Inventor collaboration and its' persistence across European regions

Gergő Tóth[1,2] - Sándor Juhász[1,3] - Zoltán Elekes[1,3] - Balázs Lengyel[1,4]

[1] Agglomeration and Social Networks Lendület Research Group, Hungarian Academy of Sciences, 1097 Budapest, Hungary

[2] Spatial Dynamics Lab, University College Dublin, Belfield, Dublin 4, Ireland

[3] Faculty of Economics and Business Administration, University of Szeged, 6722 Szeged, Hungary

[4] International Business School Budapest, 1037 Budapest, Hungary

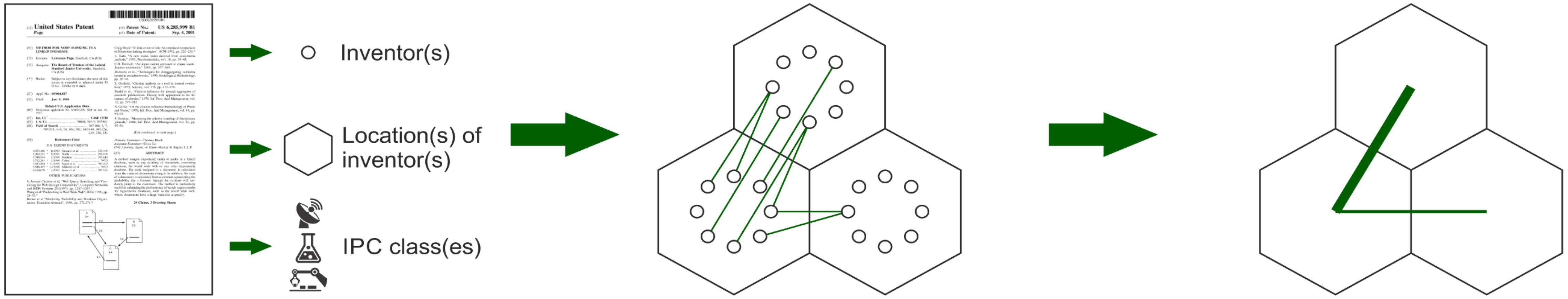
- Data
- Spatial patterns of inventor collaborations
- Multivariate gravity models
- Results
- Policy implications

- OECD REGPAT Database
- Patents authored by European inventors, and registered by European Patent Office
- EU27 and continental EFTA countries (Norway and Switzerland)
- Full data 1991-2010
- Used subsample 2006-2010
- Technological classes IPC
- NUTS3 regions
- Unique inventor ID

Step 1. Patent information from OECD RegPat

Step 2. Cross-region inventor collaboration

Step 3. Aggregation to EU regions



From the patent information we derived the cross regional inventor collaboration network. The tie weights are proportional to the number of inventor collaborations between regions in the 2006 to 2010 period.

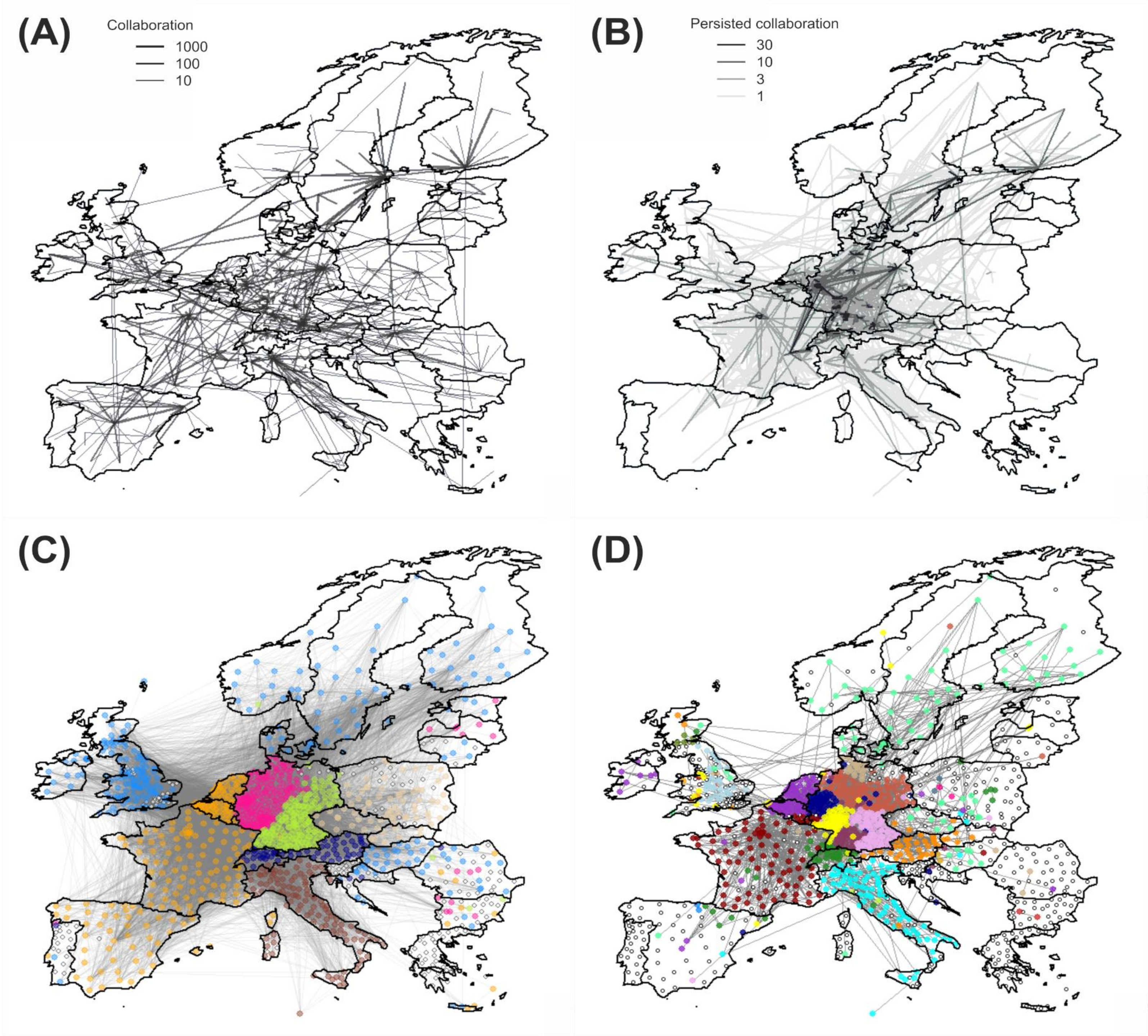
We consider an inventor-inventor tie persisted when it has occurred at least once in the 1991-2005 period, priory our investigation. (Juhász and Lengyel 2017)

This distinction is important because uncertain new connections offer access to new knowledge, whereas persistent ties represent those strong collaborations that are worth repeating despite their considerable opportunity costs. (Dahlander and McFarland 2013)

Descriptive characteristics of the co-inventor collaboration network and the persisted collaboration network.

	Collaboration	**Persisted Collaboration**
Number of Individual Collaborations	772,378	41,883
Number of Region Ties	46,857	6,200
Density of the Region Network	0.05	0.006
Number of Communities in the Region Network*	7	23
Modularity of the Region Network	0.372	0.584
Relative Modularity of the Region Network	0.329	0.379

* Communities of size 1 are excluded!



Spatial patterns of inventor collaboration and persisted collaboration networks of European regions.

(A) The maximum spanning tree of the collaboration network across NUTS3 regions in Europe reveals the importance of national centres.

(B) Most of the persisted collaborations remain within country borders, and strongest ties are concentrated within close proximity of innovative hubs.

(C) The 7 communities of the collaboration network span across countries, but are mostly concentrated in large regions.

(D) Persisted collaboration is organized into 23 smaller-scale clusters.

What are the driving factors of the probability of (persistent) co-inventorship across European regions?

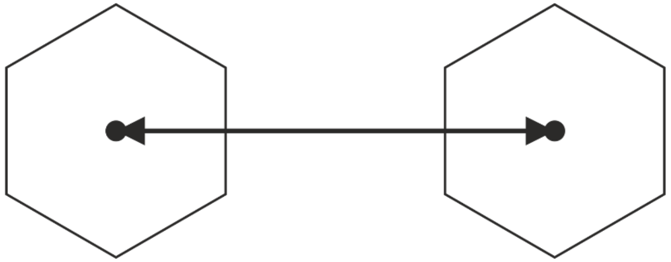
How the spatial patterns of persistent inventor collaborations differ from newly created ones?

Technological innovation is **concentrated in space** due to its increasing returns to scale, and because it is easier to share complex knowledge with partners in **geographical proximity** and through face-to-face interaction. (Jaffe, Trajtenberg and Henderson 1993; Audretsch and Feldman 1996; Balland and Rigby 2017)

Empirical findings supports these claims by showing that besides geographical proximity, the **overlap of technological portfolios** and triadic closure of partnerships also increase the probability of cross-regional innovative collaboration. (Balland 2012; Cassi and Plunket 2015)

Social relations greatly influence these phenomena by increasing externalities by facilitating the emergence of novel combinations and by enabling flows of innovation-related knowledge through connections that can bridge even great distances. (Katz and Shapiro 1985; Acemoglu, Akcigit and Kerr 2016)

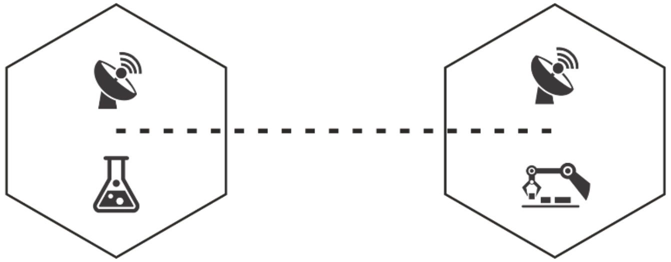
Proximity: inverse of physical distance between the centroids of region *i* and *j*.



$$\max\{\mathrm{Distance}_{\mathrm{ij}}\} - \mathrm{Distance}_{\mathrm{ij}}$$

,where Distance is the kilometres between the centroids of each region.

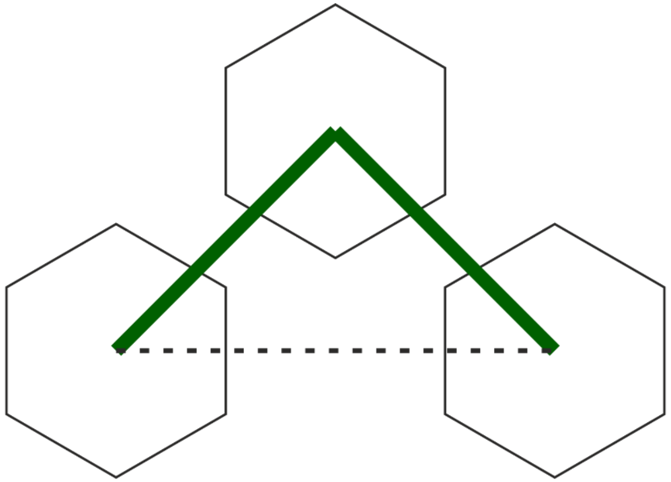
Cosine: degree of overlap between the patent class portfolios of region *i* and *j*.



$$\mathrm{Cosine}_{\mathrm{ij}} = \frac{v_i \cdot v_j}{\left|\sum_{i=1}^{M} v_i^2\right|\left|\sum_{j=1}^{M} v_j^2\right|}$$

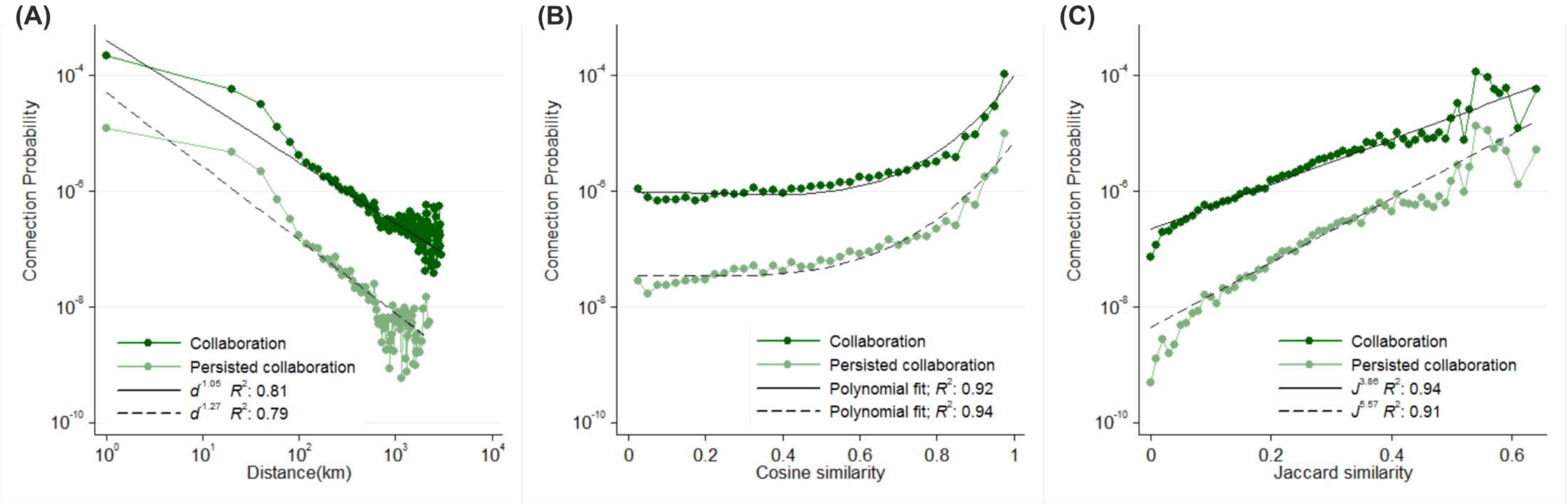
similarity of the patent portfolio vectors v_i and v_j, where zero is the case of perfectly unrelated portfolios and one represents complete similarity.

Jaccard: share of common third partners of region *i* and *j*.



$$\mathrm{Jaccard}_{\mathrm{ij}} = \frac{|\mathrm{A_i} \cap \mathrm{A_j}|}{|\mathrm{A_i} \cup \mathrm{A_j}|}$$

,where $\mathrm{A_i}$ and $\mathrm{A_j}$ refer to the underlying collaboration matrices of regions i and j. ranges from zero to one, where higher values indicate a higher share of common third partners.



The probability of collaboration and persisted collaboration as a function of region-to-region characteristics.

(A) Distance decay is smooth for geographically proximate collaboration and persisted collaboration, and follows linear decay on log-log scale with the exponents -1.05 and -1.27 for distances larger than 100 km.

(B) The overlap between technological profiles of regions, measured by cosine similarity, increases the probability of collaboration with a growing intensity as similarity rises.

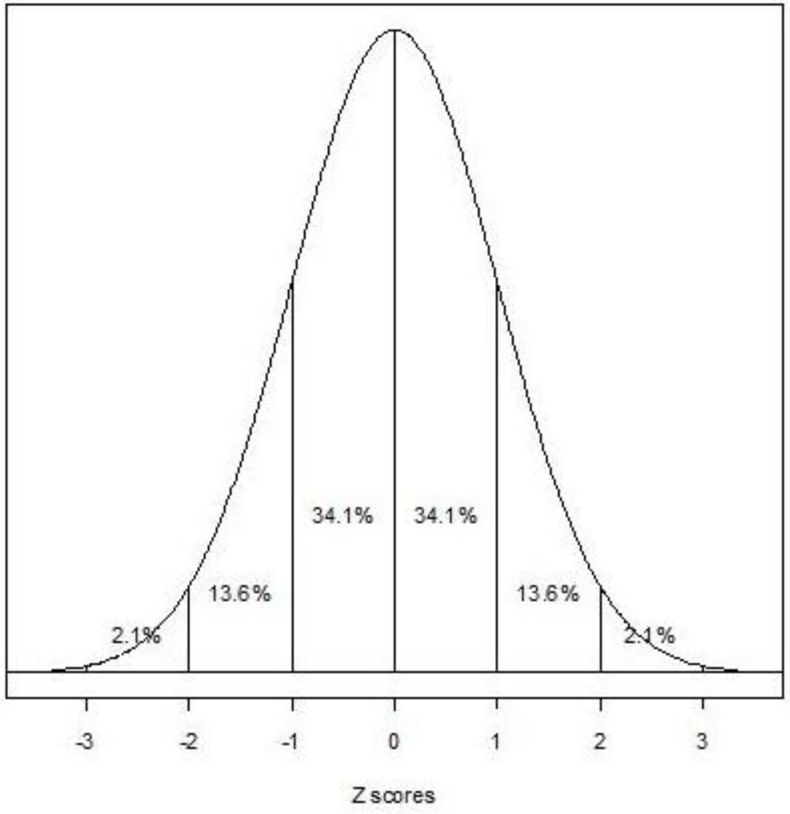
(C) The probability of collaboration grows linearly on a logarithmic scale as the share of common third partners, measured by Jaccard similarity, increases. The exponent is 3.86 for collaboration and 5.67 for persisted collaboration.

One unit change in Cosine similarity is hardly comparable with one unit change in Jaccard index or one kilometre change in geographic proximity and vice-versa.

Therefore, the three main explanatory variables have been rescaled to have a mean of zero:

$$z(x) = \frac{(x_i - \bar{x})}{\sigma_x}$$

,where $\bar{x}$ is the mean of x, and σ_x is the standard deviation of x



Co-inventor collaborations between regions can be considered as a count process. Because there are an excessive number region-region pairs with zero collaboration ties, we have to deal with the cases of missing collaborations independently as creating zero collaboration would be a distinct process from creating non-zero collaboration.

P_{ij} is the probability that the value of observation *ij* is zero: $P(Y_{ij}=0)=P_{ij}+(1-P_{ij})(1+k\lambda_{ij})^{-1/k}$

The count process is governed with probability (1-P_{ij}) by a negative binomial distribution with mean λ: $P(Y_{ij}=y_{ij})=(1-P_{ij})\frac{(y_{ij}+1/k)(k\lambda_{ij})^{y_{ij}}}{(y_{ij}+1)(1/k)(1+k\lambda_{ij})^{y_{ij}{}^{1/k}}}$, where y_{ij}=1,2..

Zero Inflated Negative Binomial Gravity Model

$$P\left(Y_{ij}=0\right)=\gamma_0+\gamma_1\theta_{ij}+\gamma_2\log(Z_i)+\gamma_3\log\left(Z_j\right)+\epsilon_{ij}$$

θ_{ij} is a dummy variable that takes the value 1 if region i and region j are in the same country, and Z_i , Z_j are a collection of region-level control variables that are commonly used in similar estimations:

- number of inventors
- population density
- log of gross value added
- aggregate number of co-inventor collaboration

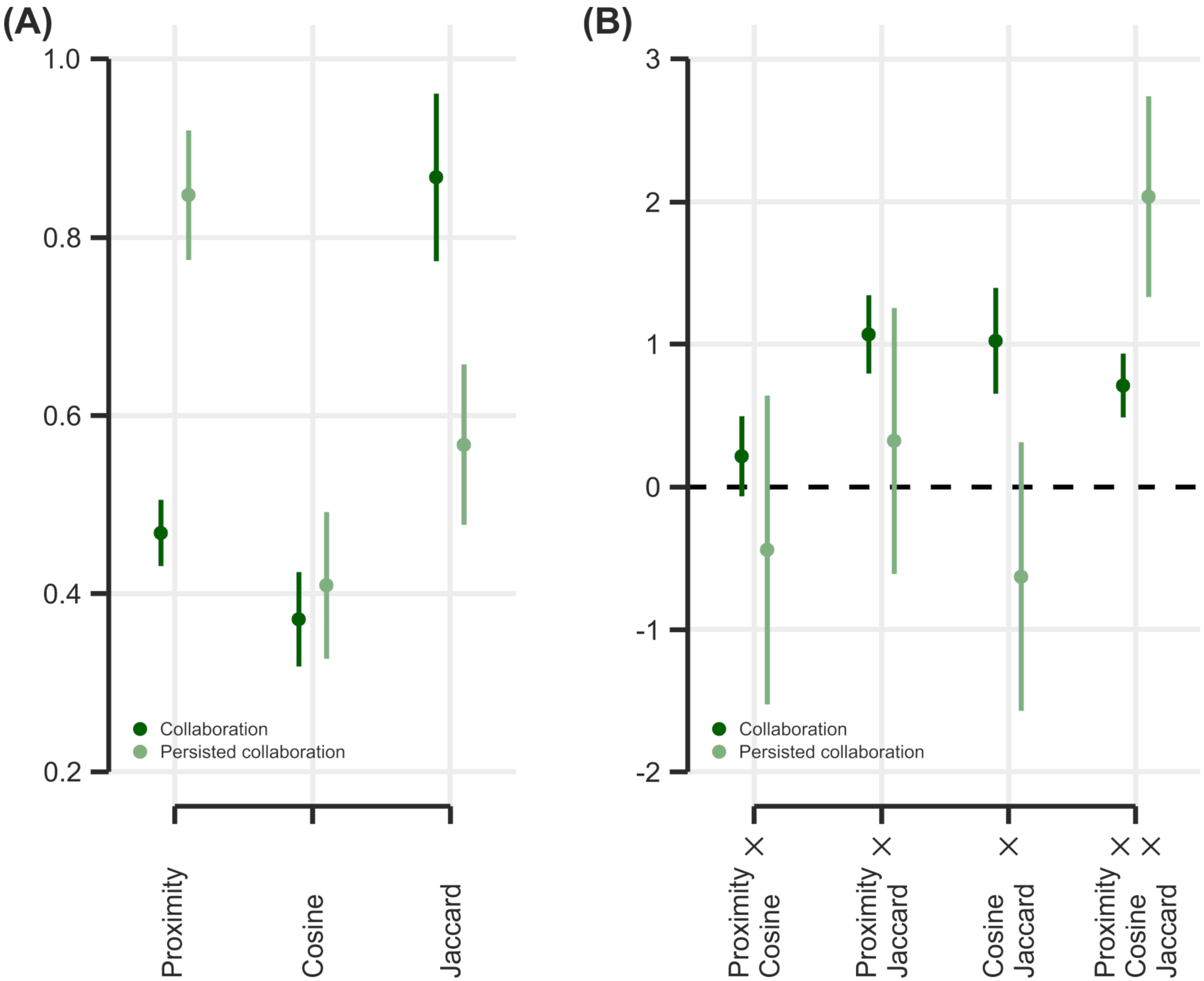
Then we estimate the number of individual ties between regions by our three main variables:

$$\log\left(Y_{ij}=y_{ij}\right)=\beta_0+\beta_1\mathrm{Proximity}_{ij}+\beta_2\mathrm{Cosine}_{ij}+\beta_3\mathrm{Jaccard}_{ij}+u_{ij}$$

	Collaboration	Persisted Collaboration
Main effects		
Cosine	0.371***	0.409***
	(0.027)	(0.042)
Proximity	0.468***	0.847***
	(0.019)	(0.037)
Jaccard	0.867***	0.567***
	(0.048)	(0.046)
Constant	-2.261***	-5.271***
	(0.141)	(0.250)

	Collaboration	Persisted Collaboration
Binary and interaction effects		
Cosine	0.094	1.075**
(reference category: low)	(0.098)	(0.424)
Proximity	0.394***	0.433
(reference category: low)	(0.109)	(0.434)
Jaccard	1.340***	2.342***
(reference category: low)	(0.119)	(0.383)
Proxy × Jaccard	1.067***	0.321
	(0.141)	(0.476)
Proxy × Cosine	0.212	-0.444
	(0.144)	(0.553)
Jaccard × Cosine	1.023***	-0.631
	(0.190)	(0.481)
Jaccard × Cosine × Proxy	0.709***	2.034***
	(0.115)	(0.359)
p	0	0
Log lik.	-226 004.8	-37 645.09
N	872 235	872 235

Zero-inflation		
Same country	-6.986***	-3.174***
(dummy variable)	(0.260)	(0.151)
Log number of connections of region *i*	-11.96***	-0.522**
	(1.180)	(0.204)
Log number of connections of region *j*	-13.89***	-0.801***
	(0.750)	(0.133)
Log number of inventors in region *i*	3.877***	-3.371***
	(0.683)	(0.381)
Log number of inventors in region *j*	11.12***	1.100***
	(1.146)	(0.269)
Log population density in region *i*	0.370***	0.0454
	(0.105)	(0.066)
Log population density in region *j*	0.710***	0.227***
	(0.078)	(0.041)
Log gross value added in region *i*	-2.040***	0.345***
	(0.119)	(0.131)
Log gross value added in region *j*	-5.137***	-0.657***
	(0.340)	(0.095)
Constant	61.01	12.32***
	(.)	(0.718)
lnalpha		
Constant	2.220***	2.160***
	(0.056)	(0.059)
p	0	0
Log lik.	-20,7048.8	-32,905.8
N	872,235	872,235



Estimation results of the multivariate gravity equation.
(A) Considering single effects only, we observe that the probability of collaboration is mostly increased by common third partners, while persistent ties gain probability if regions are geographically proximate.
(B) Interaction effects reveal that persistent ties gain extra probability if regions are geographically proximate, technologically similar and share many partners at the same time

Takeaway message and policy implications

Spatial clustering offers positive externalities, while fostering a diverse and open network of inventors helps exploring opportunities, and in avoiding lock-in.

European Research Area intended to decrease the fragmentation of the European research activities.

These are still self-organized into national innovation systems.

Repeated inter-regional collaborations are fragmented at a smaller spatial scale, regional innovation systems more adequate.

Our evidence suggests that lasting cooperations are strongly bounded by multidimensional proximity, and are likely to revert to spatial clustering.

European innovation policy could address these issues by specifically targeting persisted collaborations across national borders and larger distances.

Thank you!

toth.gergo@krtk.mta.hu
sandor.juhasz@krtk.mta.hu
zoltan.elekes@krtk.mta.hu
balazs.lengyel@krtk.mta.hu

http://anet.krtk.mta.hu

Acemoglu, D., Akcigit, U. & Kerr, W. R. Innovation network. Proc. Natl. Acad. Sci. 113, 11483–11488 (2016).

Audretsch, D. B. & Feldman, M. P. R&d spillovers and the geography of innovation and production. The Am. economic review 86, 630–640 (1996).

Balland, P.-A. Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (gnss) industry. Reg. Stud. 46, 741–756 (2012)

Balland, P.-A. & Rigby, D. The geography of complex knowledge. Econ. Geogr. 93, 1–23 (2017).

Cassi, L. & Plunket, A. Research collaboration in co-inventor networks: combining closure, bridging and proximities. Reg. Stud. 49, 936–954 (2015).

Dahlander, L. & McFarland, D. A. Ties that last: Tie formation and persistence in research collaborations over time. Adm. science quarterly 58, 69–110 (2013).

Jaffe, A. B., Trajtenberg, M. & Henderson, R. Geographic localization of knowledge spillovers as evidenced by patent citations. Q. journal Econ. 108, 577–598 (1993).

Juhasz, S. & Lengyel, B. Creation and persistence of ties in cluster knowledge networks. ´ J. Econ. Geogr. (2017).

Katz, M. L. & Shapiro, C. Network externalities, competition, and compatibility. The Am. economic review 75, 424–440 (1985).