TSE working paper series

10-160

Research Group: *Econometrics and Statistics*

May, 2010

What Does It Take to Become a Good Mathematician?

PIERRE DUBOIS, JEAN-CHARLES ROCHET AND JEAN-MARC SCHLENKER

Ecole d'économie de Toulouse (TSE) - Manufacture des Tabacs
Aile Jean-Jacques Laffont - 21, allée de Brienne - 31000 TOULOUSE
Tél : +33 (0) 5 61 12 85 89 - Fax : +33 (0) 5 61 12 86 37 - **www.TSE-fr.eu** - contact@TSE-fr.eu

Toulouse School of Economics

What Does It Take to Become a Good Mathematician?

By Pierre Dubois,[*] Jean-Charles Rochet[†] and Jean-Marc Schlenker[‡]

Using an exhaustive database on academic publications in mathematics, we study the patterns of productivity by world mathematicians over the period 1984-2006. We uncover some surprising facts, such as the absence of age related decline in productivity and the relative symmetry of international movements, rejecting the presumption of a massive "brain drain" towards the U.S. Looking at the U.S. academic market in mathematics, we analyze the determinants of success by top departments. In conformity with recent studies in other fields, we find that selection effects are much stronger than local interaction effects: the best departments are most successful in hiring the most promising mathematicians, but not necessarily at stimulating positive externalities among them. Finally we analyze the impact of career choices by mathematicians: mobility almost always pays, but early specialization does not.
JEL Classification: D85, I23, J24, L31
Keywords: faculty productivity; organization of research; peer effects in science.

[*] Toulouse School of Economics, 21 allée de Brienne, 31000 Toulouse, France. `pierre.dubois@tse-fr.eu`
[†] University of Zurich and Toulouse School of Economics. `rochet@isb.uzh.ch`
[‡] Institut de Mathématiques de Toulouse, UMR CNRS 5219, Université Toulouse III, 31062 Toulouse cedex 9, France. `jmschlenker@gmail.com`

I. Introduction

The American Mathematical Society maintains an almost exhaustive database (Mathematical Reviews) of publications in mathematical journals all over the world. Using this unique database, we study the academic output of 32574 active mathematicians over the period 1984-2006. In the first part of the article, we provide detailed descriptive statistics on academic production by mathematicians and uncover some surprising facts. For example:

- Contrarily to a widely held belief (among both scientists and lay people) the rate and quality of mathematical production does not decline rapidly with age. For mathematicians who remain scientifically active, productivity typically increases over the first 10 years, then remains almost constant until the end of their career. However there is a substantial attrition rate (i.e. mathematicians who stop publishing) at all ages.[1]
- There is a substantial variation over time of the geographical repartition of mathematical articles. For example although the U.S. are still by far the largest country in terms of mathematical production, their share has declined from 50% in 1984 to 34% in 2006. Similarly, the share of China is rapidly increasing but it is still surprisingly low (only 3.8% in 2006).
- International mobility is rather weak, and it is much more symmetric than could be expected, both in terms of numbers of mathematicians and in terms of "quality" measured by the output of the mathematicians who change countries.

In the second part of the article we perform a detailed statistical analysis of the factors that can influence the scientific production for academic mathematicians. This allows to analyze the determinants of individual productivity all along a mathematician's career taking into account unobserved "talent" of mathematicians through fixed effects. Among the important factors is of course location: the best mathematicians are (by definition) found in the best departments, but causality is not clear. Using the mobility of a sizable subset of these mathematicians (as in Kim et al. (2009)) we can separate the selection effects (hiring the most promising mathematicians) from interaction effects (stimulating positive spill-overs through exchange of expertise and feedback among colleagues).

In conformity with other recent studies (e.g. Waldinger (2009)), we find that university fixed effects (once researchers intrinsic quality is accounted for) are in general small, and are not strongly correlated with the quality of the department. A few departments have a strong positive impact on their members' productivity but in prominent examples this is largely associated with prestigious locally managed journals, which seem to publish relatively many articles from "locals".

We also analyze the impact of other characteristics of departments on the outputs of their members:

- Size does matter: large departments are good for individual productivity. However this effect is largely due to good hirings and becomes very small when authors fixed effects are incorporated.
- Having a specialized department has a negative impact on productivity when no fixed effect is used, but this impact becomes positive with fixed effects. This tends to indicate that a narrower scope lowers the quality of hiring, but that researchers fare better in a department with colleagues close to their mathematical interests.

[1]For comparable studies in other fields see Levin and Stephan (1991) and Stephan (2008).

- Looking at US universities, we find several interesting results. First, money does not seem to matter: even if the endowment per student has a strong positive impact when authors fixed effect are not used, it has a non-significant *negative* impact when these fixed effects are incorporated. This negative effect is actually significant when taking into account the fact that the university is public or private.
- Again for U.S. universities, the fact that a university is private has a small positive effect with respect to public ones. There is also a sizable positive effect of location on the east coast relative to the mid-west, the west coast standing in between the two.

Finally we analyze the impact of career decisions of individual researchers. We obtain some interesting results:

- Collaborations have a globally negative effect: the total output of mathematicians who have collaborating authors tends to be lower than the output of those who work separately. However collaborating with authors of a different specialty is positive: interdisciplinary work (within mathematics) spurs productivity.
- Mobility pays: each move increases future production.
- A high level of specialization is not a good strategy: it is correlated to a *lower* future output, in particular for young researchers. This suggests that researchers should be encouraged, especially at a younger age, to keep a broad range of interests.

II. Data description

A. The Mathematical Reviews database

The data come from the Mathematical Reviews database, which is maintained by the American Mathematical Society.[2] This database provides an almost exhaustive source of information on all publications in mathematics. It is remarkably well structured and has three features that make it particularly well suited to a statistical use:

- It provides a personal identification of each individual author, so that there is no ambiguity even when authors have the same name and initials.
- Each institution is identified by a unique institution code.
- Each article is assigned a principal code, as well as secondary codes using the "mathematical sciences classification (*M.S.C.*)". This gives a precise description of an article's main and secondary fields.

A small portion of the *Mathematical Reviews* database was used: since our focus was on "active" mathematicians, we selected the 98 journals with the highest impact factor (according to the 2006 *Journal Citation Report* in pure and applied mathematics), and compiled a list of all 129242 articles published in those journals between 1984 and 2006. Those 98 journals are the most visible in the fields of pure and applied mathematics, so that our data paints a reasonably accurate picture of the best part of mathematical research. We chose 1984 as the starting date because *Mathematical Review* only records the affiliation of authors from this date.

[2] We are grateful to *Mathematical Reviews* for allowing us to use their database in a non-standard fashion.

We then compiled a list of all 32574 mathematicians who published at least two articles in those 98 journals over this period. For these mathematicians, we compiled a list of the dates of the first and last publications of those authors in the whole database (not just the 98 journals in our list). We focused mainly, but not exclusively, on this smaller group of "active" mathematicians. The others, those with only one article in our list of 98 journals, can be of different types: mathematicians publishing few papers, mathematics PhD who have left the field, or academics of another field who have collaborated to a project published in a mathematics journal, etc.

B. Impact

For a correct assessment of the activity of the mathematicians in our list, we did not use the most standard impact factor data, which covers all scientific areas but is of limited relevance for mathematics Adler et al. (2009). Instead, the impact of each journal in the mathematical community was evaluated by its 2007 *Mathematical Citation Quotient* (*MCQ*), which is a kind of impact factor computed over five years. It is defined as the mean number of citations, within 5 years of publication, within the mathematical literature. This *MCQ* appears to be much more relevant for mathematical journals than the *impact factor* as computed by Thomson Scientific; the *impact factor* and the *MCQ* are highly correlated but the *MCQ* is less volatile. Moreover the impact factor tends to generate biases between subfields of mathematics.

However the *MCQ* is only an indication of the "quality" of a journal as evaluated by most mathematicians, rather than an absolute measure. In order to have a reasonable approximation of the importance of an article, we decided to assign to each article a weight equal to the product of its number of pages by the *square* of the *MCQ* of the journal where it was published. In this way:

- longer articles have a (linearly) higher weight,
- the weight of a page varies within our limited list of journals in a ratio of approximately 1 to 100.

This choice of weights puts a strong emphasis on a small number of very selective journals, as can be seen from the list of journals in the appendix. However, this is more in conformity with the most widely shared quality assessments within the community of mathematicians. In any case we have checked that choosing a different weight on articles does not change significantly the results of our study. There are probably some biases in the way different fields are treated (for instance journals in applied mathematics tend to have a lower *MCQ* than those in pure mathematics), but this is controlled for in our regressions by using the field of research as a control variable.

To achieve a better understanding of these differences between fields within mathematics, we used the *Mathematical Subject Classification* (*M.S.C.*) codes assigned to each article by *Mathematical Reviews*. Since this classification is quite detailed, we grouped different *M.S.C.* codes so as to obtain only 10 different areas. There are sizable differences between the properties of articles in different fields of mathematics, as seen in Table 1. We also indicate in this table the number of authors and of articles for each field, so as to give an idea of their relative numerical importance.

Finally, we compute yearly author impacts by summing these impact measures divided by the number of authors of all the articles they have published withing a year.

This way of attributing a "weight" to an article and then to the output of a researcher is to some extend arbitrary. Indeed there is a whole literature devoted to how one can measure scientific output. For instance, Hirsch Hirsch (2005) proposes to measure the

Field	Authors/paper	Pages/paper	Mean Impact	# Authors	# articles
Algebra	1.98	23.73	31.34	2463	8454
Analysis	2.07	19.70	16.46	8438	31194
DynSys	2.32	23.61	22.65	1420	4798
GeomAlg	1.84	23.36	36.43	4101	16720
GeomDiff	1.99	22.69	28.72	5814	23544
Numeric	2.30	19.75	10.17	4638	17420
Other	2.33	18.79	9.17	10265	33284
PDE	2.10	23.13	21.32	5898	25390
Physics	2.47	21.59	12.79	4455	12362
ProbaStat	2.13	18.37	8.16	10244	41721
Topo	1.87	25.13	39.09	2238	9065

TABLE 1—COMPARATIVE PROPERTIES OF ARTICLES IN DIFFERENT FIELDS

productivity of authors with the so-called h-index. Palacios-Huerta and Volij (2004) propose a new index based on citations but different from the traditional impact factor. Combes and Linnemer (2003) propose a ranking of journals in economics based on peer assessment of journal quality in addition to citations instead of a purely objective measure using citations counts. We believe that our way of weighting articles gives a result that is not too far from the heuristic assessment of many mathematicians, but we have not, at this point, tried to make this precise. It is probably *not* well adapted to other scientific areas.

III. Descriptive statistics

This section contains general data on the repartition of mathematical research (as measured by our indicators) in the world, on collaborations between regions and on mathematicians moving from one country to another. We also consider evolutions over time.

A. Countries and regions

THE WEIGHT OF DIFFERENT REGIONS OVER TIME

Rather than considering all countries separately, we considered only the countries having the biggest mathematical production, and grouped all others in a virtual region (OTH in the tables). The countries that are considered are Australia (AUS), Brazil (BR), Canada (CA), Switzerland (CH), France (FR), Germany (GER), Israel (IL), Italy (IT), Japan (JAP), the People's Republic of China (PRC), Russia (RUS, including the former U.S.S.R. before 1989), Spain (SP), Britain (UK), the United States (US).

Table 2 shows the proportion of world output coming from different countries, over time. Two striking features are the decrease in the share of the U.S. and the increase in the share of China, which however remains quite low.

COLLABORATIONS

Table 3 shows the evolution over time of the proportion of international collaborations. The evolution is particularly impressive for Russia. Before the fall of the Soviet Union in

	Country											
Years	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US
1984-1986	2.1	10.4	6.5	1.2	1.7	3.4	11.9	0.2	0.3	0.8	4.0	50.9
1987-1990	2.8	10.6	5.9	1.4	2.3	3.2	13.3	0.7	0.5	1.0	4.0	50.6
1991-1994	3.6	11.0	6.0	2.1	2.6	3.7	11.5	0.9	1.0	1.2	3.7	49.8
1995-1998	3.7	11.3	7.2	2.3	3.1	3.6	12.9	1.3	1.3	1.6	4.4	44.2
1999-2002	2.6	12.5	6.8	3.0	3.2	3.6	13.0	2.6	1.0	2.1	4.8	38.9
2003-2006	3.2	12.6	6.3	1.8	3.9	3.9	13.3	3.1	0.9	2.5	4.6	36.3

Table 2—Part of countries in the world production, over time

	Country												
Years	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US	
1984-86	0.62	0.51	0.47	0.57	0.39	0.29	0.40	0.65	0.22	0.33	0.48	0.25	0.35
1987-90	0.58	0.50	0.54	0.66	0.40	0.36	0.45	0.58	0.25	0.39	0.52	0.26	0.38
1991-94	0.60	0.50	0.54	0.62	0.48	0.31	0.46	0.51	0.55	0.36	0.53	0.28	0.40
1995-98	0.58	0.49	0.54	0.61	0.44	0.36	0.49	0.40	0.72	0.38	0.54	0.32	0.43
1999-02	0.67	0.47	0.57	0.63	0.46	0.41	0.50	0.39	0.76	0.42	0.56	0.37	0.46
2003-06	0.64	0.51	0.58	0.64	0.45	0.48	0.51	0.36	0.79	0.43	0.60	0.37	0.47
Total	0.62	0.49	0.56	0.63	0.45	0.39	0.48	0.39	0.66	0.41	0.55	0.32	0.43

Table 3—Proportion of international collaborations over time

1989, it was difficult for Russian mathematicians to collaborate with foreign colleagues. After 1989, a large proportion of the most active Russian mathematicians moved to other countries, which can account for the high proportion of collaborations between mathematicians with an affiliation in Russia and those in other countries.

The evolution of the proportion of collaborations in the U.S. is quite striking, too, since it increased markedly over the period.

International mobility

Table 4 shows the percentage of total number of years spent in different countries (when the affiliation is known) by mathematicians with a known first affiliation in a given country. The total in each row is the total number of years for which an affiliation is known for any of the mathematicians in our database with first affiliation in the corresponding country, and the total in each column is the total number of years spent in each country by mathematicians with a known affiliation.

The variable "first affiliation" is defined as follows. We know the affiliation of mathematicians for each year in which they have published a paper in one of the journals in our list. However we also have, for each mathematician in our list, the date of their first publication (in any journal, not necessarily those in our list). We define the "first affiliation" as the first affiliation that we know of, *if* it is within 3 years of the first publication. As a consequence we know the first affiliation of only 10803 of the 32437 "active" mathematicians in our database. In particular we do not know the "first affiliation" of any mathematician with first publication before 1981, since our list of articles starts in 1984. Our definition clearly creates some biases, for instance we know the first affiliation of very few mathematicians with an affiliation in Russia – this is explained by the fact that few of them published in one of the "western" and selective journals in our list early

Country of 1st affiliation	Country of current affiliation												
	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US	Total
CA	66.9	1.8	0.6	1.6	0.2		4.3	0.1		0.9	1.2	22.5	2226
FR	0.6	90.0	1.1	0.2	0.6	0.2	2.3			1.0	0.6	3.4	5403
GER	1.0	1.6	81.4		0.5	0.1	5.2	0.2		0.0	2.7	7.2	5042
IL	3.3	0.8	0.2	68.7			2.3		0.1		2.3	22.2	1204
IT	0.2	2.6	1.0		91.6		2.0			0.3	0.5	1.8	2192
JAP			1.0		0.1	92.7	2.2	0.9			0.2	2.8	2020
OTH	0.8	2.3	2.2	0.1	0.4	0.1	83.9	0.8		0.1	1.3	8.0	12341
PRC	2.6	1.1	0.8	0.1	0.2		3.8	80.5		0.1	0.6	10.2	1914
RUS	2.1	0.6	2.1		1.2		7.0		61.2		9.4	16.4	330
SP		0.7	0.3		0.0		0.9	0.2		96.6	0.4	0.8	2349
UK	2.2	1.7	2.3	0.1	1.1	0.2	11.5	1.5		0.2	68.8	10.3	3618
US	2.2	1.0	0.9	1.0	0.8	0.2	6.8	0.7	0.0	0.5	1.6	84.3	22882
Total	2455	5701	4848	1173	2384	1971	13326	1930	210	2508	3371	25921	65798

Table 4—Proportion of mathematicians in a given country, among those with their first affiliation in a given country. Totals are the number of person.year with known current country or country of first affiliation.

in their career, and those who did probably tended to emigrate.

In general, this "first affiliation" indicates the country where mathematicians have completed their PhD, or the country of their first or second year of post-doc. We only consider here mathematicians for whom we know this first affiliation.

A striking feature of Table 4 is the relatively small number of mathematicians who have changed countries: the vast majority remains in the country of their first affiliation. The only exception is the high number of moves between Canada and the U.S., in both directions.

Table 5 gives a hint on the attractivity of different regions. It shows the mean impact of mathematicians working in a given region, depending on the region where they had their first publication. Two phenomena occur here: in some cases, the first region corresponds to the place where a mathematician did his PhD (or post-doc) before coming back to his country of origin, while in others it corresponds to a "brain drain".

Table 6 has an interesting implication: in almost all regions, the mathematicians having their first affiliation outside their current region ("migrants") have a better mean impact than those with a local first affiliation ("locals"). Here again the explanation can vary between countries; in some cases it can be that the most active scientists tend do be those who went abroad to do their PhD before coming back, while some other countries actually drain the most active scientists. We have taken Japan out, since the number of mathematicians in Japan with first affiliation in another country is quite small (see Table 5 and Table 7).[3]

The proportion of mathematicians having their first affiliation in a different region is shown in Table 7.

Up to this point we have only considered the relation between the first affiliation of mathematicians and their affiliations when they publish new articles. We now concentrate

[3]More generally, the data in Table 5 and in Table 6 is more or less significant depending on the countries, as indicated in the numbers of mathematicians concerned, as seen in Table 4 and in Table 7.

Country of 1st affiliation	Country of current affiliation												
	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US	Total
CA	4.9	2.2	18.2	8.9	3.2		4.1	0.5		12.0	2.5	6.1	5.3
FR	5.1	12.4	7.8	50.9	14.6	13.9	10.8			19.7	26.5	11.3	12.5
GER	17.1	24.8	8.2		8.0	34.1	9.3	6.1		128.3	11.2	16.2	9.3
IL	11.5	38.4	10.8	8.9			3.3		9.5		41.0	19.5	12.2
IT	18.0	8.5	4.9		5.6		4.5			11.0	2.1	7.1	5.7
JAP			14.8		3.5	10.9	6.2	0.2			17.1	14.5	10.9
OTH	5.8	13.1	8.9	2.7	8.4	6.2	5.1	4.4		16.3	7.4	14.1	6.1
PRC	4.9	30.2	12.2	0.8	9.2		13.2	4.1		22.7	6.9	9.3	5.4
RUS	25.6	8.8	27.2		5.3		13.8		6.3		18.6	48.3	15.7
SP		12.1	1.6		0.8		0.6	31.5		4.7	0.9	9.0	4.7
UK	4.9	17.6	15.5	2.3	5.4	15.8	3.8	3.7		2.0	6.6	12.7	7.2
US	14.2	21.9	18.7	9.7	14.4	7.0	8.3	13.2	8.3	6.0	11.9	8.9	9.3
Total	7.7	13.1	8.9	9.3	6.6	10.9	5.7	5.1	6.4	5.3	8.1	9.4	8.5

Table 5—Mean impact of mathematicians depending on country of first affiliation and current country

Origin	Country												
	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US	Total
Locals	4.9	12.4	8.2	8.9	5.6	10.9	5.0	4.1	6.3	4.7	6.6	8.9	8.0
Migrants	11.9	17.3	13.1	10.4	11.8	10.1	6.7	8.9	8.5	11.3	12.2	13.3	10.4
Total	7.7	13.1	8.9	9.3	6.6	10.9	5.7	5.1	6.4	5.3	8.1	9.4	8.5

Table 6—Mean impact of mathematicians with first affiliation in/out the country of their current affiliation

on "real" moves between regions, defined as follows: mathematicians who have spent at least 3 years in a given region, then moved to another region and spent at least 3 years there. This excludes short moves for short-term post-doctoral positions or sabbaticals, but also changes of country immediately following the first publication.

Three striking facts emerge from those data.

1) The number of such "real" moves between countries is small, as can be seen from Table 8.

2) The numbers of "real" moves between most couples of countries are remarkably symmetric, as seen also in Table 8.

3) The "quality" of those moves, as measured by the mean impact of the mathematicians moving between countries, is also remarkably symmetric. Our data tends to indicate that the "brain-drain" phenomenon happens mostly for young researchers who move before or after their PhD or after a few years of post-doc.

Country	Proportion
CA	0.39
FR	0.15
GER	0.15
IL	0.29
IT	0.16
JAP	0.05
OTH	0.22
PRC	0.20
RUS	0.04
SP	0.10
UK	0.26
US	0.12

Table 7—Proportion of mathematicians who have migratedx

							To						
From	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US	Total
CA	0	9	5	5	3	1	38	10	1	2	16	140	230
FR	10	0	12	7	16	2	40	6	1	12	12	62	180
GER	13	37	0	2	10	13	115	5	3	1	48	114	361
IL	7	6	2	0	0	0	6	0	0	0	0	73	94
IT	3	10	5	0	0	0	19	2	0	1	6	26	72
JAP	0	2	12	0	1	0	7	5	0	0	1	18	46
OTH	35	82	100	3	18	9	0	53	8	8	65	315	696
PRC	12	5	4	0	5	2	34	0	0	0	12	43	117
RUS	2	12	14	6	5	0	28	0	0	2	17	49	135
SP	0	2	3	0	1	0	10	0	2	0	5	6	29
UK	12	18	13	1	7	1	67	14	2	3	0	91	229
US	114	65	90	88	40	23	319	42	7	17	85	0	890
Total	208	248	260	112	106	51	683	137	24	46	267	937	3,079

Table 8—Number of “real” moves from one country to another

B. Universities

General description

Table 10 shows the share of world output (measured by our indicator) for the top 30 departments by this indicator, and its variation over time. It shows notable changes in the ranking of departments, both upwards and downwards. Another feature is that the production of mathematical literature, even when measured by our quite elitist indicator, is not very concentrated. Indeed, the department ranked first produces only 1.8% of the world total output (weighted by impact) over the whole period. Moreover this concentration appears to decrease over time, since the share of the most active department was 2.25% in 1989-1994, but only 1.7% in 2001-2006.

Table 11 shows the size (yearly average number of active mathematicians) and the share, in total output, of the top author, and then the top 5 and top 10 authors (again weighted by impact). The share of the most productive author typically varies between

						To							
From	CA	FR	GER	IL	IT	JAP	OTH	PRC	RUS	SP	UK	US	Total
CA		9.6	6.5	2.6	12.4	21.3	7.4	4.6	4.2	3.2	6.3	9.1	8.3
FR	9.3		20.8	22.4	7.9	8.4	10.1	12.5	1.5	28.5	16.1	25.5	18.0
GER	23.3	15.4		13.0	9.2	17.5	11.1	5.3	3.7	4.8	12.0	16.7	13.9
IL	11.6	5.6	7.8				3.6					19.0	16.3
IT	4.7	9.9	13.1				5.6	11.6		1.7	10.6	14.5	10.4
JAP		2.4	10.6		16.6		7.8	9.3			8.3	26.8	16.1
OTH	6.9	14.0	8.9	7.1	4.6	8.0		9.5	2.2	7.6	8.6	12.8	10.9
PRC	4.2	21.2	9.6		8.4	0.6	8.7				3.2	9.3	8.3
RUS	0.6	14.3	10.0	3.2	3.3		5.1			21.3	9.1	19.1	12.0
SP		7.9	14.2		2.5		9.3		21.3		15.0	4.1	10.2
UK	9.0	12.7	15.6	40.9	12.2	21.3	5.9	11.1	1.7	7.4		15.9	11.8
US	14.2	26.9	18.8	15.7	17.4	9.4	10.7	16.1	9.7	11.3	15.0		14.6
Total	12.1	17.0	13.4	14.8	11.3	11.3	9.5	11.3	6.2	14.6	11.4	14.8	12.8

Table 9—Mean impact over lifetime of mathematicians moving from one country to another

5% and 15%, while the share of the top 10 authors varies between 23% and 70%, depending on the size of the department.

Table 12 shows the top 30 departments by total output, ranked now in terms of the average output of their researchers (among departments with a total output of at least 5000 weighted pages over the period, so as to eliminate very small departments). There is a large difference between the two rankings, since small departments with highly productive researchers (like the I.H.E.S. or the I.A.S.) have excellent rankings in terms of average output but not in terms of total output.

Where are the most active mathematicians?

Table 13 shows the ranking of the 30 top departments by total output, along with their size (mean number of active authors) and average share (%) over the period of the top 100 authors by total output, and the share of those among the top 500 by total output. The table shows relatively little concentration of the mathematicians with the highest output in the top departments, indicating again the relatively high number of departments taking part in top-level research. The last column shows the share of young mathematicians (those at most 4 years after their first publication), it gives an indication of the concentration of future active mathematicians (those which end up in our list) in the departments.

University	Rank	Share of impact	1984-88	1989-94	1995-2000	2001-06
Princeton	1	1.80%	2.15	1.96	1.61	1.70
Paris 11 (F)	2	1.73%	1.73	2.25	1.56	1.50
MIT	3	1.58%	1.94	1.94	1.44	1.30
NYU	4	1.44%	1.63	2.05	1.49	.89
Berkeley	5	1.39%	1.80	1.44	1.34	1.21
Harvard	6	1.27%	1.94	1.18	1.38	.94
Paris 6 (F)	7	1.27%	1.27	1.26	1.42	1.14
Chicago	8	1.09%	1.23	1.06	1.07	1.08
UCLA	9	.97%	1.26	1.30	.75	.79
Stanford	10	.93%	.99	1.16	.83	.84
Michigan	11	.93%	.64	.90	1.18	.85
Rutgers	12	.92%	1.09	1.08	1.04	.63
Purdue	13	.91%	1.53	1.08	.75	.64
Minnesota	14	.86%	.87	1.24	.68	.75
Maryland	15	.85%	1.26	1.06	.84	.53
IAS Princeton	16	.79%	.89	.82	.89	.63
Toronto	17	.77%	.46	.88	.82	.79
Ohio State	18	.75%	.75	.80	.82	.63
Columbia	19	.72%	1.15	.83	.62	.54
Wisconsin	20	.71%	.91	.68	.56	.77
Cornell	21	.7%	.76	.81	.78	.53
Oxford (UK)	22	.67%	.71	.6	.75	.63
Paris 7 (F)	23	.65%	1.06	.65	.69	.43
Caltech	24	.65%	.93	.45	.58	.72
SUNY Stony Brook	25	.64%	.79	1.08	.56	.33
Polytechnique (F)	26	.63%	.88	.82	.56	.45
UC San Diego	27	.62%	.84	.75	.50	.53
Hebrew U (IL)	28	.57%	.45	.41	.86	.50
Cambridge (UK)	29	.57%	.46	.50	.80	.45
Illinois at Urbana	30	.55%	.70	.50	.43	.63

Table 10—Part in the world output over time, top 30 departments

University	Rank	Size	Share 1st Author	Share 5 Authors	Share 10 authors
Princeton	1	96.9	5.9	19.5	30.3
Paris 11 (F)	2	105.5	11.8	25.1	36.2
MIT	3	123.2	11	24.1	34.2
NYU	4	102.5	6.2	19.8	29.1
Berkeley	5	146	4.1	14.6	23.4
Harvard	6	64.9	10	29.9	41.4
Paris 6 (F)	7	143.1	13.5	22	28.5
Chicago	8	81.1	7	21.4	31.8
UCLA	9	93.8	7.6	26.9	36.2
Stanford	10	106.7	6.1	21	34.3
Michigan	11	111.8	4.6	15.7	24.9
Rutgers	12	104.2	6	22.2	33.2
Purdue	13	95.5	5.9	22.6	35.9
Minnesota	14	116.5	5.4	18.3	30.3
Maryland	15	95.6	8.3	28.5	42.4
IAS Princeton	16	33.2	12	30.9	41
Toronto	17	78.7	14.5	29.8	42.2
Ohio State	18	86	7	19.5	33.1
Columbia	19	57.2	8.9	26.9	42.4
Wisconsin	20	101.3	5.9	20.4	33.5
Cornell	21	106.5	5.6	22.9	35.2
Oxford (UK)	22	80.6	8.5	28.9	40.5
Paris 7 (F)	23	49.5	6.4	23.8	38.2
Caltech	24	54.2	10	31.5	44.8
SUNY Stony Brook	25	46.1	12.2	42.6	61.3
Polytechnique (F)	26	56	7.6	19	31.3
UC San Diego	27	69.9	5.2	19.7	33.9
Hebrew U (IL)	28	64.3	10.2	35.9	51
Cambridge (UK)	29	79.8	6.2	20.7	31.8
Illinois at Urbana	30	99.2	4.4	17.4	28.3

Table 11—Size and share of main author in departments

University	Rank	Total Impact	Rank by mean	Mean Impact
Princeton	1	42522.8	4	438.7
Paris 11 (F)	2	40800.6	6	386.9
MIT	3	37408.2	13	303.6
NYU	4	34136.4	9	333.2
Berkeley	5	32770.7	30	224.4
Harvard	6	30002.3	3	462.4
Paris 6 (F)	7	29931.2	41	209.2
Chicago	8	25875.8	11	318.9
UCLA	9	22835.6	24	243.5
Stanford	10	22074	43	206.9
Michigan	11	21931.9	50	196.3
Rutgers	12	21724.7	42	208.6
Purdue	13	21477.1	29	224.9
Minnesota	14	20387.4	58	175
Maryland	15	20122.6	38	210.5
IAS Princeton	16	18592	2	559.4
Toronto	17	18157.4	27	230.7
Ohio State	18	17625.6	44	205
Columbia	19	16987.6	14	297.1
Wisconsin	20	16755.8	64	165.4
Cornell	21	16572.9	73	155.6
Oxford (UK)	22	15858.4	49	196.7
Paris 7 (F)	23	15360.3	12	310.5
Caltech	24	15294.6	17	282.3
SUNY Stony Brook	25	15116.8	10	328.1
Polytechnique (F)	26	14976.1	20	267.6
UC San Diego	27	14634.9	40	209.2
Hebrew U (IL)	28	13586.4	37	211.4
Cambridge (UK)	29	13372.4	63	167.6
Illinois at Urbana	30	13042.8	83	131.4

Table 12—Rank for total output vs rank for mean output

University	Rank	Size	Share top 100	Share top 500	Share young
Princeton	1	96.9	8.73	3.82	.76
Paris 11 (F)	2	105.5	4.23	3.21	.64
MIT	3	123.2	3.04	2.59	.96
NYU	4	102.5	4.83	2.34	.61
Berkeley	5	146	2.11	2.56	.9
Harvard	6	64.9	5.04	2.85	.55
Paris 6 (F)	7	143.1	2.01	1.51	.77
Chicago	8	81.1	2.22	2.43	.62
UCLA	9	93.8	3.36	1.24	.47
Stanford	10	106.7	2.11	1.9	.7
Michigan	11	111.8	.11	1.74	.69
Rutgers	12	104.2	2.01	1.74	.39
Purdue	13	95.5	2.44	1.59	.45
Minnesota	14	116.5	.22	.92	.73
Maryland	15	95.6	2.49	1.65	.37
IAS Princeton	16	33.2	2.39	1.28	.22
Toronto	17	78.7	1.19	1.4	.44
Ohio State	18	86	1.36	.86	.43
Columbia	19	57.2	2.49	1.42	.38
Wisconsin	20	101.3	.11	1.17	.5
Cornell	21	106.5	0	1.58	.53
Oxford (UK)	22	80.6	2.11	1.2	.56
Paris 7 (F)	23	49.5	1.3	1.21	.22
Caltech	24	54.2	3.52	1.35	.46
SUNY Stony Brook	25	46.1	4.45	1.73	.27
Polytechnique (F)	26	56	.98	.95	.36
UC San Diego	27	69.9	.22	1.12	.28
Hebrew U (IL)	28	64.3	1.95	.97	.29
Cambridge (UK)	29	79.8	.22	.36	.56
Illinois at Urbana	30	99.2	0	.09	.55

Table 13—Share of total number of years spent by very active (resp. active, resp. young) mathematicians (mean over period of study)

IV. Analysis of the scientific output of mathematicians

We now consider a more elaborate statistical analysis of the factors of scientific "productivity" for mathematicians. First we consider the effect of departments on researchers, then the effect of specific characteristics of departments, then more specific data for U.S. departments only to analyze the production of mathematicians.

For such analysis, we postulate that the output measure of mathematician i during period t, denoted y_{it} , follows a linear model :

$$y_{it} = \alpha_i + \theta_{u(i,t)} + \gamma_{f(i,t)} + \delta_t + \beta X_{it} + \varepsilon_{it} \tag{1}$$

where $u(i,t)$ is the university of author i at year t, $f(i,t)$ is the discipline of research of author i at year t, δ_t is a period effect, X_{it} are time varying characteristics of author i (mostly age, age squared and age cubed) and α_i is a fixed effect for author i capturing the effect of all unobserved characteristics of the author (fixed over time) that affect his or her productivity. Assuming that ε_{it} is mean independent of α_i, $\theta_{u(i,t)}$, $\gamma_{f(i,t)}$, δ_t, X_{it}, we can identify all parameters using Ordinary Least Squares (OLS).

Thus, $\theta_{u(i,t)}$ can be interpreted as the effect of the university or department on the output of individual i. This effect can be identified because mathematicians move from departments to departments and thus $u(i,t)$ is not fixed over time. Therefore θ_u identifies the average effect of university u on mathematicians who have been affiliated to that university in year t: by definition of the indicator $u(i,t)$, they are such that $u(i,t) = u$. Similarly $\gamma_{f(i,t)}$ can be identified because not all authors publish always in the same field and thus variations at the individual level of fields of publication allow the identification of "field effects" in addition to individual effects. One alternative model consists in assuming that there is no unobserved heterogeneity across authors and thus that $\alpha_i = \alpha$ for all i or that the deviations form the mean α_i are mean independent of all other right hand side variables of the previous equation.

Note that estimating the effect of X_{it} not taking into account the fixed effect gives a coefficient that is smaller (larger) than β (the estimator taking into account the fixed effect) if and only if X_{it} is negatively correlated with the fixed effect α_i.

A. Effects of departments on individuals

An interesting question is the impact of location: how important departments are for the scientific productivity of researchers? In order to study this question, we use the data on the yearly production of authors and regress it on department dummies. The coefficients of these dummies reflect the average output of researchers belonging to those departments over the 1984-2006 period. Then, as authors move across departments, we can introduce author fixed effects in this linear regression to separate the effect of the department itself from the average quality of the mathematicians composing this department.

Table 14 shows the effect of the 30 main mathematics departments (in terms of total output on the period) on their researchers. The first column uses no author fixed effect, and computes the mean production of researchers at a given department. It just measures the "quality" of the departement. By contrast, columns 2-4 include a fixed effect for authors, and the results given correspond to the department fixed effect, namely the impact of a department on its members' productivity. Column 2 is for the whole sample period. To study any time variation we have split the sample into two periods of similar length: column 3 corresponds to the first part (1984-1994) and column 4 to the second part (1995-2006).

Note that this analysis of the effect of departments on individuals, with fixed effects of the authors, is possible because there are relatively many moves between departments. On average, each "active" mathematician in our base had 1.87 different affiliations over his lifetime, although this varies largely between countries.

Table 14 suggests that some departments have a strong positive impact on their researchers. For instance being at N.Y.U. has a very strong positive effect (20.51). However this effect is much smaller when the journal published by NYU (CPAM) is taken out of the publications sample. The same is true for Princeton. When the "local" journal, namely *Annals of Mathematics*, is taken out of the publications sample: the coefficient for the fixed effect of Princeton falls from 12.80 to 6.54. The effect of local journals is also significative, albeit with a smaller magnitude, for Paris 11 (its fixed effect falls from 16.48 to 13.40 when its local journal, *Publications Mathématiques de l'I.H.E.S.* is not considered) and Ecole Normale Supérieure of Paris (11.39 with *Annales Scientifiques de l'E.N.S.* and only 10.24 without). This does not necessarily mean that referees and editors are more friendly to local authors, but simply that these authors are encouraged to publish in the local journal.

Another feature of Table 14 is that, generally speaking, the fixed effect of departments does not seem to be clearly decreasing between the first and the second part of our sample period. This contrasts with the findings of Kim et al. (2009) for economists (see also Agrawal and Goldfarb (2008)).

VARIABLES	(1) Mean impact	(2) Dept fixed effect	(3) Dept fixed effect	(4) Dept fixed effect
Princeton	30.95***	12.80***	2.724	21.32***
	(0.954)	(1.381)	(2.861)	(1.914)
Paris 11 (F)	25.47***	16.48***	16.46***	19.04***
	(0.925)	(1.481)	(3.396)	(1.993)
MIT	17.10***	7.899***	5.029**	5.276***
	(0.818)	(1.298)	(2.536)	(1.790)
NYU	24.72***	20.51***	20.98***	19.92***
	(0.939)	(1.487)	(2.851)	(2.235)
Berkeley	11.98***	4.742***	0.227	7.978***
	(0.779)	(1.276)	(2.657)	(1.742)
Paris 6 (F)	12.52***	7.624***	3.610	11.40***
	(0.824)	(1.372)	(3.367)	(1.835)
Chicago	19.08***	11.17***	8.770***	13.09***
	(1.031)	(1.550)	(3.003)	(2.225)
UCLA	15.86***	10.99***	8.226***	11.14***
	(0.995)	(1.616)	(3.182)	(2.236)
Stanford	11.67***	4.853***	3.970	4.597**
	(0.899)	(1.568)	(3.139)	(2.170)
Michigan	7.975***	1.174	-9.639***	2.561
	(0.895)	(1.425)	(3.164)	(1.905)
Rutgers	11.33***	5.665***	-0.728	5.397**
	(0.953)	(1.703)	(3.400)	(2.376)
Maryland	10.11***	7.535***	4.953	10.83***
	(1.002)	(1.818)	(3.518)	(2.604)
IAS Princeton	33.72***	12.20***	-3.069	22.96***
	(1.607)	(1.932)	(3.411)	(2.734)
North Carolina	6.959***	2.702*	3.458	3.156
	(0.928)	(1.564)	(3.243)	(2.170)
Oxford (UK)	7.068***	6.245***	6.990*	6.308***
	(1.043)	(1.773)	(4.046)	(2.322)
SUNY Stony Brook	21.73***	6.541***	-1.334	6.868*
	(1.449)	(2.267)	(4.155)	(3.547)
Polytechnique (F)	14.16***	4.530**	-3.056	5.525**
	(1.273)	(1.919)	(4.557)	(2.527)
Hebrew U (IL)	13.53***	6.135***	6.846	3.055
	(1.256)	(2.299)	(4.826)	(3.127)
Cambridge (UK)	5.791***	5.770***	4.053	6.636***
	(1.036)	(1.595)	(3.581)	(2.144)
Illinois at Urbana	2.410**	5.170***	3.580	5.871***
	(0.951)	(1.646)	(3.943)	(2.162)
Toulouse 3 (F)	12.88***	15.63***	48.52***	11.69***
	(1.296)	(2.161)	(7.448)	(2.510)
ENS Paris (F)	18.69***	11.39***	-1.092	13.25***
	(1.467)	(2.086)	(4.936)	(2.582)
ETH (CH)	8.578***	4.428**	8.329	3.251
	(1.242)	(2.000)	(5.180)	(2.407)
Tel Aviv (IL)	3.898***	2.040	2.676	2.552
	(1.045)	(1.944)	(4.436)	(2.549)
U Bonn (G)	9.287***	7.590***	-2.787	8.468***
	(1.273)	(1.963)	(4.252)	(2.701)
Constant	5.246***	7.673***	6.595**	16.46***
	(1.115)	(1.394)	(2.598)	(2.103)
Observations	208683	208683	71460	137223
R^2	0.091	0.387	0.515	0.441
Fixed effect	No	Yes	Yes	Yes
Period		1984 – 2006	<1995	>1994

Table 14—Fixed effects of major departments

VARIABLES	(1) Dept fixed effect	(2) Without CPAM	(3) Without Annals	(4) Without Publ. IHES	(5) Without Ann. ENS
Princeton	**12.80*****	11.75***	**6.540*****	13.42***	12.81***
	(1.381)	(1.361)	(1.312)	(1.276)	(1.382)
Paris 11 (F)	**16.48*****	16.39***	15.82***	**13.40*****	15.72***
	(1.481)	(1.460)	(1.407)	(1.368)	(1.482)
NYU	**20.51*****	**2.623***	18.75***	20.72***	20.63***
	(1.487)	(1.466)	(1.413)	(1.374)	(1.488)
IAS Princeton	**12.20*****	11.44***	**9.591*****	7.553***	12.39***
	(1.932)	(1.904)	(1.836)	(1.785)	(1.933)
Polytechnique (F)	**4.530****	3.778**	4.699***	7.067***	**2.772**
	(1.919)	(1.891)	(1.823)	(1.773)	(1.919)
ENS Paris (F)	**11.39*****	11.97***	11.31***	10.32***	**10.24*****
	(2.086)	(2.056)	(1.982)	(1.928)	(2.087)
Constant	7.673***	6.716***	7.401***	7.696***	7.567***
	(1.394)	(1.374)	(1.324)	(1.288)	(1.395)
Observations	208683	208683	208683	208683	208683
R^2	0.387	0.382	0.368	0.385	0.381
Removed	None	CPAM	Annals	IHES	E.N.S.

Table 15—Fixed effect of a selection of major departments, with some journals removed

B. Characteristics of departments

We now consider more specifically the effect of different variables, pertaining either to individual researchers or to departments. Using again simple linear regressions of yearly data, with or without fixed effects, we regress the mathematicians' outputs on some individual and university characteristics.

	(1)	(2)	(3)	(4)
VARIABLES	impact	impact	nb articles	nb articles
Univ. specialization index	4.590***	-4.252***	0.0419***	-0.110***
	(0.710)	(1.244)	(0.0157)	(0.0275)
Size	0.131***	0.0401***	0.00260***	0.00201***
	(0.00288)	(0.00504)	(6.38e-05)	(0.000111)
Stability	-11.55***	1.816	0.0148	0.0244
	(0.981)	(1.806)	(0.0217)	(0.0399)
Constant	9.727***	18.33***	0.950***	1.200***
	(1.695)	(2.047)	(0.0376)	(0.0453)
Observations	138707	138707	138707	138707
R^2	0.095	0.011	0.028	0.008
Fixed effect	No	Yes	No	Yes

TABLE 16—IMPACT, EFFECT OF VARIOUS VARIABLES, WITHOUT/WITH FIXED EFFECTS

Table 16 shows the influence of different variables on departments' fixed effects. The *University specialization index* is defined as 1 minus the sum of the squares of the proportion of the scientific output of the department in each field. It varies between 0 and 1, and is small for highly specialized departments and close to 1 for general ones. It is defined based on the subfield (*M.S.C.*, for Mathematics Subject Classification, as determined by *Mathematical Reviews*). *Closedness* measures how "open" departments are, it is the mean over its members of the proportion of their "scientific life" they will spend in this department. Other individual variables (age, year, subfield, number of co-authors) are used in the regression but not represented in the table, they are considered in Section IV.D below.

The first two columns show the effect of the variables on the "impact" of authors, as defined in Section II.B, without taking into account the fixed effects of authors in the first column, with authors fixed effects in the second column. Columns 3 and 4 are similar but measure productivity by the number of pages instead of the impact.

The coefficient of the "Stability" variable is strongly negative without fixed effects: open departments attract better researchers. However, this variable is not significant with fixed effects. The coefficient of the department size is strongly positive without fixed effects: bigger departments attract better researchers. This effect remains but is much weaker when authors fixed effects are incorporated.

The coefficient of the *University specialization index* is strongly positive without author fixed effects: departments with a wider scope tend to have better researchers, which is understandable since they can hire in a larger pool. However it is strongly negative with fixed effects: specialized departments stimulate the productivity of their researchers.

C. U.S. departments

Table 17 shows the influence of some variables that are specific to the U.S.: private vs public universities, East Coast or West Coast, endowment per student.

The effect of the endowment per student is remarkable. It is strongly positive without authors fixed effects, meaning that rich universities can attract better researchers. However it is negative (but not significantly) when authors fixed effects are taken into account. This is rather counter-intuitive since a higher endowment could imply lower teaching loads and therefore better research. Several explanations could be offered, for instance the idea that once researchers have obtained a position in a well-endowed university, they have weaker incentives to publish first-rate articles. It is also possible that the phenomenon appearing here does not extend to experimental sciences, where funding plays a bigger role than in mathematics.

It is also interesting to note that the East Coast has a significant positive effect over the Midwest (2 standard deviations). The West Coast stands in between.

Finally the effect of public universities is slightly negative but not significant. This could be attributed to higher teaching loads than in private universities. The difference becomes significant when taking into account the endowment per student, which is concentrated almost only in private universities.

VARIABLES	(1) Author Impact	(2) Author Impact
Nb. of coauthors	-7.116***	-5.350***
	(0.272)	(0.341)
Age	1.723***	0.741***
	(0.115)	(0.182)
Age^2	-0.0678***	-0.0310***
	(0.00584)	(0.00818)
Age^3	0.000685***	0.000245**
	(8.17e-05)	(0.000113)
Univ. Specialization Index	-7.292	-5.955
	(4.845)	(9.510)
Size of University	0.0845***	0.0605***
	(0.00658)	(0.0127)
Closedness	-58.20***	-12.79
	(4.833)	(9.282)
Private University	3.959***	2.146*
	(0.517)	(1.140)
Endowment per Student	3.688***	-2.618**
	(0.622)	(1.152)
East Coast University	3.163***	3.291***
	(0.503)	(1.196)
West Coast University	1.762***	1.677
	(0.631)	(1.462)
Constant	37.66***	30.85***
	(5.809)	(10.04)
Observations	37320	37320
R^2	0.126	0.428
Fixed effect	No	Yes
Year Effects	Yes	Yes
Discipline Effects	Yes	

Table 17—The determinants of mathematicians' scientific output, without/with fixed effects, US only

D. Individuals

Table 18 shows the effects on mathematicians' scientific productivity of their current number of co-authors, and their current number of co-authors from different fields. It also shows the impact of authors' personal histories: total number of co-authors, number of subfields (or mathematical subject classifications) in which they have published, and the number of institutions where they have held a position. More precisely, we have defined the number of co-authors (at a given time period) as simply the number of total coauthors who co-signed a publication with the author (co-authorship with the same coauthor for several publications in a year are counted). The number of co-authors having a different specialty (measured by the *M.S.C.* of articles) is as before the number of coauthors but whose main specialty measured over the life-cycle is different of the main specialty (also measure over the life cycle) of the author. The number of past moves between two departments is the number of times that the author has moved of affiliation since is first affiliations (where changes of affiliation are identified only when the author publishes). The number of past *M.S.C.* codes of each author is the number of *M.S.C.* codes of his articles in his past publications records and the number of past co-authors is measured simply with past articles published.

A remarkable fact is that the current number of co-authors has a negative impact[4], with or without fixed effects. This suggests that generally speaking, collaboration does not spur productivity: the output of a group of researchers, measured in terms of weighted pages published, is lower than it would have been if each of them had worked separately.

The number of past co-authors, however, has a positive impact. Moreover, the number of co-authors from different fields also has a positive impact. One interpretation is that collaboration with colleagues with a closely related competence is detrimental to the total output as considered here (shared between co-authors), but collaborating with mathematicians with a different main field is useful.

Having published in the past different subfields and in different sub-subfields also has a positive impact. With fixed effects, the number of past subfields seems dominant. The number of past institutions is also clearly positive, which is consistent with our previous observations on the number of past subfields and of past co-authors: having been exposed to a wider spectrum of mathematical ideas has a positive effect on mathematical output.

[4]Recall that in the individual's output, the impact of each paper is shared between the authors.

VARIABLES	(1) Output	(2) Output	(3) Output	(4) Output
Nb co-authors	-3.571***	-3.409***	-3.410***	-3.412***
	(0.104)	(0.159)	(0.159)	(0.158)
Co-authors from other fields	1.183***	2.076***	2.096***	2.108***
	(0.166)	(0.236)	(0.236)	(0.234)
Past moves	2.115***	0.998***	1.048***	0.994***
	(0.0969)	(0.219)	(0.219)	(0.218)
Past fields	1.137***	-0.0551	0.934***	
	(0.152)	(0.317)	(0.212)	
Past MSC	1.100***	0.929***		0.897***
	(0.108)	(0.221)		(0.147)
Past co-authors	0.420***	0.235***	0.297***	0.239***
	(0.0282)	(0.0573)	(0.0553)	(0.0570)
Constant	4.806***	12.54***	11.96***	13.80***
	(1.464)	(2.716)	(2.713)	(1.763)
Observations	108447	108447	108447	108929
R^2	0.100	0.475	0.475	0.475
Fixed Effect	No	Yes	Yes	Yes

TABLE 18—EFFECT OF INDIVIDUAL VARIABLES ON MATHEMATICIAN'S OUTPUT

E. Variations between countries

There are many cross country differences in the organization of universities and academic research. Among the key differences are the age at which long-term or permanent positions can be obtained, the grant systems, and more generally the nature of the incentives given to scientists, as well as the degree of mobility between institutions, the amount of teaching, etc. Rather than considering each of these differences separately, we only look at their impact on the profile of scientific productivity of mathematicians as a function of their age.

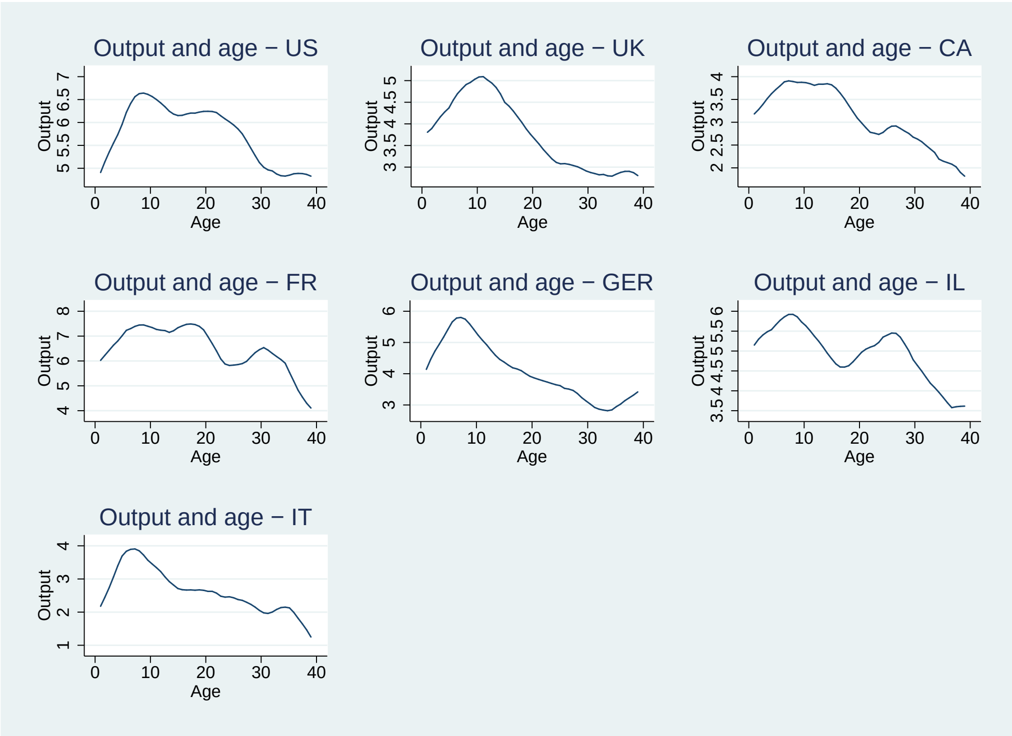

FIGURE 1. VARIATION OF PRODUCTIVITY WITH AGE, SELECTED COUNTRIES

Figure 1 shows how "age" – the number of years after the first publication – influences the production, depending on the country. Obviously, there are important differences in the mean scientific output of researchers between countries. However the evolution of this output also varies strongly between countries. The graphs indicate that some countries are better than other at helping their researchers to remain active. Among striking facts is the difference between the U.S. and Canada, where the output of mathematicians tends to decrease much faster.

V. Conclusion

The analysis presented here might have some interesting consequences for individual researchers, departments, or in terms of scientific policy.

A. How to detect a promising mathematician?

There are some indicators of future success for a young mathematician, among which, obviously, the quality of his publications. However our results tend to indicate some less obvious criteria, which can be measured over the first few years of activity, and are correlated with a higher future scientific output.

Among those, we can note a wider spectrum of interests. The capability to collaborate with colleagues with different mathematical interests could also be a positive indication. In other terms, a strong focus on one area, which is sometimes presented as a way for young scientists to gain a head start, could be counter-productive in the longer term.

The number of collaborations, on the other hand, has more complex implications. Having a large number of past co-authors appears positively correlated to the output, but a large number of current co-authors appears to have a negative impact.

B. How to make a maths department better?

A higher level of mobility appears to be a way to improve both the quality of a department and the scientific output of its members. On the other hand, encouraging members of a department to collaborate more does not appear to be efficient, except if the collaboration is with colleagues from different areas. This suggests that reading groups or seminars bringing together mathematicians with different specialties could be a way to broaden their interests and to improve their output.

Concentration on some subfields has mixed effects: it appears to lower the output of the department through the hiring of less productive mathematicians, but allows to get better papers from mathematicians of given talent.

C. How to improve mathematical research on a large scale?

Here again a high level of mobility seems to have positive effects. By contrast, allocating large subsidies to some departments appears to be useless: it may attract the more active researchers to the richer departments, but does not increase their output when taking in account authors fixed effects.

An important question, for which we do not have a definite answer, is how important it is to train young researchers in the more active departments. One problem here is that it is difficult to distinguish the quality of the training from the intrinsic "talent" of mathematicians.

D. Beyond mathematics

An extension of our results to other scientific areas than mathematics would probably be hazardous. There are many differences between sciences: for instance the importance of funding is fundamentally different between experimental fields and the more theoretical ones. We believe that it would be interesting to check to what extent our findings for mathematics are also valid in other disciplines.

More on the data

Tables A.1 and A.2 contains the list of journals used here. With each journal we list the total number of pages published in the sample period (variable "tpages"), the number of articles (variable "nart"), the 2007 *M.C.Q.*, the mean number of pages by article, and the mean number of authors by article. The code used for each journal (in the first column) should make it easy, for those familiar with the mathematical literature, to identify each journal.

REFERENCES

Adler, Robert, John Ewing, and Peter Taylor (2009), "Citation statistics: A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)." *Statist. Sci.*, 24, 1–14, URL `http://dx.doi.org/10.1214/09-STS285`.

Agrawal, A. and A. Goldfarb (2008), "Restructuring research: communication costs and the democratization of university innovation." *American Economic Review*, 98, 1578–1590.

Combes, P.P. and L. Linnemer (2003), "Where are the economists who publish? Publication concentration and rankings in Europe based on cumulative publications." *Journal of the European Economic Association*, 1, 1250–1308.

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

Kim, E.H., A. Morse, and L. Zingales (2009), "Are elite universities losing their competitive edge?" *Journal of Financial Economics*.

Levin, S.G. and P.E. Stephan (1991), "Research productivity over the life cycle: evidence for academic scientists." *The American Economic Review*, 81, 114–132.

Lotka, A.J. (1926), "The frequency distribution of scientific productivity." *Journal of the Washington Academy of Science*, 16, 317–323.

Palacios-Huerta, I. and O. Volij (2004), "The measurement of intellectual influence." *Econometrica*, 72, 963–977.

Stephan, P. (2008), "Job Market Effects on Scientific Productivity." *Conferences on New Political Economy*, 25, 11–29.

Waldinger, Fabian (2009), "Peer effects in science - evidence from the dismissal of scientists in nazi germany." CEP Discussion Papers dp0910, Centre for Economic Performance, LSE, URL `http://ideas.repec.org/p/cep/cepdps/dp0910.html`.

journal	Mean				
	tpages	nart	mcq	pages	nb aut
Acha	15953	819	0.83	19.48	2.50
Acmms	12581	756	0.49	16.64	2.5
Acta	23678	585	2.14	40.48	2.09
AdvApplProba	43739	2360	0.36	18.53	2.07
AdvCompMath	20798	961	0.63	21.64	2.35
Advances	94806	2711	1.04	34.97	2.06
Ajm	48824	1691	1.03	28.87	1.89
AnnApplProba	45783	1790	0.81	25.58	2.22
AnnProba	79936	3396	0.89	23.54	1.99
AnnStat	87197	4140	0.75	21.06	2.03
Annals	63255	1605	1.98	39.41	1.98
Arkiv	12884	710	0.64	18.15	1.79
Asens	27757	846	1.19	32.81	1.78
Bams	13372	1044	2.03	12.81	1.97
Bernoulli	19542	915	0.40	21.36	2.23
Biometrika	36867	3622	0.38	10.18	2.16
Cmh	26054	1285	0.92	20.28	1.90
CombProbaComput	16808	1127	0.34	14.91	2.31
Combinatorica	23545	1654	0.44	14.24	2.31
Commpde	69948	2495	0.94	28.04	1.90
Compositio	52333	2233	0.76	23.44	1.76
Computcomp	10917	470	0.33	23.23	2.84
Constr	22313	1115	0.71	20.01	2.12
Cpam	59562	1844	1.67	32.30	2.18
Crelle	89626	3568	0.91	25.12	1.85
Dcds	34491	2249	0.40	15.34	2.21
Dcg	38155	2291	0.50	16.65	2.44
Duke	91063	3093	1.38	29.44	1.93
DynSys	4707	232	0.33	20.29	2.31
Econometrica	42090	1832	0.70	22.97	2.04
ElecJComb	22490	1613	0.44	13.94	2.20
ElectrCommunProba	3201	331	0.63	9.67	2.03
ElectronJProba	14833	500	0.55	29.67	2.25
Expo	12656	674	0.44	18.78	1.68
FinancStoch	8686	419	0.78	20.73	2.29
Gafa	29987	960	1.17	31.24	1.97
Geotopo	16950	461	1.28	36.77	1.95
Ihes	16881	330	2.71	51.15	2.02
IhpProba	26998	1145	0.69	23.58	2.01
Ihpan	29750	1130	1.26	26.33	2.10
Imajna	27521	1401	0.63	19.64	2.17
Indiana	54789	2254	0.91	24.31	2.02
InfinDimAnal	10403	561	0.62	18.54	2.15
Inventiones	93613	3220	1.94	29.07	1.87
Inverse	54049	3569	0.81	15.14	2.42
Irmn	41359	1978	0.95	20.91	1.95
JAmStatAssoc	44051	4339	0.47	10.15	2.37
JCombThA	49950	3414	0.54	14.63	2.11
JCombThB	38601	2501	0.63	15.43	2.26

Table A.1—Journals in the database, A–J

journal	Mean				
	tpages	nart	mcq	pages	nb aut
JGeomPhys	45937	2209	0.38	20.80	2.19
JRStatSocB	17309	947	0.59	18.28	2.52
JStatPlan	95233	6402	0.28	14.88	2.18
JTheorProba	28841	1392	0.38	20.72	1.95
Jag	17880	655	0.78	27.30	1.83
Jams	34388	1000	2.54	34.39	2.11
Jcryptol	10009	475	0.36	21.07	2.56
Jde	149292	5822	0.91	25.64	2.03
Jdg	52365	1632	1.19	32.09	1.82
Jems	6895	230	1.42	29.98	2.25
Jfa	161178	5732	0.97	28.12	2.08
Jlms	50910	3667	0.65	13.88	1.94
Jmaa	268584	17941	0.47	14.97	2.05
Jmpa	36215	1255	1.03	28.86	2.18
Jnls	18661	637	0.75	29.30	2.44
Mams	107327	984	1.58	109.07	2.11
Mathann	82398	4237	0.96	19.45	1.81
Mathcomput	69607	4263	0.68	16.33	2.12
Mathprog	72525	3659	0.65	19.82	2.32
Mathz	75035	4401	0.62	17.05	1.80
Mrl	18007	1503	0.75	11.98	1.98
Nonlinearity	62933	3168	0.55	19.87	2.33
NumLinAlgA	17725	1000	0.39	17.73	2.49
Numermath	77775	3603	0.75	21.59	2.20
PTRF	64525	2709	0.92	23.82	1.94
Physicad	130332	7418	0.33	17.57	2.64
Plms	53111	1892	0.99	28.07	2.01
PublMath	15861	906	0.41	17.51	1.83
Qjm	18410	1319	0.54	13.96	1.91
Random	27551	1343	0.57	20.51	2.38
Rmibero	23016	771	0.71	29.85	2.04
ScandJStat	20070	1327	0.29	15.12	2.10
Siamco	80559	3795	0.71	21.23	2.14
Siamjam	89115	4291	0.49	20.77	2.41
Siamjc	80871	4130	0.40	19.58	2.71
Siamjmaa	44987	2737	0.67	16.44	2.33
Siamjsc	65780	3299	0.61	19.94	2.57
Siamma	69467	3636	0.91	19.11	2.08
Siamna	93630	4513	0.79	20.75	2.27
Siamopti	38298	1897	1.08	20.19	2.42
Siamrev	22594	1008	1.01	22.41	2.18
SochProcAppl	61820	3320	0.57	18.62	1.98
StatSci	11825	605	0.23	19.55	2.13
StatSinica	29945	1711	0.23	17.50	2.33
StudAM	31616	1287	0.31	24.57	2.22
Tams	187528	8770	0.83	21.38	1.93
Topology	34157	1732	0.82	19.72	1.85
Total	85963	4584	0.74	20.83	2.13

TABLE A.2—JOURNALS IN THE DATABASE, J–Z