

Access to the scientific literature in India

Patrick Gaulé^{1 2}

¹Chaire en Economie et Management de l'Innovation
Ecole Polytechnique Fédérale de Lausanne
patrick.gaule@epfl.ch

²Department of Economics
University of Geneva

June 15, 2009

Introduction

- Early studies (Lawrence 2001, Eysenbach 2006 and many others) had found that papers that were freely available received more citations, suggesting that open access increases the **diffusion** of scientific articles.
- However, **self-selection** of higher quality articles into open access introduces a spurious correlation between open access and citations (Gaulé & Maystre 2008).
- Davis and coauthors (2008) randomized articles into open access and found no open access citation advantage.

Introduction (2)

- In this paper, I search for an effect of access restrictions where it is most likely to be found: developing countries
- I do a citation analysis using backward citations rather than forward citations (as in previous studies)
- But let me show you some survey evidence first...

Survey of Indian biologists

- I contacted by email the corresponding authors of papers published by scientists based in India in 2007 in biology (2231 individuals)
- Small number of straightforward questions in the web survey
- 377 answers (response rate 16.8%), including a lot of comments

Informal file sharing practices

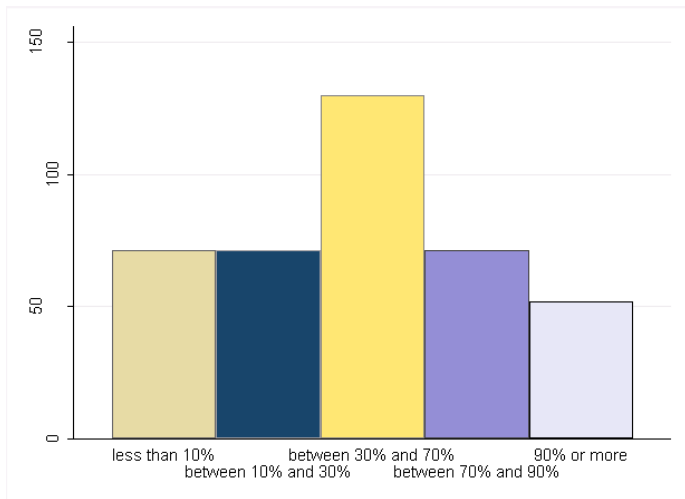
"In the last three months, have you... requested a copy of an article from an author?"

		No	Yes	
asked a friend to send you a copy of a paper that you could not access yourself?"	No	16%	25%	41%
	Yes	13.5%	45.5%	59%
		29.5%	70.5 %	100 %

Sample: 377 biologists based in India

Heterogeneity in access

Figure: Journal access as self-reported by biologists in India



An observation in my dataset

AU Goswami, BN ; Venugopal, V ; Sengupta, D; Madhusoodanan, MS;
Xavier, PK

TI Increasing trend of extreme rain events over India in a warming
environment

SO SCIENCE

AB Against a backdrop of rising global surface temperature, the stability
of the Indian monsoon rainfall over the past century has been a puzzle. ()

RP Goswami, BN, Indian Inst Trop Meteorol, Doctor Homi Bhabha Rd,
Pune 411008, Maharashtra, India.

EM goswami@tropmet.res.in

CR ALEXANDER LV, 2006, **J GEOPHYS RES-ATMOS**, V111, ARTN
D05109

ALLEN MR, 2002, **NATURE**, V419, P224, DOI 10.1038/nature01092

BRUNETTI M, 2001, **INT J CLIMATOL**, V21, P299

Hypotheses

I am studying the length and composition of the reference lists of authors based in Switzerland and India.

	Expected signs		
	Number references	Share expensive	Share OA
India	-	-	+
IIT	+	+	-
Collaboration_OECD	+	+	-

Controls: citing journal fixed effects (should control for both field narrowly defined and quality)

- publication data from ISI web of science;
 - ▶ all publications by Swiss and Indian authors in Science and Engineering in 2007 for a total of 43150 articles
 - ▶ I have the reference/backward citations for each article, i.e., more than 1.25 million citations
- journal data:
 - ▶ data from www.journalprices.com by McAfee and Bergstrom: information on price per article for around 4000 journals
 - ▶ journal coded as expensive if price per article exceeds the mean price in the journal database (USD 14).
 - ▶ manual coding of whether the journal is in open access (316 open access journals)

Results

	OLS (I)	OLS (II)	OLS (III)
	Number references	Share expensive	Share OA
India	-1.691a [0.215]	-0.0083a [0.00143]	0.0056a [0.0005]
IIT	0.790a [0.297]	0.0095a [0.002]	-0.0027a [0.0076]
Collaboration_OECD	0.736a [0.229]	0.00248 [0.00152]	-0.0005 [0.00058]
constant	29.958a [0.204]	0.113a [0.00136]	0.106a [0.0005]
Citing journal fixed effects	yes	yes	yes
Observations	43150	43021	43201
R-squared	0.05	0.0081	0.0049

Notes: Robust standard errors in brackets. c at 10%; b at 5%; a at 1%

Differences across fields

Table 4 - Results by field

	All fields (I)	Biology (II)	Medicine (III)	Physics (IV)	Engineering (V)	Chemistry (VI)
	Number references	Number references	Number references	Number references	Number references	Number references
India	-1.862a [0.217]	-3.27a [0.603]	-3.17a [0.473]	-0.673 [0.463]	-0.705 [0.49]	0.028 [0.552]
IIT	0.817a [0.298]	0.708 [1.107]	5.465b [2.622]	1.125 [0.697]	0.483 [0.453]	0.459 [0.618]
Collaboration_OECD	0.553 [0.385]	0.114 [0.62]	0.461 [0.485]	1.57 [0.518]	1.391c [0.54]	-0.26 [0.622]
# authors	0.101a [0.026]	0.087 [0.082]	0.2a [0.057]	-0.025 [0.043]	0.163 [0.069]	0.0053 [0.093]
constant	30.094a [0.203]	39.072a [0.586]	29.44a [0.435]	28.62a [0.433]	23.4a [0.487]	30.46a [0.628]
Citing journal fixed effects	yes	yes	yes	yes	yes	yes
Observations	43150	5898	8941	7536	6835	7518
R-squared	0.05	0.067	0.079	0.028	0.0179	0.0007

Notes: Robust standard errors in brackets.

c significant at 10%; b significant at 5%; a significant at 1%

Conclusion

- Evidence from citation data shows significant discrepancies in access to the literature between Swiss and Indian scientists
- In the long run, having all scientific publications freely available to the world from the day of publication of goals may be desirable.
- In the short run, however, it is more important to make scientific publications freely available for developing countries because this is where the problem really is.