Exploring the world of Economics through RePEc data

Sultan Orazbayev

Independent

15 October 2017

Exploring the world of Economics through RePEc data*

Sultan Orazbayev [†]

October 15, 2017

Preliminary version – comments are welcome.[‡]

Abstract

This document describes the data available through RePEc and related services: CitEc, CollEc, EDIRC, IDEAS, Genealogy and EconPapers. The document is purely descriptive, and is intended as a guide to some of the data available through RePEc on authors, institutions, collaborations, and networks.

Keywords: economics; economists; sociology of science; collaboration networks; geography of science; gender.

JEL codes: A11, Z00.

1 Introduction

Research Papers in Economics (RePEc) is a popular resource for economists - as of April 2016 there were 61,097 registered users, of which 47,129 users claimed authorship of at least one work (journal publication, working paper, book or a chapter in a book, or software).[1] In addition to claiming authorship, users can submit their affiliation and education history, provide lists of references for publications, subscribe to newsletters covering different fields of Economics, upload and

---

†contact@econpoint.com.

‡This is a working copy of the document, so there can be small discrepancies between the tables.

[1]Users can claim authorship of a work without prior approval. This self-censoring approach is easily justified by the low cost of verifying the claim and the high cost of discovering a false claim.

share links to their publications, keep track of the downloads to their papers and participate in various rankings (by age, cohort, gender, country and others).

RePEc is run by a team of volunteers and freely shares the available information. Krichel and Zimmermann (2009) explain the challenges of providing quality bibliographic information and give a glimpse into the history of development of RePEc and related services from the early 1990s. The information on the website is clearly organised with key summary statistics provided on the relevant web-pages. The purpose of this document is to describe and summarise most of the available data, as well as provide additional information that may be useful to researchers thinking of using RePEc data. Information is organised in the following sections: authors, papers, journals and working paper series, organisations (institutions), collaborations, and networks. The appendix provides technical details about processing of the data.

2 Authors

This section provides an overview of the registered users, with an emphasis on authors. The majority of RePEc users are economists that have published at least one work – 8 out of 10 registered users have authored at least one work included in RePEc's database, see Table 1 summarises information on the authors. About a half of the registered authors have claimed 10 or more works, and there are more than a thousand prolific authors with more than 100 claimed works each.

Table 1: Number of authors and collaborators among RePEc users.

	Registered users		
	total	as percentage of	
	count	authors	all users
Registered users	61,097		100.0
Authored 1+ works	47,129	100.0	77.1
Authored 10+ works	22,703	48.2	37.2
Authored 100+ works	1,102	2.3	1.8
Collaborated on 1+ works	37,796	80.2	61.9
Collaborated on 10+ works	13,599	28.9	22.3
Collaborated on 100+ works	256	0.5	0.4

Source: own calculations based on RePEc data. Notes: the number of authored works is adjusted for related works (see Section 3.1); not all works are claimed by all of their authors, so the number of collaborations is likely to be underestimated.

The rapid decline in the number of claimed works can also be seen on Figure 1, there are fewer than 100 authors (less than 1% of the registered users with publications) that have more than 200 journal publications.

Authors that publish more on their own also tend to collaborate more with the others, although there is a significant heterogeneity in proclivity to collaborate

Figure 1: Distribution of number of works per author.
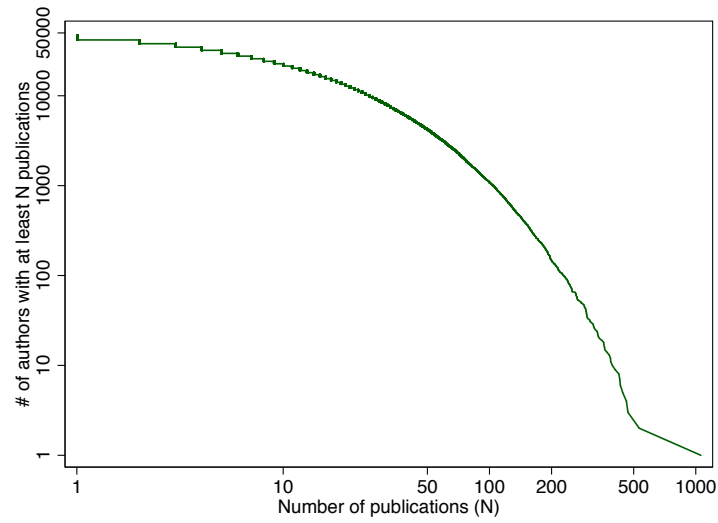
(Figures 2 and 3). Pearson correlation between numbers of solo and joint works by an author is 0.83, between numbers of joint and total works is 0.46, suggesting a possible link between an author's own productivity and collaborations.

Figure 2: Solo vs. joint academic output.
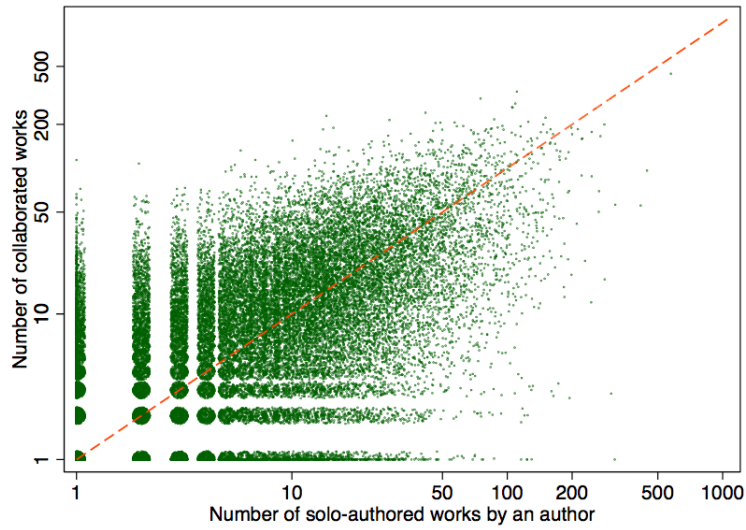
Source: own calculations based on RePEc data. Notes: the number of works is adjusted for related works (see Section 3.1); points are plotted with jitter added to differentiate observations at the frequently-occurring values.

Figure 3: Author collaboration and total academic output.

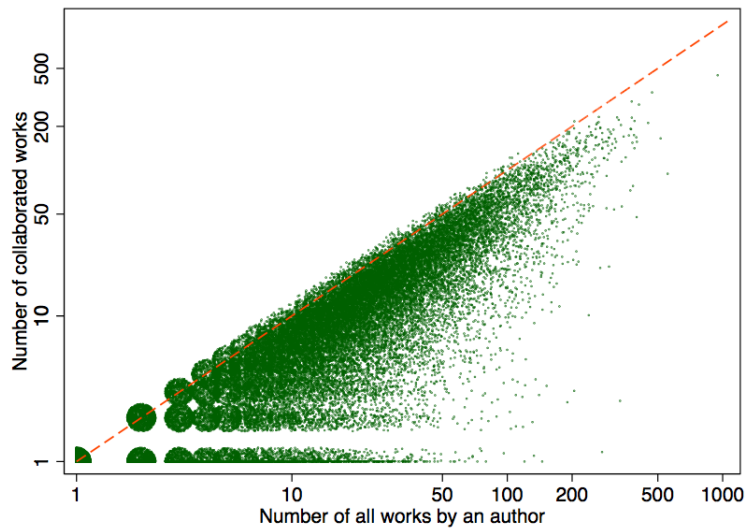Source: own calculations based on RePEc data. Notes: the number of works is adjusted for related works (see Section 3.1); points above the 45-degree line are purely due to the jitter added to show the number of points at the frequently-occurring values.

2.1   Genealogy

An interesting source of information on social ties between economists is RePEc Genealogy, similar to the Mathematics Genealogy Project (MGP). RePEc Genealogy tracks graduation information of registered authors (specifically, year and institution awarding the PhD degree, as well as the advisor(s)). As of April 2016 there are 9,676 registered users, and as can be seen in Table 2 they are positively selected in terms of their output and collaboration activity - they represent about 20% of the RePEc authors, but claimed almost half of all the works and nearly two-thirds of all collaborated works.

Table 2: RePEc Genealogy summary statistics.

	authors	As % of all all RePEc data
Registered authors	9,514	20.2
Claimed works	277,949	44.1
Collaborated works	126,136	62.9

Source: own calculations based on RePEc data. Notes: 162 users have Genealogy profiles but did not claim any works, these users are not included in the table above; the number of authored works is adjusted for related works (see Section 3.1); not all works are claimed by all of their authors, so the number of collaborations is likely to be underestimated.

Most of the authors with genealogy information are relatively young, see Figure 4. The genealogy links are continuously updated, but there is a bias towards users with publications and so the decline in the registered students from 2007 to 2016 is likely to disappear in a few years.

The data also allows establishing social ties among collaborators based on advisors and graduating department information. Specifically, it is possible to infer whether an academic work is coauthored by an advisor and a (former) student, close classmates (who share the same advisor and studied at the same department within 4 years of each other's graduation), by students who graduated from the same university within a 4-year period, or by alumni (graduated from the same department). These categories are not mutually exclusive.

The information on collaborations by the social ties between coauthors is presented in Table 3. The number of works for which social ties information can be identified is relatively small, less than a quarter of all collaborated works. However, once the RePEc Genealogy dataset expands, the data can become very useful for exploring the importance of personal ties for collaboration, citations, and other measures of academic performance.[2]

---

[2]For example, see Head et al. (2015) who explore the importance of personal ties among mathematicians for the diffusion of knowledge.

Figure 4: Graduation of new students over time.
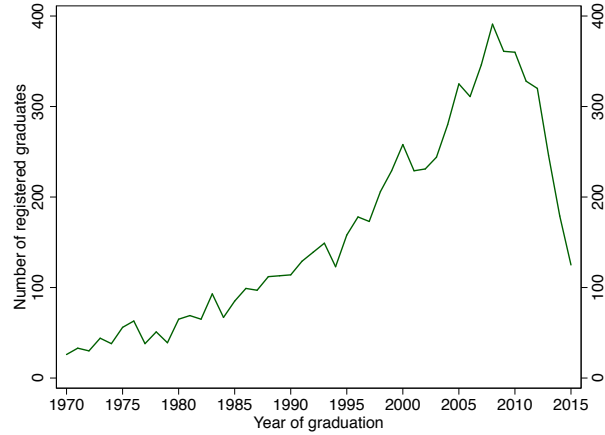
Source: own calculations based on RePEc data. Notes: the decline in registered graduates after 2014 does not indicate a decline in the actual number of graduates since not all graduates register on the site.

Table 3: Collaboration and social ties.

| | Number of unique works | | |
| | | as % of | |
	Count	(1)	(2)
All collaborated works (1)	200,411	100.0	
No social ties information available	154,798	77.2	
Social ties information is available (2)	45,613	22.8	100.0
Collaborations by social ties (not mutually exclusive categories)			
Advisor and student	13,626	6.8	29.9
Close classmates	1,839	0.9	4.0
Studies overlapped	4,730	2.4	10.4
Shared advisor	3,913	2.0	8.6
Alumni	9,202	4.6	20.2
Supervised a common student	3,180	1.6	7.0
No genealogy-based social ties	26,252	13.1	57.6

Source: own calculations based on RePEc data. Notes: the number of authored works is adjusted for related works (see Section 3.1); not all works are claimed by all of their authors, so the number of collaborations is likely to be underestimated; the total number of works with identifiable social ties differs from the number of collaborated works with an author that has genealogy information (in Table 2) due to lack of information on coauthors' genealogy; for works that have more than 2 authors a social tie is counted if there is a tie between any two coauthors. Social ties are defined as follows: close classmates – studied at the same institution within 4 years of each other's graduation and had a thesis advisor in common; studies overlapped – authors studied at the same institution within 4 years of each other's graduation; alumni – authors graduated from the same institution; supervised a common student – the authors supervised a student in common.

6

2.2 Gender

Among various RePEc rankings there is a table containing rankings of female economists. To identify gender of an author, RePEc team applies an algorithm based on a user's name with manual adjustments, where identified. This information can also be used to examine gender distribution of registered users and their collaborations, Tables 4 and 5 show distributions of users and collaborations by gender of the authors.

Table 4: User statistics by gender.

	Number of			Share of total		
	users	authors	collaborators	users	authors	collaborators
Female	11,183	8,557	6,997	18.2	18.2	18.5
Male	47,187	37,752	30,207	76.8	80.1	79.9
Unknown	3,106	820	592	5.1	1.7	1.6
Total	61,476	47,129	37,796	100.0	100.0	100.0

Source: own calculations based on RePEc data. Notes: the number of authored works is adjusted for related works (see Section 3.1); not all works are claimed by all of their authors, so the number of collaborations is likely to be underestimated.

Table 5: Collaboration statistics by gender.

	Collaborated works		
		Share of	
	Count	total	identified
Female-only	5,182	2.6	2.6
Male-only	145,092	72.4	73.6
Mixed	46,967	23.4	23.8
Can't identify	3,170	1.6	
Total	200,411	100.0	100.0

Source: own calculations based on RePEc data. Notes: the number of authored works is adjusted for related works (see Section 3.1); not all works are claimed by all of their authors, so the number of collaborations is likely to be underestimated; a work is considered identified with respect to gender if gender of at least two authors is known.

3 Papers

RePEc's service EconPapers provides information on almost 2 million works by economists, including roughly 700 thousand working papers and 1.2 million journal publications. Apart from papers, EconPapers provides information on software items and books (including chapters), but the main mode of communication by

Table 6: Paper statistics for EconPapers.

	Type of work	
	Journal publications	Working papers
Registered	1202287	699993
Claimed by an author	345821	413415
Claimed by an author and has at least one citation	217852	210653

Source: own calculations based on RePEc data. Notes: these numbers are not adjusted for related works; citations are not adjusted for self-citations.

volume of publications and citations in Economics is journal publications. This section provides information on journal publications and working papers.

3.1   Related works

The working papers uploaded to RePEc remain on the website even if the work is published in a journal at a later stage. The citations to and from such working papers can distort citation counts, and also can artificially inflate number of works by an author. Moreover, it is not uncommon for multi-author works to be circulated in several working paper series. Removing such 'related works' from the sample is an important step in obtaining a better understanding of the sample, especially for the working papers sample. EconPapers, a RePEc service, identifies related works, and this information can be used to clean the data sample.

The procedure used was to collect all related papers into a paper-specific bundle. Out of all papers in this bundle only one will be kept. If one of the papers in the bundle is a journal publication, then it would be kept. If there is no journal publication, then the latest version (by calendar year) of the working paper is kept. Based on information about almost 300 thousand related papers, roughly 170 thousand papers were removed from the sample because they represent either an earlier version of a working paper, or the working paper was published in a journal. The detailes are presented in Table 7.

Table 7: Removing related works from the sample.

Number of all related papers	293276
Number of bundles of related papers	121565
Number of old versions of related papers	171711
Number of latest versions of related papers	121565

Source: own calculations based on RePEc data.

Unless specified otherwise, all further calculations are based on the sample which is adjusted for related works.

Figure 5: Effect of correction for related works.
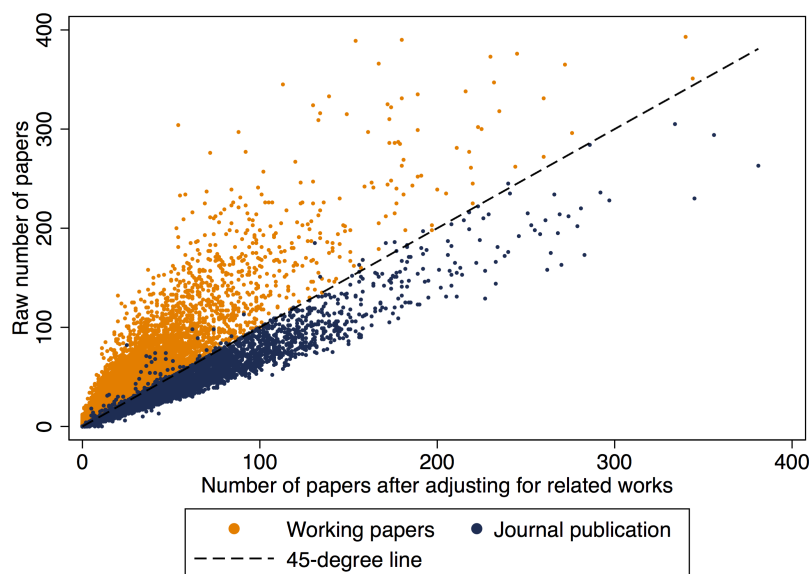
Source: own calculations based on RePEc data. Notes: paper counts are adjusted for related works; total number of publications exceeds unique papers due to co-authorships.

3.2 JEL codes

Information on JEL codes is available for many, but not all papers. Figure 6 uses the available information. For a more complete treatment see Rath and Wohlrabe (2015) who fill-in some of the missing JEL codes using information from EconLit and examine changes in popularity of specific JEL codes over time. Also, an interesting study is Krichel (2007), where the most central topics in Economics are identified using JEL code information for each paper. A network of JEL codes is constructed by creating a link between JEL codes if a single paper reports relevance of both JEL codes, then a set of centrality measures is calculated to determine post 'central' JEL codes.

Figure 6: Distribution of JEL codes.
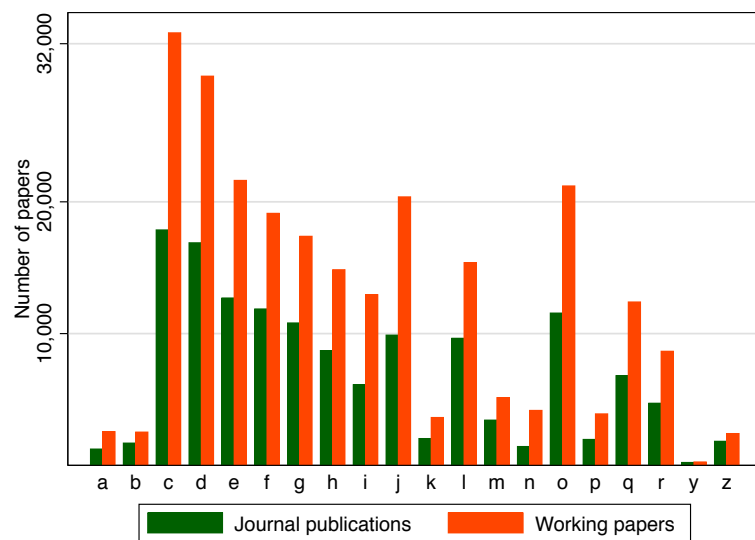
Source: own calculations based on RePEc data. Notes: paper counts are adjusted for related works; the number of papers will exceed actual papers due to multiple JEL codes per paper; all JEL codes were aggregated to the one-symbol level (multiple instances of codes under the same one-symbol level are counted as one).

4   Citations

RePEc service CitEc tracks references for papers registered on RePEc. Processing of references is automatic, but registered users also can contribute reference lists for the papers (not necessarily the papers for which they claim authorship). CitEc has processed approximately 20 million references, of which about 8 million can be tracked to a cited paper on RePEc.

By combining information on cited-citing papers and claimed authorship it is possible to calculate citation counts that exclude self-citations. It must be noted that this can be done only using the papers for which authors have been identified (and registered on RePEc). Adjustment for self-citation makes little difference for the highly cited papers, but can result in a significant drop in citation count for the others - the largest absolute drop is by 95 citations for a paper that had 126 gross citations, and the largest relative drop is 100% for an article that has been exclusively self-cited 37 times. Table 8 shows summary statistics for papers that are exclusively self-cited, conditional on being cited at least once and on being able to identify authorship of the citing paper.

Table 8: Summary statistics for papers that are exclusively self-cited.

		Gross citation count				
	Median	Mean	St. dev.	Min	Max	N
Journal publications	1	1.61	1.34	1	37	18554
Working papers	1	1.60	1.50	1	51	19574

Source: own calculations based on RePEc data. Notes: these numbers may exaggerate self-citation because these citation counts do not included papers with unidentified authors.

However, when looking at the sample as a whole, correcting for self-citations doesn't change the relative ranking of the papers among each other. Table 9 shows correlations between gross and net (without self-citations) citations depending on the number of gross citations. Thus, gross citations can be a useful proxy for net citations, when self-citations cannot be identified.

Figure 7: The effect of adjusting same-country citations.
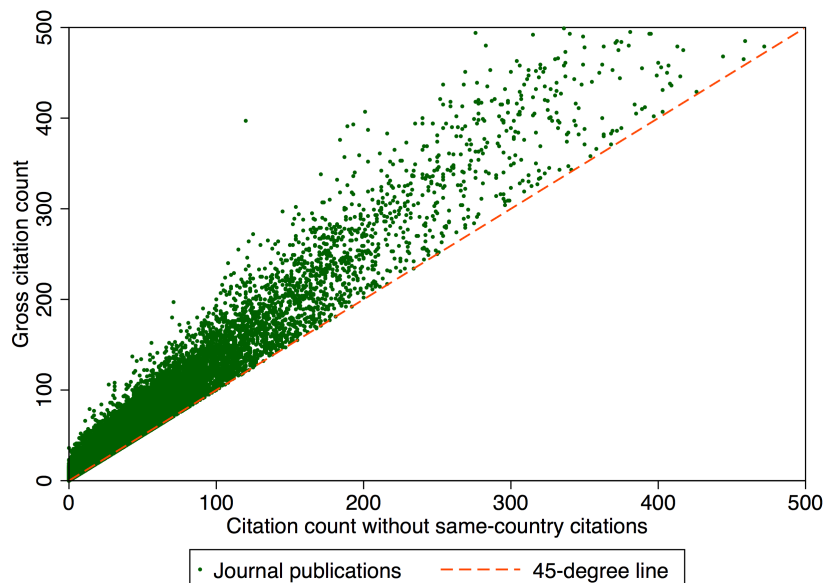
Table 9: Correlation between gross and net citations.

	Journal publications	Working papers
All papers	1.00	0.99
Sample	324426	241652
Papers with 100+ gross citations	1.00	1.00
Sample	3748	202
Papers with 10-99 gross citations	0.98	0.97
Sample	53089	10944
Papers with 0-9 gross citations	0.93	0.91
Sample	267589	230506

As would be expected, journal publications have higher citation counts, see Table 10. Citation counts initially increase with the number of authors, but greater collaboration teams do not necessarily produce highly-cited papers.

Identifying authors of the citing papers also allows comparing countries of the cited-citing papers' authors. This is a way of measuring the relevance and diffusion of a paper in countries other than those of the authors'. The 'different country' citation counts were calculated for each paper as follows: after attach-

Table 10: Summary citation statistics by type of work and number of authors.

# of authors	Median	Mean	St. dev.	Min	Max	N
		Journal publications				
1	0	5.47	28.77	0	2814	222781
2	2	11.30	44.19	0	3485	81468
3	3	13.70	48.55	0	2292	17264
4	3	15.61	61.37	0	1714	2240
5	1	6.41	16.08	0	157	368
6	0	8.59	25.92	0	219	105
7	0	3.25	8.15	0	55	69
8	0	2.41	5.71	0	25	46
9	0	10.42	30.06	0	129	33
10+	0	4.73	12.21	0	70	52
All items	1	7.44	34.92	0	3485	324426
		Working papers				
1	0	1.14	5.16	0	373	157243
2	0	2.42	9.36	0	484	65386
3	0	2.99	9.55	0	282	15943
4	1	3.86	11.06	0	210	2325
5	0	4.17	14.53	0	188	446
6	0	3.72	13.57	0	150	148
7	1	3.22	6.22	0	33	64
8	0	6.89	18.79	0	82	35
9	0	10.82	29.76	0	100	11
10+	0	1.29	3.27	0	20	51
All items	0	1.64	7.03	0	484	241652

Source: own calculations based on RePEc data. Notes: these numbers are adjusted for related works and exclude self-citations; sample includes only papers claimed by users.

Figure 8: The effect of adjusting self-citations.
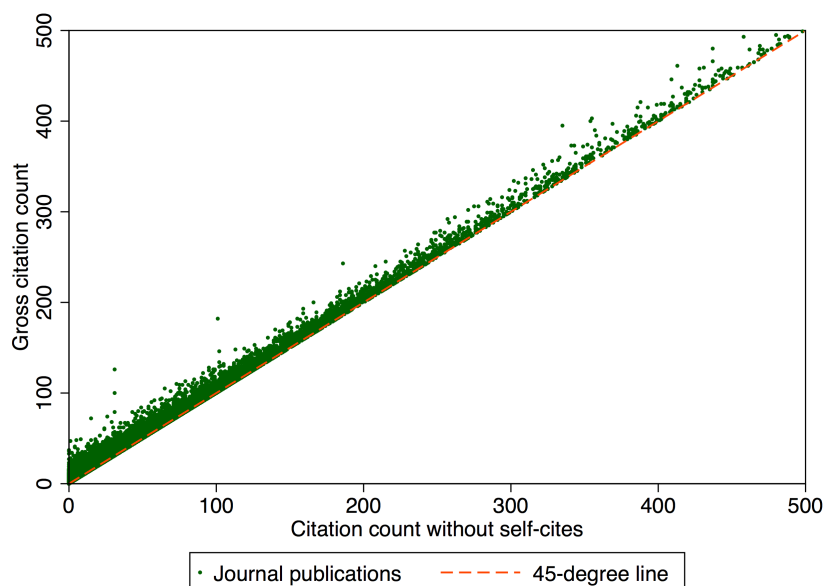
Source: own calculations based on RePEc data. Notes: paper counts are adjusted for related works.

ing the author-country information to the cited and citing papers, a citation was counted as coming from a different country if at least one country of the citing paper's country list was different from the cited paper's country list. Tables 11, 12 and 13 show summary statistics based on 'different country' citation counts. The numbers suggest that gross citation counts can be a useful proxy for 'different country' citation counts.

Table 11: Summary statistics for papers that are cited exclusively from the authors' countries.

| | Gross citation count | | | | | |
	Median	Mean	St. dev.	Min	Max	N
Journal publications	1	1.80	1.52	1	36	31579
Working papers	1	1.75	1.61	1	41	29515

Source: own calculations based on RePEc data. Notes: these numbers may exaggerate same-country citation because citation counts do not include papers with unidentified authors.

14

Table 12: Correlation between gross and 'different country' citations.

	Journal publications	Working papers
All papers	0.99	0.96
Sample	324426	241652
Papers with 100+ gross citations	0.98	0.92
Sample	3748	202
Papers with 10-99 gross citations	0.93	0.90
Sample	53089	10944
Papers with 0-9 gross citations	0.87	0.84
Sample	267589	230506

Source: own calculations based on RePEc data. Notes: these numbers are adjusted for related works; sample includes only papers claimed by users.

Table 13: Summary citation statistics by type of work and number of countries.

# of countries	Median	Mean	St. dev.	Min	Max	N
Journal publications						
1	0	5.30	26.63	0	2543	285156
2	2	9.94	34.93	0	3072	36147
3	3	11.58	34.66	0	1213	2786
4	2	7.18	14.52	0	94	184
5	0	5.41	11.73	0	55	75
6	0	1.77	6.51	0	41	56
7	0	5.00	10.92	0	41	16
8	8	31.00	44.24	3	82	3
9	11	26.00	35.00	1	66	3
All items	0	5.87	27.78	0	3072	324426
Working papers						
1	0	1.09	5.16	0	368	208539
2	0	2.26	7.62	0	267	30272
3	1	2.78	8.40	0	192	2572
4	1	5.04	10.76	0	76	194
5	0	3.24	10.12	0	64	41
6	0	1.50	2.79	0	10	18
7	1	4.14	8.43	0	23	7
8	0	0.80	1.30	0	3	5
9	59	59.00	.	59	59	1
10+	0	0.67	1.15	0	2	3
All items	0	1.26	5.59	0	368	241652

Source: own calculations based on RePEc data. Notes: these numbers are adjusted for related works and include citations only if at least one country from the citing paper's country list is different from the countries in the cited paper's country list; sample includes only papers claimed by users.

The summary statistics by number of authors and number of countries suggest that citations on average increase for larger and geographically-diverse collaboration teams. Table 14 shows the statistics for more geographically diverse teams conditional on the number of collaborators. For collaboration teams of 2-3 authors, having more geographically diverse team has a positive impact on the net and 'different country' citations, while for larger teams having authors spread across different countries is correlated with lower citation counts. A possible interpretation is that coordination costs increase rapidly with the number of different geographical locations, which hampers the quality of the collaborative product.

Table 14: Summary citation statistics by the country diversity of the authors.

# of authors	countries	Net citations Mean	St. dev.	Median	N	'Different country' citations Mean	St. dev.	Median	N
				Journal publications					
1	1	5.47	28.77	0	222781	4.38	23.20	0	222781
2	1	10.91	45.26	2	53227	8.43	35.69	1	53227
2	2	12.02	42.08	3	28241	9.56	34.09	2	28241
3	1	13.01	52.67	2	8144	9.96	40.06	1	8144
3	2	13.98	44.12	3	6859	10.89	33.66	3	6859
3	3	15.34	45.86	4	2261	11.84	36.49	3	2261
4	1	10.31	32.45	1	793	7.63	23.70	1	793
4	2	21.10	88.37	4	888	15.50	62.11	3	888
4	3	15.05	36.25	4	442	11.10	26.92	3	442
4	4	11.92	19.15	5	117	8.78	15.22	4	117
5+	1	3.06	10.77	0	211	2.00	6.91	0	211
5+	2	7.93	22.04	1	159	5.86	15.40	0	159
5+	3	9.16	17.36	3	83	7.08	13.78	2	83
5+	4	6.75	20.71	1	67	4.39	12.85	1	67
5+	5+	6.97	18.56	0	153	4.94	12.73	0	153
				Working papers					
1	1	1.14	5.16	0	157243	0.89	4.30	0	157243
2	1	2.25	9.18	0	42369	1.67	7.12	0	42369
2	2	2.74	9.68	0	23017	2.15	7.64	0	23017
3	1	2.65	9.34	0	7697	1.99	7.35	0	7697
3	2	3.31	9.61	1	6163	2.51	7.29	0	6163
3	3	3.32	10.13	1	2083	2.56	7.61	1	2083
4	1	2.68	9.59	0	928	1.83	6.32	0	928
4	2	4.35	11.45	1	881	3.26	8.71	1	881
4	3	4.74	12.37	1	398	3.58	9.98	1	398
4	4	6.47	13.16	2	118	4.68	8.93	2	118
5+	1	2.45	9.47	0	302	1.63	6.20	0	302
5+	2	4.14	13.70	0	211	2.86	8.79	0	211
5+	3	5.79	21.05	1	91	4.23	15.07	0	91
5+	4	7.16	17.10	1	76	5.59	13.15	1	76
5+	5+	4.75	14.74	1	75	3.39	10.32	0	75

Source: own calculations based on RePEc data. Notes: these numbers are adjusted for related works and include citations only if at least one country from the citing paper's country list is different from the countries in the cited paper's country list; sample includes only papers claimed by users.

5 Journals and working paper series

Using information on the cited-citing papers it is possible to compute simple impact factor for each journal and working paper series (see Zimmermann (2013)). RePEc provides up-to-date rankings on its website. The historical ranking/impact factor information is also available. The rankings based on simple impact factor were calculated using the downloaded data, the following tables contain historical rankings for the top 30 journals (in 2015).

Table 15: Top 30 journals by simple 10-year impact factor over 2005-2015.

Journal title	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005
The Quarterly Journal of Economics	45.9	47.9	44.9	56.2	57.5	58.9	56.9	50.7	51.3	45.3	36.7
Journal of Political Economy	44.2	44.2	38.9	37.6	32.1	29.3	26.5	26.2	24.0	22.1	19.2
Journal of Economic Literature	44.0	47.7	45.9	51.6	52.4	55.2	61.1	54.7	52.5	47.7	36.9
Journal of Finance	30.0	28.4	25.8	26.4	25.2	25.1	24.2	22.6	23.2	20.8	19.6
Econometrica	28.3	26.9	27.2	26.0	25.3	24.7	22.4	20.4	18.7	16.6	14.5
Review of Economic Studies	28.1	23.9	22.6	21.9	20.4	18.6	17.3	15.7	13.9	11.9	10.4
Journal of Economic Perspectives	26.7	24.8	22.2	22.1	22.1	23.4	20.1	19.2	19.9	17.3	17.8
American Economic Review	24.3	23.4	21.8	20.8	21.4	21.2	20.9	19.2	17.4	16.5	15.2
American Economic Journal: Macroeconomics	23.1	19.8	16.7	13.4	.	.	.	.	.	.	.
Journal of Economic Growth	22.8	31.2	32.4	32.4	28.9	28.7	24.9	21.1	18.6	19.4	15.9
The Review of Economics and Statistics	21.9	21.7	21.3	21.3	20.1	18.7	16.8	15.2	13.4	12.3	10.4
Review of Financial Studies	21.9	18.6	15.9	14.1	13.4	13.4	14.2	13.9	14.0	12.6	11.6
Journal of Monetary Economics	21.2	20.3	21.0	20.8	19.4	19.0	18.6	17.3	16.4	14.9	13.5
Journal of International Economics	21.0	20.0	19.4	18.1	18.2	19.1	18.1	17.3	16.4	15.1	13.2
Annals of Economics and Finance	20.7	16.5	3.1	2.9	.	.	.	.	.	.	.
Brookings Papers on Economic Activity	20.6	19.3	23.1	21.2	22.5	23.4	22.3	25.5	23.4	24.4	26.3
The Journal of Business	20.2	16.8	15.0	14.0	12.6	12.2	10.9	9.5	8.5	7.8	8.6
Journal of Financial Economics	19.8	18.1	17.1	17.0	16.3	16.7	15.6	14.9	14.4	13.4	12.0
Journal of the European Economic Association	19.5	17.8	16.5	14.5	12.8	11.4	10.1	9.0	7.5	5.4	4.2
Experimental Economics	19.3	17.4	14.4	12.3	10.5	8.9	7.8	7.0	5.7	4.9	4.1
Journal of Labor Economics	19.1	17.4	17.2	15.7	15.4	14.3	13.1	12.5	11.6	10.4	9.0
Economic Policy	18.5	17.2	15.1	15.3	16.2	18.3	16.8	16.3	15.2	14.7	13.8
Economic Journal	18.3	17.9	16.5	16.9	15.8	16.0	15.3	14.9	14.1	12.6	10.9
Journal of Applied Econometrics	16.8	15.0	14.9	13.8	13.5	12.7	11.3	9.8	8.9	8.2	7.8
American Economic Journal: Applied Economics	16.5	13.1	9.7	7.4	5.7	.	.	.	.	.	.
Proceedings of the Federal Reserve Bank of San Francisco	16.3	18.3	22.4	21.8	18.6	17.6	15.1	17.2	16.0	13.3	11.0
Journal of Development Economics	16.2	16.2	14.8	13.5	12.5	11.9	11.1	10.3	9.3	8.8	7.9
Journal of Money	15.7	15.1	14.1	13.4	12.6	11.4	10.7	9.8	9.6	8.9	9.0
IMF Economic Review	15.4	14.5	12.8	11.5	11.7	11.1	11.4	13.8	14.1	13.0	12.7
Journal of Business and Economic Statistics	15.4	14.0	12.3	16.9	15.6	13.7	12.7	12.1	11.4	10.8	11.2

Source: own calculations based on CitEc data (as of Jan. 1, 2016).

Table 16: Top 30 journals by simple 5-year impact factor over 2005-2015.

Journal title	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005
Annals of Economics and Finance	32.9	26.3	2.6	.	.	.	.	.	.	.	.
Journal of Economic Literature	27.9	27.6	26.7	22.3	18.0	14.6	24.0	25.7	31.8	27.0	19.9
The Quarterly Journal of Economics	25.6	25.8	21.5	21.2	20.0	22.3	28.4	22.7	33.1	30.5	26.0
Journal of Political Economy	20.2	19.2	18.4	20.5	17.2	24.1	22.7	18.2	15.9	12.5	9.2
American Economic Journal: Macroeconomics	19.1	19.8	16.7	13.4	12.1	.	.	.	.	.	.
Review of Economic Studies	16.7	14.2	13.6	13.7	14.3	14.5	12.4	12.3	10.8	8.4	6.8
Journal of the European Economic Association	15.8	13.7	11.0	9.6	9.4	9.4	9.0	9.0	7.5	5.4	4.2
VOPROSY ECONOMIKI	15.1	13.7	11.3	12.2	6.9	.	.	.	.	.	.
American Economic Review	14.8	13.5	12.4	11.4	11.1	11.3	11.5	10.8	9.9	10.0	9.7
Econometrica	14.8	15.6	13.8	13.9	13.2	12.6	11.4	11.6	10.6	9.5	9.3
American Economic Journal: Applied Economics	14.6	13.1	9.7	7.4	5.7	3.5	.	.	.	.	.
Journal of Finance	14.2	13.2	12.7	11.4	11.4	12.0	11.7	10.5	12.4	11.6	12.1
IMF Economic Review	13.3	13.4	10.1	7.3	6.6	7.4	7.5	7.2	6.9	7.0	7.1
Journal of Economic Perspectives	12.7	12.3	9.4	11.2	12.0	11.6	10.4	10.1	10.8	9.4	10.0
Journal of Economic Growth	12.6	12.3	13.4	15.7	13.1	11.6	18.3	16.7	16.2	12.9	11.2
Journal of International Economics	12.5	11.0	10.5	10.6	10.2	10.9	10.9	11.4	10.5	10.2	10.2
Brookings Papers on Economic Activity	12.4	12.5	14.0	9.8	7.7	8.4	8.1	14.9	14.0	11.8	11.3
The Review of Economics and Statistics	12.3	11.2	12.2	12.5	12.9	12.6	12.4	10.9	10.3	8.9	7.8
Review of Financial Studies	12.2	14.9	12.9	10.4	8.8	8.1	7.0	7.6	7.8	8.2	8.4
American Economic Journal: Economic Policy	12.2	9.6	6.9	5.6	4.0	.	.	.	.	.	.
Journal of Financial Intermediation	11.3	9.2	9.3	8.4	6.3	4.6	7.4	7.6	8.7	6.5	7.1
Economic Journal	11.3	10.8	9.8	9.2	8.2	8.6	8.8	8.8	9.9	8.6	8.2
Economic Policy	11.2	11.6	10.2	8.6	9.3	9.0	8.2	7.0	7.6	9.8	12.3
Journal of Monetary Economics	11.2	11.8	10.8	11.5	10.3	11.1	10.3	12.1	12.8	11.6	9.9
Journal of Applied Econometrics	10.7	8.9	8.5	9.6	8.5	8.3	7.2	7.0	6.2	6.0	5.8
Review of Economic Dynamics	10.6	10.1	10.1	8.4	6.9	7.6	7.2	6.5	5.4	4.5	3.9
Journal of Labor Economics	10.6	9.3	8.7	8.6	9.4	8.1	7.9	7.7	6.9	7.2	6.5
Journal of Development Economics	10.3	10.9	9.5	9.2	8.2	7.6	7.1	6.8	6.2	6.1	5.4
Journal of Business and Economic Statistics	10.1	9.0	8.2	7.9	7.6	6.7	6.3	5.4	11.7	10.4	9.0
Journal of Financial Economics	9.8	8.5	8.4	8.9	9.2	9.5	9.0	8.4	8.9	8.1	8.2

Source: own calculations based on CitEc data (as of Jan. 1, 2016).

19

Table 17: Top 30 journals by simple 3-year impact factor over 2005-2015.

Journal title	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006	2005
Annals of Economics and Finance	46.0	38.6	2.1	.	.	.	.	.	.	.	.
The Quarterly Journal of Economics	17.0	16.8	14.9	16.0	14.5	14.0	12.2	12.6	19.9	17.2	21.4
Review of Economic Studies	15.6	10.9	9.4	8.9	10.0	10.1	10.5	11.6	8.4	7.4	5.8
Journal of the European Economic Association	14.1	14.7	10.6	7.6	6.4	6.4	7.6	7.8	6.9	5.4	4.2
Journal of Economic Literature	13.6	17.0	14.4	14.2	16.2	14.5	8.7	3.5	20.4	20.0	15.3
American Economic Journal: Macroeconomics	12.8	10.6	12.8	13.4	12.1	12.4	.	.	.	.	.
Journal of Political Economy	12.5	15.0	12.2	13.3	12.5	13.9	11.5	18.1	14.6	10.5	7.7
American Economic Review	11.3	10.7	8.5	8.1	8.5	7.5	7.5	7.5	7.4	7.0	6.2
NBER Macroeconomics Annual	10.8	.	.	.	.	.	.	.	.	.	.
Journal of Finance	10.3	9.4	7.9	6.9	6.9	6.9	7.6	7.7	7.1	6.3	8.8
American Economic Journal: Applied Economics	10.0	9.7	8.4	7.4	5.7	3.5	.	.	.	.	.
Journal of International Economics	9.8	8.8	7.7	6.6	6.9	7.6	7.1	8.0	9.0	8.5	6.7
American Economic Journal: Economic Policy	9.8	8.4	6.5	5.6	4.0	.	.	.	.	.	.
Journal of Economic Growth	9.6	8.2	10.6	9.8	9.3	11.1	8.1	6.7	13.7	11.1	9.7
Econometrica	9.5	10.9	9.8	9.6	8.2	9.5	9.0	8.1	6.2	6.4	6.4
The Review of Economics and Statistics	9.4	8.1	8.1	8.1	10.1	10.7	9.7	8.4	7.8	6.8	6.6
Review	9.2	6.5	2.9	2.5	3.5	3.3	3.2	3.3	4.1	3.6	3.2
Journal of Economic Perspectives	8.4	7.6	5.7	5.7	4.9	7.8	6.9	6.4	6.2	5.5	6.3
Journal of African Economies	8.2	6.2	2.4	2.8	2.8	2.5	2.9	2.7	2.1	2.0	2.0
Journal of Monetary Economics	8.2	8.9	7.0	7.4	7.0	7.8	7.0	8.1	7.7	9.8	8.8
Journal of Labor Economics	7.8	6.8	6.3	6.2	5.1	5.4	6.2	5.2	5.0	5.5	4.8
Scottish Journal of Political Economy	7.6	7.3	2.1	1.8	2.7	2.5	2.7	2.1	2.7	3.6	3.6
IMF Economic Review	7.5	8.5	7.9	8.6	6.6	5.6	3.6	4.6	6.5	5.4	5.6
Economic Journal	7.5	8.5	6.7	5.9	6.3	6.5	6.1	5.9	6.3	6.1	6.4
Journal of Money	7.4	6.5	7.1	6.7	6.6	7.4	6.9	7.2	7.0	5.9	5.4
Journal of Development Economics	7.2	8.8	6.7	6.9	5.6	6.1	5.9	5.0	4.4	4.2	3.9
Journal of Business and Economic Statistics	7.2	7.4	6.0	6.0	5.7	5.2	4.8	4.1	3.9	3.4	13.0
Journal of Applied Econometrics	7.1	6.0	8.0	7.4	6.2	6.2	5.2	5.4	5.1	5.0	4.1
Economic Policy	7.1	7.9	6.7	7.1	6.1	5.1	6.3	5.8	4.9	4.0	5.4
Journal of Financial Economics	7.0	5.8	5.2	4.7	6.0	6.3	6.3	6.5	5.9	6.0	5.6

Source: own calculations based on CitEc data (as of Jan. 1, 2016).

20

6 Economic research organisations

Data on economic research organisations is collected by EDIRC (Economics Departments, Institutions and Research Centres), a RePEc service. EDIRC maintains data on names, locations and website links for 13,954 institutions, which include departments at universities, institutes, research centres, central banks, think-tanks, economic associations and societies.[3] Table 18 shows that the registered institutions vary in size, however many of the Economics departments are well represented in the data: about 70% of the faculty members of the top 30 US Economics departments are registered on RePEc (Hamermesh, 2015).

Table 18: Number of users per organisation.

| | Registered user type: | |
	author	non-author
Organisations with 1+ users	6,972	3,571
Organisations with 10+ users	1,403	75
Organisations with 100+ users	29	3
Users per organisation:		
Median	3	1
Mean	7.5	2.5
90th percentile	18	4
99th percentile	65	15
Max	669	669

Source: own calculations based on RePEc data. Notes: organisation refers to departments, institutions and research centres registered with EDIRC (associations and societies are excluded); statistics related to the number of users per organisation are calculated conditional on an organisation having at least one affiliated user.

Registered users can claim affiliation with any organisation[4] and can specify the shares of their affiliation to any particular institution. Most of the registered authors are affiliated with at just one organisation, however some users have multiple affiliations, see Table 19.

Combing EDIRC data with affiliation information provided by users, it is possible to geolocate registered authors based on their last-known affiliation.[5] In fact, EDIRC sorts organisations based on their country (this information is used, for example, to calculate country-based rankings), but to get a richer set of geographical details all of the institution addresses were geocoded using Google Maps API.

---

[3]EDIRC holds information on 846 associations and societies, however there is just one registered user (non-author) that indicated a society as their main affiliation. This makes it easier to assign geographical location to each author, since membership in an association or society often does not imply residence in the relevant country.

[4]If the organisation is not registered on EDIRC, the authors can provide basic details, such as the name of the organisation and its postal address.

[5]A panel dataset containing author-specific affiliations over time is under construction.

Table 19: Number of affiliations per registered user.

| | Registered user type: | |
	author	non-author
Total registered users	47,129	13,968
Unaffiliated users	5,947	5,795
Affiliated users	41,182	8,173
As % of total	87	59
Affiliations per user:		
Median	1	1
Mean	1.2	1.6
90th percentile	2	3
99th percentile	4	5
Max	9	12

Source: own calculations based on RePEc data. Notes: associations and societies are excluded from the sample; statistics related to the number of affiliations per user are calculated conditional on user having an affiliation.

Table 20 provides some statistics for the top 50 countries (by the number of registered organisations and users). There is a heterogeneity in 'density' of authors per institution, suggesting some cross-country differences in connectedness of the authors across and within institutions. For example, in Czech Republic there are about 17 authors per registered organisation, in Netherlands – 14, while in Taiwan and South Korea there are 3 authors per institution, on average. These are highly aggregated statistics, of course, but this heterogeneity in density of authors could also have implications for collaboration network of the economists (see Section 8). Also, the number of non-author users, who registered on RePEc but did not claim authorship of any works, is very high in some countries, for example in Russia, China, India and Pakistan.

Table 20: Top 50 countries sorted by the number of registered authors.

	# of users:		Organi-	Authors per organisation:		Authors as %
	authors	non-authors	sations	average	maximum	of the world total
United States	10,911	1,034	1,429	7.6	490	18.6
Germany	4,307	246	415	10.4	669	7.4
United Kingdom	4,156	459	389	10.7	296	7.1
France	3,724	233	301	12.4	165	6.4
Italy	3,213	176	316	10.2	176	5.5
Spain	2,373	155	301	7.9	110	4.1
Canada	1,657	119	193	8.6	108	2.8
Netherlands	1,592	100	117	13.6	144	2.7
Australia	1,447	124	155	9.3	66	2.5
Russia	1,099	1,670	110	10.0	148	1.9
Romania	1,034	88	106	9.8	83	1.8
Belgium	1,004	68	100	10.0	70	1.7
Switzerland	970	114	117	8.3	76	1.7
Japan	939	68	175	5.4	51	1.6
China	832	674	182	4.6	60	1.4
Sweden	816	28	91	9.0	68	1.4
Portugal	796	86	90	8.8	57	1.4
Turkey	736	173	172	4.3	121	1.3
Brazil	629	124	99	6.4	49	1.1
India	616	431	172	3.6	29	1.1
Colombia	616	94	74	8.3	87	1.1
Poland	508	124	76	6.7	79	0.9
Denmark	480	41	50	9.6	58	0.8
Austria	429	28	56	7.7	55	0.7
Czech Republic	419	57	24	17.5	91	0.7
Norway	411	33	42	9.8	53	0.7
Greece	404	40	60	6.7	36	0.7
Chile	395	59	52	7.6	50	0.7
South Africa	335	52	49	6.8	60	0.6
Argentina	331	63	66	5.0	40	0.6
Pakistan	319	274	64	5.0	78	0.5
Mexico	280	58	72	3.9	33	0.5
South Korea	271	49	91	3.0	26	0.5
Hungary	265	27	49	5.4	43	0.5
Ireland	256	13	23	11.1	48	0.4
Finland	234	22	40	5.8	22	0.4
New Zealand	232	17	37	6.3	30	0.4
Malaysia	204	117	41	5.0	27	0.3
Peru	185	53	25	7.4	43	0.3
Tunisia	175	72	29	6.0	27	0.3
Singapore	174	19	24	7.2	44	0.3
Taiwan	170	30	63	2.7	16	0.3
Israel	165	16	29	5.7	29	0.3
Indonesia	161	171	41	3.9	29	0.3
Luxembourg	143	6	13	11.0	31	0.2
Uruguay	117	9	13	9.0	38	0.2
Hong Kong	115	16	29	4.0	21	0.2
Bulgaria	114	38	21	5.4	40	0.2
Ukraine	113	161	37	3.1	29	0.2
Unidentified	48	20	26	1.8	11	0.1
Unaffiliated	5,947	5,795	.	.	.	10.2

Source: own calculations based on RePEc data. Notes: non-authors are the registered users that did not claim any works on RePEc; only organisations with at least one registered author are included in the table (associations and societies are excluded); authors with multiple affiliations are attributed to each of the affiliations; unaffiliated author count includes authors affiliated with organisations not registered on EDIRC.

Figure 9: Geographic distribution of the economists.
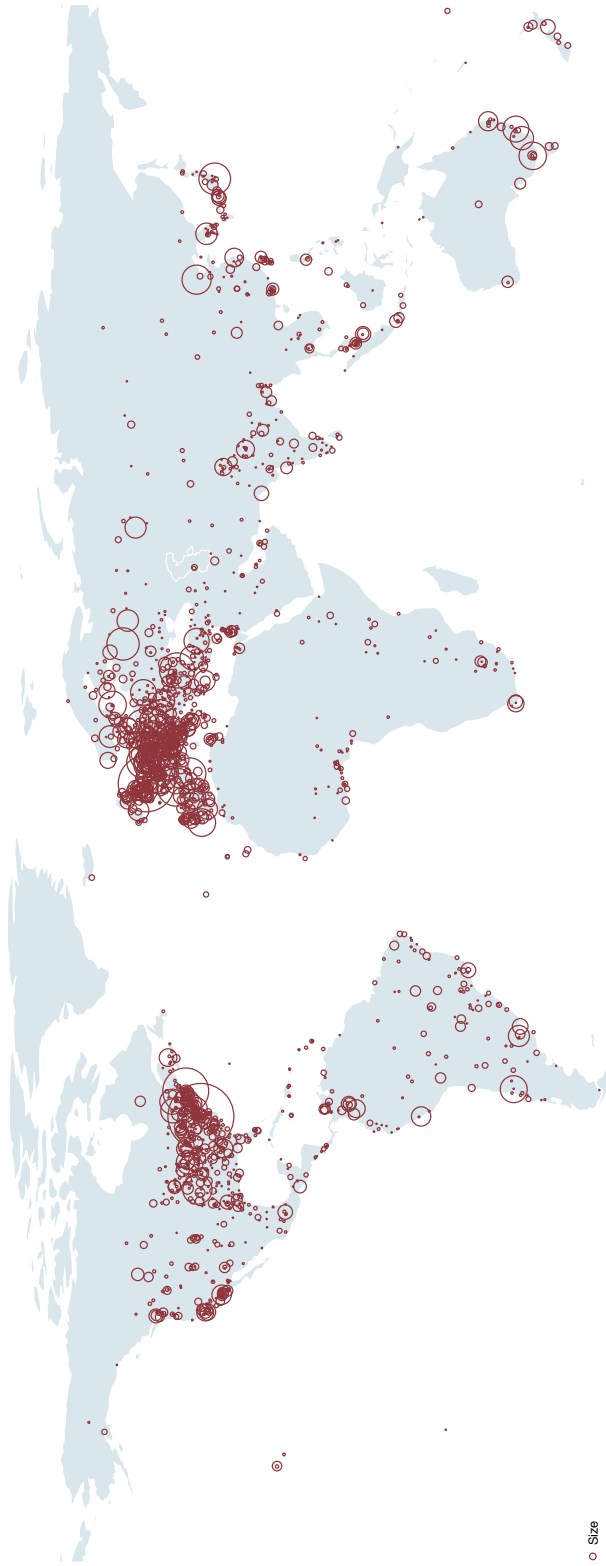
Source: own calculations based on RePEc data. Notes: the diameter of each circle is proportional to the number of authors registered on RePEc at a given location; the location coordinates were rounded to the first decimal point precision and then aggregated for simpler visual representation.

7 Collaborations

7.1 Domestic and international collaboration

Multiple affiliations of an author present a challenge if one is interested in identifying their main country of residence. For some authors, multiple affiliations are all based in the same country, so it is simple to assign the main country of residence. For authors with affiliations in multiple countries the following approach is taken. Firstly, if the author provided shares of affiliation, then the shares are aggregated to the country level and the country with the highest share is considered to be the main country of residence. Secondly, if the author did not provide shares or the previous calculation results in a tie, then the country of the first affiliation listed is used.[6] Table 21 provides the number of users for whom a country of residence could be assigned.

Most of the collaboration is domestic - both authors' main residence is in the same country. International collaboration comes at the 11th rank if collaborations are ranked according to the number of joint journal publications, see Table 22. Looking at the international country-pairs in Table 23, largest collaboration flows are between developed countries.

On average, international collaboration leads to papers that are cited more, see Table 24.

Unlike in other areas of science, in Economics the average international collaboration distance has been declining, see 10. It is possible that this observation is partly due to use of aggregated, country-level distances. If however this is indeed the underlying pattern, then this suggests that international collaboration is being concentrated between relatively proximate countries, possibly aided by migration of economists into these research core countries from other countries.

---

[6]The assumption here is that users provide their main affiliation first. This may not always be the case, but there is no further information to make a better inference.

Table 21: Identifying country of residence.

	# of users
Authored 1+ works	45819
Have affiliation information	42740
of which: have multiple affiliations	3308
Don't have affiliation information	3079
Can be assigned a country unambiguously	44591
of which: require order assumption to assign a country	1228
of which: collaborated on at least 1 paper	34449

Source: own calculations based on RePEc data. Note: the number of users that have affiliation information includes those users that provided affiliation to institutions which are not included in EDIRC database, so the numbers will differ slightly from Table 19.

Figure 10: Average international collaboration distance.
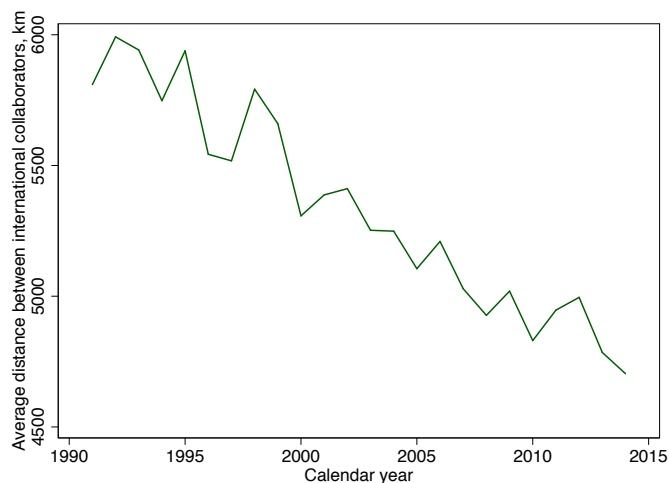
Source: own calculations based on RePEc data. Notes: the sample includes only journal publications; distances are the average country-level distances from CEPII dataset.

Table 22: Top 50 country-level collaboration pairs.

Country 1	Country 2	Journal publications	Working papers
United States	United States	58354	101054
Germany	Germany	18116	26756
United Kingdom	United Kingdom	12212	21278
France	France	11922	30256
Italy	Italy	10870	22352
Spain	Spain	7910	12046
Australia	Australia	5662	7600
Netherlands	Netherlands	4836	8624
Romania	Romania	4828	1026
Canada	Canada	4546	7826
United Kingdom	United States	4007	8860
Japan	Japan	2278	4240
Switzerland	Switzerland	2206	4520
Russia	Russia	2154	1244
Canada	United States	2122	3554
Austria	Austria	2034	5764
Sweden	Sweden	1858	5096
Germany	United States	1781	4610
France	United States	1746	4247
Greece	Greece	1668	1796
Belgium	Belgium	1606	5856
Norway	Norway	1596	3276
Turkey	Turkey	1482	2126
Germany	United Kingdom	1440	3086
Portugal	Portugal	1340	2860
Italy	United Kingdom	1239	2837
Australia	United States	1197	1847
China	China	1168	862
Italy	United States	1109	2769
Germany	Switzerland	1096	1729
Colombia	Colombia	1060	3380
France	United Kingdom	1024	1952
Denmark	Denmark	1014	2040
Spain	United Kingdom	964	1781
China	United States	949	1177
Netherlands	United States	930	1475
Czech Republic	Czech Republic	846	1332
Switzerland	United States	831	1907
Brazil	Brazil	822	2814
Chile	Chile	820	1618
Australia	United Kingdom	777	1313
Malaysia	Malaysia	734	732
Poland	Poland	712	1786
Spain	United States	712	1642
France	Germany	655	1478
Netherlands	United Kingdom	638	975
Ireland	Ireland	610	1432
India	India	608	950
France	Italy	599	1640
Belgium	France	586	1708

Source: own calculations based on RePEc data. Note: the paper counts were not adjusted for related works; country pairs are sorted alphabetically.

Table 23: Top 50 country-level international collaboration pairs.

Country 1	Country 2	Journal publications	Working papers
United Kingdom	United States	4007	8860
Canada	United States	2122	3554
Germany	United States	1781	4610
France	United States	1746	4247
Germany	United Kingdom	1440	3086
Italy	United Kingdom	1239	2837
Australia	United States	1197	1847
Italy	United States	1109	2769
Germany	Switzerland	1096	1729
France	United Kingdom	1024	1952
Spain	United Kingdom	964	1781
China	United States	949	1177
Netherlands	United States	930	1475
Switzerland	United States	831	1907
Australia	United Kingdom	777	1313
Spain	United States	712	1642
France	Germany	655	1478
Netherlands	United Kingdom	638	975
France	Italy	599	1640
Belgium	France	586	1708
Japan	United States	584	989
Israel	United States	511	939
Germany	Italy	509	1332
France	Switzerland	506	1169
Switzerland	United Kingdom	503	883
Germany	Netherlands	500	1136
Canada	France	496	1271
Austria	Germany	496	875
Greece	United Kingdom	491	476
Canada	United Kingdom	477	781
Sweden	United States	415	1008
Belgium	Germany	390	1137
France	Spain	388	716
Germany	Spain	369	709
Belgium	United States	368	923
Chile	United States	366	831
Italy	Spain	361	943
Singapore	United States	323	500
Australia	Canada	320	378
New Zealand	United States	313	505
France	Netherlands	308	593
South Korea	United States	299	447
Belgium	Netherlands	276	636
Denmark	United States	273	625
Turkey	United States	268	473
India	United States	267	452
Norway	United States	265	679
Italy	Netherlands	263	653
Brazil	United States	261	517
Hong Kong	United States	249	483

Source: own calculations based on RePEc data. Note: the paper counts were not adjusted for related works; country pairs are sorted alphabetically.

Table 24: Summary statistics for citations by type of collaboration.

	Median	Mean	St. dev.	Min	Max	N
Journal publications						
Domestic	1	6.71	33.42	0	2814	285156
International	3	12.74	43.98	0	3485	39270
Working papers						
Domestic	0	1.43	6.42	0	484	208539
International	0	2.99	9.92	0	376	33113

Source: own calculations based on RePEc data. Note: the citation counts were adjusted for self-citations and related works.

8 Networks

This section contains information on networks among users, based on journal publications, working papers, or both, and presented at the global, country and ego levels. Network statistics were calculated using *netsis* program in Stata, see Miura (2012).

8.1 Global network

Table 25 contains descriptive statistics on the global network of economists, where the links between authors are made on the basis of joint publication of journal articles, working papers, or both. The descriptive statistics indicate a somewhat closer network than described in Goyal et al. (2006). Apart from the difference in time periods covered, the samples are different, in the sense that RePEc users are self-selected (most likely positively). For example, average distance and number of isolates are lower than in Goyal et al. (2006), indicating that the registered users tend to collaborate with someone who is already a part of RePEc network. Also, the degree (number of co-authors) is higher for RePEc users than in the Goyal et al. (2006) sample, suggesting that the more collaborative authors are interested in registering on RePEc.

Table 25: Global research network.

| | Research networks based on | | |
	journal publications	working papers	all works
Unique authors	28299	30380	34449
Giant component	26108	28494	32537
As share of total	0.92	0.94	0.94
Isolates	2191	1886	1912
As share of total	0.08	0.06	0.06
Degree:			
Median	3.00	3.00	4.00
Mean	4.88	5.56	6.20
St. dev.	5.53	6.33	7.21
Max	82	77	95
Distance:			
Median	6.19	5.93	5.70
Mean	6.38	6.07	5.87
St. dev.	0.98	0.81	0.86
Min	4.45	4.28	4.13
Max	14.57	13.84	13.77
Clustering coefficient (overall)	0.20	0.20	0.19
Clustering coefficient (average)	0.24	0.27	0.26

Source: own calculations based on RePEc data using *netsis* program in Stata. Note: the paper counts were not adjusted for related works; single-author publications are not included in the sample.

8.2 Domestic networks

Domestic collaboration networks differ along several dimensions, see Table 26. Generally, smaller countries have higher clustering, for example Austria or Russia, where most of the registered economists are concentrated in Moscow and St. Petersburg. The giant component is large in almost every country, notable exceptions are Russia, China and India. Since the giant component is calculated based on domestic collaborations only, this indicates that many economists in these countries are either producing solo works or collaborating internationally, rather than domestically. The median degree is below 4 in every country except the United States, suggesting that the degree of domestic collaboration is much higher in the United States. The average distance within the domestic networks varies between 2 and roughly 6, suggesting that in some countries it takes more handshakes to reach another economist within the giant component of the domestic network. As expected, the average distance is smaller in countries that have a relatively smaller giant component.

Table 26: Domestic network statistics for the top 30 countries by number of collaborators.

Country	N	Percentage of isolates	Degree				Distance					Clustering coeff.	
			Median	Mean	St. dev.	Max	Median	Mean	St. dev.	Min	Max	Average	Overall
United States	7316	0.03	4	5.53	5.49	56.00	5.24	5.36	0.71	3.77	10.12	0.19	0.13
Germany	2179	0.07	3	4.92	5.95	63.00	4.67	4.75	0.75	3.02	9.01	0.27	0.25
France	2171	0.08	3	4.38	4.43	42.00	5.01	5.19	0.89	3.51	10.34	0.23	0.19
United Kingdom	2054	0.10	2	3.77	3.87	37.00	5.16	5.30	0.83	3.69	9.17	0.20	0.17
Italy	2000	0.08	3	4.38	4.03	36.00	5.07	5.23	0.88	3.61	11.50	0.26	0.21
Spain	1406	0.15	3	3.45	2.81	22.00	5.60	5.73	0.94	3.89	10.61	0.27	0.25
Canada	927	0.12	2	3.14	2.65	20.00	5.36	5.49	0.87	3.81	9.50	0.16	0.17
Australia	891	0.07	3	3.79	3.43	24.00	5.17	5.27	0.81	3.77	8.54	0.23	0.20
Netherlands	756	0.05	3	4.22	4.03	34.00	4.56	4.67	0.69	3.31	7.65	0.26	0.18
Romania	659	0.37	2	2.79	2.29	22.00	4.58	4.76	1.04	3.06	8.20	0.30	0.33
Japan	562	0.10	2	3.35	3.19	18.00	4.88	5.00	0.82	3.39	9.23	0.20	0.21
Switzerland	487	0.18	2	2.91	2.48	14.00	5.78	6.02	1.14	4.02	10.30	0.23	0.27
Belgium	465	0.10	2	3.55	3.46	21.00	4.51	4.63	0.81	3.07	8.29	0.31	0.23
Sweden	442	0.12	3	3.57	3.03	18.00	4.64	4.76	0.85	3.17	9.05	0.24	0.20
Russia	441	0.54	1	2.86	3.80	30.00	2.71	2.93	0.81	1.77	5.16	0.27	0.42
Turkey	434	0.33	2	3.22	3.11	21.00	3.23	3.39	0.67	2.27	5.94	0.24	0.24
Brazil	409	0.12	2	3.14	2.72	19.00	4.61	4.76	0.82	3.29	8.24	0.19	0.18
Portugal	396	0.18	2	2.73	2.25	13.00	5.19	5.42	0.99	3.51	8.62	0.22	0.25
Colombia	330	0.11	2.5	3.78	3.53	28.00	4.12	4.22	0.81	2.65	6.90	0.27	0.20
Greece	308	0.22	2	2.84	2.36	16.00	4.17	4.49	1.10	2.68	9.88	0.25	0.21
China	272	0.56	1	2.14	2.10	18.00	3.04	3.01	0.64	1.87	5.07	0.23	0.23
India	267	0.66	1	2.06	1.68	15.00	2.10	2.18	0.58	1.32	4.26	0.16	0.20
Norway	261	0.11	3	4.02	3.24	22.00	4.05	4.28	1.00	2.78	8.20	0.27	0.23
Poland	260	0.40	2	2.93	2.47	16.00	3.14	3.23	0.70	1.95	5.29	0.35	0.38
Chile	247	0.06	3	3.49	2.79	16.00	4.78	5.30	1.62	3.51	12.44	0.19	0.17
Austria	246	0.07	3.5	6.05	6.03	27.00	3.53	3.73	0.82	2.46	8.13	0.34	0.45
Denmark	235	0.11	2	3.22	2.90	17.00	4.44	4.58	0.83	3.15	7.72	0.23	0.22
Czech Republic	215	0.29	2	3.39	3.47	24.00	2.69	2.82	0.58	1.82	4.58	0.24	0.25
Pakistan	204	0.26	2	3.17	3.00	20.00	2.92	3.06	0.56	1.98	4.93	0.32	0.25
South Africa	159	0.13	2	3.32	3.37	27.00	3.66	3.78	0.66	2.60	5.42	0.30	0.21

Source: own calculations based on RePEc data using *netsis* program in Stata. Note: both journal publication and working papers were used to determine collaborations; all calculations are unweighed; single-author publications are not included in the sample.

References

Goyal, Sanjeev, Marco J. van der Leij, and José Luis Moraga-González (2006). "Economics: An emerging small world". In: *Journal of Political Economy* 114.2, pp. 403–412.

Hamermesh, Daniel S. (2015). "Citations in Economics: Measurement, Uses and Impacts". NBER working paper 21754.

Head, Keith, Yao Amber Li, and Asier Minondo (2015). *Geography, Ties, and Knowledge Flows: Evidence from Citations in Mathematics*. Tech. rep. HKUST Institute for Emerging Market Studies.

Krichel, Thomas (2007). "What is the central matter of Economics? An investigation using JEL code centrality". Mimeo.

Krichel, Thomas and Christian Zimmermann (2009). "The Economics of Open Bibliographic Data Provision". In: *Economic Analysis and Policy* 39.1, pp. 143–152.

Miura, Hirotaka (2012). "Stata graph library for network analysis". In: *Stata Journal* 12.1, pp. 94–129.

Natural Earth (2016). *Natural Earth data.* http://www.naturalearthdata.com.

Pisati, Maurizio (2008). "SPMAP: Stata module to visualize spatial data". Mimeo.

Rath, Katharina and Klaus Wohlrabe (2015). "Trends in economics publications represented by JEL categories between 2007 and 2013". In: *Applied Economics Letters*, pp. 1–4.

Zimmermann, Christian (2013). "Academic Rankings with RePEc". In: *Econometrics* 1.3, pp. 249–280.

A   Technical details

A.1   Accessing the data

The data described in this document can be accessed for free. The information on how to access the data is available via the following link:

https://ideas.repec.org/getdata.html.

The information on country of affiliation for users is available via EDIRC (Economic Departments, Institutions and Research Centers), a specialized service. However, for the purposes of this document all of the addresses were geocoded using Google Maps API.

A.2   Recommended order of processing

RePEc data is distributed across a large number of files, so combining them into one dataset requires several independent steps. A recommended way to process the key information is given below.

1. A single academic work might exist in several versions, usually a working paper and the resulting journal publication. RePEc maintains information on all versions of an academic work, so without adjusting for such related works the number of publications and citation counts are likely to be inflated. So the recommended step in processing the data is to adjust the paper counts by combining related works. The information on related works is maintained by EconPapers, please visit the URL in Section A.1 to obtain the file.

2. Author profiles are the next step in processing of the data. By combining information from each author's profile it is possible to construct the collaboration network, identify author's affiliation and various author-specific statistics (e.g. number of journal publications and working papers). To get a more accurate representation of each author's profile it's best to correct for the information on related works from the previous step.

3. Citation information is the third step and it allows calculating additional author-, collaboration- or affiliation-specific statistics. For technical reasons, there are a few instances of multiple identifiers assigned to the same work. Luckily these identifiers are different only by one non-alphanumeric symbol at the end of the 'RePEc handle'. There are two ways of dealing with these: the simplest approach is to identify these instances and delete that one extra

character, while a more robust way is to identify these cases as 'related works' and add their information to the list in the first step.[7]

Some points to keep in mind:

- In the related works file there are some observations for translated versions of articles. These are identified as related works, as they should be, which makes identification of the 'final version' of an academic work a bit tricky. This identification does not matter for author-specific statistics (e.g. their overall citation count or h-index), but it matters for proper attribution of citations to the correct journal or working paper series.

A.3   Additional data used

The map diagrams were created using software described in Pisati (2008), with the data files taken from Natural Earth (2016).

---

[7]This may seem to violate the order of the steps presented, however this is merely a technical glitch which applies to a tiny fraction of the overall dataset: for the data processed in February 2016, there were 61 such identifiers out of roughly 650'000 RePEc handles in the citations file. Finally, this is likely to be fixed at some point in the future.