Bibliometric tools in economics and some examples of their use

Pierre-Philippe Combes

University of Lyon

Sciences Po

How it started

- Late 1990's, early 2000's, a number of claims that
  - Standard bibliometric approaches were not possible to implement in social sciences (lack of data),
  - Would make no sense because ignoring plenty of supports (books,...),
  - Each ranking methodology would lead to a different ranking.
- My first co-author on that, Laurent Linnemer and me, but also a number of academic instances (French ministery of higher education, European Economic Association,...) were a bit annoyed by these claims.
- Overview of a number of things we did on that,
  - From the least interesting: Sensitivity of rankings to index choice,
  - To more interesting stuffs:
    - Determinants of promotions in economics (skills vs networks),
    - Gender differences in promotion (discrimination vs application),
    - Research efficiency determinants (individual skills vs departement effects).

First papers: Are rankings affected by index choice?

- General answer: Not really.

  Even less at the department level than at the individual level.

- First series: Papers on French economists and departments in 1998.

  (Revue Économique, 2001, Annales d'Économie et Statistiques, 2003).

  Also at the European level + top 50 US deparments

  (Journal of the European Economic Association, 2003).

- Second wave: Reports for the French Ministry of Higher Education (2009, 2010, 2011).

  Access to administrative data for all academics in economics in France

  in 2008 with all their past positions since the 80's

  and we gathered publications records.

First papers: Are rankings affected by index choice?

- Role of considering in the publication index:
  - The journal's quality vs only counting the number of papers,
  - The number of co-authors $(1/n)$ or not,

    in economics, 50% are written alone of publications are written alone,

    less than 4% written by 4 or more people.

    + alphabetical order $\Rightarrow$ Simple.
  - The paper's number of pages or not,
  - The publication period (last five years, all, all time-discounted, all per year of carreer,...).

Journal ranking

- First series of rankings: Based on the 1218 Econlit journals,

  with peer assessment of quality/weights (Tirole, Laffont, LL and me,...).
  - Not really serious (this is still what CNRS and AERES do,...),
  - But WoS impossible because only 304 journals are classified in 'Business and Economics'.

- Second series of rankings:

  New journal ranking for the 1218 journals in Econlit.
  - Working paper not published but used for promotions

    in a number of North-American departments,
  - Best explains academics' salaries in the University of California

    economics departments (Gibson et al., 2014, Economic inquiry).

Journal ranking

- Index proposed, Weighted average of three indexes:

$$CL_j = 0.5\, I_1 + 0.25\, I_2 + 0.25\, I_3,$$

  where:

  - I_1: JCR (Wos) citation impact factor,
  - I_2: JCR citation impact factor "within field" (JEL code lettre),
  - I_3: A Google Scholar citation index.

- Details:
  - Ad hoc weights.
  - Correction of the impact factor by the share of their "economist" authors (each author is economist at x% where x is his share of papers in the Business and Economics category of WoS).

- Only computable for the 304 WoS journals in economics.

Journal ranking

- But one can compute for any author in Econlit (world level) some publication scores according to the three underlying indexes I_1, I_2 and I_3 (and a number of other GS indexes).
- Then for any of the journal j, one can sum its authors' score to obtain a journal score, $I_{i,j}^{Authors}$.

Journal ranking

- For the 304 WoS index, we regress their CL index and their author's score:

$$CL_j = \alpha + \beta_1 I_{1,j}^{Authors} + \beta_2 I_{2,j}^{Authors} + \beta_3 I_{3,j}^{Authors} + \varepsilon_j.$$

  Actually, slightly more complex (polynomial function, also other GS indexes).

- Finally we obtain can predict for any of the 1218 Econlit journals its CL index through its authors' scores (in WoS journals and GS):

$$\widehat{CL}_j = \hat{\alpha} + \hat{\beta}_1 I_{1,j}^{Authors} + \hat{\beta}_2 I_{2,j}^{Authors} + \hat{\beta}_3 I_{3,j}^{Authors}.$$

- Correlation between $\widehat{CL}_j$ and CL_j for the 304 WoS journals: 0.97.

- Remark: Then we only rank the journals according to this score and we convexify in an ad hoc way, slightly (CLm) and largely (CLh).

Journal ranking: Top 30 Journals

Journal	*clm*	*clh*
quarterly journal of economics	100,0	100,0
american economic review	98,1	94,4
journal of political economy	96,2	89,1
econometrica	95,7	87,7
review of economic studies	81,0	53,1
journal of financial economics	80,6	52,4
journal of monetary economics	75,8	43,6
review of economics and statistics	74,1	40,7
journal of economic theory	72,8	38,5
journal of finance	72,2	37,6
journal of econometrics	68,6	32,3
economic journal	64,5	26,8
rand journal of economics	63,7	25,8
journal of public economics	62,0	23,9
journal of international economics	61,5	23,3
journal of the european economic association	57,0	18,5
european economic review	55,2	16,8
journal of labor economics	55,1	16,7
international economic review	54,7	16,4
games and economic behavior	54,1	15,8
review of financial studies	49,1	11,8
journal of business and economic statistics	48,1	11,1
journal of health economics	43,9	8,5
journal of development economics	42,7	7,8
journal of human resources	42,2	7,5
journal of money credit and banking	41,9	7,3
journal of law and economics	40,7	6,8
journal of accounting and economics	40,5	6,6
journal of urban economics	40,0	6,4

Sensitivity of rankings using different criteria

- Little sensitivity of rankings using different criteria.
- Slightly more for individual ranking

  and when based on different periods of time.
- Rank correlation for French departments:

	E11	E1n	Ep1	Epn	CLm11	CLm1n	CLmp1	CLmpn	CLh11	CLh1n	CLhp1	CLhpn
E11	1	0.96	0.95	0.91	0.93	0.93	0.92	0.92	0.81	0.81	0.81	0.81
E1n		1	0.98	0.90	0.88	0.93	0.87	0.92	0.76	0.78	0.75	0.78
Ep1			1	0.91	0.87	0.91	0.89	0.93	0.75	0.77	0.76	0.78
Epn				1	0.97	0.98	0.98	0.99	0.93	0.95	0.94	0.95
CLm11					1	0.98	0.99	0.97	0.96	0.95	0.95	0.95
CLm1n						1	0.97	0.99	0.93	0.94	0.92	0.94
CLmp1							1	0.98	0.95	0.94	0.95	0.95
CLmpn								1	0.92	0.94	0.93	0.94
CLh11									1	0.99	1	0.99
CLh1n										1	0.99	1
CLhp1											1	0.99
CLhpn												1

Individual citations vs journal quality based rankings?

- Paper with Clément Bosquet, Scientometrics 2013.
- Comparison of rankings based on CL journal indexes and on individual GS citations records.
- Both still largely correlated (not a surprise) but less than within the family of journal-based rankings:

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Quantity (1)	1	0.89	0.70	0.32	0.43	0.61	0.60
Quality total score (2)		1	0.93	0.72	0.77	0.64	0.64
Top quality total score (3)			1	0.85	0.94	0.59	0.60
Average quality (4)				1	0.93	0.41	0.40
Average top quality (5)					1	0.47	0.48
Total citations (6)						1	0.95
G-index (7)							1

- Then we also assess the role of some individual characteristics on publications but extended in a later paper, see below.

Promotions: Skills or networks?

- Paper with Laurent Linnemer and Mickael Visser, Labour Economics 2008 .
- Determinants of success at the "agrégation du supérieur".
- Demographic, publication, and network effects.
- Networks, four variables tested:
  - A-link: PhD advisor in the jury
    (5% of the candidates, from 17% in 1984 to 3% in 2003),
  - T-link: PhD done in one of the universities of the jury members
    (40% of the candidates),
  - P-link: Assistant Prof position in one of the universities of the jury members
    (18% of the candidates),
  - P-link: PhD advisor co-author of one of the jury members
    (3% of the candidates).

Promotions: Skills or network?

Variable	Coeff.	(Std. err.)	Coeff.	(Std. err.)
Publication variables				
Number of pub.	0.178**	(0.022)	0.178**	(0.024)
Quality of pub.	0.103**	(0.023)	0.092**	(0.024)
Network variables				
A-Link	0.706**	(0.225)	0.891**	(0.233)
T-Link	0.230	(0.152)	0.106	(0.190)
P-Link	0.433**	(0.163)	0.391*	(0.193)
IA-Link	-0.095	(0.390)	-0.513	(0.403)
Control variables				
Age			-0.492**	(0.139)
Age squared			0.005**	(0.002)
Female			-0.179	(0.163)
French			-0.025	(0.307)
Academic			0.401	(0.225)
Ph.D. abroad			1.111**	(0.371)
Position in Paris			0.496**	(0.167)
Position in top-6 univ.			0.067	(0.185)
Ph.D. from top-6 univ.			-0.220	(0.197)
First time			-0.230	(0.158)
Ph.D. adv. nb. pub.			0.016**	(0.004)
Ph.D. adv. qual. pub.			0.048	(0.026)
Number of observations	993		993	

Promotions: Skills or networks?

		(A-Link,P-Link)			Quantile nb. pub.					Quantile qual. pub.				
Year	Ref.	(1,0)	(0,1)	(1,1)	.1th	.25th	.5th	.75th	.9th	.1th	.25th	.5th	.75th	.9th
1984	0.174	0.371	0.249	0.497	0.143	0.143	0.172	0.202	0.262	0.146	0.146	0.163	0.193	0.228
1987	0.217	0.457	0.311	0.582	0.168	0.168	0.200	0.247	0.315	0.176	0.176	0.198	0.257	0.323
1989	0.192	0.405	0.277	0.526	0.157	0.157	0.183	0.211	0.294	0.165	0.165	0.189	0.209	0.245
1991	0.245	0.483	0.334	0.618	0.194	0.194	0.228	0.267	0.359	0.219	0.219	0.216	0.247	0.300
1993	0.187	0.400	0.262	0.522	0.145	0.145	0.165	0.232	0.269	0.166	0.166	0.176	0.188	0.215
1995	0.156	0.345	0.228	0.468	0.120	0.120	0.142	0.198	0.222	0.140	0.140	0.164	0.176	0.185
1997	0.124	0.277	0.174	0.386	0.079	0.100	0.122	0.163	0.191	0.106	0.112	0.120	0.134	0.154
1999	0.130	0.289	0.184	0.391	0.082	0.098	0.113	0.159	0.223	0.107	0.119	0.128	0.141	0.155
2001	0.149	0.338	0.220	0.460	0.094	0.108	0.157	0.177	0.274	0.126	0.142	0.157	0.167	0.189
2003	0.084	0.191	0.125	0.267	0.049	0.057	0.079	0.090	0.157	0.069	0.076	0.084	0.093	0.097
Av.	0.166	0.356	0.236	0.472	0.123	0.129	0.156	0.195	0.257	0.142	0.146	0.160	0.180	0.209

Why female academic economists are less promoted than males?

- Paper with Clément Bosquet and Cecilia Garcia-Penalosa, Journal of Scandinavian Economics, forthcoming.
- Demographic and publication determinants of promotion for French academics.

  Both for university and full research positions.
- Decomposition of the probability to be promoted in
  - The probability to apply to promotion
  - And the probability to be promoted conditional on applying,

  controlling for a number of individual characteristics and publication records.

Unconditional probability to be promoted

	Likelihood to						
	hold a rank A position			be promoted (potential candidates)			
				University		CNRS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Woman	-0.233[a]	-0.045[a]	-0.043[a]	-0.009[a]	-0.010[a]	-0.029[a]	-0.023[b]
	(0.019)	(0.016)	(0.016)	(0.003)	(0.003)	(0.010)	(0.011)
Age		0.011[a]	0.011[a]		-0.009[a]		0.020[a]
		(0.002)	(0.002)		(0.001)		(0.004)
Age^2		0.000[b]	0.000[b]		0.000[a]		-0.000[a]
		(0.000)	(0.000)		(0.000)		(0.000)
Publisher(Pub)		0.316[a]	0.317[a]		0.041[a]		0.032[b]
		(0.021)	(0.026)		(0.007)		(0.015)
Pub*Quantity		0.179[a]	0.182[a]		0.015[a]		0.037[a]
		(0.009)	(0.010)		(0.003)		(0.008)
Pub*Quality		0.034[a]	0.033[a]		0.009[a]		0.011[a]
		(0.004)	(0.005)		(0.002)		(0.004)
CNRS		-0.114[a]	-0.110[a]		-0.038[a]		
		(0.024)	(0.028)		(0.005)		
Woman*Pub			0.001				
			(0.045)				
Woman*Pub*Quantity			-0.024				
			(0.021)				
Woman*Pub*Quality			0.007				
			(0.011)				
Woman*CNRS			-0.013				
			(0.057)				
Int. Department					0.006		-0.002
					(0.004)		(0.011)
Ile de France					0.009[b]		0.018
					(0.004)		(0.013)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.048	0.392	0.393	0.003	0.051	0.010	0.077
Observations	17,467	17,467	17,467	8,085	8,085	1,132	1,132

Probability to apply

	University		CNRS		Diff. in diff.	Diff. across dept. University	Diff. across dept. CNRS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Woman	-0.030^a	-0.033^a	-0.093^b	-0.071^b	-0.028^a	-0.033^a	-0.071^b
	(0.008)	(0.007)	(0.040)	(0.036)	(0.008)	(0.008)	(0.036)
Age		-0.021^a		0.084^a	-0.008^a	-0.021^a	0.084^a
		(0.003)		(0.012)	(0.003)	(0.003)	(0.012)
Age^2		0.000^a		-0.002^a	0.000	0.000^a	-0.002^a
		(0.000)		(0.000)	(0.000)	(0.000)	(0.000)
Publisher(Pub)		0.114^a		0.118^a	0.124^a	0.114^a	0.117^a
		(0.015)		(0.045)	(0.015)	(0.015)	(0.045)
Pub*Quantity		0.046^a		0.108^a	0.046^a	0.046^a	0.107^a
		(0.007)		(0.028)	(0.007)	(0.007)	(0.028)
Pub*Quality		0.014^a		0.022^b	0.015^a	0.014^a	0.022^b
		(0.004)		(0.010)	(0.004)	(0.004)	(0.010)
Int. Department		0.004		0.020	0.007	0.005	0.016
		(0.010)		(0.048)	(0.010)	(0.010)	(0.050)
Ile de France		0.019^b		0.005	0.020^b	0.007	0.020
		(0.009)		(0.043)	(0.009)	(0.011)	(0.058)
CNRS		-0.127^a				-0.127^a	
		(0.011)				(0.011)	
Woman*CNRS					-0.056		
					(0.036)		
Woman*Ile de France						0.032^b	-0.041
						(0.016)	(0.077)
Interacted terms	No	No	No	No	Yes	No	No
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.005	0.110	0.033	0.194	0.116	0.110	0.195
Observations	8,085	8,085	1,132	1,132	9,217	8,085	1,132

Probability to be promoted conditional on applying

	University			CNRS		
	(1)	(2)	(3)	(4)	(5)	(6)
Woman	-0.029	-0.035	-0.046	-0.068	-0.079	-0.063
	(0.033)	(0.032)	(0.032)	(0.059)	(0.053)	(0.061)
Age		-0.032^{b}	-0.032^{b}		-0.048	-0.042
		(0.013)	(0.014)		(0.039)	(0.039)
Age^2		0.001	0.001		0.001	0.001
		(0.000)	(0.000)		(0.001)	(0.001)
Publisher(Pub)		0.175^{a}	0.157^{a}		0.004	0.028
		(0.046)	(0.048)		(0.124)	(0.137)
Pub*Quantity		0.074^{a}	0.074^{a}		0.063^{c}	0.075^{c}
		(0.020)	(0.020)		(0.038)	(0.041)
Pub*Quality		0.057^{a}	0.053^{a}		0.041^{a}	0.039^{b}
		(0.009)	(0.009)		(0.015)	(0.018)
Int. Department			0.093^{a}			-0.010
			(0.036)			(0.069)
Ile de France			0.023			0.067
			(0.033)			(0.074)
Pos. other than univ.			0.001			
			(0.036)			
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.030	0.150	0.164	0.043	0.192	0.197
Observations	781	781	781	198	198	198

What makes an academic productive?

- Paper with Clément Bosquet, Journal of Urban Economics, 2017.
- What makes an academic productive?
  - Its individual characteristics?

    Demographic (age, gender, status)

    vs research (field, number and location of co-authors,...)
  - Or its department?

    Size, field, composition (Assistant prof vs full prof, women,...).
- Two concerns:
  - Spatial sorting,
  - Reverse causality.
- Same French administrative data set as before.

What makes an academic productive?

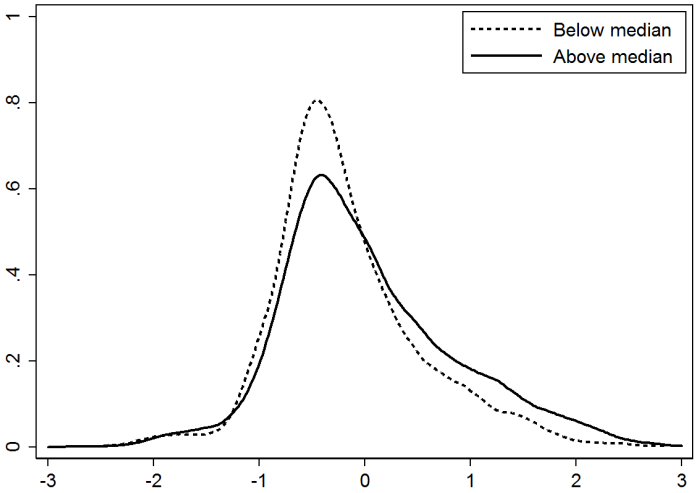
	Publishing		Quantity		Quality		Top quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Individual characteristics								
Women	-0.016[a]		-0.119[a]		-0.067[a]		-0.270[a]	
	(0.001)		(0.009)		(0.009)		(0.024)	
Age	-0.005[a]		-0.035[a]		-0.022[a]		-0.093[a]	
	(0.000)		(0.003)		(0.003)		(0.009)	
Age square	0.000[a]	-0.000[a]	0.000[a]	-0.000	0.000[b]	-0.000[a]	0.000[a]	-0.001[a]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rank A	0.044[a]		0.218[a]		0.136[a]		0.542[a]	
	(0.001)		(0.009)		(0.009)		(0.024)	
Authors per publication			-0.948[a]	-0.925[a]	0.186[a]	0.192[a]	0.508[a]	0.539[a]
			(0.011)	(0.015)	(0.010)	(0.013)	(0.027)	(0.034)
Individual diversity			-0.096[a]	-0.130[a]	0.013[c]	0.003	0.109[a]	0.024
			(0.007)	(0.009)	(0.007)	(0.008)	(0.018)	(0.020)
Non-USA connection			0.376[a]	0.193[a]	0.307[a]	0.084[a]	1.128[a]	0.342[a]
			(0.011)	(0.014)	(0.011)	(0.012)	(0.029)	(0.031)
USA connection			0.408[a]	0.223[a]	0.509[a]	0.209[a]	1.604[a]	0.611[a]
			(0.015)	(0.019)	(0.014)	(0.016)	(0.038)	(0.042)
Dep.-field characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects								
Field-time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department-time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position	Yes	No	Yes	No	Yes	No	Yes	No
Individual	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.07	0.13	0.33	0.54	0.37	0.65	0.46	0.72
Observations	758,790	424,044	38,836	38,836	38,836	38,836	38,836	38,836

What makes an academic productive?

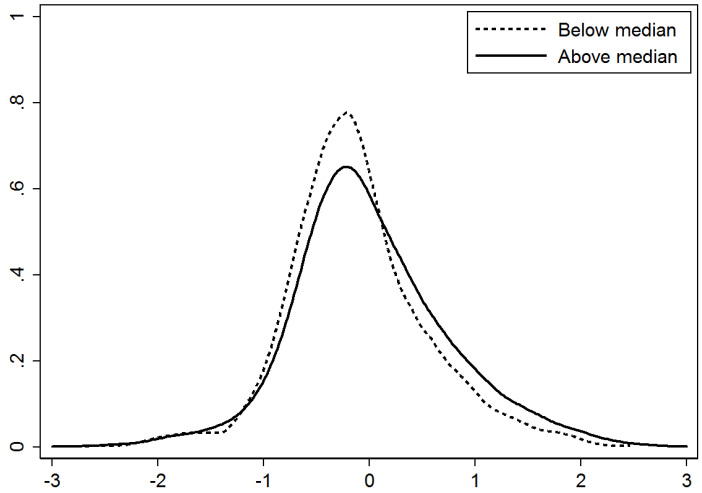
	Publishing		Quantity		Quality		Top quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Field presence	0.063^{a}	0.122^{a}	0.345^{a}	0.334^{a}	0.114^{a}	0.087^{a}	0.359^{a}	0.318^{a}
	(0.001)	(0.002)	(0.022)	(0.022)	(0.021)	(0.019)	(0.055)	(0.048)
Specialisation	0.014^{a}	0.024^{a}	0.098^{a}	0.088^{a}	0.036^{a}	0.020^{a}	0.132^{a}	0.084^{a}
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)	(0.003)	(0.009)	(0.008)
Size	0.003^{a}	−0.002	0.009	0.000	0.034^{a}	−0.013	0.055	−0.033
	(0.001)	(0.002)	(0.011)	(0.015)	(0.013)	(0.014)	(0.035)	(0.035)
% women	0.009	0.008	−0.041	0.157	−0.002	0.068	0.011	−0.110
	(0.007)	(0.014)	(0.083)	(0.120)	(0.097)	(0.109)	(0.264)	(0.276)
Average age	0.001^{a}	0.002^{a}	−0.002	0.002	−0.006^{c}	−0.005	−0.026^{a}	−0.025^{a}
	(0.000)	(0.001)	(0.003)	(0.004)	(0.003)	(0.004)	(0.009)	(0.009)
% rank A	−0.013^{b}	−0.037^{a}	0.101	−0.149	0.286^{a}	−0.022	1.061^{a}	−0.150
	(0.005)	(0.011)	(0.067)	(0.097)	(0.079)	(0.089)	(0.214)	(0.225)
Diversity	0.001	−0.011^{a}	−0.073^{a}	−0.059^{b}	0.043^{c}	−0.027	0.054	−0.043
	(0.002)	(0.003)	(0.020)	(0.027)	(0.023)	(0.024)	(0.064)	(0.062)
Research access	0.001^{b}	−0.003^{b}	0.025^{a}	−0.012	0.036^{a}	0.001	0.122^{a}	0.033
	(0.001)	(0.001)	(0.007)	(0.010)	(0.008)	(0.009)	(0.022)	(0.024)
Heterogeneity	−0.022^{a}	−0.021^{a}	0.000	−0.014	0.098^{a}	0.025	0.382^{a}	0.141^{c}
	(0.002)	(0.004)	(0.026)	(0.034)	(0.031)	(0.032)	(0.084)	(0.079)
USA connections	0.139^{a}	0.094^{a}	−0.270	0.048	1.052^{a}	0.251	3.080^{a}	0.700
	(0.017)	(0.025)	(0.181)	(0.219)	(0.220)	(0.205)	(0.597)	(0.515)
Non-USA connections	0.162^{a}	0.031	0.267^{c}	0.192	0.302^{c}	−0.289^{c}	1.324^{a}	−0.549
	(0.013)	(0.020)	(0.144)	(0.177)	(0.175)	(0.166)	(0.474)	(0.416)
Positions' shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE in 1st step	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.62	0.72	0.69	0.58	0.56	0.49	0.56	0.62
OLS within-time R^2	0.47	0.15	0.08	0.09	0.22	0.04	0.28	0.06
Observations	1208	1208	1208	1208	1208	1208	1208	1208

What makes an academic productive?

Figure: Distribution of the (detrended logarithm of) individual publication quality in departments above and below median field presence



Panel (a): Gross publication quality



Panel (b): Net of individual characteristics

- Field presence impact: +40% of publications in that field.
- Doubling the share of other colleagues' publication in the field: +6% of publications in that field.