(3) Size of Dept. > 10
Visible Citations	0.280	0.314	0.212
	(0.035)	(0.039)	(0.035)
Invisible Citations	0.062	0.085	0.060
	(0.021)	(0.020)	(0.021)
Subject Fixed Effects	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001
Observations	27,315	17,066	22,753
R^2	0.153	0.136	0.135
Dependent Variable Mean	50.40	56.56	54.97

Notes: The table reports the estimates of Equation (1) for alternative subsamples. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist i. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. In column (1), we use the full sample, i.e., it is equivalent to column (3) in of Table 3, Specification 1. Column (2) reports results for the subsample of scientists who have more than zero citations. Column (3) reports results for the subsample of scientists who are employed at departments with at least ten scientists. Standard errors are clustered at the department level.

C.3 Ruling out Alternative Explanations

In this section, we show that neither differences in the quality of citing journals nor differential timing of citations biases our findings (Tables C.6 and C.7). Figure C.2 illustrates the variation used in these tests.

Figure C.2: Illustration of Variation Used in Additional Tests

(a) Specification 1

	Citations in Journal A	Citations in Journal B	Citations in Journal C
1956			
1957		1	
1958			
1959	1		1
1960			
1961	1	1	
1962			1
1963	1		
1964			
1965		1	
1966		3	
1967	2		
1968			
1969			1

(b) Alternative Explanation 1

	Citations in Journal A	Citations in Journal B	Citations in Journal C
1956			
1957		1	
1958			
1959	1		1
1960			
1961	1	1	
1962			1
1963	1		
1964			
1965		1	
1966		3	
1967	2		
1968			
1969			1

(c) Alternative Explanation 2

	Citations in Journal A	Citations in Journal B	Citations in Journal C
1956			
1957		1	
1958			
1959	1		1
1960			
1961	1	1	
1962			1
1963	1		
1964			
1965		1	
1966		3	
1967	2		
1968			
1969			1

(d) Specification 2

	Citations in Journal A	Citations in Journal B	Citations in Journal C
1956			
1957		1	
1958			
1959	1		1
1960			
1961	1	1	
1962			1
1963	1		
1964			
1965		1	
1966		3	
1967	2		
1968			
1969			1

Notes: The four panels illustrate the sets of citations used for testing the alternative explanations in Appendix C.3 and for Specification 2 in Section II.C. As in Table 2, these tables illustrate citations for a hypothetical scientist. Panel (a) illustrates the variation used in Specification 1, see Table 3). Numbers in dark blue cells indicate citations that were visible in the SCI because the citation occurred in a journal and year (1961, or 1964-69) that was indexed by the SCI. Numbers in light blue cells indicate citations that were invisible, but are observable today. Panel (b) illustrates the variation used in testing Alternative Explanation 1, i.e., where citations are counted from a consistent set of journals (see Table C.6). We disregard citations in journals that were not indexed by the first SCI in 1961 (here: journals B and C), and focus only on citations in journals that were indexed by the 1961 SCI (here: journal A). Numbers in dark blue cells indicate citations that were visible in the SCI, i.e., citations from 1961, or 1964-69. Numbers in light blue cells indicate citations that were invisible because they came from years not covered by the SCI. Panel (c) illustrates the variation used in testing Alternative Explanation 2, i.e., where citation are counted in years in which the SCI was published (see Table C.7). We disregard citations from years in which the SCI was not published, and focus only on citations in years that were covered by the SCI, i.e., citations from 1961, or 1964-69. Numbers in dark blue cells indicate citations that were visible in the SCI, because they came from journals indexed by the SCI. Numbers in light blue cells indicate citations that were invisible because they came from journals not indexed by the SCI. Panel (d) illustrates the variation used in Specification 2, equivalent to Table 4 in the main paper.

Table C.6: Alternative Explanation 1: Citations From Consistent Set of Journals

	Dependent Variable: Department Rank				
	(1)	(2)	(3)	(4)	(5)
Visible Citations	0.289	0.299	0.260	0.228	0.219
	(0.034)	(0.030)	(0.033)	(0.033)	(0.034)
Invisible Citations	0.109	0.075	0.067	0.069	0.066
	(0.022)	(0.020)	(0.021)	(0.023)	(0.024)
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year		Yes			
Publications by Year × Subject			Yes	Yes	Yes
Publications by Journal				Yes	
Publications by Journal × Subject					Yes
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Observations	27,315	27,315	27,315	27,315	27,315
R^2	0.129	0.131	0.147	0.228	0.257
Dependent Variable Mean	50.40	50.40	50.40	50.40	50.40

Notes: The table reports the estimates of Equation (1), where individual citation counts are based only on the restricted set of citing journals that were indexed in the 1961 edition of the SCI. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist *i*. The explanatory variable *Visible Citations* measures scientist *i*'s individual rank in the distribution of visible citations in the restricted set of citing journals. *Invisible Citations* measures scientist *i*'s individual rank in the distribution of *in*visible citations in the restricted set of citing journals. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Publications by Year* separately measure the number of scientist *i*'s publications in each year between 1956 and 1969. *Publications by Journal* separately measure the number of scientist *i*'s publications in each journal (e.g., *Nature*). Standard errors are clustered at the department level.

Table C.7: Alternative Explanation 2: Citations Only From Years With SCI

	Dependent Variable: Department Rank				
	(1)	(2)	(3)	(4)	(5)
Visible Citations	0.342	0.347	0.302	0.275	0.263
	(0.039)	(0.035)	(0.040)	(0.040)	(0.041)
Invisible Citations	0.066	0.047	0.046	0.033	0.037
	(0.017)	(0.014)	(0.014)	(0.015)	(0.015)
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year		Yes			
Publications by Year × Subject			Yes	Yes	Yes
Publications by Journal				Yes	
Publications by Journal × Subject					Yes
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Observations	27,315	27,315	27,315	27,315	27,315
R^2	0.137	0.140	0.153	0.232	0.260
Dependent Variable Mean	50.40	50.40	50.40	50.40	50.40

Notes: The table reports the estimates of Equation (1), where individual citation counts are based only the restricted set of citations from years when the SCI was available, i.e., 1961 and 1964-1969. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist *i*. The explanatory variable *Visible Citations* measures scientist *i*'s individual rank in the distribution of visible citations in the restricted citation years. *Invisible Citations* measures scientist *i*'s individual rank in the distribution of *in*visible citations in the restricted citation years. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Publications by Year* separately measure the number of scientist *i*'s publications in each year between 1956 and 1969. *Publications by Journal* separately measure the number of scientist *i*'s publications in each journal (e.g., *Nature*). Standard errors are clustered at the department level.

While the test for Alternative Explanation 2 in Table C.7 considers only citations in years in which the SCI was published, one might still be concerned that even in this subset of citations, visible citations, on average, come from later years. If later citations are more important for career outcomes in 1969, this might still bias the results.

We address this concern by repeating the robustness test for smaller time windows within the years covered by the SCI. In Table C.8, we present the results for five different regressions in which we only count visible and invisible citations within five three-year windows (1961 and 1964-1965, 1964-1966, 1965-1968, 1966-1968, and 1967-1969). This enables us to abstract from the timing of citations and consider almost exclusively across-journal variation in visibility. We show that the difference between visible and invisible citations remains unchanged. Furthermore, the actual time window of measuring visible and invisible citations only has a small impact on the estimates. This corroborates the finding in Table C.7, that the timing of visible and invisible citations does not drive our results.

Table C.8: Alternative Explanation 2: Restricted Time Windows

	Dependent Variable: Department Rank				
	(1)	(2)	(3)	(4)	(5)
Citation Years:	1961, 1964-65	1964-66	1965-67	1966-68	1967-69
Visible Citations	0.278	0.293	0.302	0.305	0.302
	(0.038)	(0.039)	(0.039)	(0.039)	(0.039)
Invisible Citations	0.050	0.040	0.054	0.072	0.085
	(0.013)	(0.013)	(0.015)	(0.016)	(0.016)
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes	Yes
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Observations	27,315	27,315	27,315	27,315	27,315
R^2	0.141	0.145	0.147	0.149	0.150
Dependent Variable Mean	50.40	50.40	50.40	50.40	50.40

Notes: The table reports the estimates of Equation (1), where individual citation counts are based on restricted sets of citations from years when the SCI was available. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist *i*. The explanatory variable *Visible Citations* measures scientist *i*'s individual rank in the distribution of visible citations in the restricted citation years. *Invisible Citations* measures scientist *i*'s individual rank in the distribution of *in*visible citations in the restricted citation years. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. In column (1), visible and invisible citation counts are based on the years 1961 and 1964-65; in column (2) 1964-66; in column (3) 1965-67; in column (4) 1966-68; and in column (5) 1967-69. *Publications by Year* separately measure the number of scientist *i*'s publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

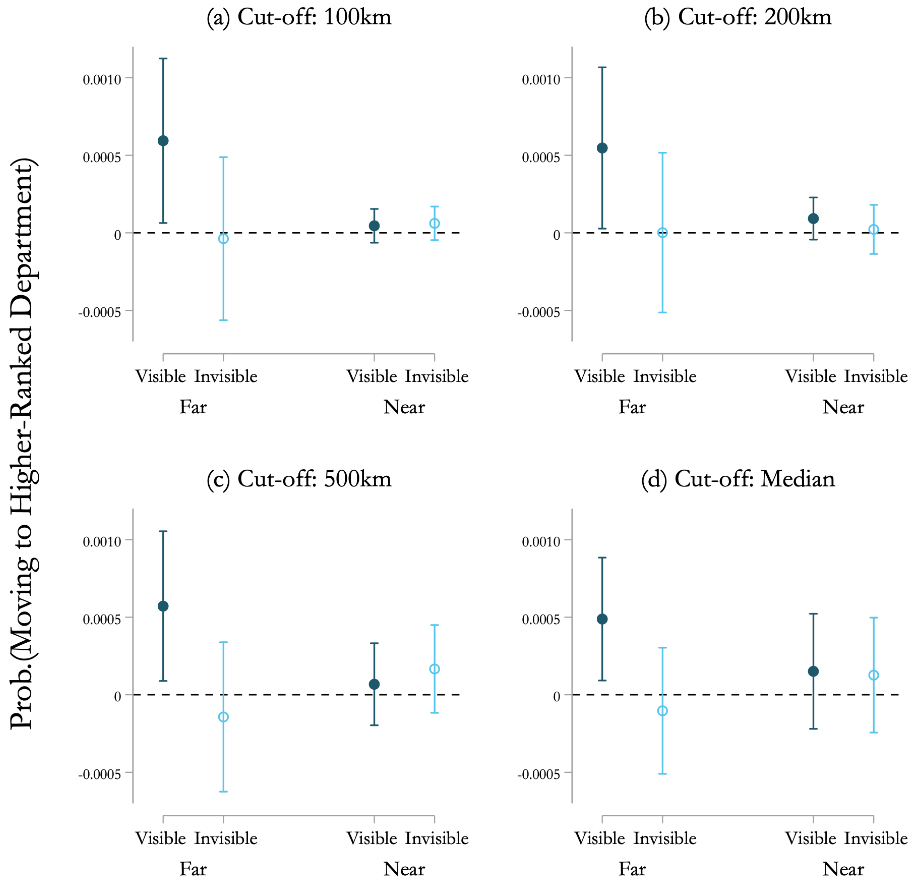
C.4 Additional Findings

Table C.9: Moving to Higher-Ranked Department by Geographic Distance

Dependent Variable: Moving to Higher-Ranked Department by Geographic Distance					
	(1)	(2)	(3)	(4)	(5)
Panel A: New Department Far					
Visible Citations	0.0007	0.0006	0.0006	0.0008	0.0006
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Invisible Citations	-0.0001	0.0000	-0.0000	-0.0004	-0.0003
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
P-value (Visible = Invisible)	0.097	0.227	0.220	0.070	0.154
Observations	6,478	6,478	6,478	6,478	6,478
R^2	0.013	0.017	0.036	0.332	0.398
Dependent Variable Mean	0.042	0.042	0.042	0.042	0.042
Panel B: New Department Near					
Visible Citations	0.0000	0.0000	0.0000	0.0000	0.0001
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Invisible Citations	0.0000	0.0000	0.0001	0.0000	-0.0000
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
P-value (Visible = Invisible)	0.952	0.797	0.873	0.976	0.742
Observations	6,478	6,478	6,478	6,478	6,478
R^2	0.001	0.003	0.021	0.321	0.442
Dependent Variable Mean	0.004	0.004	0.004	0.004	0.004
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year		Yes			
Publications by Year × Subject			Yes	Yes	Yes
Publications by Journal				Yes	
Publications by Journal × Subject					Yes

Notes: The table reports the estimates from variants of Equation (3) with different dependent variables: in Panel A, an indicator for moving to a higher-ranked department that was far from scientist i's department; in Panel B, an indicator for moving to a higher-ranked department that was close to scientist i's department. The cut-off between near and far departments is 100km. These regressions use the sample of scientists observed in 1956 and 1969. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst scientist. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. *Publications by Journal* separately measure the number of scientist i's publications in each journal (e.g., *Nature*). Standard errors are clustered at the department level.

Figure C.3: Moving to Higher-Ranked Departments by Geographic Distance - Alternative Cutoffs



Notes: The figure plots coefficients and 95 percent confidence intervals from variants of Equation (3). Each panel reports results from two regressions with alternative dependent variables: (i) an indicator for moving to a higher-ranked department that was far from scientist i's department; (ii) an indicator for moving to a higher-ranked department that was close to scientist i's department. In panel (a) the cut-off between near and far departments is 100km; in panel (b) 200km; in panel (c) 300km; and in panel (d) 837km, which is the median distance of moves.

Table C.10: Moving to Higher-Ranked Department by Citation Distance

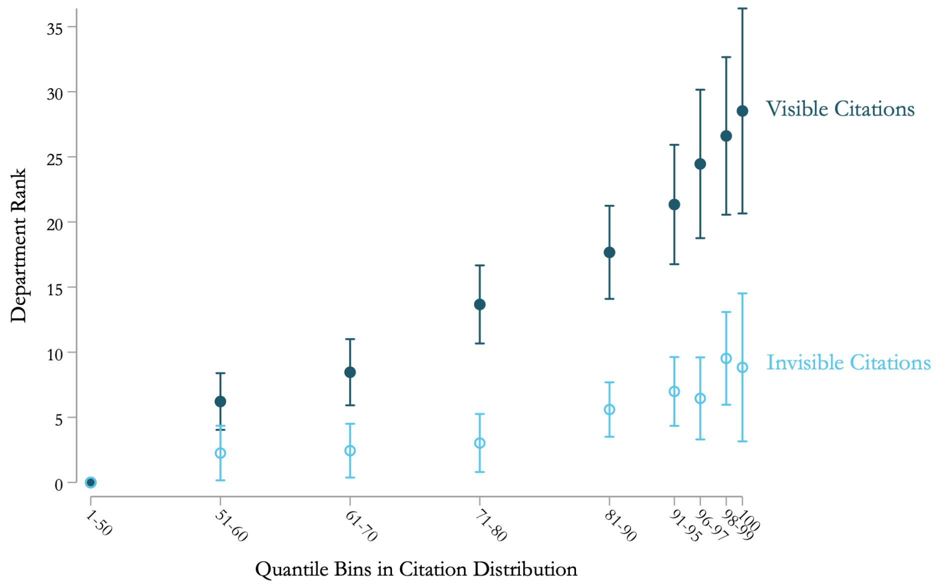
	Dependent Variable: Moving to Higher-Ranked Department by Citation Distance				
	(1)	(2)	(3)	(4)	(5)
Panel A: Not Cited In New Department Before SCI					
Visible Citations	0.0007	0.0007	0.0006	0.0008	0.0007
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Invisible Citations	-0.0004	-0.0002	-0.0002	-0.0005	-0.0005
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
P-value (Visible = Invisible)	0.027	0.082	0.110	0.034	0.053
Observations	6,478	6,478	6,478	6,478	6,478
R^2	0.008	0.012	0.026	0.294	0.360
Dependent Variable Mean	0.035	0.035	0.035	0.035	0.035
Panel B: Cited In New Department Before SCI					
Visible Citations	0.0001	-0.0000	0.0000	0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Invisible Citations	0.0004	0.0003	0.0002	0.0002	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
P-value (Visible = Invisible)	0.019	0.051	0.209	0.333	0.208
Observations	6,478	6,478	6,478	6,478	6,478
R^2	0.020	0.030	0.060	0.439	0.533
Dependent Variable Mean	0.011	0.011	0.011	0.011	0.011
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year		Yes			
Publications by Year × Subject			Yes	Yes	Yes
Publications by Journal				Yes	
Publications by Journal × Subject					Yes

Notes: The table reports the estimates from variants of Equation (3) with different dependent variables: in Panel A, an indicator for moving to a higher-ranked department where scientist i's papers were not cited before 1963; in Panel B, an indicator for moving to a higher-ranked department where scientist i's papers were cited before 1963. These regressions use the sample of scientists observed in 1956 and 1969. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst scientist. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. *Publications by Journal* separately measure the number of scientist i's publications in each journal (e.g., *Nature*). Standard errors are clustered at the department level.

D Additional Findings: Heterogeneity Analysis

D.1 Heterogeneous Effect in Non-Parametric Analysis

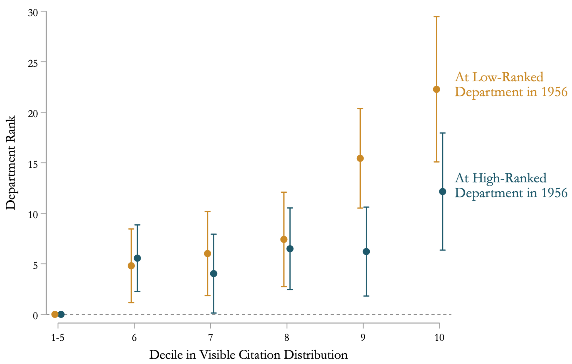
Figure D.1: Heterogeneous Effects by Percentile Rank



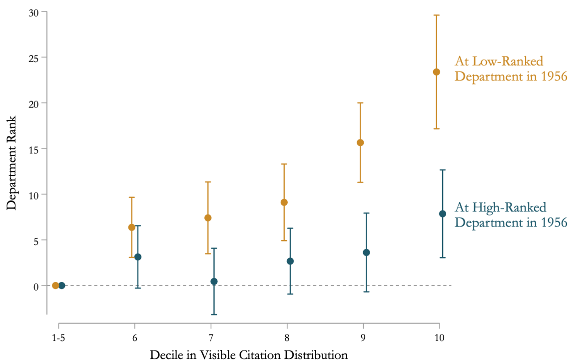
Notes: The figure plots coefficients $\hat{\delta}_q$ (dark blue) and $\hat{\theta}_q$ (light blue) and 95 percent confidence intervals from a variant of Equation (5). It differs from Figure 9 in that it splits up the 10th decile into smaller percentile bins.

Figure D.2: Heterogeneous Effects for Peripheral Scientists

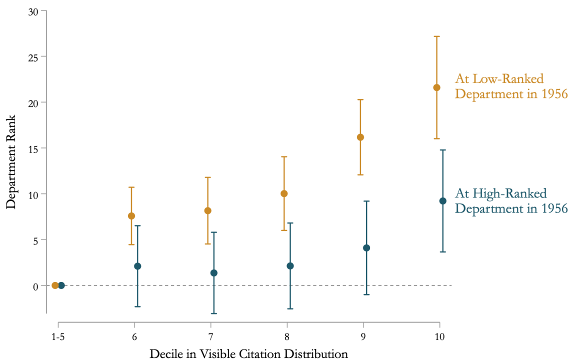
(a) Cutoff: 60th percentile



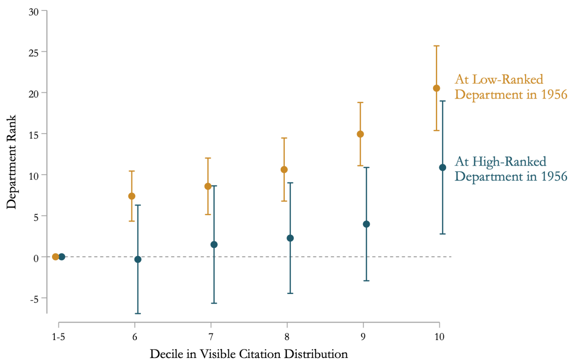
(b) Cutoff: 70th percentile



(c) Cutoff: 80th percentile



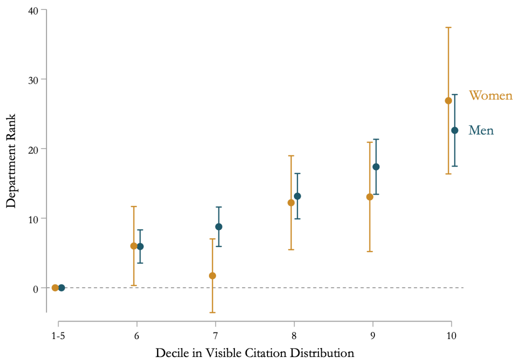
(d) Cutoff: 90th percentile



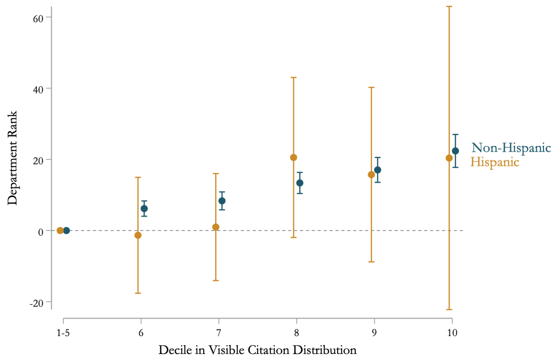
Notes: The figure plots coefficients $\hat{\delta}_q^H$ (orange) and $\hat{\delta}_q^L$ (blue) and 95 percent confidence intervals from Equation (6) for alternative cutoffs of high and low-ranked departments. In panel (a) we define low-ranked departments as those below the 60th percentile of the department ranking in 1956. In panel (b) we define low-ranked departments as those below the 70th percentile of the department ranking in 1956. In panel (c) we define low-ranked departments as those below the 80th percentile of the department ranking in 1956. In panel (d) we define low-ranked departments as those below the 90th percentile of the department ranking in 1956.

Figure D.3: Heterogenous Effects for Minority Scientists

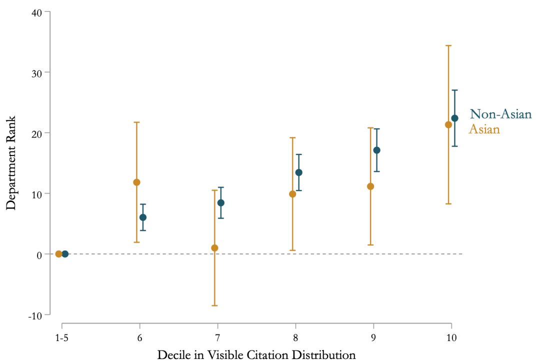
(a) Female Academics



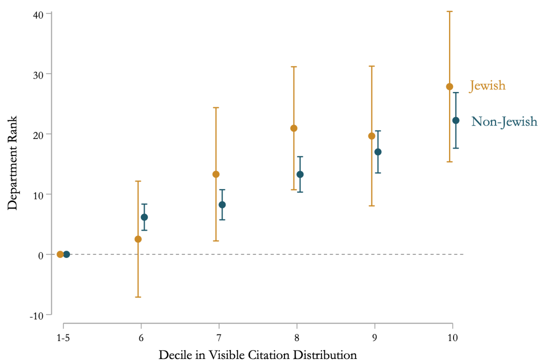
(b) Academics with Hispanic Names



(c) Academics with Asian Names

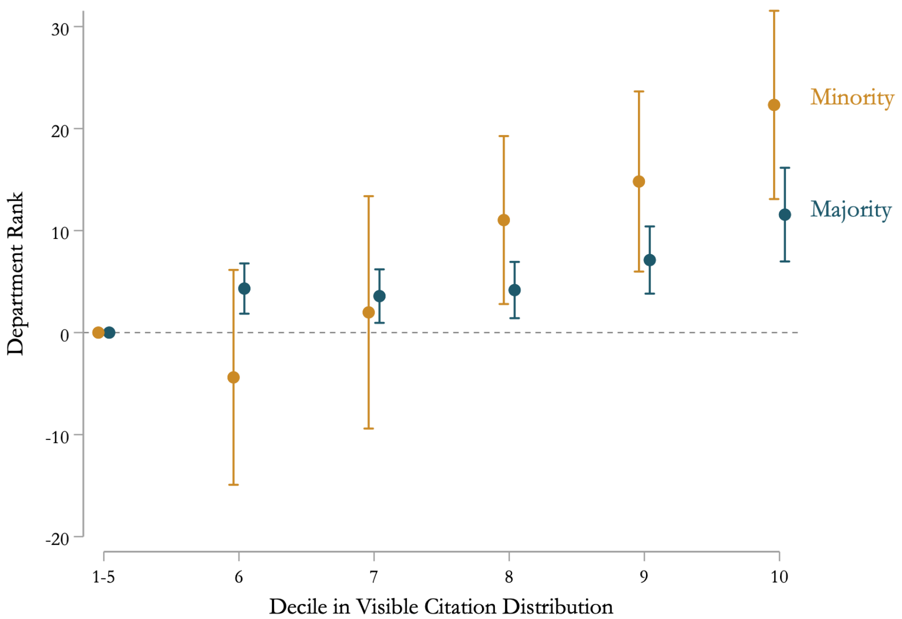


(d) Academics with Jewish Names



Notes: The figure plots coefficients $\hat{\delta}_q^M$ (blue) and $\hat{\delta}_q^m$ (orange) and 95 percent confidence intervals from Equation (7). Panel (a) plots separate sets of coefficients for women (orange) and men (blue). Panel (b) plots separate sets of coefficients for Hispanics (orange) and Non-Hispanics (blue). Panel (c) plots separate sets of coefficients for Asians (orange) and Non-Asians (blue). Panel (d) plots separate sets of coefficients for Jewish (orange) and Non-Jewish scientists (blue).

Figure D.4: Heterogenous Effects for Minority and Majority Scientists (Controlling For Department Rank in 1956)



Notes: The figure plots coefficients $\hat{\delta}_q^M$ (blue) and $\hat{\delta}_q^m$ (orange) and 95 percent confidence intervals from a variant of Equation (7), while controlling for the department rank of scientist in 1956. As a result, the sample is restricted to scientists who appear in both 1956 and 1969. The p-value for the test that the coefficients for the tenth decile are the same among minority and majority scientists is 0.034.

D.2 Heterogeneous Effect on Assortative Matching

In Sections III.B and III.C, we perform heterogeneity analyses for scientists at low-ranked departments and for minority scientists, respectively. These are based on a non-parametric regression as outlined in Equations (6) and (7). Below, we report additional results on the heterogeneous effect of citation metrics on assortative matching based on a variant of the main specification (Equation (1)):

$$\begin{aligned} Dep.\ Rank_i = {} & \delta \cdot Visible\,Citations_i + \delta^I \cdot Visible\,Citations_i \times Indicator_i \\ & + \theta \cdot Invisible\,Citations_i + \theta^I \cdot Invisible\,Citations_i \times Indicator_i \\ & + \omega \cdot Indicator_i + \pi \cdot Publications_i + Subject\,FE + \epsilon_i \end{aligned} \tag{D.1}$$

$Indicator_i$ takes value 1 if scientist i is a member of a specific subgroup of scientists. In Table D.1, we report results for peripheral scientists, i.e., where the indicator captures whether a scientist was working at a low-ranked department in 1956. In Table D.2, we report results for minority scientists, i.e., where the indicator captures whether the scientist was part of a minority group.

Table D.1: Heterogeneous Effect on Assortative Matching for Peripheral Scientists

	Dependent Variable: Department Rank				
	(1)	(2)	(3)	(4)	(5)
Definition of Low-Ranked Department:	Below 60	Below 70	Below 75	Below 80	Below 90
Visible Citations	0.168	0.112	0.088	0.119	0.176
	(0.043)	(0.038)	(0.040)	(0.047)	(0.070)
Invisible Citations	-0.001	-0.011	-0.008	-0.025	-0.074
	(0.035)	(0.035)	(0.036)	(0.042)	(0.058)
Visible Citations × Indicator	0.075	0.138	0.169	0.151	0.100
	(0.059)	(0.050)	(0.052)	(0.057)	(0.076)
Invisible Citations × Indicator	0.071	0.097	0.099	0.121	0.191
	(0.054)	(0.052)	(0.051)	(0.053)	(0.064)
Indicator	-36.700	-41.744	-43.410	-42.901	-40.917
	(3.488)	(3.273)	(3.368)	(3.688)	(5.275)
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes	Yes
Observations	6,374	6,374	6,374	6,374	6,374
R^2	0.394	0.367	0.351	0.319	0.240
Dependent Variable Mean	59.47	59.47	59.47	59.47	59.47

Notes: The table reports the estimates of Equation (D.1), where the indicator captures whether scientist i was working at a low-ranked department in 1956. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist i. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Indicator* is equal to one if scientist i worked at a low-ranked department in 1956. Thus, the sample used in this analysis is all scientists who appear in our data in both 1956 and 1969. We define low-ranked departments as those below a specific percentile in the 1956 department ranking. The different columns report estimates using different definitions of low-ranked department: 60th percentile in column (1), 70th percentile in (2), 75th percentile in column (3), 80th percentile in column (4), and 90th percentile in column (5). *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

Table D.2: Heterogeneous Effect on Assortative Matching for Minority Scientists

	Dependent Variable: Department Rank					
	(1)	(2)	(3)	(4)	(5)	(6)
Group Indicator:	Main	Female	Asian	Hispanic	Jewish	Any Minority
Visible Citations	0.280	0.285	0.281	0.280	0.279	0.270
	(0.035)	(0.040)	(0.035)	(0.035)	(0.035)	(0.033)
Invisible Citations	0.062	0.049	0.063	0.062	0.063	0.064
	(0.021)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)
Visible Citations × Indicator		-0.053	-0.050	0.068	0.049	0.020
		(0.050)	(0.076)	(0.181)	(0.088)	(0.044)
Invisible Citations × Indicator		-0.050	-0.043	0.035	-0.050	-0.039
		(0.055)	(0.084)	(0.179)	(0.087)	(0.043)
Indicator		-2.871	2.452	-5.042	5.754	-5.772
		(2.472)	(3.262)	(5.556)	(3.352)	(2.632)
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Publications by Year × Subject	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,315	24,529	27,315	27,315	27,315	27,315
R^2	0.153	0.162	0.153	0.153	0.154	0.159
Dependent Variable Mean	50.40	48.08	50.40	50.40	50.40	50.40

Notes: The table reports the estimates of Equation (D.1), where the indicator captures whether scientist i is part of a minority group. The dependent variable is the department rank in 1969, based on the leave-out mean of citations in the department of scientist i. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. We transform ranks into percentiles, where 100 is the best and 1 the worst department/scientist. *Indicator* is equal to one if scientist i is part of a minority group. Column (1) reports estimates of the main specification for reference (see column (3) in Table 3, Specification 1). Columns (2)-(5) report estimates from regressions where the indicator captures if scientist i is part of a minority group: female in column (2), Asian in column (3), Hispanic in column (4), and Jewish in column (5). Column (6) reports the estimates from a regression where the indicator equals one if scientist i is part of any one of these subgroups. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. Standard errors are clustered at the department level.

E Additional Findings: Career Outcomes

Table E.1: Receiving an NSF Grant

	Dependent Variable: Receiving NSF Grant				
	(1)	(2)	(3)	(4)	(5)
Specification 1: Visible vs. Invisible Citations					
Visible Citations	0.0013	0.0013	0.0009	0.0008	0.0007
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Invisible Citations	-0.0001	-0.0001	-0.0002	-0.0001	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
P-value (Visible = Invisible)	< 0.001	< 0.001	< 0.001	< 0.001	0.001
R^2	0.066	0.067	0.107	0.221	0.268
Specification 2: Visible vs. Pseudo-Visible vs. Invisible Citations					
Visible Citations	0.0014	0.0014	0.0009	0.0008	0.0007
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Pseudo-Visible Citations	-0.0005	-0.0005	-0.0005	-0.0003	-0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Invisible Citations (SCI years)	0.0001	0.0001	0.0001	0.0002	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Invisible Citations (non-SCI years)	0.0003	0.0003	0.0003	0.0002	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
P-value (Visible = Pseudo-Visible)	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
P-value (Visible = Invisible (SCI))	< 0.001	< 0.001	< 0.001	0.001	0.001
P-value (Visible = Invisible (non-SCI))	< 0.001	< 0.001	< 0.001	0.001	0.003
P-value (Pseudo-Vis. = Invis. (SCI) = Invis. (non-SCI))	< 0.001	< 0.001	< 0.001	0.021	0.067
R^2	0.067	0.068	0.108	0.222	0.268
Subject Fixed Effects	Yes	Yes	Yes	Yes	Yes
Publications by Year		Yes			
Publications by Year × Subject			Yes	Yes	Yes
Publications by Journal				Yes	
Publications by Journal × Subject					Yes
Observations	27,315	27,315	27,315	27,315	27,315
Dependent Variable Mean	0.039	0.039	0.039	0.039	0.039

Notes: The table reports the estimates of Equation (8) in the first panel and of Equation (9) in the second panel. The dependent variable is an indicator equal to one if scientist i received an NSF grant in 1969. These regressions use the sample of scientists observed in 1969, including medicine. The explanatory variable *Visible Citations* measures scientist i's individual rank in the distribution of visible citations. *Invisible Citations* measures scientist i's individual rank in the distribution of *in*visible citations. *Pseudo-Visible Citations* measures scientist i's individual rank in the distribution of pseudo-visible citations (citations in journals indexed in the SCI in 1961, but for years not covered in the SCI, i.e., 1956-1960 and 1962-1963). *Invisible Citations (SCI years)* measures scientist i's individual rank in the distribution of *in*visible citations in SCI years (1961 and 1964-1969). *Invisible Citations (non-SCI years)* measures scientist i's individual rank in the distribution of *in*visible citations in non-SCI years (citations in journals not indexed in the SCI in 1961 and in years that were not covered, i.e., 1956-1960 and 1962-1963). We transform ranks into percentiles, where 100 is the best and 1 the worst scientist. *Publications by Year* separately measure the number of scientist i's publications in each year between 1956 and 1969. *Publications by Journal* separately measure the number of scientist i's publications in each journal (e.g., *Nature*). Standard errors are clustered at the department level.

References

Benetti, S., E. Locke, and R. Mattheis (2023). Our Crowd: Discrimination and Identity in Jewish Migration to the US. *Working Paper*.

Cartter, A. M. (1966). *An Assessment of Quality in Graduate Education.*

Cole, J. R. and J. A. Lipton (1977). The Reputations of American Medical Schools. *Social Forces 55*(3), 662–684.

Ellison, G. (2013). How does the Market use Citation Data? The Hirsch Index in Economics. *American Economic Journal: Applied Economics 5*(3), 63–90.

Garfield, E. (1963). *Science Citation Index 1961*, Volume 1. Philadelphia: Institute for Scientific Information, Inc.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences 102*(46), 16569–16572.

Iaria, A., C. Schwarz, and F. Waldinger (2022). Gender Gaps in Academia: Global Evidence Over the Twentieth Century. *Available at SSRN 4150221*.

Name Census (2023a). Most Common Asian and Pacific Islander Last Names in the United States. `https://namecensus.com/last-names/common-asian-and-pacific-islander-surnames/`.

Name Census (2023b). Most Common Hispanic Last Names in the United States. `https://namecensus.com/last-names/common-hispanic-surnames/`.

Roose, K. D. and C. J. Andersen (1970). *A Rating of Graduate Programs.* Washington: American Council on Education.