The network structure of success: Evidence from an empirical study of European patents

Alex Stivala[1] Alessandro Lomi[1,2]

[1]Università della Svizzera italiana, Lugano, Switzerland
[2]The University of Exeter Business School, Exeter, UK

Fifth Annual Australian Social Network Analysis Conference (ASNAC 2020). On-line conference hosted by the University of Western Australia. November 26-27, 2020.

Introduction

- One measure of the "success" of a patent is the number of citations it receives from other patents.
- These are known as "forward citations", and is just the in-degree in the citation network.
- Innovation involves the combination of knowledge in different ways.
- But not all possible combinations of knowledge are equally likely to succeed. So what factors contribute to success?
- We will use the ideas of *categorical contrast* and *niche width* (Hannan et al., 2007; Kovács and Hannan, 2010, 2015), as well as a new measure of technology class boundary crossing, to try to answer this question.
- We will use both negative binomial regression and ERGM, as appropriate, to test hypotheses.

Contrast (Kovács and Hannan, 2010)

- The *contrast* of a category captures the idea of sharpness or fuzziness of category boundaries:
  - A category has high contrast (sharp boundaries) if it is seldom assigned low or moderate levels of category membership.
  - A category has lower contrast (fuzzier boundaries) as partial membership is more common.
- A technology class that is seldom assigned together with other classes to a patent has high contrast.
- A technology class that is frequently assigned together with other classes to a patent has low contrast.
- Contrast is defined as the average grade-of-membership (GoM) in a category, for those with nonzero GoM.
  - When the category membership is binary (as in patent technology classes), then for each patent GoM is just 0 if the patent does not have that class, and $1/K_p$ when it does, where K_p is the number of categories assigned to patent p.

Niche width (Hannan et al., 2007; Kovács and Hannan, 2010)

- Niche width captures the idea of breadth:
  - A patent with high niche width spans many categories (technology classes); it is generalist.
  - A patent with a single technology class has a niche width of 0; it is specialized.
- The *niche width* of a patent is the Simpson diversity index of the GoM vector.
- Equivalently, $1 - H$ where H is the Herfindahl concentration index.
- For binary memberships as used here, niche width is just $1 - 1/K_p$.

Assigned technology classes or cited technology classes?

- Patents are assigned technology classes by the patent office.
- In our data, multiple classes can be assigned.
- So GoM can be defined in two ways:
  - By the set of technology classes assigned to a patent.
  - By the set of technology classes assigned to the patents cited by a patent.
- The latter is claimed to better capture the combination of knowledge by a patent (Gruber et al., 2013; Ferguson and Carnabuci, 2017).
- We will use both.
- When niche width is defined by classes of cited patents, it is the same as the "originality" of Trajtenberg et al. (1997); Hall et al. (2001).

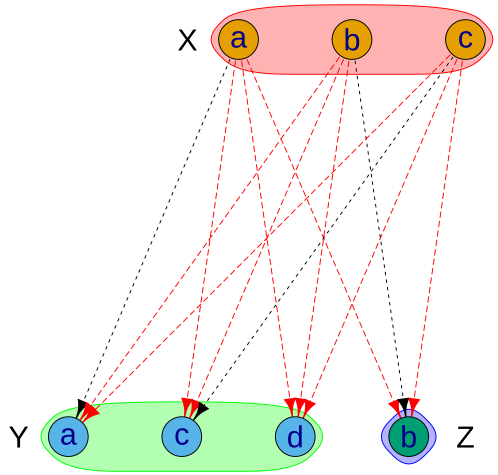
Class crossing ratio I

- Niche width is a monotonic function of the number of technology classes, so it captures just *breadth* and not *diversity* as such.
- We define the *class crossing ratio* to capture a particular idea of diversity or "boundary crossing":
  - Consider each citation as an arc between each of the classes in the citing patent to each of the classes in the cited patents.
  - The class crossing ratio is the ratio of the number of these virtual arcs which join different classes, to the total number of virtual arcs.
- So class crossing ratio is high when a patent cites patents that have different technology classes than those it is assigned itself.

Class crossing ratio II

- This is conceptually different from the *typicality* measure of Ferguson and Carnabuci (2017) which measures similarity among sets of technology classes assigned to the cited patents only, with a Jaccard index.
- It is also different from Jaccard similarity between classes of citing patent and union of classes of cited patents.

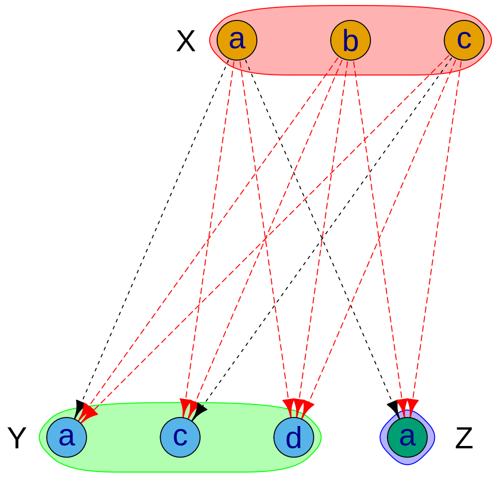
Class crossing ratio illustration 1



Class crossing ratio $= 9/12 = 0.75$

$J(X, Y \cup Z) = 3/4 = 0.75$

Class crossing ratio illustration 2



Class crossing ratio $= 9/12 = 0.75$

$J(X, Y \cup Z) = 2/4 = 0.5$

Hypotheses I

H0 **Success (citations received) increases with breadth.**
- This is measured by niche width.
- "... the positive association between recombinant breadth and citation impact is one of the most frequently replicated findings in innovation research..." (Ferguson and Carnabuci, 2017, p. 134).

H1 **Success (citations received) increases with diversity.**
- Compare with Uzzi et al. (2013), the highest-impact science has atypical combinations grounded in conventional combinations; and
- Ferguson and Carnabuci (2017), patents with "more typical" combinations receiver fewer citations.
- Instead we measure technology class *diversity* or "boundary crossing" here with class crossing ratio.

H2 **Success increases with maximum contrast of technology classes.**

Hypotheses II

- Higher contrast categories are easier to interpret; lower contrast can lead to confusion about categories (Hannan et al., 2007; Kovács and Hannan, 2010, 2015).

H3 **But spanning high contrast categories makes success less likely.**
- Membership in more than one high-contrast category can also lead to confusion (Kovács and Hannan, 2010, 2015).
- This can be tested by a negative effect for secondary contrast, that is, the second-largest contrast (Kovács and Hannan, 2015).

H4 **Patents with high maximum contrast are unlikely to cite other patents with high maximum contrast.**
- A patent with a very sharply defined category (rarely combined with other categories) is more likely to cite patents with less sharply defined categories, combining knowledge from categories that are more often combined.

Hypotheses III

H5 **(Geographical knowledge spillover): citations are more likely to be geographically localized.**
- Jaffe et al. (1993); Thompson and Fox-Kean (2005); Henderson et al. (2005); Stivala et al. (2019a).

Data source

- The patent data is from the Information Retrieval Facility `https://www.ir-facility.org/`
- We used the MAREC (Matrixware Research Collection), of over 19 million patents from 1976 – 2008. `https://www.ir-facility.org/prototypes/marec`
- Specifically we used patents (applications and granted) from the European Patent Office (EPO).
- We extracted bibliographic data for 1 933 231 unique patents from the full text XML data.
- From this a 1 933 231 node citation network is built.
- 149 instances of self-loops are removed.
- Including nodes for patents cited from patents in that data (but for which we have no data other than a unique identifier), a 4 903 886 node citation network is built.
- But this larger network has no attribute data for 61% of the nodes.

Patent technology classifications

- The International Patent Classification (IPC) scheme is hierarchical.
- The highest level is Section (of which there are 8).
- There are then 120 classes and 600 subclasses.
- E.g. Section H is "Electricity" and class H01 is "basic electric elements".
- We will use Section and Class levels.
- Note that the EPO (unlike the USPTO data e.g. from NBER) allows multiple sections and classes to be assigned to a patent.
- Also the EPO assigns classes based on the entire application, not just the "claims" so is determined objectively by the examiner (Gruber et al., 2013).

Summary statistics of the patent data

Statistic	N	Mean	St. Dev.	Min	Max
Forward citations	1933231	0.573	1.448	0	76
App. Year [base 1978]	1933231	18.442	7.297	0	30
Niche width	1928684	0.236	0.282	0.000	0.929
Max. contrast	1928684	0.659	0.064	0.305	0.812
Secondary contrast	817292	0.586	0.071	0.305	0.766
Contrast share	1928684	0.779	0.265	0.087	1.000
Contrast variance	817292	0.006	0.006	0.000	0.086
Num. classes	1933231	1.595	0.841	1	14
Num. subclasses	1933231	1.934	1.190	1	20
Backward citations (subgraph)	1933231	0.573	1.029	0	117
Cited max. contrast	650656	0.666	0.060	0.383	0.812
Cited secondary contrast	374032	0.599	0.070	0.305	0.766
Cited contrast variance	452945	0.004	0.005	0.000	0.086
Cited contrast share	650656	0.680	0.289	0.080	1.000
Class crossing ratio	650511	0.414	0.311	0.000	1.000
Cited niche width	650866	0.325	0.293	0.000	0.923
Num. sections	1933231	1.370	0.579	1	7
Backward citations (all)	1933231	3.251	2.911	1	142

There are 8 technology sections (highest level IPC classification), and at the next level, 123 technology classes. A patent can be assigned multiple classes and multiple sections.

Summary statistics of IPC sections

IPC Section	Description	N
A	Human necessities	405804
B	Performing operations; transporting	497492
C	Chemistry; metallurgy	464874
D	Textiles; paper	54695
E	Fixed constructions	78438
F	Mechanical engineering; lighting; heating ...	227017
G	Physics	477022
H	Electricity	438685
Y	General ...	0

Note that a patent need not be assigned to only a single section; the sections are not mutually exclusive.

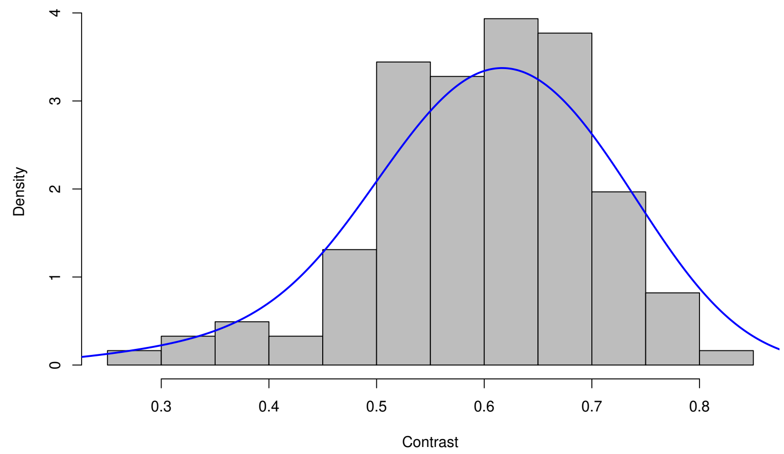
Summary statistics of the patent citation network

Description	N	Components	Giant component	Mean degree	Density
EPO (full)	4903886	746741	3789545	2.30	0.0000002
EPO (subgraph)	1933231	1119794	673306	1.15	0.0000003

Description	Reciprocity	Clustering coefficient	Assortativity coefficient
EPO (full)	0.0005	0.03125	0.08300
EPO (subgraph)	0.0025	0.07862	0.13231

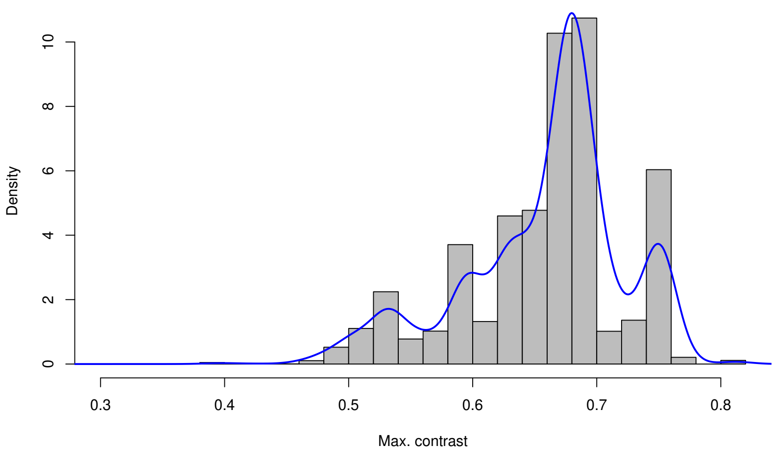
The “full” network is the network containing not only patents in the data set, but also nodes representing patents outside the data set, but which are cited by a patent in the data set. The “subgraph” network is the network induced by only those nodes in the data set itself.

Distribution of contrast values of technology classes

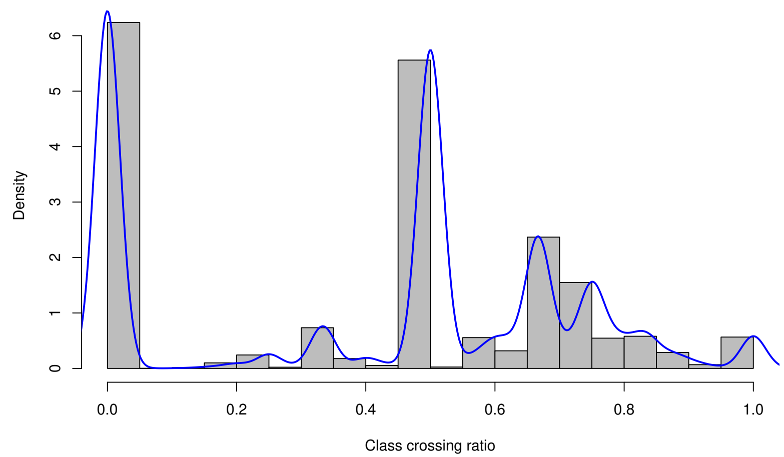


The highest value of contrast (0.812) is for A43 (footwear), and the lowest value (0.250) is for C99 (chemistry; metallurgy).

Distribution of maximum contrast value of patents

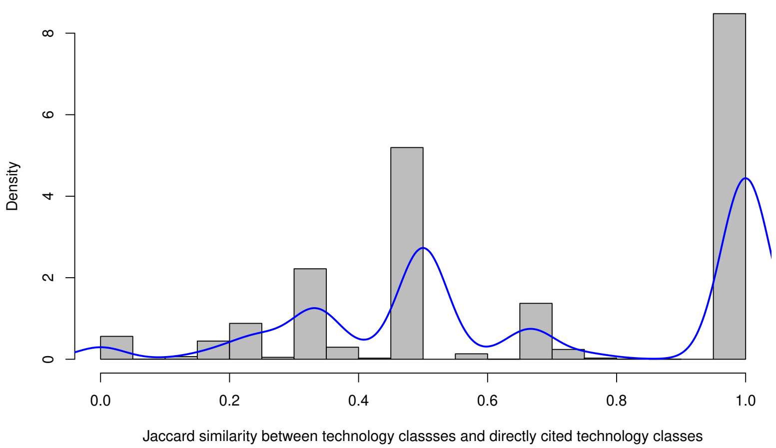


Distribution of class crossing ratio of patents



The class crossing ratio of a patent is the number of backward citations that represent a direct citation from a class assigned to the patent, to a different class in the cited patent, divided by the total number of possible class citations (to both the same or different classes).

Distribution of technology class Jaccard similarity



Distribution of the Jaccard similarity between the sets of technology classes assigned to a patent, and the union of the sets of technology classes assigned to the backward citations (directly cited patents) of the patent.
$N = 650511$, median $= 0.667$, mean $= 0.674$, sd $= 0.307$.

Negative binomial models, citations as response variable I

	Model 1	Model 2	Model 3
App. Year [base 1978]	$-0.12\ (0.00)^{***}$	$-0.12\ (0.00)^{***}$	$-0.12\ (0.00)^{***}$
Section A	$-0.24\ (0.01)^{***}$	$-0.31\ (0.01)^{***}$	$-0.31\ (0.01)^{***}$
Section B	$0.04\ (0.00)^{***}$	$-0.05\ (0.01)^{***}$	$-0.04\ (0.01)^{***}$
Section C	$0.25\ (0.00)^{***}$	$0.15\ (0.01)^{***}$	$0.14\ (0.01)^{***}$
Section D	$0.07\ (0.01)^{***}$	$-0.00\ (0.01)$	$-0.01\ (0.01)$
Section E	$-0.39\ (0.01)^{***}$	$-0.46\ (0.01)^{***}$	$-0.45\ (0.01)^{***}$
Section F	$-0.07\ (0.01)^{***}$	$-0.15\ (0.01)^{***}$	$-0.14\ (0.01)^{***}$
Section G	$0.19\ (0.00)^{***}$	$0.11\ (0.01)^{***}$	$0.10\ (0.01)^{***}$
Section H	$0.17\ (0.01)^{***}$	$0.09\ (0.01)^{***}$	$0.09\ (0.01)^{***}$
Pub. Language German	$-0.29\ (0.00)^{***}$	$-0.29\ (0.00)^{***}$	$-0.31\ (0.00)^{***}$
Pub. Language French	$-0.31\ (0.01)^{***}$	$-0.31\ (0.01)^{***}$	$-0.32\ (0.01)^{***}$
Backward citations (all)	$0.17\ (0.00)^{***}$	$0.17\ (0.00)^{***}$	$0.17\ (0.00)^{***}$
Max. contrast	$-2.36\ (0.44)^{***}$	$-2.70\ (0.44)^{***}$	$-2.68\ (0.44)^{***}$
Max. contrast^2	$3.45\ (0.34)^{***}$	$3.65\ (0.34)^{***}$	$3.61\ (0.35)^{***}$
Niche width		$0.22\ (0.01)^{***}$	$0.23\ (0.01)^{***}$
Appplicant Switzerland			$-0.05\ (0.02)^{**}$
Inventor Switzerland			$-0.07\ (0.03)^{**}$
Appplicant Switzerland $\times$ Inventor Switzerland			$0.27\ (0.03)^{***}$
Cited max. contrast			
Cited max. contrast^2			
Cited niche width			
AIC	3331171.47	3330604.41	3248519.42
BIC	3331371.01	3330816.43	3248768.46
Log Likelihood	-1665569.73	-1665285.20	-1624239.71
Deviance	1181391.34	1181445.64	1157693.65
Num. obs.	1927639	1927639	1889616

Negative binomial models, citations as response variable II

	Model 4	Model 5	Model 6
App. Year [base 1978]	$-0.11\ (0.00)^{***}$	$-0.11\ (0.00)^{***}$	$-0.11\ (0.00)^{***}$
Section A	$-0.14\ (0.01)^{***}$	$-0.14\ (0.01)^{***}$	$-0.14\ (0.01)^{***}$
Section B	$-0.00\ (0.01)$	$-0.01\ (0.01)$	$-0.00\ (0.01)$
Section C	$0.11\ (0.01)^{***}$	$0.10\ (0.01)^{***}$	$0.10\ (0.01)^{***}$
Section D	$0.01\ (0.01)$	$0.01\ (0.01)$	$0.01\ (0.01)$
Section E	$-0.35\ (0.01)^{***}$	$-0.35\ (0.01)^{***}$	$-0.35\ (0.01)^{***}$
Section F	$-0.04\ (0.01)^{***}$	$-0.05\ (0.01)^{***}$	$-0.04\ (0.01)^{***}$
Section G	$0.06\ (0.01)^{***}$	$0.06\ (0.01)^{***}$	$0.05\ (0.01)^{***}$
Section H	$-0.03\ (0.01)^{***}$	$-0.03\ (0.01)^{***}$	$-0.03\ (0.01)^{***}$
Pub. Language German	$-0.34\ (0.01)^{***}$	$-0.34\ (0.01)^{***}$	$-0.36\ (0.01)^{***}$
Pub. Language French	$-0.34\ (0.01)^{***}$	$-0.34\ (0.01)^{***}$	$-0.34\ (0.01)^{***}$
Backward citations (all)	$0.04\ (0.00)^{***}$	$0.04\ (0.00)^{***}$	$0.04\ (0.00)^{***}$
Max. contrast	$-2.48\ (0.72)^{***}$	$-2.63\ (0.71)^{***}$	$-2.63\ (0.72)^{***}$
Max. contrast^2	$2.88\ (0.57)^{***}$	$3.21\ (0.57)^{***}$	$3.19\ (0.58)^{***}$
Niche width	$0.23\ (0.01)^{***}$	$0.18\ (0.01)^{***}$	$0.18\ (0.01)^{***}$
Appplicant Switzerland			$-0.07\ (0.02)^{**}$
Inventor Switzerland			$-0.05\ (0.03)$
Appplicant Switzerland $\times$ Inventor Switzerland			$0.23\ (0.04)^{***}$
Cited max. contrast	$0.01\ (0.75)$	$-0.02\ (0.75)$	$0.01\ (0.76)$
Cited max. contrast^2	$0.78\ (0.59)$	$0.55\ (0.59)$	$0.53\ (0.60)$
Cited niche width		$0.11\ (0.01)^{***}$	$0.11\ (0.01)^{***}$
AIC	1615185.10	1615025.57	1579868.81
BIC	1615401.42	1615253.28	1580130.28
Log Likelihood	-807573.55	-807492.79	-789911.40
Deviance	548718.44	548738.71	538346.76
Num. obs.	650434	650434	639387

Negative binomial models with secondary contrast I

	Model 1	Model 2	Model 3
App. Year [base 1978]	$-0.10\ (0.00)^{***}$	$-0.10\ (0.00)^{***}$	$-0.10\ (0.00)^{***}$
Section A	$-0.19\ (0.01)^{***}$	$-0.21\ (0.01)^{***}$	$-0.21\ (0.01)^{***}$
Section B	$0.02\ (0.01)^{**}$	$-0.00\ (0.01)$	$-0.00\ (0.01)$
Section C	$0.11\ (0.01)^{***}$	$0.07\ (0.01)^{***}$	$0.07\ (0.01)^{***}$
Section D	$0.07\ (0.02)^{***}$	$0.04\ (0.02)^{*}$	$0.04\ (0.02)^{*}$
Section E	$-0.33\ (0.02)^{***}$	$-0.34\ (0.02)^{***}$	$-0.34\ (0.02)^{***}$
Section F	$0.03\ (0.01)^{**}$	$0.01\ (0.01)$	$0.01\ (0.01)$
Section G	$0.08\ (0.01)^{***}$	$0.06\ (0.01)^{***}$	$0.05\ (0.01)^{***}$
Section H	$-0.02\ (0.01)$	$-0.04\ (0.01)^{***}$	$-0.04\ (0.01)^{***}$
Pub. Language German	$-0.27\ (0.01)^{***}$	$-0.27\ (0.01)^{***}$	$-0.29\ (0.01)^{***}$
Pub. Language French	$-0.26\ (0.01)^{***}$	$-0.26\ (0.01)^{***}$	$-0.27\ (0.01)^{***}$
Backward citations (subgraph)	$0.16\ (0.00)^{***}$	$0.16\ (0.00)^{***}$	$0.16\ (0.00)^{***}$
Max. contrast	$-1.00\ (0.97)$	$-1.43\ (0.98)$	$-1.67\ (0.99)$
Max. contrast^2	$2.39\ (0.76)^{**}$	$2.73\ (0.76)^{***}$	$2.90\ (0.77)^{***}$
Class crossing ratio	$3.01\ (0.25)^{***}$	$2.39\ (0.26)^{***}$	$2.41\ (0.27)^{***}$
Class crossing ratio^2	$-2.34\ (0.19)^{***}$	$-2.00\ (0.19)^{***}$	$-2.01\ (0.19)^{***}$
Secondary contrast	$-5.73\ (0.77)^{***}$	$-6.03\ (0.77)^{***}$	$-5.77\ (0.78)^{***}$
Secondary contrast^2	$5.14\ (0.67)^{***}$	$5.25\ (0.67)^{***}$	$5.03\ (0.68)^{***}$
Niche width		$0.49\ (0.05)^{***}$	$0.50\ (0.05)^{***}$
Appplicant Switzerland			$-0.04\ (0.03)$
Inventor Switzerland			$-0.08\ (0.05)$
Appplicant Switzerland $\times$ Inventor Switzerland			$0.23\ (0.06)^{***}$
Cited max. contrast			
Cited max. contrast^2			
Cited secondary contrast			
Cited secondary contrast^2			
Cited niche width			
AIC	761913.27	761804.30	745025.12
BIC	762124.45	762026.05	745278.11
Log Likelihood	-380936.63	-380881.15	-372488.56
Deviance	251830.34	251837.63	247074.61
Num. obs.	284767	284767	279728

Negative binomial models with secondary contrast II

	Model 4	Model 5	Model 6
App. Year [base 1978]	−0.10 (0.00)***	−0.10 (0.00)***	−0.10 (0.00)***
Section A	−0.21 (0.01)***	−0.21 (0.01)***	−0.21 (0.01)***
Section B	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Section C	0.07 (0.01)***	0.07 (0.01)***	0.06 (0.01)***
Section D	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
Section E	−0.34 (0.02)***	−0.35 (0.02)***	−0.35 (0.02)***
Section F	0.02 (0.01)	0.02 (0.01)	0.03 (0.01)*
Section G	0.06 (0.01)***	0.06 (0.01)***	0.06 (0.01)***
Section H	−0.05 (0.01)***	−0.05 (0.01)***	−0.05 (0.01)***
Pub. Language German	−0.26 (0.01)***	−0.26 (0.01)***	−0.28 (0.01)***
Pub. Language French	−0.26 (0.02)***	−0.26 (0.02)***	−0.26 (0.02)***
Backward citations (subgraph)	0.14 (0.00)***	0.14 (0.00)***	0.14 (0.00)***
Max. contrast	−0.45 (1.42)	−0.22 (1.43)	−0.54 (1.45)
Max. contrast^2	1.61 (1.12)	1.49 (1.13)	1.75 (1.14)
Class crossing ratio	1.76 (0.32)***	1.36 (0.32)***	1.35 (0.33)***
Class crossing ratio^2	−1.66 (0.23)***	−1.46 (0.24)***	−1.46 (0.24)***
Secondary contrast	−4.76 (0.95)***	−4.99 (0.95)***	−4.85 (0.96)***
Secondary contrast^2	4.13 (0.83)***	4.46 (0.83)***	4.35 (0.84)***
Niche width	0.64 (0.06)***	0.62 (0.06)***	0.63 (0.06)***
Appplicant Switzerland			−0.04 (0.04)
Inventor Switzerland			−0.07 (0.06)
Appplicant Switzerland × Inventor Switzerland			0.25 (0.07)***
Cited max. contrast	−3.32 (1.54)*	−3.63 (1.55)*	−3.44 (1.57)*
Cited max. contrast^2	3.09 (1.20)*	3.29 (1.21)**	3.13 (1.23)*
Cited secondary contrast	−1.47 (1.07)	−1.52 (1.07)	−1.76 (1.08)
Cited secondary contrast^2	1.31 (0.92)	1.15 (0.92)	1.35 (0.93)
Cited niche width		0.31 (0.05)***	0.31 (0.05)***
AIC	597762.41	597712.34	584577.96
BIC	598019.71	597979.92	584875.90
Log Likelihood	−298856.21	−298830.17	−292259.98
Deviance	195603.94	195596.40	191914.04
Num. obs.	217890	217890	214014

Negative binomial models with secondary contrast III

ERGM conditional estimation, 4 903 886 node network I

Effect	Model 1	Model 2	Model 3
Arc	−12.831 (−13.152, −12.509)	−13.367 (−13.656, −13.079)	−13.188 (−13.501, −12.876)
Isolates	3.236 (2.888, 3.583)	3.292 (3.069, 3.514)	3.144 (2.927, 3.362)
Sink	0.936 (0.771, 1.100)	0.764 (0.584, 0.944)	0.604 (0.437, 0.771)
Source	−0.471 (−0.553, −0.389)	−0.424 (−0.448, −0.401)	−0.417 (−0.460, −0.373)
Popularity spread (AinS)	1.135 (1.016, 1.254)	1.021 (0.985, 1.056)	1.054 (0.954, 1.154)
Activity spread (AoutS)	−0.129 (−0.163, −0.095)	0.119 (0.080, 0.158)	0.260 (0.211, 0.309)
Two-path (A2P-T)	0.018 (0.009, 0.028)	0.024 (0.014, 0.034)	0.032 (0.023, 0.042)
Shared popularity (A2P-D)	0.029 (0.018, 0.040)	0.027 (0.018, 0.037)	0.027 (0.017, 0.037)
Shared activity (A2P-U)	0.048 (0.032, 0.064)	0.035 (0.031, 0.040)	0.025 (0.019, 0.032)
Sender App. Year [base 1978]	0.473 (0.458, 0.488)	0.450 (0.430, 0.470)	0.474 (0.454, 0.493)
Receiver App. Year [base 1978]	−0.532 (−0.551, −0.513)	−0.500 (−0.524, −0.476)	−0.512 (−0.536, −0.487)
DiffSign App. Year	1.910 (1.713, 2.107)	2.132 (2.015, 2.249)	2.118 (2.007, 2.230)
AbsDiff App. Year	−0.673 (−0.704, −0.642)	−0.614 (−0.644, −0.584)	−0.625 (−0.657, −0.593)
Jaccard similarity Applicant countries	0.825 (0.652, 0.998)	0.783 (0.605, 0.960)	0.760 (0.588, 0.932)
Jaccard similarity Inventor countries	0.552 (0.388, 0.717)	0.495 (0.365, 0.626)	0.474 (0.315, 0.632)
Jaccard similarity Sections	4.061 (3.696, 4.426)	1.449 (1.337, 1.561)	1.337 (1.179, 1.496)
Matching Pub. Language	0.216 (0.124, 0.309)	0.174 (0.099, 0.249)	0.103 (0.039, 0.166)

ERGM conditional estimation, 4 903 886 node network II

Sender Max. contrast	−2.874 (−3.169, −2.580)	−3.649 (−3.944, −3.355)	−6.471 (−6.734, −6.208)
Sender Max. contrast^2	−0.036 (−0.320, 0.247)	0.376 (0.184, 0.568)	2.499 (2.355, 2.644)
Receiver Max. contrast	−7.182 (−7.636, −6.728)	−6.953 (−7.230, −6.676)	−9.947 (−10.309, −9.586)
Receiver Max. contrast^2	5.552 (5.098, 6.005)	4.538 (4.287, 4.789)	6.379 (6.036, 6.722)
Jaccard similarity Classes	—	5.215 (4.919, 5.510)	6.466 (6.141, 6.791)
DiffSign Max. contrast	0.005 (−0.007, 0.017)	—	—
AbsDiff Max. contrast	−17.307 (−18.407, −16.207)	—	—
Sender Niche width	—	—	1.780 (1.724, 1.836)
Receiver Niche width	—	—	2.181 (1.937, 2.425)
Sender Secondary contrast	—	—	—
Sender Secondary contrast^2	—	—	—
Receiver Secondary contrast	—	—	—
Receiver Secondary contrast^2	—	—	—
Converged runs	20	20	20
Total runs	20	20	20

ERGM conditional estimation, 4 903 886 node network III

Effect	Model 4
Arc	−12.952 (−13.332, −12.573)
Isolates	3.164 (2.928,3.401)
Sink	0.648 (0.460,0.835)
Source	−0.425 (−0.500, −0.350)
Popularity spread (AinS)	1.061 (0.975,1.148)
Activity spread (AoutS)	0.207 (0.158,0.255)
Two-path (A2P-T)	0.030 (0.018,0.041)
Shared popularity (A2P-D)	0.028 (0.016,0.039)
Shared activity (A2P-U)	0.027 (0.017,0.037)
Sender App. Year [base 1978]	0.468 (0.446,0.490)
Receiver App. Year [base 1978]	−0.507 (−0.535, −0.479)
DiffSign App. Year	2.107 (1.959,2.255)
AbsDiff App. Year	−0.623 (−0.658, −0.589)
Jaccard similarity Applicant countries	0.739 (0.562,0.916)
Jaccard similarity Inventor countries	0.471 (0.326,0.617)
Jaccard similarity Sections	1.317 (1.149,1.485)
Matching Pub. Language	0.111 (0.025,0.197)

ERGM conditional estimation, 4 903 886 node network IV

Sender Max. contrast	−5.497 (−5.831, −5.162)
Sender Max. contrast^2	0.772 (0.586,0.958)
Receiver Max. contrast	−8.115 (−8.459, −7.771)
Receiver Max. contrast^2	3.496 (3.224,3.769)
Jaccard similarity Classes	6.570 (6.219,6.921)
DiffSign Max. contrast	—
AbsDiff Max. contrast	—
Sender Niche width	1.614 (1.374,1.854)
Receiver Niche width	2.071 (1.823,2.320)
Sender Secondary contrast	−3.218 (−3.444, −2.991)
Sender Secondary contrast^2	5.695 (5.395,5.994)
Receiver Secondary contrast	−3.834 (−4.106, −3.563)
Receiver Secondary contrast^2	6.676 (6.131,7.221)
Converged runs	20
Total runs	20

Results for hypotheses I

H0 **Success (citations received) increases with breadth.**

- Confirmed by significant positive niche width estimate in negative binomial models.
- Note also significant positive backward citation effect in negative binomial models: another (cruder) measure of breadth, the number of citations a patent makes.
- Also confirmed in ERGM by significant positive receiver effect for niche width.

H1 **Success (citations received) increases with diversity.**

- We included a quadratic term for for diversity, as was done for max. contrast (following Kovács and Hannan (2010) who find a quadratic relationship for max. contrast).
- Partly confirmed: there is a quadratic relationship between class crossing ratio and success, with success increasing with class crossing ratio up to a point, after which it negatively affects success.

Results for hypotheses II

H2 **Success increases with maximum contrast of technology classes.**

- Partly confirmed: there is a quadratic relationship between success and max. contrast, with success decreasing with maximum contrast up to a point, but increasing thereafter.
- This applies for both maximum contrast of a patent's classes, and of maximum contrast of its cited patents' classes.
- The ERGM models also show a similar pattern with the Receiver effect on max. contrast.

H3 **But spanning high contrast categories makes success less likely.**

- Partly confirmed: there is a quadratic relationship between success and secondary contrast, with success decreasing with secondary contrast only up to a point, after which it increases.
- There is a similar pattern in the ERGM for the Receiver effect for secondary contrast.

H4 **Patents with high maximum contrast are unlikely to cite other patents with high maximum contrast.**

Results for hypotheses III

- Contradicted: In the ERGM model the effect for heterophily (AbsDiff) on max. contrast is negative and significant.
- DiffSign is not significant.
- It seems that, contrary to H4, there is significant homophily on max. contrast.
- Is this a poor test of H4, as it is confounded by patents citing patents with the same technology class?
  - Positive significant Jaccard similarity of technology class sets in all models in which it is included (unsurprising: patents cite other patents in the same technology classes).
  - Note ERGM parameter estimation does not converge well with both Jaccard similarity of technology classes and the AbsDiff effect for max. contrast included.

H5 **(Geographical knowledge spillover): citations are more likely to be geographically localized.**

- Confirmed: The effect for Jaccard similarity is positive and significant for both applicant countries and inventor countries in all ERGM models.

Acknowledgments

- This work was funded by Swiss National Science Foundation NRP 75 Big Data project 167326 "The Global Structure of Knowledge Networks: Data, Models and Empirical Results".
- We thank Mr Manajit Chakraborty and Prof. Fabio Crestani for assisting with access to patent data.
- We used the high performance computing cluster at the Institute of Computational Science, Università della Svizzera italiana, for all data processing and statistical computations.

Unpublished work

- This is unpublished work (as of November 2020).
- Details including methods and references are in the "hidden bonus slides" after this one.
- I will make these slides available on my website:
- `https://sites.google.com/site/alexdstivala/home/conferences`

Hidden bonus slides

CPC technology sections

- A Human necessities
- B Performing operations; transporting
- C Chemistry; metallurgy
- D Textiles; paper
- E Fixed constructions
- F Mechanical engineering; lighting; heating; weapons; blasting engines or pumps
- G Physics
- H Electricity
- Y General tagging of new technological developments ...

https:
//www.epo.org/searching-for-patents/helpful-resources/first-time-here/classification/cpc.html

Jaccard similarity

The Jaccard similarity $0 \leq J(A, B) \leq 1$ between two sets is the size of their intersection over the size of their union:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

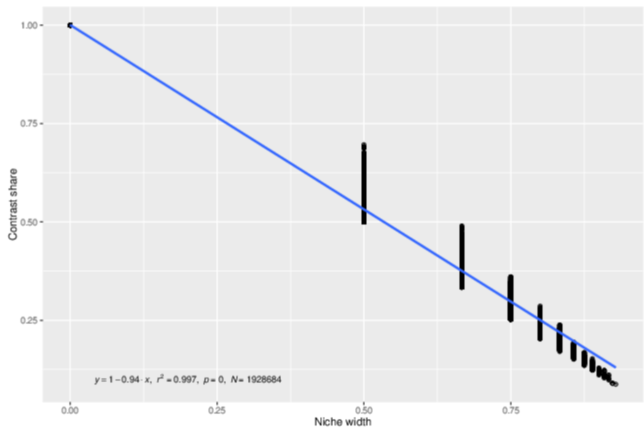
If $|A \cup B| = 0$ i.e. A and B are both empty, then define $J(A, B) = 1$.

Class crossing ratio example

- Assume patent X has classes a,b,c and it cites patent Y with classes a,c,d and patent Z with class b only
- We consider the total of $3 \times 3 + 3 \times 1 = 12$ virtual ties (a–a, a–c, a–d, b–a, b–c, ... , c–b)
- Of these 12 virtual ties 9 are "boundary crossing" (a–c, a–d, b–a, ..., but not a–a, c–c, b–b, ...)
- So we would give it a boundary crossing score of $9/12 = 0.75$
- (In R we can do this using the vector outer product.)
- Note that this is like a kind of generalized E-I index (Krackhardt and Stern, 1988)
- Although it is in $[0, 1]$ not $[-1, +1]$ — to make it more like E-I index we would have the numerator as (mismatching - matching) not just mismatching, applicable to sets of categories on nodes, rather than just a simple nodal categorical variable.

Contrast share

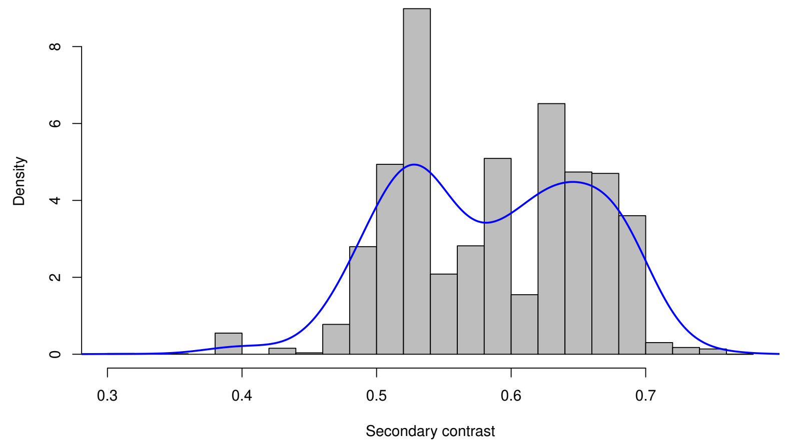
- *Contrast share* is the ratio of the maximum contrast of assigned categories to their sum (Kovács and Hannan, 2010).
- In our data, contrast share is highly inversely correlated with niche width, so we use only niche width.



Summary statistics of publication languages

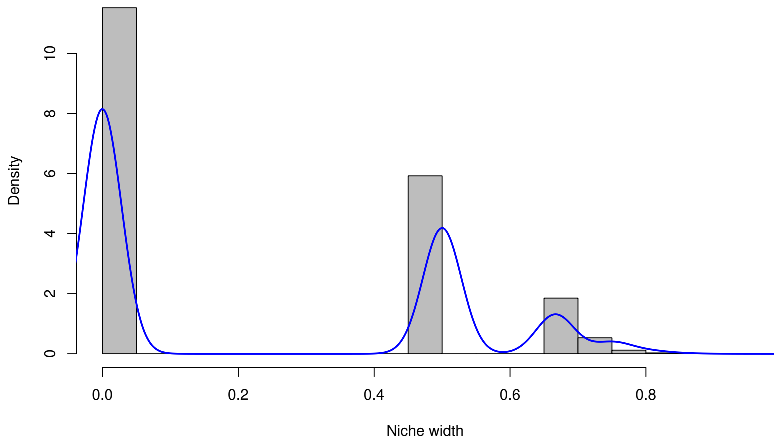
Language	N
English	1355416
German	435373
French	141397
NA	1045

Distribution of secondary contrast value of patents

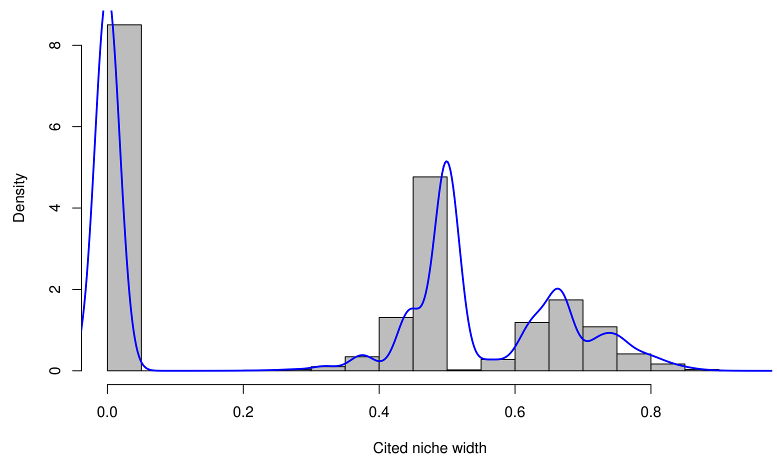


For each patent, the second-largest contrast of the classes it is assigned.

Distribution of niche width values of patents

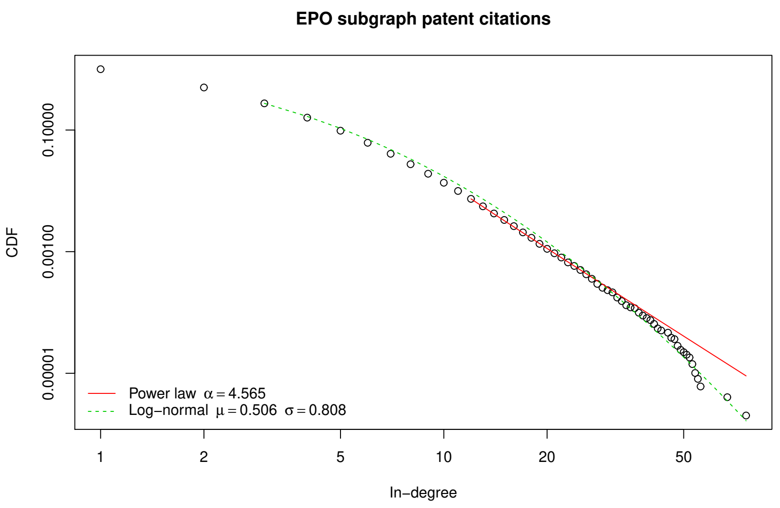


Distribution of cited niche width values of patents



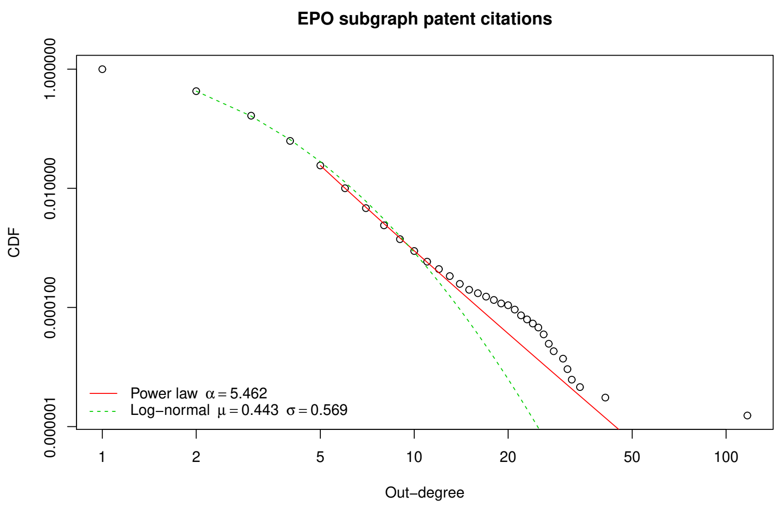
The niche width defined over the classes of the directed cited patents of a patent, rather than the classes assigned to the patent itself.

Citation network in-degree distribution



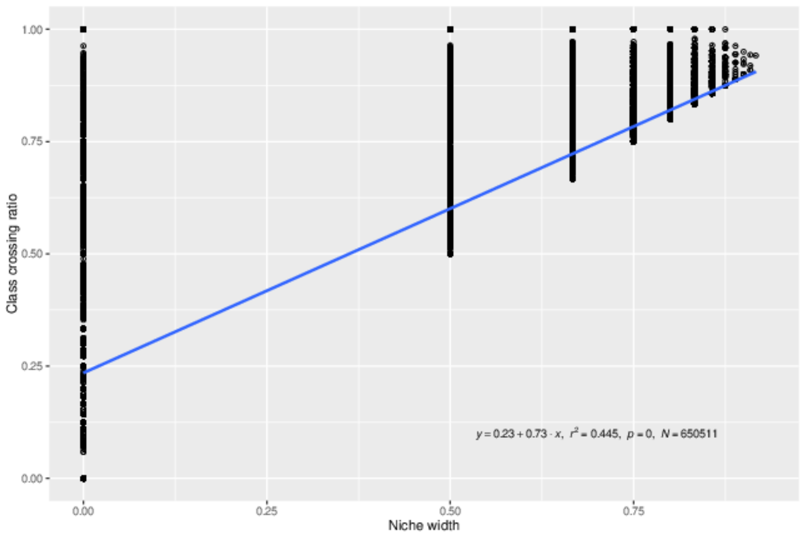
The in-degree distribution is consistent with neither a power law ($p < 0.05$) nor a log-normal distribution ($p < 0.05$).

Citation network out-degree distribution

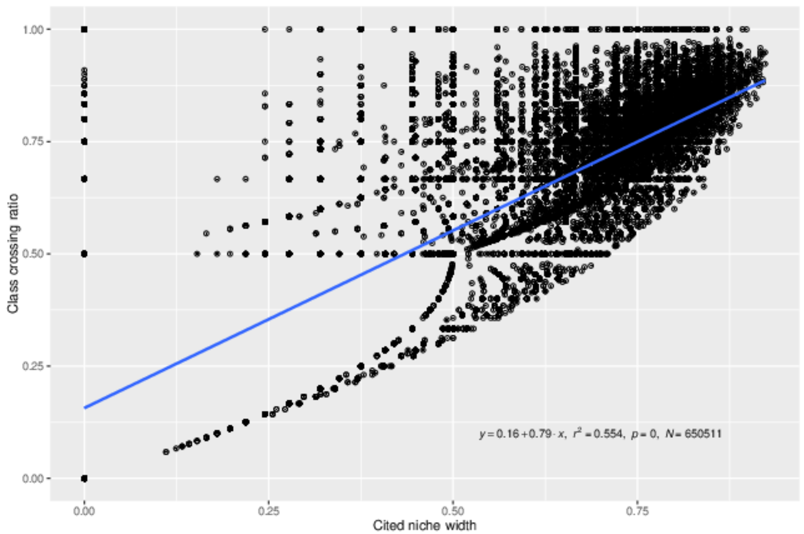


The out-degree distribution is consistent with neither a power law ($p < 0.01$) nor a log-normal distribution ($p < 0.001$).

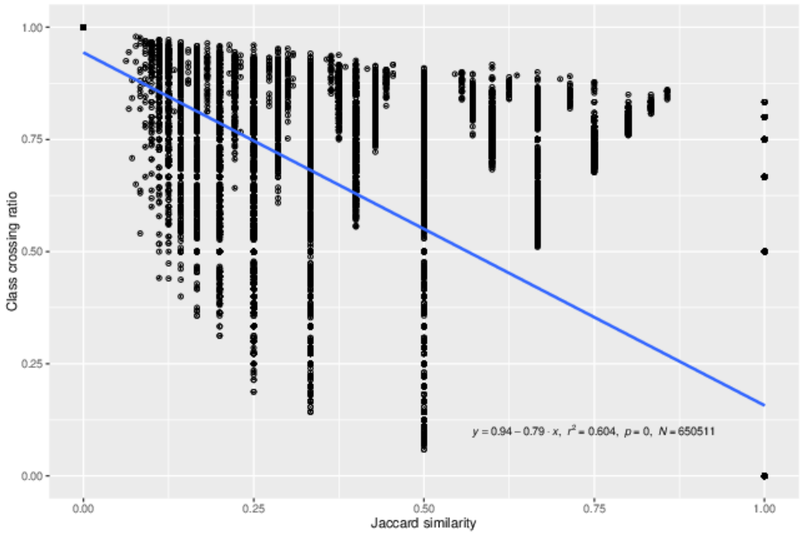
Linear correlation between niche width and class crossing ratio of patents



Linear correlation between cited niche width and class crossing ratio of patents



Linear correlation between class crossing ratio and Jaccard similarity between technology classes and union of directly cited technology classes



Methods I

- Power law and log-normal distributions were fitted using the methods of Clauset et al. (2009) implemented in the poweRlaw package (Gillespie, 2015).
- Negative binomial regression models were estimated using the MASS (Venables and Ripley, 2002) and formatted with the texreg (Leifeld, 2013) packages in R (R Core Team, 2016). Robust standard errors (Hinkley, 1977; MacKinnon and White, 1985) were estimated with the sandwich (Zeileis, 2004, 2006) and lmtest (Zeileis and Hothorn, 2002) packages in R. Residual diagnostics from the DHARMa R package (Hartig, 2019).
- ERGM models were estimated with EstimNetDirected (Byshkin et al., 2018; Borisenko et al., 2020; Stivala et al., 2019b).

Methods II

- The ERGM DiffSign parameter to control for citation temporal direction was introduced by Graham et al. (2018); McLevey et al. (2018) and also used in Stivala et al. (2019a).
- In the full 4.9 million node network, only 1.9 million nodes represent patents in the data set. The remaining 3 million nodes (61% of the nodes) represent patents cited by one of those in the data set, but for which we have no data.
- An ERGM model with NA for all values on those 3 million nodes does not converge (unlike the 3.7 million node NBER patent citation network where only 27% of the nodes have no data in Stivala et al. (2019a)).
- So conditional estimation based on snowball sampling structure (Pattison et al., 2013; Stivala et al., 2016) was used. The 1.9 million nodes (39%) with data are treated as wave 0 (seeds) and the remaining 3 million nodes treated as wave 1, and estimation is conditional on this structure.

Negative binomial models with class crossing ratio I

	Model 1	Model 2	Model 3
App. Year [base 1978]	$-0.13\ (0.00)^{***}$	$-0.11\ (0.00)^{***}$	$-0.11\ (0.00)^{***}$
Section A	$-0.20\ (0.01)^{***}$	$-0.06\ (0.01)^{***}$	$-0.13\ (0.01)^{***}$
Section B	$0.12\ (0.00)^{***}$	$0.11\ (0.01)^{***}$	$0.03\ (0.01)^{**}$
Section C	$0.06\ (0.00)^{***}$	$0.14\ (0.01)^{***}$	$0.04\ (0.01)^{***}$
Section D	$0.08\ (0.01)^{***}$	$0.10\ (0.01)^{***}$	$0.02\ (0.01)$
Section E	$-0.21\ (0.01)^{***}$	$-0.21\ (0.01)^{***}$	$-0.28\ (0.01)^{***}$
Section F	$0.09\ (0.01)^{***}$	$0.09\ (0.01)^{***}$	$0.00\ (0.01)$
Section G	$0.13\ (0.00)^{***}$	$0.13\ (0.01)^{***}$	$0.05\ (0.01)^{***}$
Section H	$0.14\ (0.01)^{***}$	$0.06\ (0.01)^{***}$	$-0.02\ (0.01)^{*}$
Pub. Language German	$-0.25\ (0.00)^{***}$	$-0.33\ (0.01)^{***}$	$-0.33\ (0.01)^{***}$
Pub. Language French	$-0.27\ (0.01)^{***}$	$-0.33\ (0.01)^{***}$	$-0.33\ (0.01)^{***}$
Backward citations (subgraph)	$0.43\ (0.00)^{***}$	$0.16\ (0.00)^{***}$	$0.17\ (0.00)^{***}$
Max. contrast	$-1.74\ (0.43)^{***}$	$-3.34\ (0.56)^{***}$	$-4.04\ (0.56)^{***}$
Max. contrast^2	$2.67\ (0.34)^{***}$	$4.01\ (0.44)^{***}$	$4.43\ (0.44)^{***}$
Class crossing ratio		$0.33\ (0.02)^{***}$	$0.18\ (0.03)^{***}$
Class crossing ratio^2		$-0.48\ (0.03)^{***}$	$-0.42\ (0.03)^{***}$
Niche width			$0.30\ (0.02)^{***}$
Cited max. contrast			
Cited max. contrast^2			
Cited niche width			
Appplicant Switzerland			
Inventor Switzerland			
Appplicant Switzerland × Inventor Switzerland			
AIC	3318050.97	1610355.84	1609898.86
BIC	3318250.52	1610560.78	1610115.18
Log Likelihood	−1659009.49	−805159.92	−804930.43
Deviance	1199294.95	549422.32	549407.66
Num. obs.	1927639	650434	650434

Negative binomial models with class crossing ratio II

	Model 4	Model 5	Model 6
App. Year [base 1978]	−0.11 (0.00)***	−0.11 (0.00)***	−0.11 (0.00)***
Section A	−0.13 (0.01)***	−0.13 (0.01)***	−0.13 (0.01)***
Section B	0.03 (0.01)***	0.02 (0.01)**	0.03 (0.01)***
Section C	0.04 (0.01)***	0.03 (0.01)***	0.03 (0.01)**
Section D	0.02 (0.01)	0.02 (0.01)	0.01 (0.01)
Section E	−0.27 (0.01)***	−0.28 (0.01)***	−0.28 (0.01)***
Section F	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)
Section G	0.05 (0.01)***	0.05 (0.01)***	0.05 (0.01)***
Section H	−0.02 (0.01)*	−0.02 (0.01)*	−0.02 (0.01)*
Pub. Language German	−0.33 (0.01)***	−0.33 (0.01)***	−0.34 (0.01)***
Pub. Language French	−0.33 (0.01)***	−0.33 (0.01)***	−0.33 (0.01)***
Backward citations (subgraph)	0.16 (0.00)***	0.16 (0.00)***	0.16 (0.00)***
Max. contrast	−2.81 (0.73)***	−3.18 (0.73)***	−3.21 (0.74)***
Max. contrast^2	3.14 (0.58)***	3.59 (0.59)***	3.60 (0.59)***
Class crossing ratio	0.14 (0.03)***	−0.11 (0.03)***	−0.11 (0.03)***
Class crossing ratio^2	−0.41 (0.03)***	−0.30 (0.03)***	−0.30 (0.03)***
Niche width	0.34 (0.02)***	0.38 (0.02)***	0.38 (0.02)***
Cited max. contrast	−0.69 (0.76)	−0.98 (0.77)	−0.91 (0.78)
Cited max. contrast^2	1.01 (0.60)	1.00 (0.61)	0.94 (0.62)
Cited niche width		0.20 (0.01)***	0.20 (0.01)***
Appplicant Switzerland			−0.06 (0.02)**
Inventor Switzerland			−0.04 (0.03)
Appplicant Switzerland × Inventor Switzerland			0.21 (0.04)***
AIC	1609786.42	1609545.75	1574445.58
BIC	1610025.52	1609796.23	1574729.79
Log Likelihood	−804872.21	−804750.87	−787197.79
Deviance	549418.03	549427.11	539036.16
Num. obs.	650434	650434	639387

Negative binomial models using cited contrast only I

	Model 1	Model 2	Model 3
App. Year [base 1978]	−0.11 (0.00)***	−0.11 (0.00)***	−0.11 (0.00)***
Section A	0.04 (0.01)***	−0.01 (0.01)	−0.01 (0.01)
Section B	0.15 (0.01)***	0.13 (0.01)***	0.14 (0.01)***
Section C	0.07 (0.01)***	0.12 (0.01)***	0.13 (0.01)***
Section D	0.10 (0.01)***	0.07 (0.02)***	0.08 (0.02)***
Section E	−0.04 (0.01)**	−0.14 (0.02)***	−0.14 (0.02)***
Section F	0.11 (0.01)***	0.13 (0.01)***	0.13 (0.01)***
Section G	0.17 (0.01)***	0.17 (0.01)***	0.17 (0.01)***
Section H	0.21 (0.01)***	0.14 (0.01)***	0.14 (0.01)***
Pub. Language German	−0.34 (0.01)***	−0.32 (0.01)***	−0.32 (0.01)***
Pub. Language French	−0.33 (0.01)***	−0.32 (0.01)***	−0.32 (0.01)***
Backward citations (subgraph)	0.17 (0.00)***	0.15 (0.00)***	0.15 (0.00)***
Class crossing ratio	0.32 (0.02)***		
Class crossing ratio^2	−0.49 (0.03)***		
Cited max. contrast		−1.22 (0.96)	−1.05 (0.96)
Cited max. contrast^2		1.94 (0.74)**	1.80 (0.74)*
Cited secondary contrast		−3.88 (0.76)***	−3.77 (0.76)***
Cited secondary contrast^2		3.25 (0.65)***	3.24 (0.65)***
Cited niche width			−0.12 (0.03)***
Appplicant Switzerland			
Inventor Switzerland			
Appplicant Switzerland × Inventor Switzerland			
AIC	1611861.29	964173.87	964153.35
BIC	1612043.45	964368.85	964359.16
Log Likelihood	−805914.64	−482068.94	−482057.67
Deviance	549299.64	322525.69	322527.82
Num. obs.	650434	373983	373983

Negative binomial models using cited contrast only II

	Model 4	Model 5
App. Year [base 1978]	−0.11 (0.00)***	−0.11 (0.00)***
Section A	−0.01 (0.01)	0.01 (0.01)
Section B	0.14 (0.01)***	0.17 (0.01)***
Section C	0.12 (0.01)***	0.14 (0.01)***
Section D	0.07 (0.02)***	0.10 (0.02)***
Section E	−0.14 (0.02)***	−0.11 (0.02)***
Section F	0.13 (0.01)***	0.16 (0.01)***
Section G	0.17 (0.01)***	0.19 (0.01)***
Section H	0.14 (0.01)***	0.17 (0.01)***
Pub. Language German	−0.33 (0.01)***	−0.33 (0.01)***
Pub. Language French	−0.32 (0.01)***	−0.32 (0.01)***
Backward citations (subgraph)	0.15 (0.00)***	0.15 (0.00)***
Class crossing ratio		0.40 (0.11)***
Class crossing ratio^2		−0.58 (0.09)***
Cited max. contrast	−0.97 (0.97)	−1.01 (0.97)
Cited max. contrast^2	1.74 (0.75)*	1.73 (0.75)*
Cited secondary contrast	−3.86 (0.77)***	−4.03 (0.78)***
Cited secondary contrast^2	3.30 (0.66)***	3.43 (0.66)***
Cited niche width	−0.12 (0.03)***	0.10 (0.04)**
Appplicant Switzerland	−0.06 (0.03)	−0.05 (0.03)
Inventor Switzerland	−0.05 (0.04)	−0.05 (0.04)
Appplicant Switzerland × Inventor Switzerland	0.23 (0.06)***	0.23 (0.06)***
AIC	943423.07	943090.97
BIC	943661.00	943350.52
Log Likelihood	−471689.54	−471521.49
Deviance	316546.11	316510.52
Num. obs.	367615	367532

ERGM results, 1 933 231 node network I

Effect	Model 1	Model 2	Model 3
Arc	−13.638 (−13.896, −13.380)	−13.932 (−14.224, −13.639)	−13.417 (−13.703, −13.131)
Isolates	−0.182 (−0.253, −0.111)	0.046 (−0.009, 0.101)	0.087 (0.023, 0.151)
Sink	−0.763 (−0.848, −0.679)	−0.486 (−0.541, −0.430)	−0.490 (−0.559, −0.421)
Source	−0.225 (−0.290, −0.159)	−0.223 (−0.269, −0.176)	−0.222 (−0.285, −0.160)
Popularity spread (AinS)	0.784 (0.697, 0.870)	0.757 (0.684, 0.831)	0.775 (0.685, 0.865)
Activity spread (AoutS)	1.238 (1.096, 1.381)	0.841 (0.744, 0.937)	0.847 (0.728, 0.966)
Two-path (A2P-T)	−0.003 (−0.016, 0.010)	−0.023 (−0.041, −0.005)	−0.029 (−0.046, −0.012)
Shared popularity (A2P-D)	−0.213 (−0.246, −0.180)	−0.119 (−0.146, −0.091)	−0.120 (−0.149, −0.092)
Shared activity (A2P-U)	0.074 (0.055, 0.092)	0.062 (0.047, 0.078)	0.057 (0.038, 0.076)
Sender App. Year [base 1978]	0.454 (0.442, 0.465)	0.417 (0.402, 0.432)	0.449 (0.431, 0.466)
Receiver App. Year [base 1978]	−0.523 (−0.540, −0.505)	−0.505 (−0.525, −0.486)	−0.532 (−0.554, −0.509)
DiffSign App. Year	1.872 (1.741, 2.003)	2.032 (1.916, 2.148)	2.050 (1.937, 2.164)
AbsDiff App. Year	−0.625 (−0.650, −0.599)	−0.600 (−0.624, −0.576)	−0.629 (−0.659, −0.600)
Jaccard similarity Applicant countries	0.756 (0.582, 0.931)	0.808 (0.646, 0.970)	0.786 (0.615, 0.957)
Jaccard similarity Inventor countries	0.586 (0.432, 0.739)	0.573 (0.443, 0.702)	0.551 (0.399, 0.704)
Jaccard similarity Sections	3.837 (3.518, 4.156)	1.501 (1.360, 1.643)	1.402 (1.269, 1.535)
Matching Pub. Language	0.102 (0.050, 0.154)	0.044 (0.004, 0.083)	−0.025 (−0.061, 0.011)

ERGM results, 1 933 231 node network II

Sender Max. contrast	−1.409 (−1.596, −1.221)	−0.975 (−1.383, −0.567)	−3.547 (−3.849, −3.245)
Sender Max. contrast^2	−0.788 (−0.946, −0.630)	−1.375 (−1.762, −0.988)	0.668 (0.490, 0.847)
Receiver Max. contrast	−6.515 (−6.802, −6.229)	−5.204 (−5.433, −4.975)	−8.099 (−8.373, −7.825)
Receiver Max. contrast^2	5.169 (4.917, 5.420)	3.303 (3.108, 3.497)	5.067 (4.788, 5.346)
Jaccard similarity Classes	—	4.563 (4.308, 4.817)	5.802 (5.523, 6.080)
DiffSign Max. contrast	0.008 (−0.001, 0.018)	—	—
AbsDiff Max. contrast	−15.999 (−17.996, −14.002)	—	—
Sender Niche width	—	—	1.487 (1.424, 1.551)
Receiver Niche width	—	—	1.978 (1.798, 2.159)
Sender Secondary contrast	—	—	—
Sender Secondary contrast^2	—	—	—
Receiver Secondary contrast	—	—	—
Receiver Secondary contrast^2	—	—	—
Converged runs	20	20	20
Total runs	20	20	20

ERGM results, 1 933 231 node network III

Effect	Model 4
Arc	−13.241 (−13.577, −12.906)
Isolates	0.063 (−0.003, 0.130)
Sink	−0.483 (−0.573, −0.393)
Source	−0.252 (−0.324, −0.179)
Popularity spread (AinS)	0.799 (0.710, 0.888)
Activity spread (AoutS)	0.834 (0.721, 0.947)
Two-path (A2P-T)	−0.022 (−0.041, −0.003)
Shared popularity (A2P-D)	−0.107 (−0.136, −0.077)
Shared activity (A2P-U)	0.058 (0.038, 0.078)
Sender App. Year [base 1978]	0.433 (0.416, 0.449)
Receiver App. Year [base 1978]	−0.514 (−0.535, −0.492)
DiffSign App. Year	2.046 (1.904, 2.189)
AbsDiff App. Year	−0.609 (−0.639, −0.579)
Jaccard similarity Applicant countries	0.764 (0.597, 0.931)
Jaccard similarity Inventor countries	0.540 (0.382, 0.699)
Jaccard similarity Sections	1.392 (1.259, 1.525)
Matching Pub. Language	−0.016 (−0.051, 0.020)

ERGM results, 1 933 231 node network IV

Sender Max. contrast	−2.529 (−2.965, −2.093)
Sender Max. contrast^2	−1.325 (−1.736, −0.914)
Receiver Max. contrast	−6.258 (−6.603, −5.914)
Receiver Max. contrast^2	2.104 (1.910, 2.299)
Jaccard similarity Classes	5.907 (5.647, 6.167)
DiffSign Max. contrast	—
AbsDiff Max. contrast	—
Sender Niche width	1.253 (1.108, 1.399)
Receiver Niche width	1.726 (1.539, 1.914)
Sender Secondary contrast	−4.322 (−4.497, −4.147)
Sender Secondary contrast^2	7.709 (7.216, 8.203)
Receiver Secondary contrast	−4.578 (−4.798, −4.359)
Receiver Secondary contrast^2	8.102 (7.661, 8.544)
Converged runs	20
Total runs	20

References I

A. Borisenko, M. Byshkin, and A. Lomi. A simple algorithm for scalable Monte Carlo inference. *arXiv preprint arXiv:1901.00533v4*, 2020.

M. Byshkin, A. Stivala, A. Mira, G. Robins, and A. Lomi. Fast maximum likelihood estimation via equilibrium expectation for large network data. *Scientific Reports*, 8:11509, 2018.

A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4): 661–703, 2009.

J.-P. Ferguson and G. Carnabuci. Risky recombinations: Institutional gatekeeping in the innovation process. *Organization Science*, 28(1):133–151, 2017.

C. S. Gillespie. Fitting heavy tailed distributions: The poweRlaw package. *Journal of Statistical Software*, 64(2), 2015.

A. Graham, P. Browne, J. Barrett, and J. McLevey. Modelling directed acyclic graphs in exponential random graph models. Talk presented at INSNA Sunbelt XXXVIII, Utrecht, The Netherlands, June 26–July 1, 2018, June 2018.

M. Gruber, D. Harhoff, and K. Hoisl. Knowledge recombination across technological boundaries: Scientists vs. engineers. *Management Science*, 59(4):837–851, 2013.

B. H. Hall, A. B. Jaffe, and M. Trajtenberg. The NBER patent citation data file: Lessons, insights and methodological tools, 2001. National Bureau of Economic Research Working Paper 8498. `http://www.nber.org/papers/w8498`.

M. T. Hannan, L. Pólos, and G. R. Carroll. *Logics of organization theory: Audiences, codes, and ecologies*. Princeton University Press, Princeton, NJ, 2007.

F. Hartig. *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*, 2019. URL `https://CRAN.R-project.org/package=DHARMa`. R package version 0.2.6.

R. Henderson, A. Jaffe, and M. Trajtenberg. Patent citations and the geography of knowledge spillovers: A reassessment: Comment. *American Economic Review*, 95(1):461–464, 2005.

D. V. Hinkley. Jackknifing in unbalanced situations. *Technometrics*, 19(3):285–292, 1977.

A. B. Jaffe, M. Trajtenberg, and R. Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598, 1993.

References II

B. Kovács and M. T. Hannan. The consequences of category spanning depend on contrast. *Research in the Sociology of Organizations*, 31:175–201, 2010.

B. Kovács and M. T. Hannan. Conceptual spaces and the consequences of category spanning. *Sociological Science*, 2:252–286, 2015.

D. Krackhardt and R. N. Stern. Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, 51:123–140, 1988.

P. Leifeld. texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software*, 55(8):1–24, 2013. URL `http://www.jstatsoft.org/v55/i08/`.

J. G. MacKinnon and H. White. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3):305–325, 1985.

J. McLevey, A. V. Graham, R. McIlroy-Young, P. Browne, and K. S. Plaisance. Interdisciplinarity and insularity in the diffusion of knowledge: an analysis of disciplinary boundaries between philosophy of science and the sciences. *Scientometrics*, 117(1):331–349, 2018.

P. E. Pattison, G. L. Robins, T. A. B. Snijders, and P. Wang. Conditional estimation of exponential random graph models from snowball sampling designs. *Journal of Mathematical Psychology*, 57:284–296, 2013.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL `https://www.R-project.org/`.

A. Stivala, A. Palangkaraya, D. Lusher, G. Robins, and A. Lomi. ERGM parameter estimation of very large directed networks: implementation, example, and application to the geography of knowledge spillovers. Talk presented at INSNA Sunbelt XXXIX, Montréal, Canada, June 18–23, 2019, June 2019a. URL `https://sites.google.com/site/alexdstivala/home/conferences`.

A. Stivala, G. Robins, and A. Lomi. Exponential random graph model parameter estimation for very large directed networks. *arXiv preprint arXiv:1904.08063v3*, 2019b.

A. D. Stivala, P. Wang, J. L. Koskinen, G. L. Robins, and D. Rolls. Snowball sampling for estimating exponential random graph models for large networks. *Social Networks*, 46:167–188, 2016.

P. Thompson and M. Fox-Kean. Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, 95(1):450–460, 2005.

References III

M. Trajtenberg, R. Henderson, and A. Jaffe. University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1):19–50, 1997.

B. Uzzi, S. Mukherjee, M. Stringer, and B. Jones. Atypical combinations and scientific impact. *Science*, 342(6157): 468–472, 2013.

W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, New York, fourth edition, 2002. URL `http://www.stats.ox.ac.uk/pub/MASS4`. ISBN 0-387-95457-0.

A. Zeileis. Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, 11(10):1–17, 2004. ISSN 1548-7660. doi: 10.18637/jss.v011.i10. URL `https://www.jstatsoft.org/v011/i10`.

A. Zeileis. Object-oriented computation of sandwich estimators. *Journal of Statistical Software*, 16(9):1–16, 2006. ISSN 1548-7660. doi: 10.18637/jss.v016.i09. URL `https://www.jstatsoft.org/v016/i09`.

A. Zeileis and T. Hothorn. Diagnostic checking in regression relationships. *R News*, 2(3):7–10, 2002.