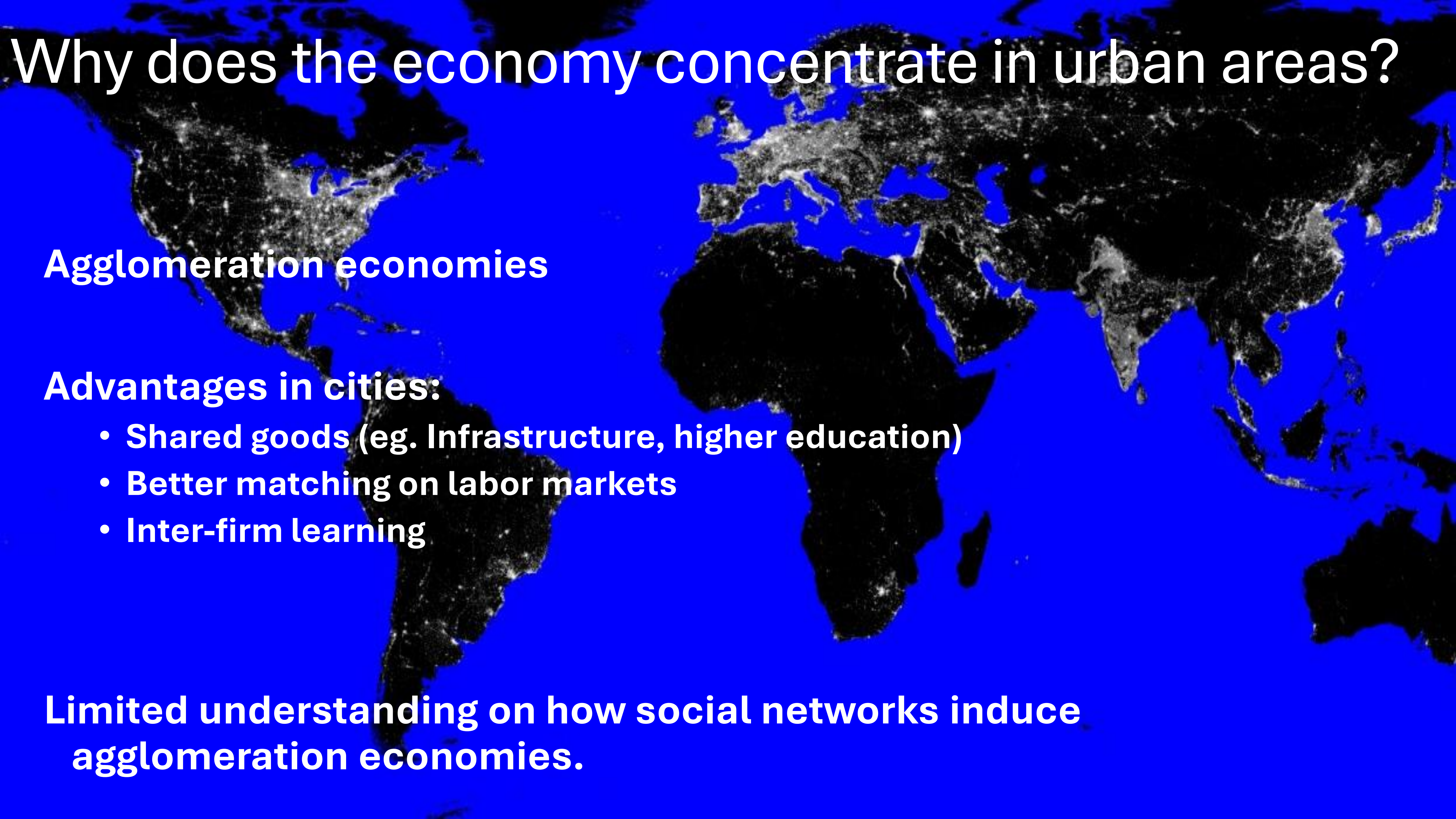
Social Networks and Economic Geography

Balázs Lengyel
lengyel.balazs@krtk.hun-ren.hu

Class 4: Regional development

February 18 2025



Why does the economy concentrate in urban areas?

Agglomeration economies

Advantages in cities:

- **Shared goods (eg. Infrastructure, higher education)**
- **Better matching on labor markets**
- **Inter-firm learning**

Limited understanding on how social networks induce agglomeration economies.

We live in a connected world

Diffusion of viruses and information is fast

Brockman D., Helbing D (2013) Science

The global society is segregated

Why do we observe growing inequalities if diffusion of ideas and knowledge in social networks is very quick?

Prosperous areas next to a slum in Mumbai.

Network fragmentation and inequalities

Small-world social networks are fragmented by groups.

One of the strongest forces of group formation is prosperity.

Diffusion of innovation is unequal across network fragments.

Currarini, Pin, Jackson (2009) Econometrica

Topics today

1. Labor mobility, co-worker networks, and economic development

2. Innovation and inequality in regional networks

Topics today

1. Labor mobility, co-worker networks, and economic development

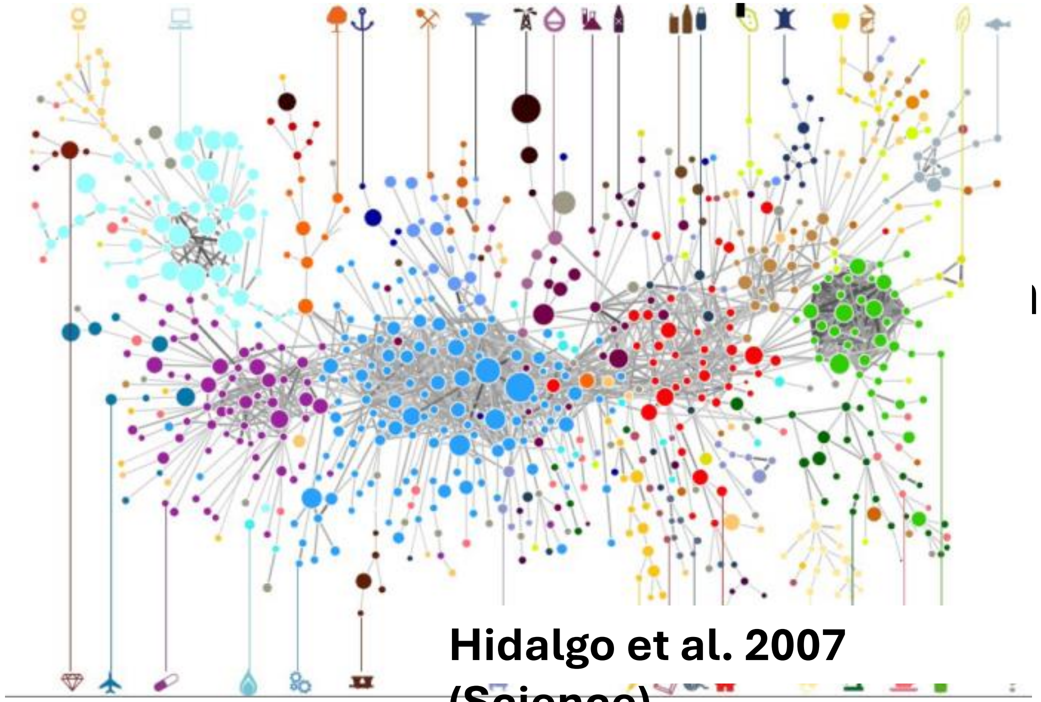
2. Innovation and inequality in regional networks

Economic progress and urban success

- Social interaction: population density

- Learning in the city: industry structure

- Related knowledge is easier to learn but contains less novelty



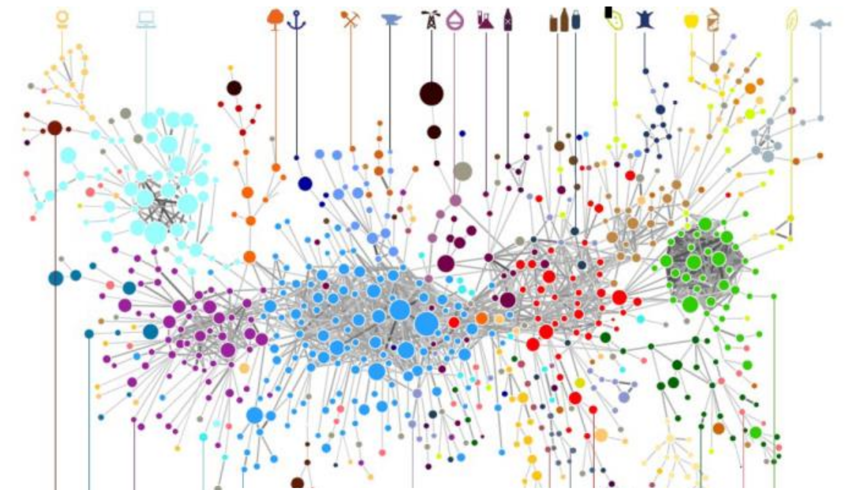
Hidalgo et al. 2007 (Science)

Human mobility and labor mobility

González et al. (Nature 2008)

- Labor mobility: PEOPLE CHANGE JOBS
- Direct effect
  - New employee brings new knowledge to the firm
- Indirect effect
  - Previous colleagues tend to communicate or even work together
- Urban success (Silicon Valley vs Route 128)

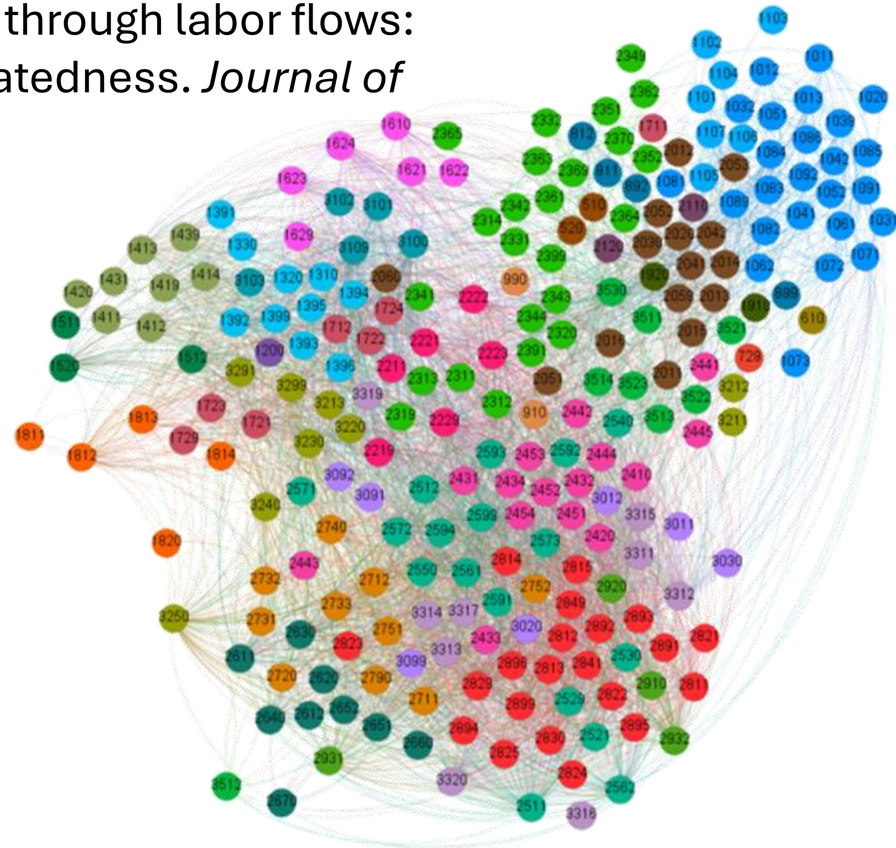
How does interaction between firms and industries induce economic success of cities?



Hidalgo et al. 2007 (Science)

Direct effect in the firm: How does the knowledge of mobile employees induce firm performance?

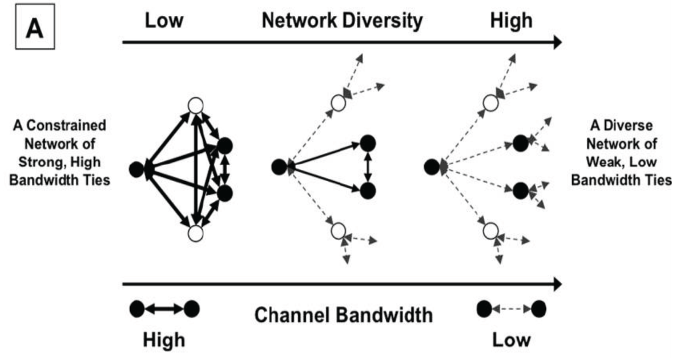
Csáfordi, Zs., Lőrincz, L., Lengyel, B., Kiss, K.M. (2020) Productivity spillovers through labor flows: The effect of the productivity gap, foreign-owned firms, and technological relatedness. *Journal of Technology Transfer*



Data:
Hungarian firms, 2003-2010
Company: revenues, costs, industry, size
Employees: education, wage, gender, age

Total factor productivity: all revenues over all cost

Indirect effect in the firm: How does the networks of mobile employees induce firm performance?



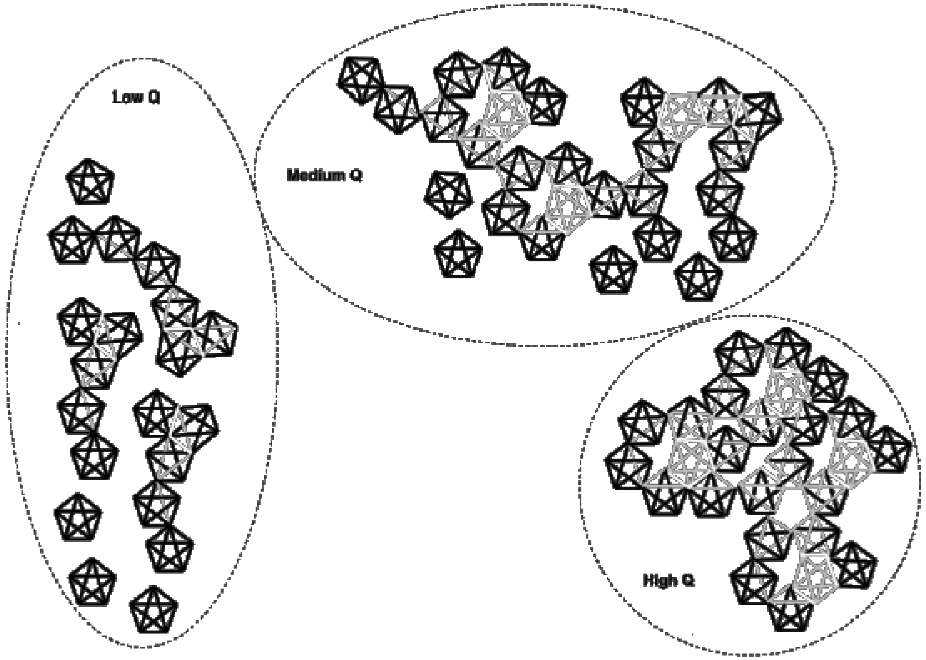
Tóth, G., Lengyel, B. (2019) Inter-firm inventor movements and the optimal structure of co-inventor networks. *The Journal of Technology Transfer*.

Data: European patents in ICT, 1977-2010

Inventor mobility
author a patent for SOURCE FIRM , then later for DESTINATION FIRM

Dependent variable
Citations to the patents of the DESTINATION FIRM

Explanatory variables
Ego network of mobile inventors
Collaboration networks in the DESTINATION FIRM



Indirect effect in the region: How does social ties created by labor mobility induce urban performance?

Lengyel, B., Eriksson, R.H. (2017) Co-worker networks, labor mobility and productivity growth in regions. *Journal of Economic Geography*. https://doi.org/10.1093/jeg/lbw027

Eriksson, R.H., Lengyel, B. (2017) Co-worker networks and agglomeration externalities. Under revision for *Economic Geography*.

The importance of workplaces

- Connections developed at workplaces:
  - Represesent both social and cognitive proximity (Storper &Venables 2004).
  - Often maintaned and vital for continued knowledge inputs (Agrawal et al, 2006; Boschma & Frenken 2011, Breschi & Lissoni, 2009, Dahl & Pedersen, 2004)

- Labor economics assumption: employees in two firms know each other if they have worked at the same place in the same time of their career previously.

- The co-worker network is an important field of information spreading
  - because it speeds up job search (Calvo-Armengol and Jackson 2004, AER; Granovetter 1995, UCP);
  - and increases the wage of a new employee (Hensvik and Nordström Skans 2016, JLE).

Problem: employees might not know each other at large workplaces.

Three further assumptions

1. The probability of employee-employee connections at a workplace are inversely proportional to the size of the workplace.

2. Employees are more likely to build connections to similar employees (Currarini et al. 2009, Econometrica, Granovetter 1995, UCP; Kossinets and Watts 2006 , Science; McPherson et al. 2001, ARS).

3. People increase the strength of their relationship when spending time together (Marsden and Campbell, 1984) and social ties loose strength over time after last contact (Burt, 2003).

Probability of co-worker tie creation

$$P_{ij} = \frac{\ln N}{N} + \sum_{G=1}^{M} \left(\frac{\ln N_m}{N_m} / \frac{N_m}{N} \right) \times \delta_{ij}$$

where G $\in \{1, 2, \dots M\}$ denotes employee characteristics, *N* is the size of the workplace, N_m is the group size with characteristics *m*, and δ_{ij} equals 1, if *i* and *j* are similar according to *m*, otherwise 0.

Probability of ties are inversely proportional to the size of the workplace (Erdős and Rényi 1959).

Co-workers are more likely to know each other if they are similar (Currarini et al. 2009, Econometrica, Granovetter 1995, UCP; Kossinets and Watts 2006 , Science; McPherson et al. 2001, ARS).

Similarity adds more to probability if there are few similar co-workers in a given characteristic.

Data

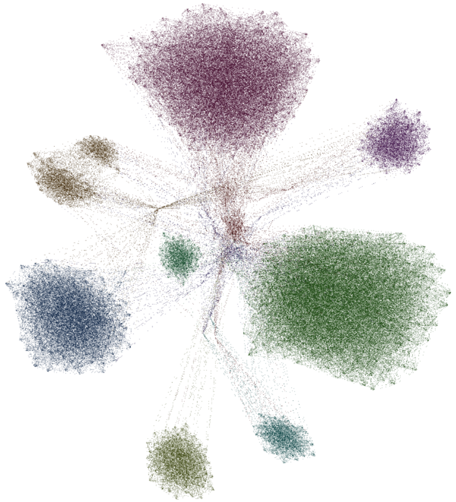
- Matched employer-employee data stored in public administration databases of every employee in Sweden matched by their social security number.
  - 1990-2008
  - ID of the employee and plant at *t*
  - Empoyee characteristics: wage, direction of graduation (6), age (3), sex (2)
  - Plant characteristics: number of employees, net turnover

		1990	2008
Total number of employees	Employees	2,628,306	3,824,182
	Plants	254,445	402,610
Employees with BA degree or above	Employees	366,336	785,578
	Plants	52,872	113,441

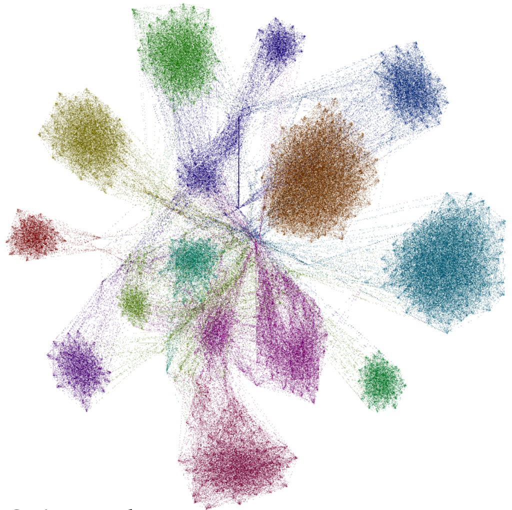
Alternative tie formation (large plants)

(Lengyel & Eriksson, 2016)

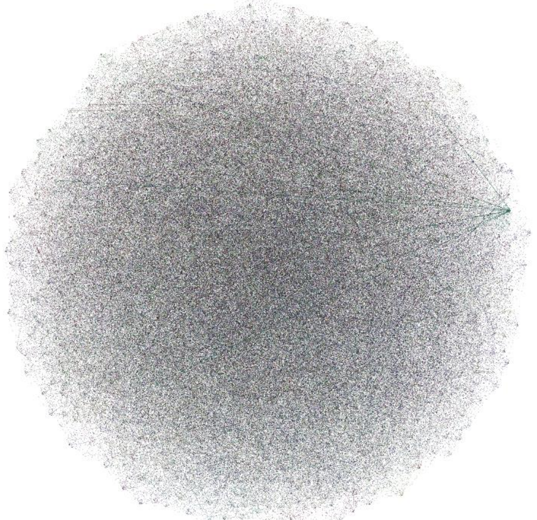
50: Edu, sex, age



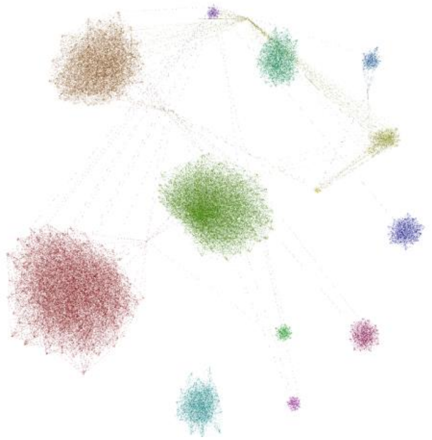
50: Edu, sex, age, wage



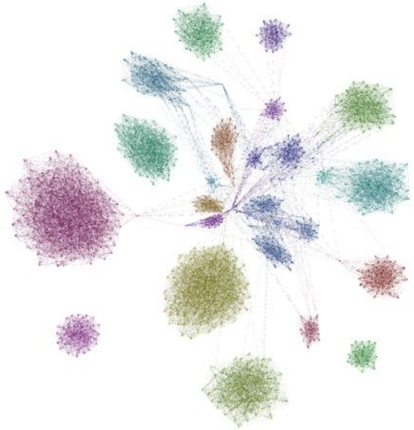
50: Random



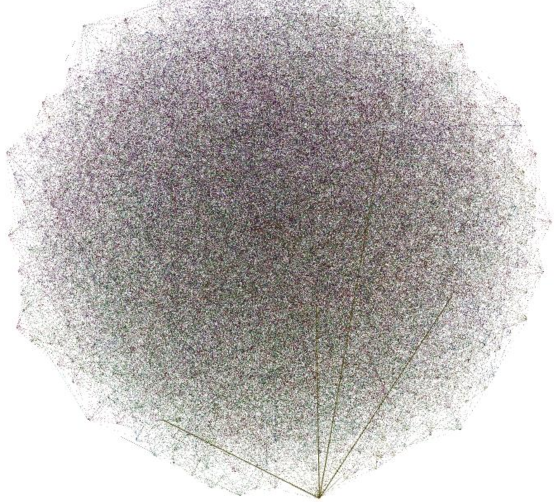
25: Edu, sex, age



25: Edu, sex, age, wage



25: Random

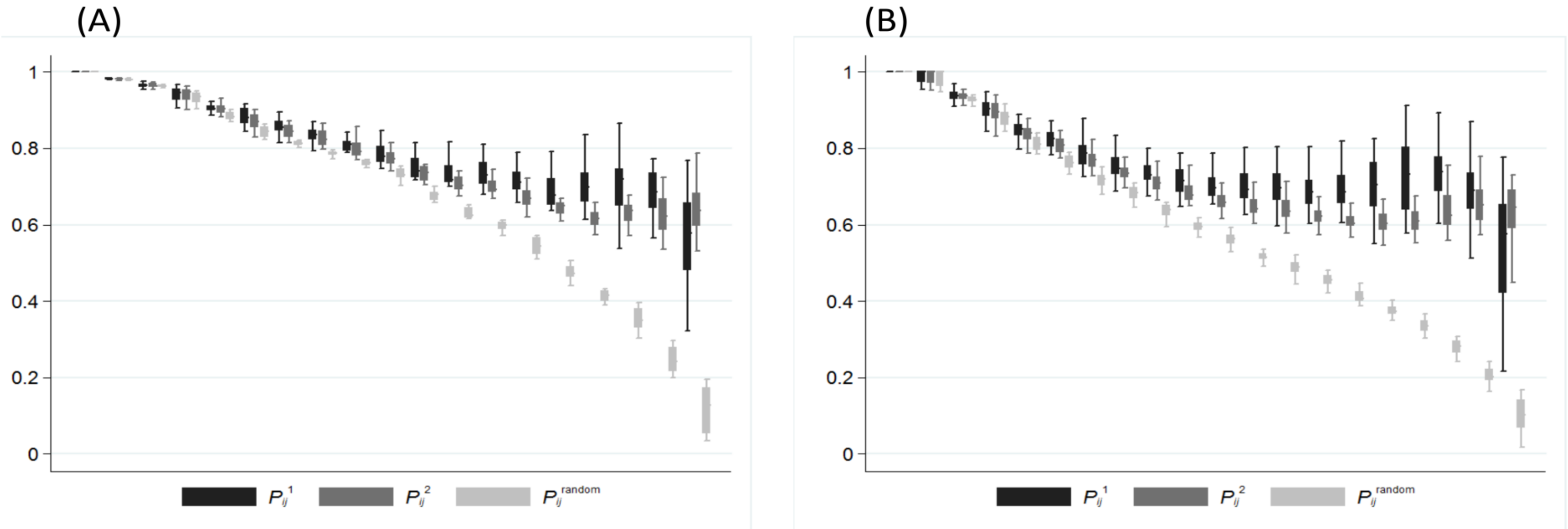


Two versions of tie probabilities

(Lengyel & Eriksson, 2016)

$$P_{ij}^{1} = N \times \left(\frac{\ln N}{N^2} + \frac{\ln N_e}{N_e^2} \times \delta_{ij,e} + \frac{\ln N_a}{N_a^2} \times \delta_{ij,a} + \frac{\ln N_g}{N_g^2} \times \delta_{ij,g} \right)$$

$$P_{ij}^{2} = N \times \left(\frac{\ln N}{N^2} + \frac{\ln N_w}{N_w^2} \times \delta_{ij,w} + \frac{\ln N_e}{N_e^2} \times \delta_{ij,e} + \frac{\ln N_a}{N_a^2} \times \delta_{ij,a} + \frac{\ln N_g}{N_g^2} \times \delta_{ij,g} \right)$$



Triadic closure is higher than in random networks.

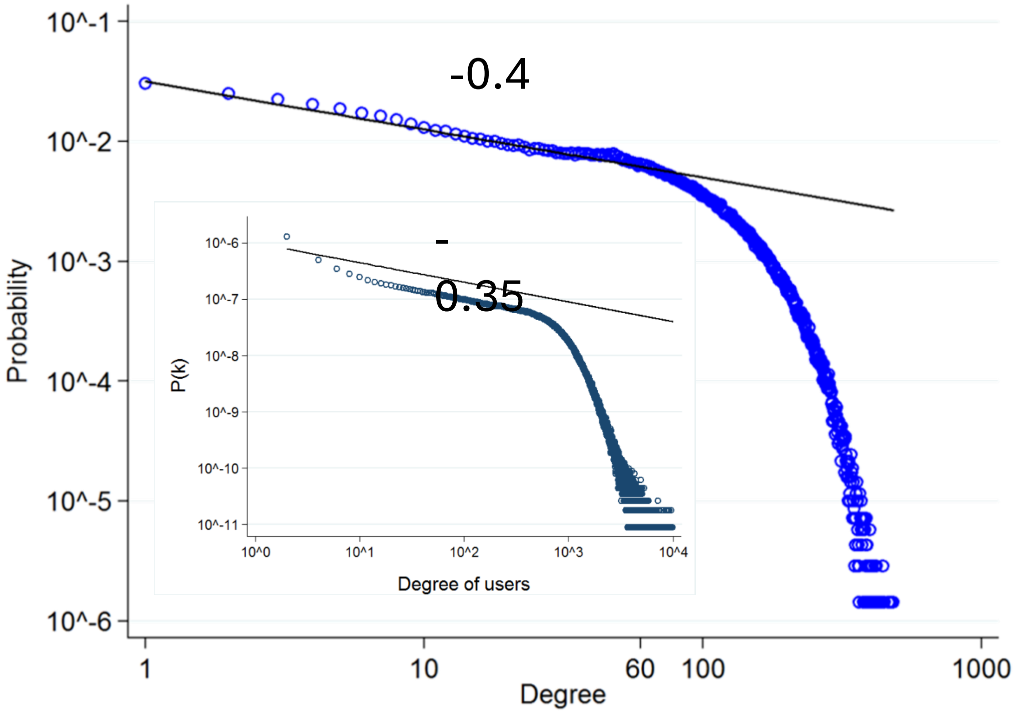
Selection from networks

- Only consider ties at large plants (>25 or >50) because these approaches do not differ from random networks in small plants.
- Creat networks where every worker can establish 50 or 25 ties per year. We track all ties after the creation over the full period. Exclude those when (1) peers work at the same plant, (2) either of them is out of Sweden, (3) veither of them is more than 65.
- pVAR models with income per capita and network density on the level of functional regions.

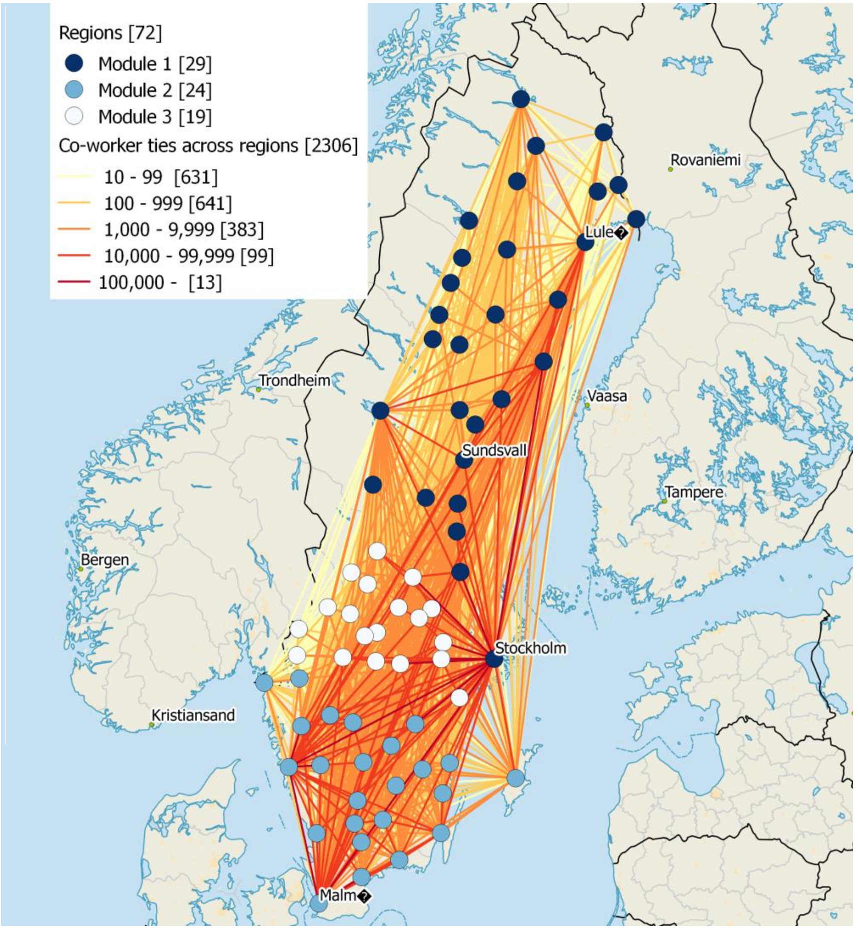
	P^1_{ij} 50 ties	P^2_{ij} 50 ties	Random 50 ties	P^1_{ij} 25 ties	P^2_{ij} 25 ties	Random 25 ties
			First differenced regional income per capita (RegProd)			
L.RegProd	0.333**	1.018	0.363***	0.456***	0.793***	0.270*
	(0.164)	(0.912)	(0.126)	(0.139)	(0.209)	(0.147)
L2.RegProd	0.166**	0.059	0.187***	0.067	-0.085	0.101
	(0.073)	(0.193)	(0.072)	(0.081)	(0.098)	(0.062)
L3.RegProd	0.165*	0.362	0.162*	0.073	0.040	0.098
	(0.086)	(0.298)	(0.087)	(0.098)	(0.085)	(0.068)
L.D_t	0.034**	-0.070	0.029	0.028	0.043**	0.023
	(0.016)	(0.126)	(0.018)	(0.021)	(0.020)	(0.018)
L2.D_t	0.002	-0.010	-0.003	0.002	0.005	0.004
	(0.003)	(0.016)	(0.003)	(0.002)	(0.005)	(0.003)
L3.D_t	0.002	0.009	0.002	0.003	0.000	0.003
	(0.004)	(0.009)	(0.002)	(0.003)	(0.004)	(0.003)
Hansen J	10.351	3.838	10.017	2.323	6.903	6.549
N	614	603	609	735	732	734

The co-worker network in Sweden

Resembles other large social networks

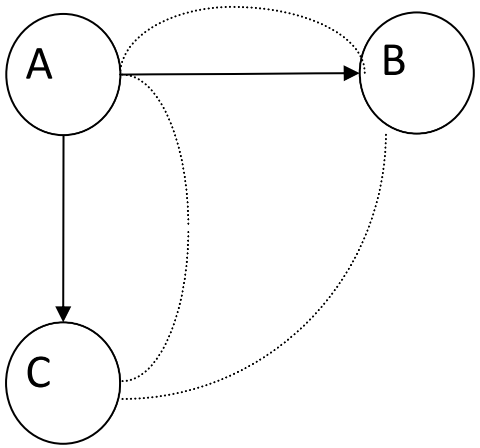


Is spatially concentrated



Causal link between co-worker network formation and regional growth

- The network forms dynamically and there are less mobility links than co-worker links.



Network density

$$D_c = \frac{2 \times L}{N_{reg} \times (N_{reg} - 1) - \sum_k N_k \times (N_k - 1)};$$

We can decompose network density into a direct-mobility and to an indirect-mobility component:

$$D_c = \sum_{ab}^{l} \frac{2 \times L_{ab}}{N_a \times N_b} \times \frac{N_a \times N_b}{N_{reg} \times (N_{reg} - 1) - \sum_a N_a \times (N_a - 1)} \times \delta_{ab}^{l};$$

pVAR regression

- Productivity growth in the region is induced not only by A-B and A-C links, but also by B-C links.
- Labor mobility in itself does not have a causal effect on growth.

Industry-region network

Larger industry-regions are more connected

Co-located industries are more strongly connected

Empirical strategy

- Within (FE) regressions on industry-regions (region-clustered S.E.-s and year dummies)
- Dependent variable: Gross income per capita (productivity)
- Years: 1995 to 2008
- Control variables: SPEC (log plants in IR), HC (share of bachelors), REGSIZE (tot emp – IR emp), IO (LQ of upward-linked industries), PLANTSIZE (average plant size)
- Explanatory variables constructed from the network
  - *Within* industry-regions
  - *Between* industry-regions *in the same* region
  - *Between* industry-regions *arcoss* regions

Within industry-regions

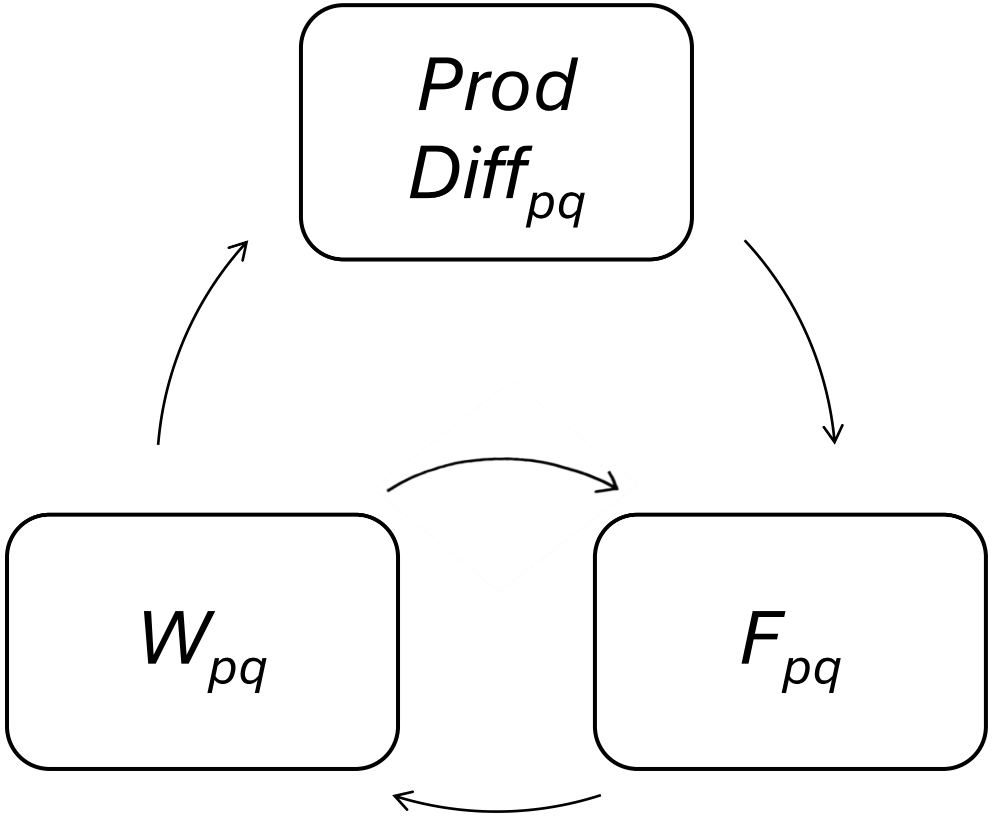
- DENSITY: Share of observed links compared to the number of possible links between firms within IR (MAR)

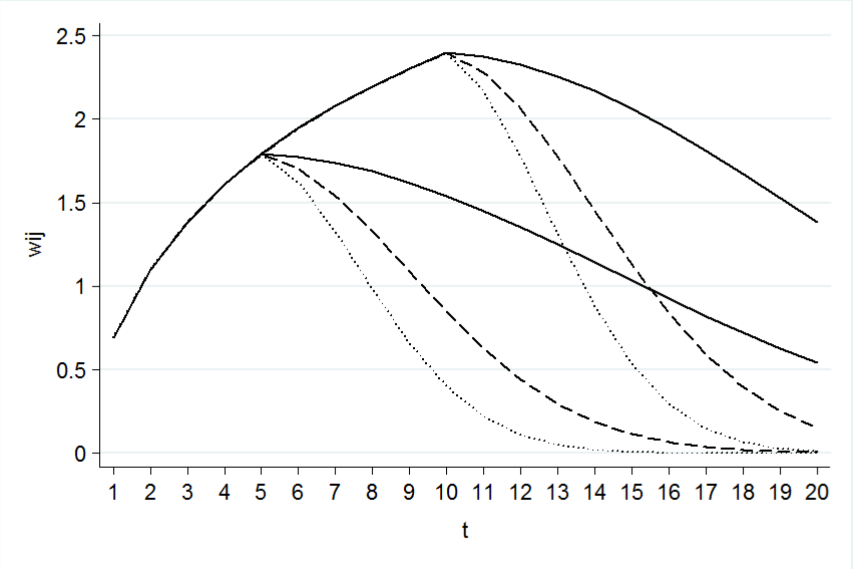
$$DENS_{pt} = \frac{2 \times L_{pt}}{N_{pt} \times (N_{pt} - 1)}$$

- TRANS (Triadic closure): The number of triangles among all possible triangles. (MAR)

$$\text{Trans}_{pt} = \frac{\sum_k \#\{lo \in pt | l \neq o,\ l \in N_k(pt),\ o \in N_k(pt)\}}{\sum_k \#\{lo | l \neq o,\ l \in N_k(pt),\ o \in N_k(pt)\}}$$

Aggregation to industries in regions: endogeneity





$$W_{pq,t} = \beta \sum_{u=t-1}^{t-n} F_{pq,u} \times e^{u\lambda} + \varepsilon_{pq,t}$$

- More productive and thus higher income attracts more employees, which creates new co-worker ties to other industries.
- Co-worker ties may induce new mobility.
- We estimate tie weights by previous labor flows and use the remaining error term to denote tie strength between industries.

Strength of relationships

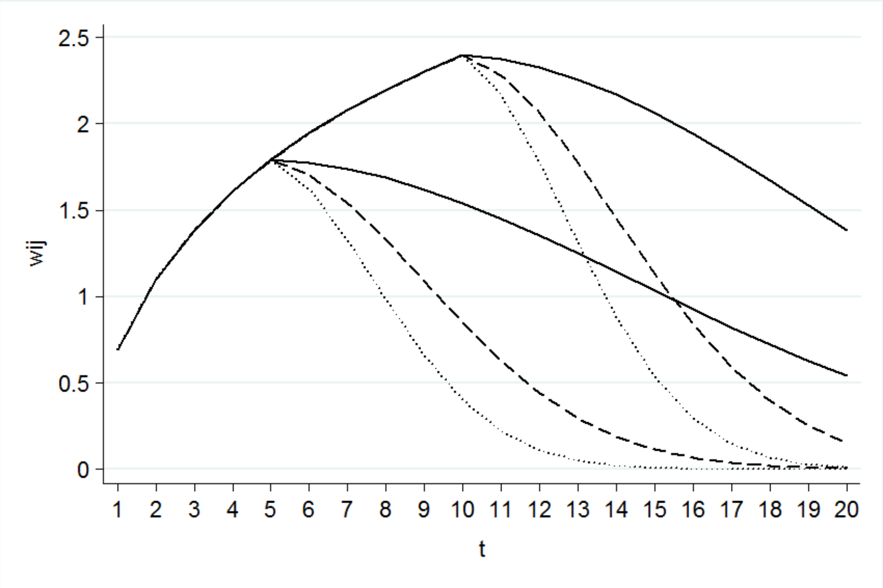
Increase when spending time together:

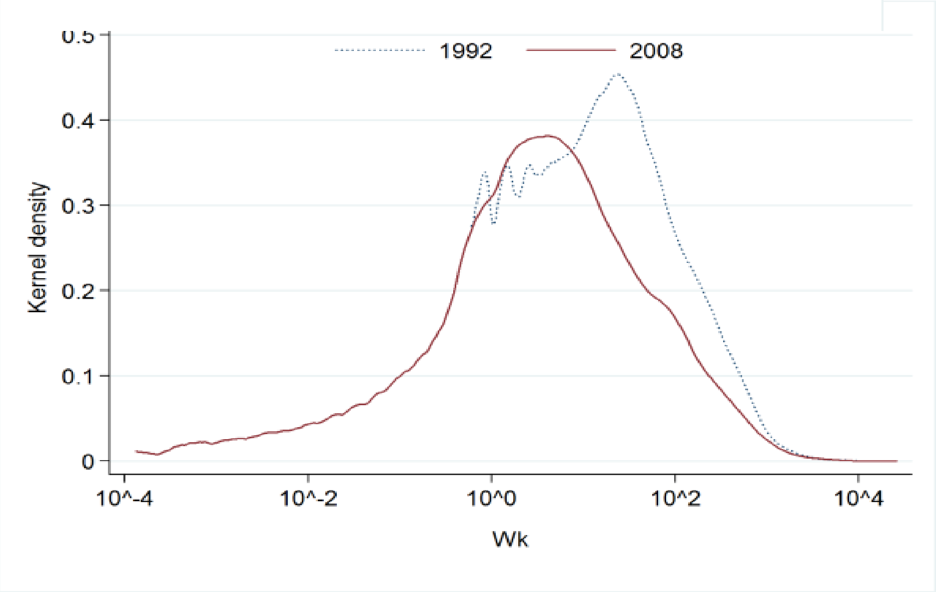
$$W_{ij}^{0} = \ln(t_l - t_f + 1)$$

Decrease exponentially as time passes after the last contact

$$W_{ij}^{t} = W_{ij}^{0} \times e^{-\lambda t}$$

The introduction of time log-normalizes degree distribution on the plant level





Across industries within regions

- REL: Logged tie strength to skill-related industries / logged tie strength to unrelated industries. (JACOBS)

$$REL_{pt} = \log\left(\frac{\sum_{p\in s}^{q\in s} W_{pq}\{pq \in s|\mathrm{RR_{pq}} \geq 0\}}{\sum_{p\in s}^{r\in s} W_{pr}\{pr \in s|\mathrm{RR_{pr}} < 0\}}\right)$$

Across industries in different regions

- DIVOUT: Diversity (entropy) of links normalized by the number of links of industry-region *p* at year *t outside* the region ("PIPELINE")

$$DIV(OUT)_{pt} = \frac{-\sum_{p \in r, \in t}^{q \notin r, \notin t} \frac{w_{pq}}{w_p} \times \log\left(\frac{w_{pq}}{w_p}\right)}{\log(L_p)}$$

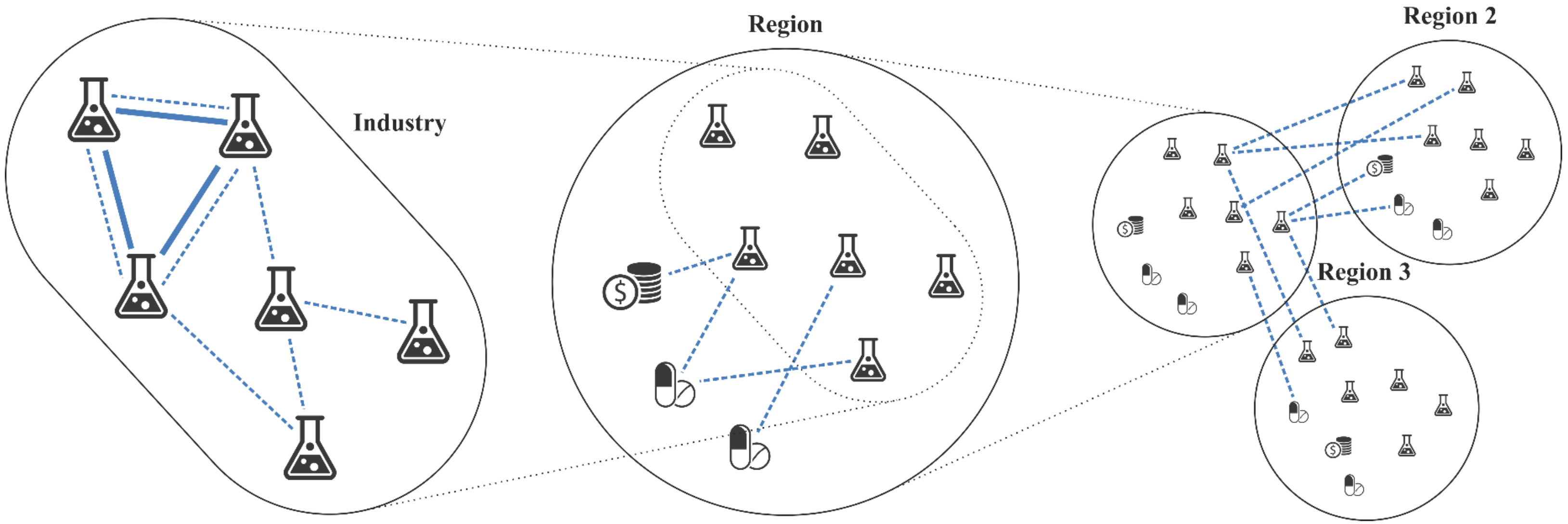
- DIVOUTSAME: Diversity (entropy) of inter-regional links to same industry (”PIPELINE”)

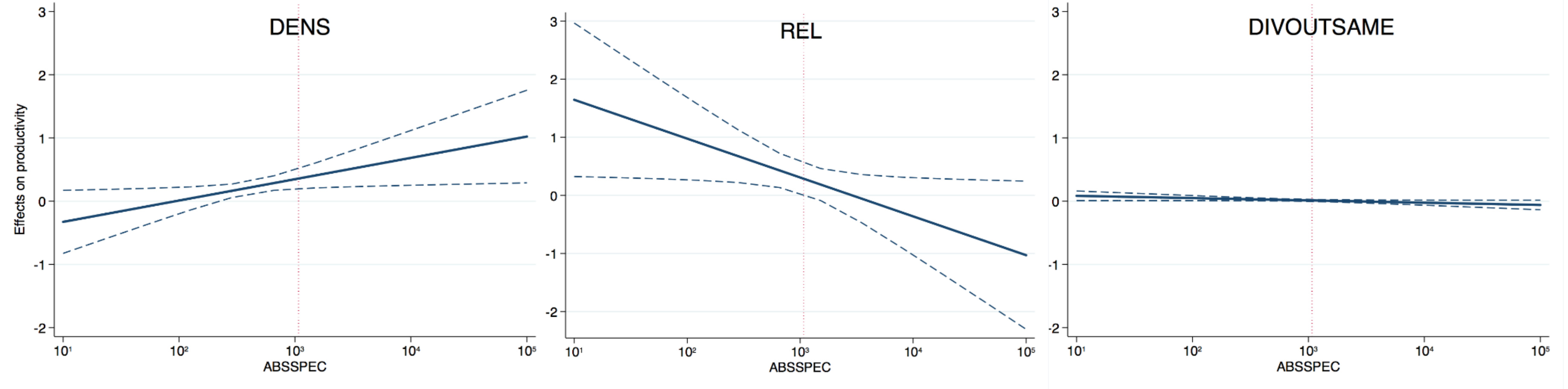
$$DIV(OUTSAME)_{pt} = \frac{-\sum_{p \in r, \in t}^{q \notin r, \in t} \frac{w_{pq}}{w_p} \times \log\left(\frac{w_{pq}}{w_p}\right)}{\log(L_p)}$$

Co-worker networks in geography

The effect of co-worker networks on regional growth of income depends on region size

(ABSSPEC)





Eriksson and Lengyel (2019) Economic Geography


BASED ON THE BEST-SELLING NOVEL BY THE AUTHOR OF
THE DA VINCI CODE AND ANGELS & DEMONS
TOM HANKS
FELICITY JONES
INFERNO
IS UPON US


JENNIFER LAWRENCE
JOEL EDGERTON
MATTHIAS SCHOENAERTS
JEREMY IRONS
RED SPARROW
MARCH 2


IT BEGINS
DUNE
IN THEATERS HBO MAX
OCTOBER 22


THEY SAY NO ONE CAN SAVE THE WORLD.
MEET NO ONE.
RYAN REYNOLDS
A MICHAEL BAY FILM
6 UNDERGROUND
DECEMBER 13 | NETFLIX


WILL SMITH
WILL SMITH
WHO WILL SAVE YOU FROM YOURSELF
GEMINI MAN
IN IMAX, 3D & 2D CINEMAS
COMING SOON


RYAN GOSLING
HARRISON FORD
BLADE RUNNER 2049
10.6.17


CHALAMET EDGERTON
THE KING
ALL HAIL
IN SELECT THEATERS AND ON
NETFLIX | THIS FALL


YOU DON'T NEED A PAST
TO HAVE A FUTURE
VIN DIESEL
BLOODSHOT
BASED ON THE BEST-SELLING COMIC BOOK
SUPERHEROES ARE GETTING AN UPGRADE
MARCH 13


WELCOME TO THE DAY AFTER JUDGMENT DAY
TERMINATOR
DARK FATE


THE WITCHER

Local capabilities and interregional connections







Knowledge sourcing in diversification

1. Relatedness of local capacities (Hidalgo et al., 2007, Neffke et al., 2011)

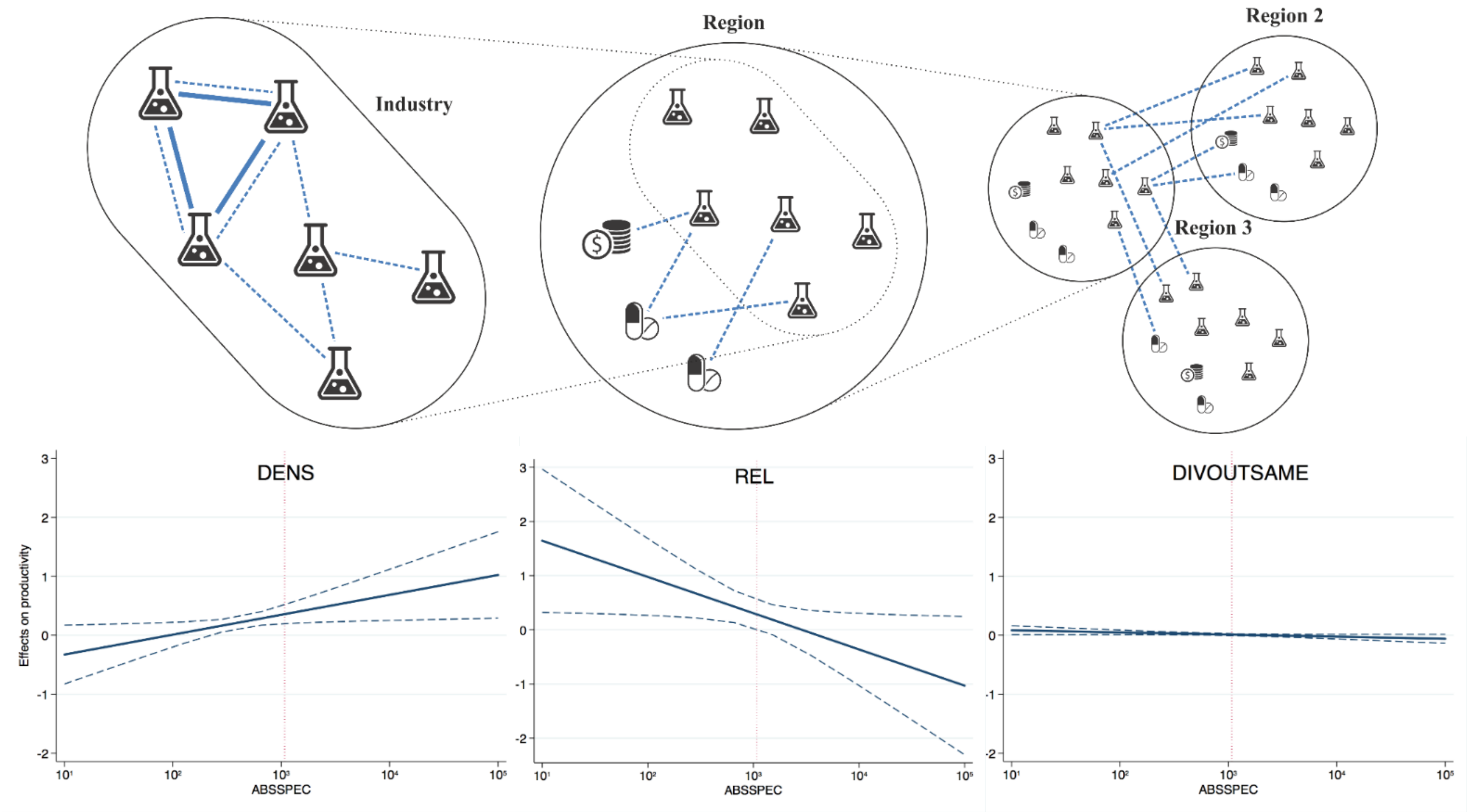
2. Interregional connections
    - Trade (Boschma and Iammarino, 2009)
    - Social networks as important channels (Boschma and Frenken, 2011)

3. Interregional networks substitute local capabilities in productivity growth (Eriksson and Lengyel, 2019)

4. R&D collaborations with regions that are specialized in complementary technologies favours technlogical diversification (Balland and Boschma, 2021).

Knowledge sourcing in diversification

- The effect of co-worker networks on regional growth of income depends on region size
- (ABSSPEC)



Eriksson and Lengyel (2019) Economic Geography

Question 1

Can interregional networks

substitute

local capabilities in diversification?

Lőrincz et al. (2020) Applied Network Science

Firms as drivers of diversification

Agents of structural change: mobile entrepreneurs (Neffke et al., 2018); foreign-owned firms (Elekes et al., 2019), R&D investments (Crescenzi et al., 2022).

Interregional knowledge sourcing dominantly happens within boundaries of multilocation companies (Frigon and Rigby, 2022; Zhang and Rigby, 2022).

Multilocation companies intensify collaboration to source from diversity (Alcacer and Zhao, 2012).

Firms as drivers of diversification

Exploitation of the local knowledge base favors related diversification (Wang and Zhao, 2018).

Agents of structural change and co-worker networks literature suggest that regions can source unrelated knowledge from the diversity of regions that multilocation companies connect (Neffke et al., 2018; Elekes et al., 2019; Eriksson and Lengyel, 2019; Crescenzi et al., 2022).

Question 2

Do internal connections

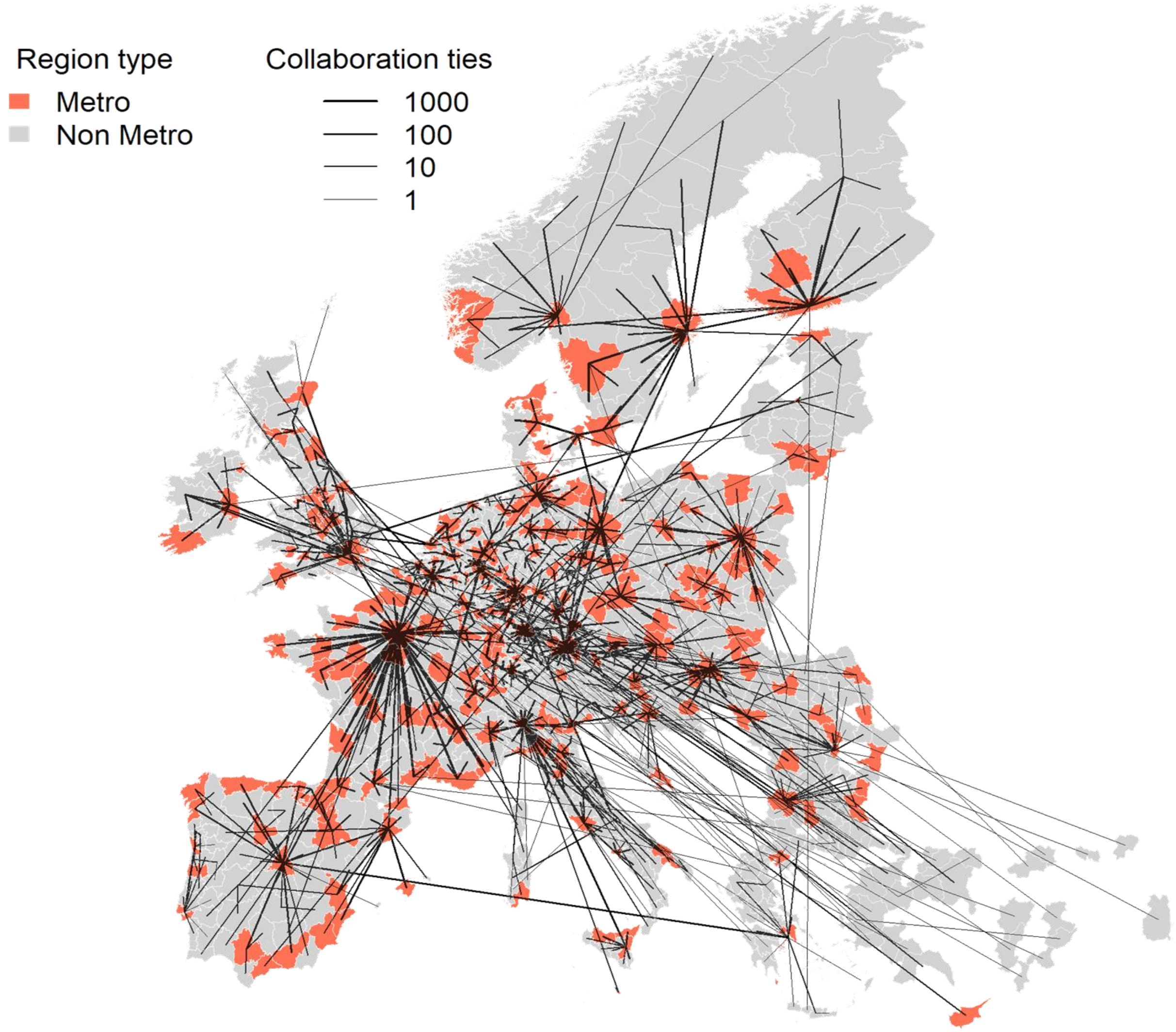
of multilocation firms across regions

favors diversification

into related or unrelated new activities?

Lőrincz et al. (2020) Applied Network Science

Data



- EPO PATSTAT
- 1981-2015, five-year periods
- 4-digit CPC codes
- NUTS3 regions (EUROSTAT classificaiton of metro/nonmetro)
- Fractionally split patents to regions by inventor addresses

Probability of ENTRY

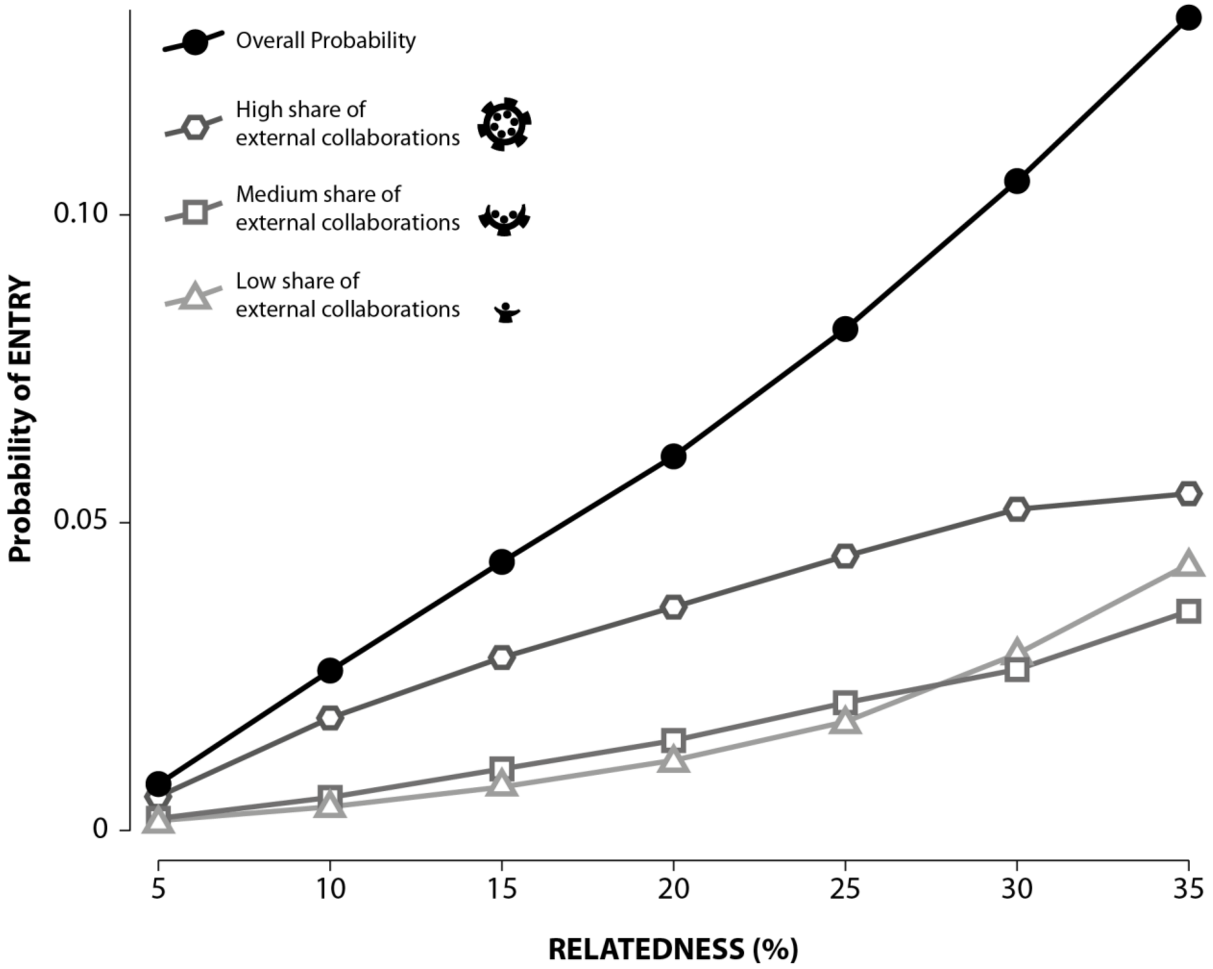
$$ENTRY_{i,r,t} = \begin{cases} 1\ if\ RTA_{i,r,t-1} < 1\ and\ RTA_{i,r,t} > 1 \\ 0\ otherwise \end{cases}$$

$$RTA_{i,r,t} = \frac{patents_{r,t} \frac{i}{\Sigma_i\ patents_{r,t}}(i)}{\Sigma_r\ patents_{r,t} \frac{i}{\Sigma_r \Sigma_i\ patents_{r,t}}(i)}$$

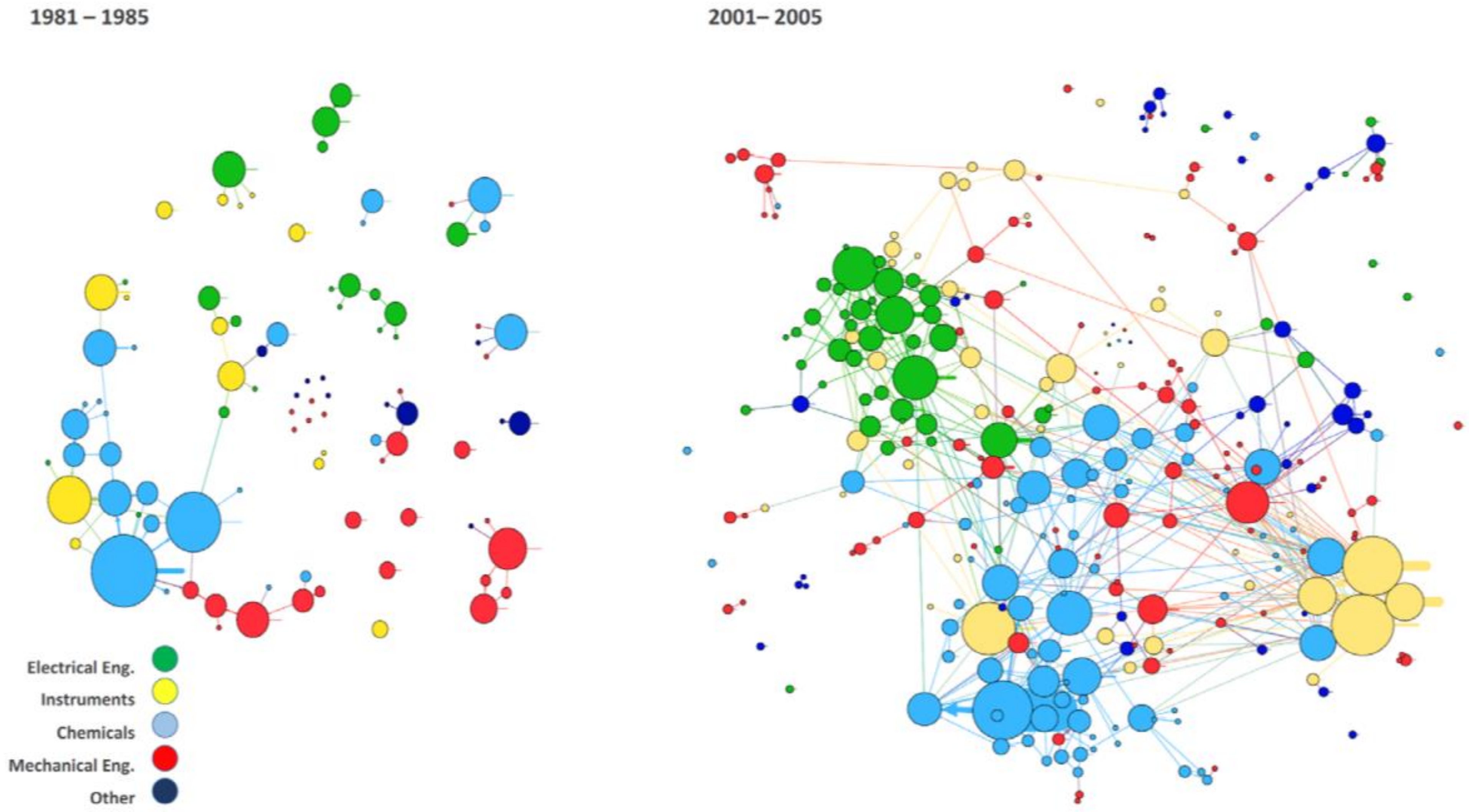
$$RELATEDNESS_{i,r,t} = \frac{\sum_{j \in r} x_j\ \varphi_{ij,t}}{\sum_j \quad \varphi_{ij,t}} \times 100$$

$$\varphi_{ij,t} = min\{P, \left(R \middle| TAx_{i,t} RTAx_{j,t}\right) \{P\left(R \middle| TAx_{j,t} RTAx_{i,t}\right)\}\}$$

$$EXTCOLLAB_{i,r,t} = \frac{EL_{i,r,t}}{EL_{i,r,t} + IL_{i,r,t}}$$



The knowledge space



Kogler and Whittle (2018) Edward Elgar Handbook on Geographies of Regions and Territories

Regression analysis

Regress the probability of ENTRY on NETWORK variables and controllers.

- Region-technology fixed effects (time fixed effects).
- LPM fit (logit).
- Lagged independent variables.

$$P\left(ENTRY_{i,r,t} = 1\right) = \alpha + \beta X_{i,r,t-1} + \gamma Z_{i,r,t-1} + \mu_{i,r,t} + \varepsilon_{i,r,t}$$

$$DIVERSITY_{i,r,t} = \frac{-\sum l_{rq} \times log\left(l_{rq}\right)}{log(k)}$$

$$INTENSITY_{i,r,t} = \frac{\sum_{q} L_{f\in i,r=f\in q}}{\sum_{q,f\in i,r=f\in q} N_{f\in i,r} \times N_{f\in q}}$$

Explanatory and control variables

Variable Name	Description
RELATEDNESS	Relatedness density (Hidalgo et al. 2007) calculated on the knowledge-space network. Measure of how close an emerging technology *i* is to other existing technologies in the region at time *t*.
EMPLOYMENT	Total number of employees in a region.
FIRMS	The number of firms in a region and CPC class that have been assigned at least one patent until *t*.
NEWFIRMS	The number of firms in a region and CPC class that have been assigned a first patent since *t*-1.
MULTILOC	The share of firms in the region and CPC class that employ inventors in other regions as well.
INVPERFIRM	The number of inventors per number of firms that are assignees of patents in the region and CPC class at period *t*.
EXTCOLLAB	The ratio of interregional collaborations among all collaborations of the inventors living in the region and patenting in a CPC class in period *t*.
DENSITY	The ratio of observed co-inventor collaboration among all possible collaborations the inventors in a specific CPC class could have within the region.
DIVERSITY	Entropy of the aggregated co-inventor connections from a CPC class and the region to other regions at period *t*.
INTENSITY	The ratio of observed links within the firm to other regions among all possible such links. This measure is calculated at the CPC and region level at period *t*.

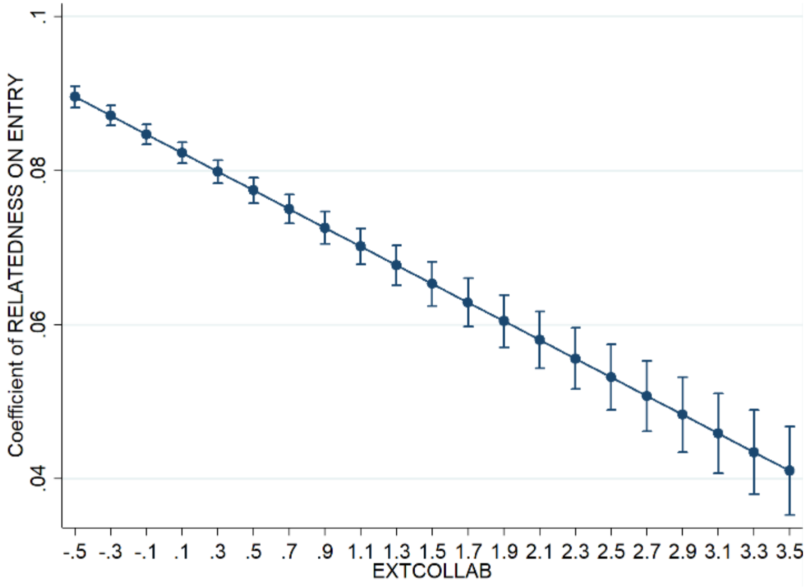
	(1) LPM	(2) LPM	(3) LPM	(4) LPM	(5) LPM	(6) LPM	(7) LPM	(8) LPM	(9) LOGIT	(10) LOGIT
	ALL REGIONS					METRO	NONMETRO	ALL REGIONS		
RELATEDNESS	0.084***	0.084***	0.087***	0.084***	0.087***	0.084***	0.089***	0.072***	0.726***	0.541***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.010)
EMPLOYMENT	0.025***	0.025***	0.024***	0.025***	0.024***	0.033***	0.016***	0.000	0.834***	0.350***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.022)	(0.021)
INVPERFIRM	-0.007**	-0.005*	-0.003	-0.007**	-0.005*	-0.006*	-0.003	-0.003	-0.004	0.015
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.009)	(0.009)
FIRMS	-0.023***	-0.023***	-0.013***	-0.024***	-0.014***	-0.012***	-0.016***	-0.009***	-0.082***	-0.054***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.011)	(0.011)
NEWFIRMS	0.014***	0.014***	0.010**	0.015***	0.006	0.005	0.004	0.002	0.065***	0.015
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.007)	(0.003)	(0.014)	(0.014)
MULTILOC	0.003***	0.003***	0.003***	0.003***	0.003***	0.004***	0.003***	0.003***	0.002	-0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.002)	(0.003)
DENSITY	0.003**	0.003**	0.003**	0.003**	0.003**	0.003*	0.002	0.001	-0.008*	-0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.004)	(0.004)

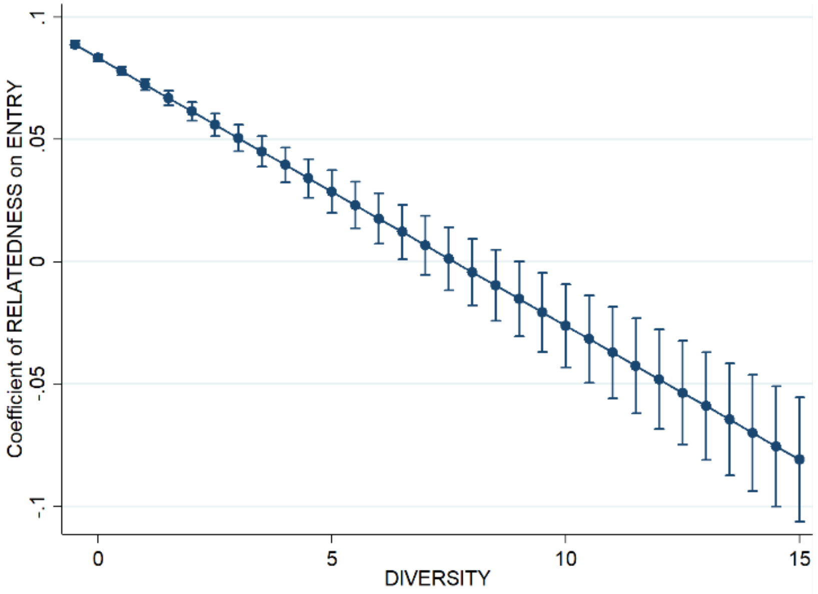
	(1) LPM	(2) LPM	(3) LPM	(4) LPM	(5) LPM	(6) LPM	(7) LPM	(8) LPM	(9) LOGIT	(10) LOGIT
	ALL REGIONS					METRO	NONMETRO	ALL REGIONS		
RELATEDNESS	0.084***	0.084***	0.087***	0.084***	0.087***	0.084***	0.089***	0.072***	0.726***	0.541***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.010)
EXTCOLLAB		0.002*	0.030***							
		(0.001)	(0.002)							
RELATEDNESS × EXTCOLLAB			-0.029***							
			(0.002)							
DIVERSITY				0.001	0.021***	0.023***	0.002	0.018***	0.153***	0.038**
				(0.001)	(0.003)	(0.004)	(0.007)	(0.003)	(0.014)	(0.014)
INTENSITY				0.001	0.023***	0.030***	0.021***	0.005	0.208***	0.017
				(0.001)	(0.003)	(0.004)	(0.005)	(0.003)	(0.012)	(0.012)
RELATEDNESS × DIVERSITY					-0.018***	-0.019***	0.005	-0.009**	-0.178***	-0.072***
					(0.003)	(0.003)	(0.008)	(0.003)	(0.013)	(0.012)
RELATEDNESS × INTENSITY					-0.023***	-0.027***	-0.024***	-0.006*	-0.230***	-0.049***
					(0.002)	(0.003)	(0.004)	(0.003)	(0.011)	(0.011)
CONSTANT	0.134***	0.135***	0.136***	0.134***	0.136***	0.122***	0.135***	0.228***		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.009)		
Period FE	NO	NO	NO	NO	NO	NO	NO	YES	NO	YES
N	2,418,740	2,418,740	2,418,740	2,418,740	2,418,740	613,604	1,805,136	2,418,740	762,490	762,490
R-sq	0.017	0.017	0.017	0.017	0.017	0.016	0.018	0.022		
AIC	150,348.7	150,329.5	149,423.6	150,348.5	149,314.0	254,731.8	-169,679.4	138,328.0	520,371.7	511,581.8
ll	-75,167.3	-75,156.8	-74,702.8	-75,165.2	-74,659.5	-127,354.9	84,850.7	-69,148.0	-260,174.9	-255,774.9

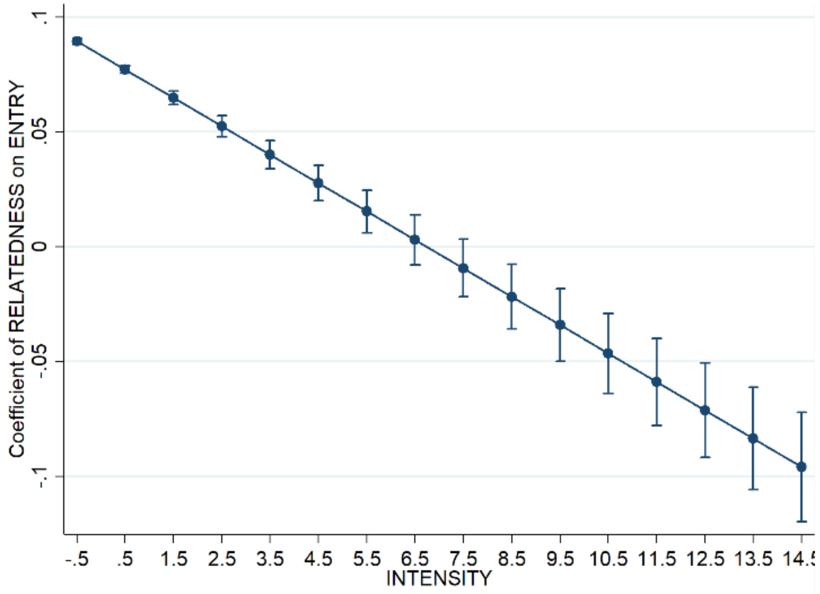
Interregional collaboration as substitution for relatedness?

Coefficients of RELATEDNESS at levels of EXTCOLLAB, DIVERSITY, INTENSITY.

- EXTCOLLAB: RELATEDNESS becomes smaller but remains positive.
- DIVERSITY, INTENSITY: RELATEDNESS becomes negative at their high levels.







Lőrincz et al. (2020) Applied Network Science
Lőrincz et al. (2020) Applied Network Science

Topics today

1. Labor mobility, co-worker networks, and economic development

2. Innovation and inequality in regional networks

Research Policy 52 (2023) 104886

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Atypical combinations of technologies in regional co-inventor networks

Milad Abbasiharofteh [a,b,c,*], Dieter F. Kogler [d,e,f], Balázs Lengyel [g,h,i]

[a] Department of Economic Geography, Faculty of Spatial Sciences, University of Groningen, 9747 AD Groningen, the Netherlands
[b] Department of Human Geography and Spatial Planning, Utrecht University, 3584 CB Utrecht, the Netherlands
[c] Department of Structural Change, Leibniz Institute of Agricultural Development in Transition Economies, 06120 Halle, Germany
[d] Spatial Dynamics Lab, University College Dublin, D04 V1W8 Dublin, Ireland
[e] Insight Centre for Data Analytics, University College Dublin, D04 V1W8 Dublin, Ireland
[f] School of Architecture, Planning & Environmental Policy, University College Dublin, D04 V1W8 Dublin, Ireland
[g] Agglomeration and Social Networks Lendület Research Group, Centre for Economic- and Regional Studies, Eötvös Loránd Research Network, 1097 Budapest, Hungary
[h] Laboratory for Networks, Technology and Innovation, Corvinus Institute for Advanced Studies, Corvinus University of Budapest, 1093 Budapest, Hungary
[i] Institute of Data Analytics and Information Systems, Corvinus University of Budapest, 1093 Budapest, Hungary

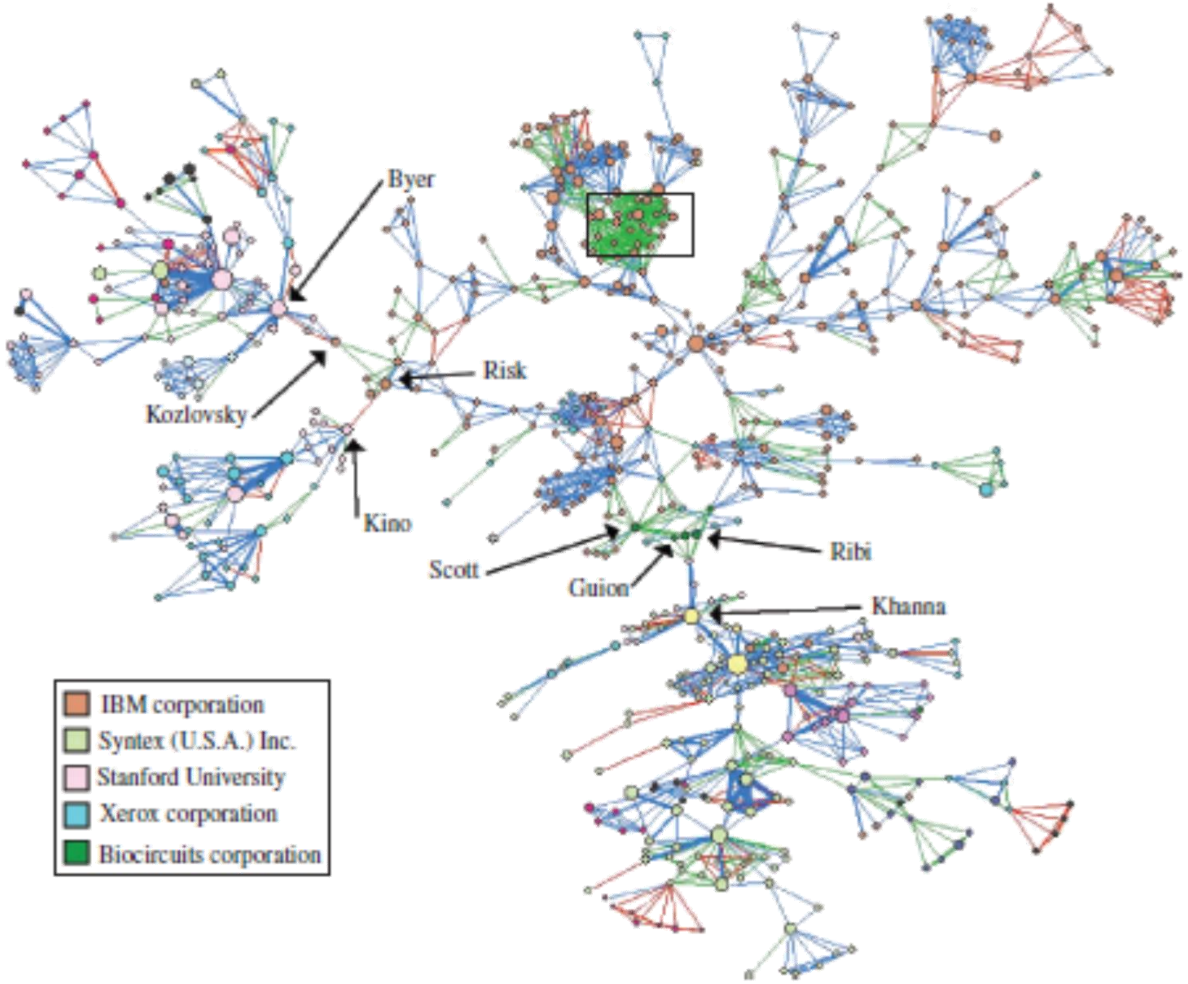
Organization Science
Vol. 18, No. 6, November–December 2007, pp. 938–954
ISSN 1047-7039 | EISSN 1526-5455 | 07 | 1806 | 0938

informs
DOI 10.1287/orsc.1070.0289

Small Worlds and Regional Innovation

Lee Fleming
Harvard Business School, Morgan Hall 485, Boston, Massachusetts 02163, lfleming@hbs.edu
Charles King III
Greylock McKinnon Associates, One Memorial Drive, Suite 1410, Cambridge, Massachusetts 02142, and
Pleiades Consulting Group, Inc., P.O. Box 711, Lincoln, Massachusetts 01773, cking@gma-us.com
Adam I. Juda
Google, Inc., 76 Ninth Avenue, 4th Floor, New York, New York 10011, juda@goggle.com

Short access in the co-inventor network speeds up information flow and induces the combination of knowlegde.



The role of knowledge: diversity vs specialization

Radical novelty needs diverse knowledge (Florida et al. 2017, Glaeser et al. 1992).

Specialization can also facilitate innovation (Beaudry and Schiffauerova, 2009; Lobo and Strumsky, 2008) when critical masses of experts specialized in distinct pieces of knowledge (Castaldi et al., 2015) are connected through knowledge transfer mechanisms (Berkes and Gaetani, 2020).

Boston biotechnology example (Cooke 2002, Powell et al. 1996): distinct local critical masses in engineering and biology connected by social ties.

The role of knowledge: network dynamics

Similar specialization in knowledge increases link probability.

- Local collaboration networks can become too cohesive and locked-in into technologies – Grabher (1993), Boschma & Frenken (2010), Giuliani (2013), Balland et al. (2016)

Inter-regional links can introduce novelty

- Inter-regional links combined with dense local networks
  - – theory: Bathelt et al. (2004), Glückler (2007);
  - – empirical evidence: Breschi and Lenzi (2016), Eriksson and Lengyel (2019); Kogler et al. (2023)
- Local variation – Juhász and Lengyel (2018)

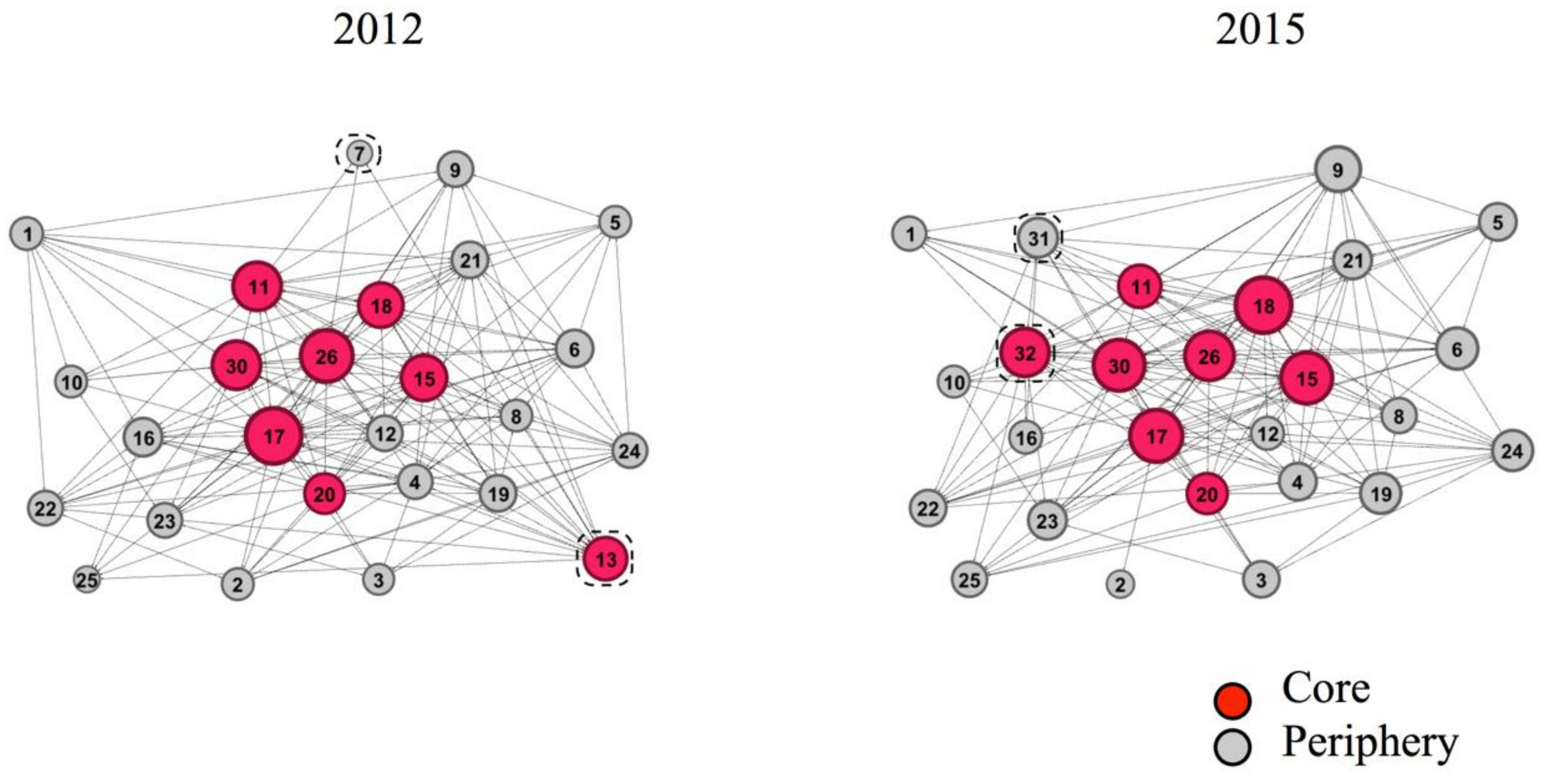
Closure and variation in networks

CONTEXT

The printing and paper product cluster of Kecskemét, Hungary

DATA

Face-to-face interviews
in 2012 and 2015

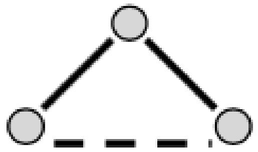
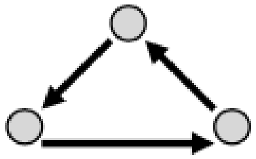


METHODOLOGY

Stochastic Actor Oriented Models [RSiena]

[Shortly..] detects the influential effect of network change in structural, dyadic and individual level

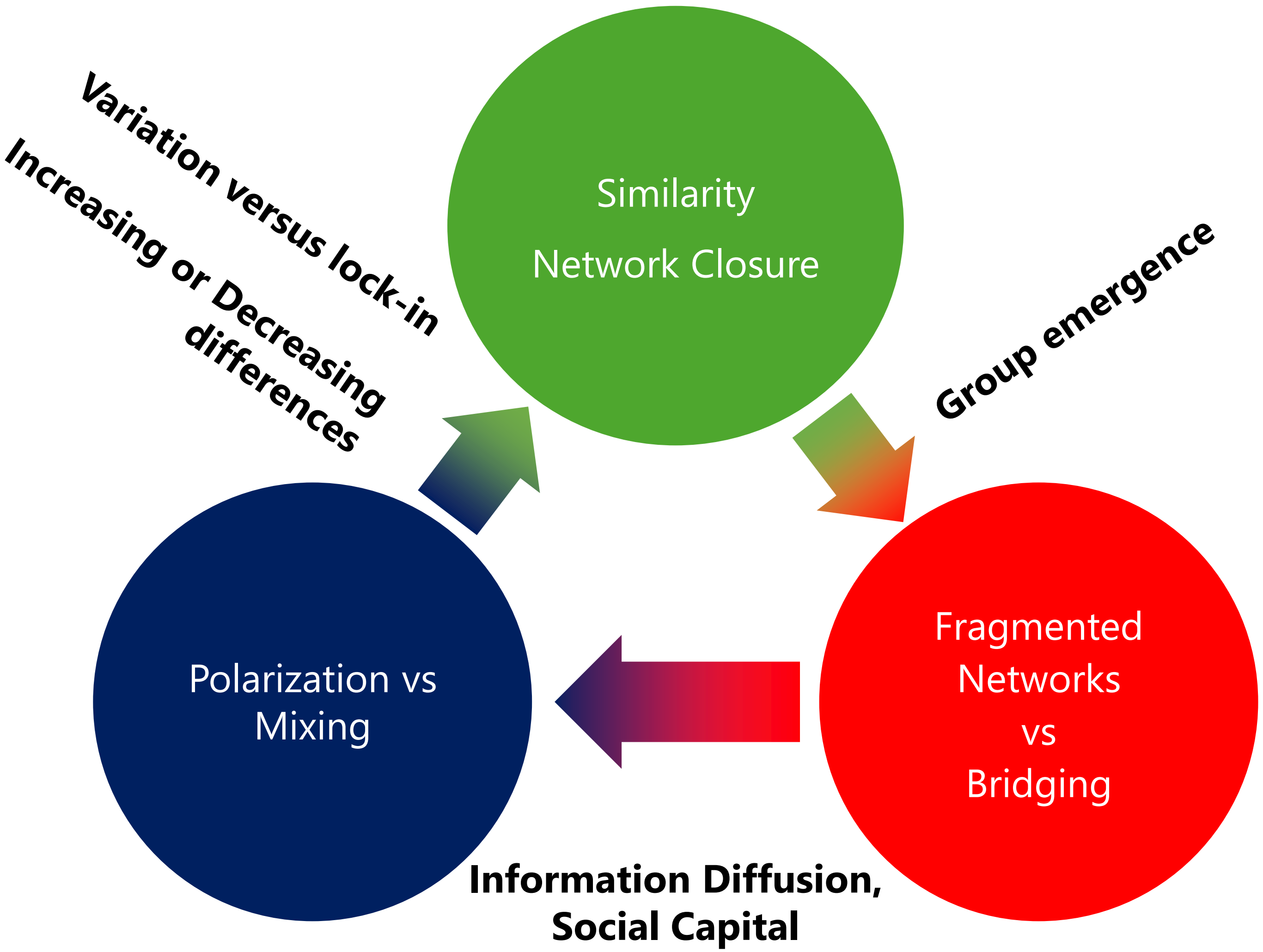
	Evaluation		Number of ties	**Creation**		Number of ties	**Endowment**		Number of ties
	t	$t+1$		t	$t+1$		t	$t+1$	
Creation	$i \quad j$	$i \rightarrow j$	71	$i \quad j$	$i \rightarrow j$	71			
Persistence	$i \rightarrow j$	$i \rightarrow j$	110				$i \rightarrow j$	$i \rightarrow j$	110
Termination	$i \rightarrow j$	$i \quad j$	113				$i \rightarrow j$	$i \quad j$	113
No ties	$i \quad j$	$i \quad j$	462	$i \quad j$	$i \quad j$	462			
Odds ratio			181/575			71/462			110/113

	Structural variables		
	Description	Formula	Visualization
Triadic closure	Tendency toward triadic closure when two knowledge ties existed in the previous period	$T_i = \sum_{j,h} x_{ij} x_{ih} x_{jh}$	
Reciprocity	Tendency of mutual knowledge exchange	$R_i = \sum_{j} x_{ij} x_{ji}$	
Cyclicity	Tendency of knowledge exchange in cycles	$C_i = \sum_{j,h} x_{ij} x_{jh} x_{hi}$	
Density	Overall tendency of actors to ask advices	$D_i = \sum_{j} x_{ij}$	
	Dyadic variables		
Geographical proximity	Physical distance of firms subtracted from the maximum distance in the sample		
Cognitive proximity	Number of digits two firms share in common in their NACE 4 codes		
Triadic closure X cognitive proximity	Number of common third partners multiplied by the number of digits two firms share in common in their NACE 4 codes		
	Firm level variables		
External knowledge ties	Number of knowledge linkages outside the region		
Age	Number of years since establishment		
Ownership	Equals *1* if foreign and *0* if domestic		
Employment	Total number of employees		

RESULTS

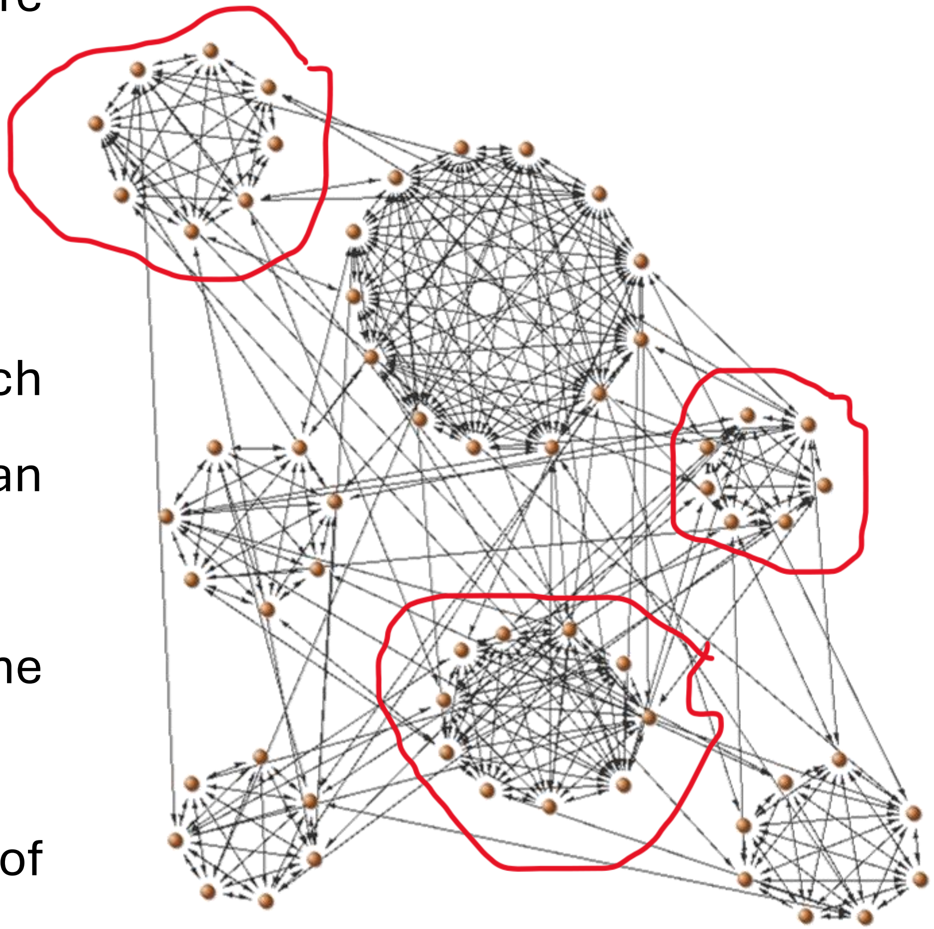
	Evaluation		Creation		Endowment	
	Model (1)	Model (2)	Model (1)	Model (2)	Model (1)	Model (2)
Triadic closure	0.191***	0.218***	0.422***	0.462***	-0.103	0.003
	(0.032)	(0.036)	(0.059)	(0.072)	(0.108)	(0.131)
Geographical proximity	0.031	0.043	0.172*	0.259**	-0.103	-0.086
	(0.041)	(0.043)	(0.104)	(0.123)	(0.081)	(0.078)
Cognitive proximity	0.111**	0.276***	0.062	0.361***	0.194**	0.448***
	(0.050)	(0.077)	(0.092)	(0.137)	(0.083)	(0.146)
Triadic closure X Cognitive proximity		-0.049***		-0.141***		-0.055**
		(0.017)		(0.044)		(0.027)
Firm level controls	Yes	Yes	Yes	Yes	Yes	Yes
Network level controls	Yes	Yes	Yes	Yes	Yes	Yes
Iteration steps	3898	4194	4141	4194	4191	4194
Convergence t-ratios	< 0.07	< 0.03	< 0.03	< 0.05	< 0.04	< 0.07

A Network Dynamics Framework



Problem: the impact of knowledge and network dynamics are difficult to disentangle.
Solution: network communities

Network communities: cohesive segments of networks that are loosely connected to other communities.

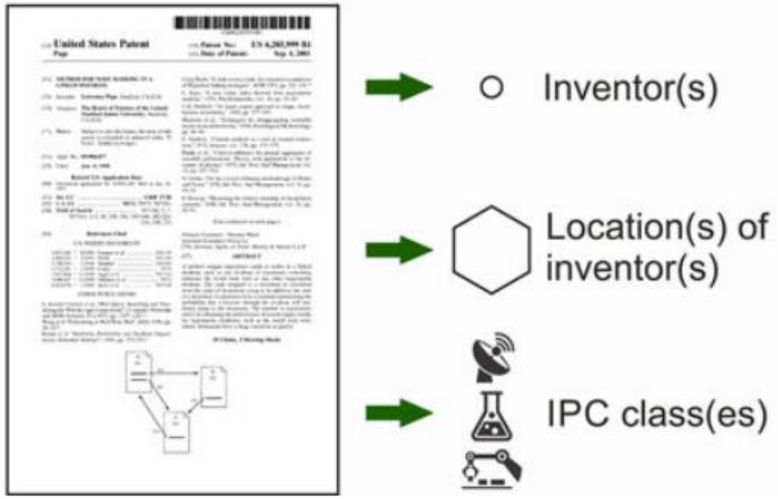


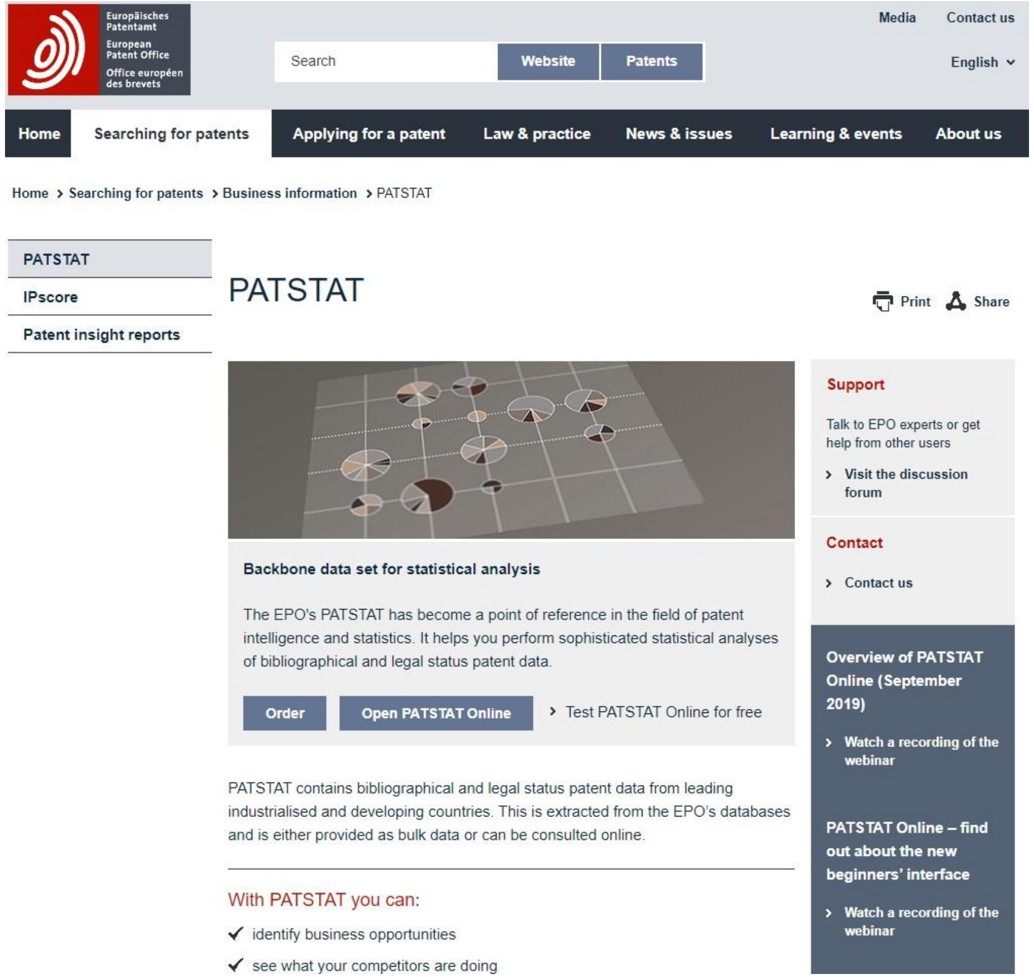
Advantages:

- Small-world networks can be decomposed to communities such that bridges are inter-community ties (Girwan and Newman 2002)
- Community detection only considers the structure of the network (Fortunato 2010) – Triadic Closure effect
- One can measure the technological specialization of communities – Technological Proximity effect

Data from European Patents

- The European Patent Office (EPO) PATSTAT database
- Creating co-inventor network in seven non-overlapping 5-year time-windows for each NUTS2 region





Tóth et al. (2021) European Planning Studies

Atypical patents as radical innovation

SHARE

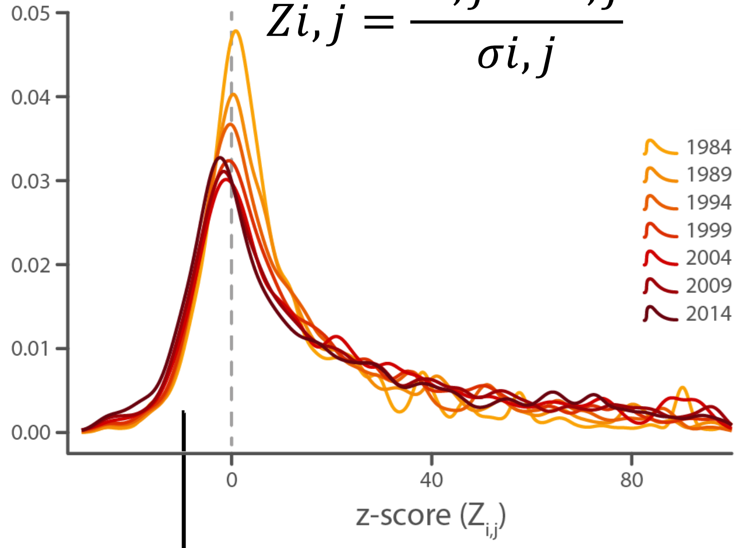
REPORT

Atypical Combinations and Scientific Impact

Brian Uzzi[1,2], Satyam Mukherjee[1,2], Michael Stringer[2,3], Ben Jones[1,4,*]
+ See all authors and affiliations

Science 25 Oct 2013:
Vol. 342, Issue 6157, pp. 468-472
DOI: 10.1126/science.1240474

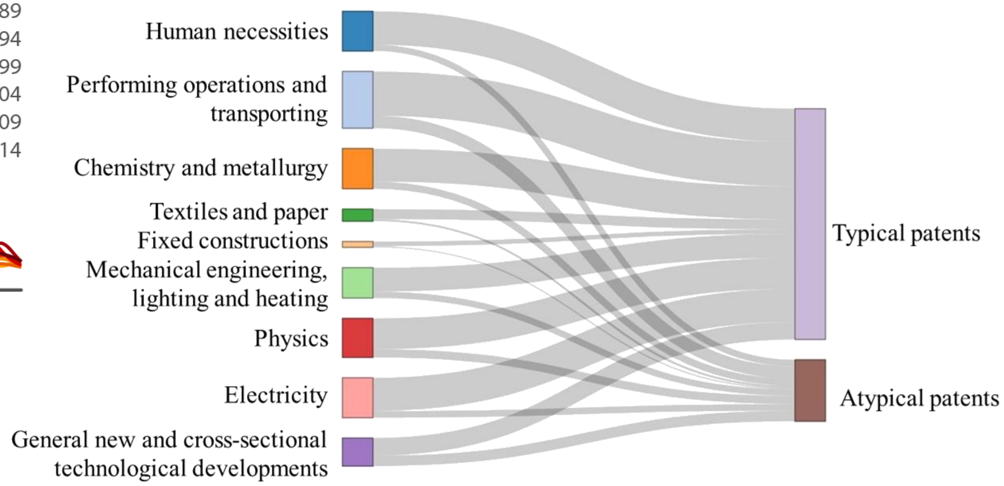
$$Zi,j = \frac{oi,j - ei,j}{\sigma i,j}$$

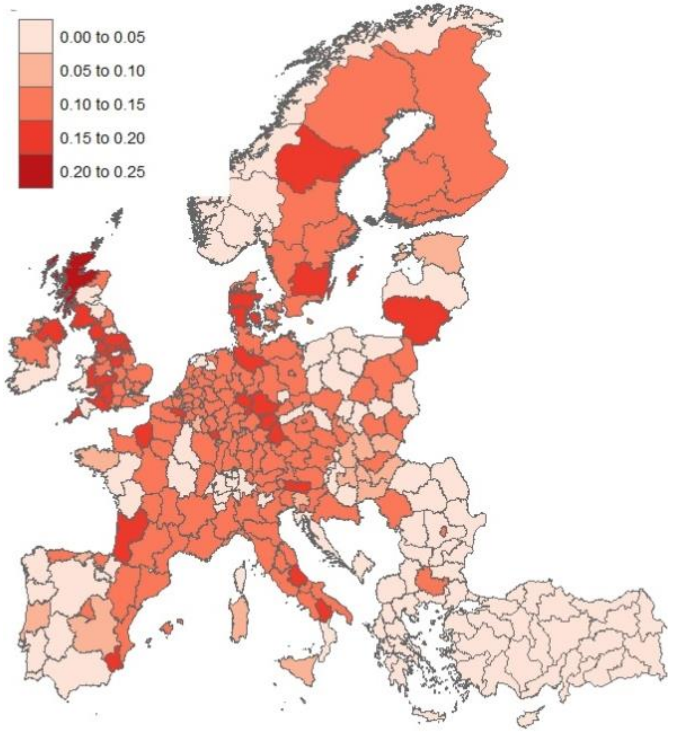


< 30%

Negative values represent atypical combination of classes

Atypical patents concentrate in cities (Mewes 2019).

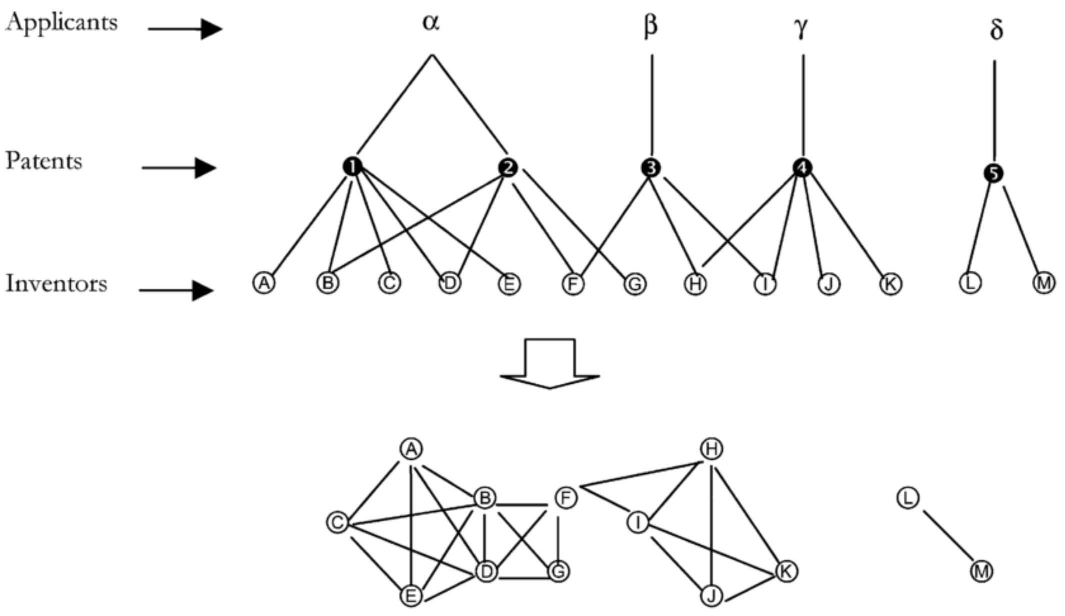


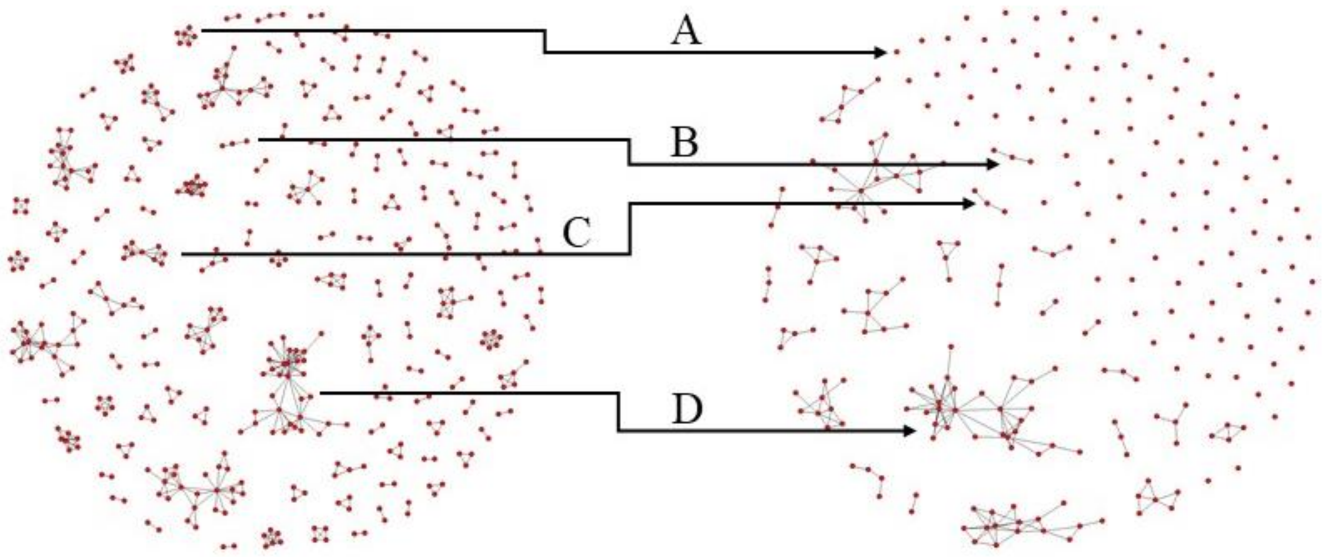


Dependent variable: share of atypical patents in EU regions, to avoid the effect of scale.

Network creation

We handle structurally equivalent inventors as one node in the network.





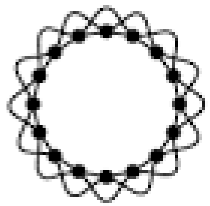
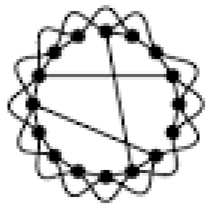
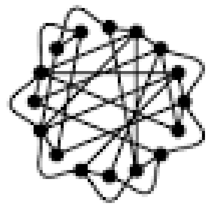
Breschi et al. (2021) Research Policy, Lucena-Piquero and Vicente (2019) Research Policy

Network variables 1

Small-worldness (Neal, 2018)

1. Average path length / compared to a random network
2. Clustering / compared to lattice

Small worldness ranges from 0 to 1.

			
Network	Lattice, Ordered	Small World	Random, Disordered
Clustering Coefficient	High	High	Low
Mean Path Length	Long	Short	Short

$$\omega = \frac{L_r}{L} - \frac{C}{C_l}$$

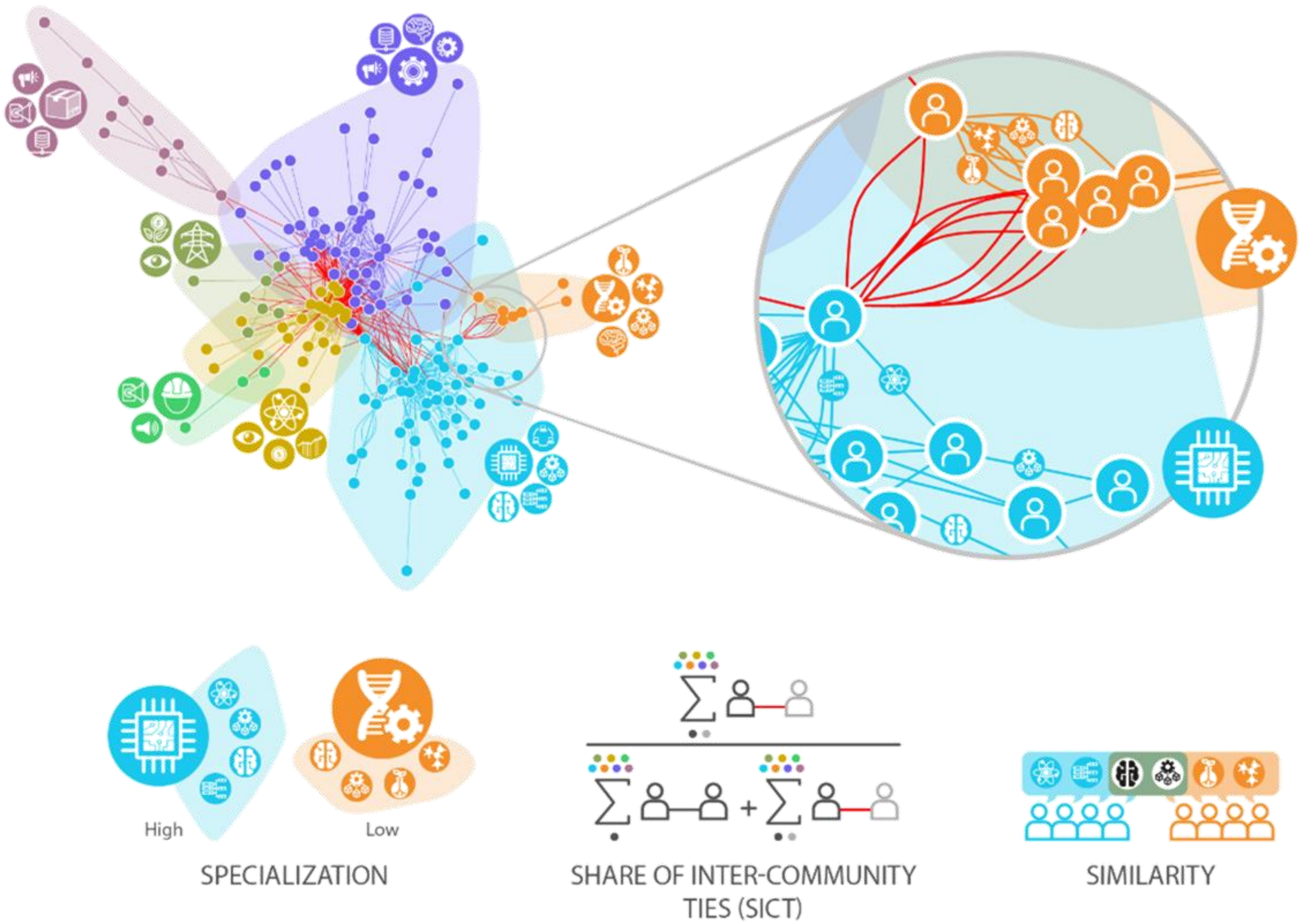
$$SMALLWORLDNESS = 1 - |\omega|$$

Network variables 2

Network communities are detected by the Louvain method

Regional-level indicators

1. Specialization: median level of technology HHI in network communities (not correlated with community and inventor numbers)
2. SICT: share of inter-community ties compared to randomly rewired networks
3. Similarity: median level of technology correlations (Spearman) across connected communities



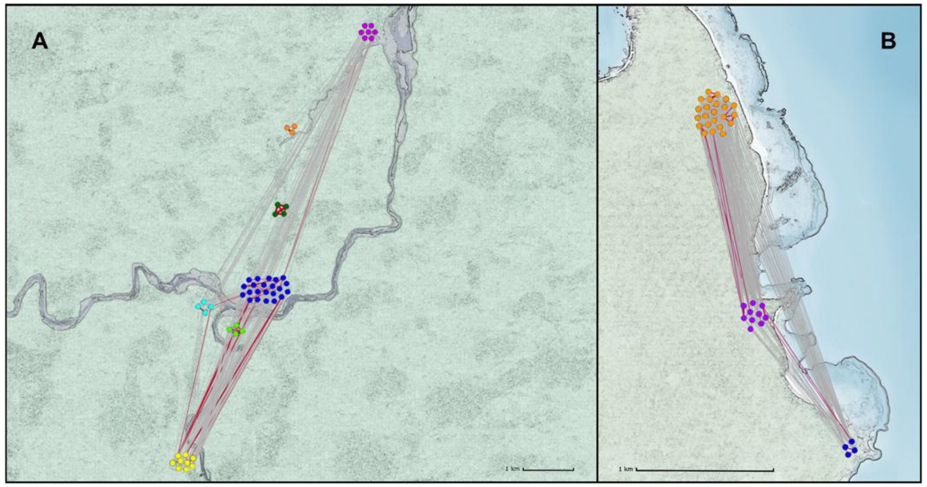
Results

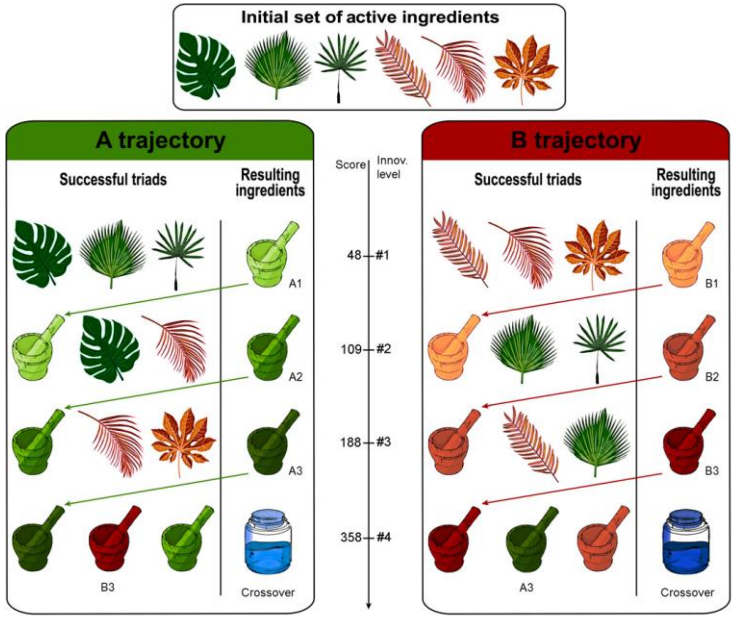
$$Y_{r,t} = \alpha + \beta_1 SMALLWORLDNESS_{r,t-1} + \beta_2 SPECIALIZATION_{r,t-1} + \beta_3 SICT_{r,t-1} + \beta_4\, SIMILARITY_{r,t-1} + \beta_5 N_{r,t-1} + \beta_6 Z_{r,t-1} + \varphi_t + \mu_r + \varepsilon_{r,t}$$

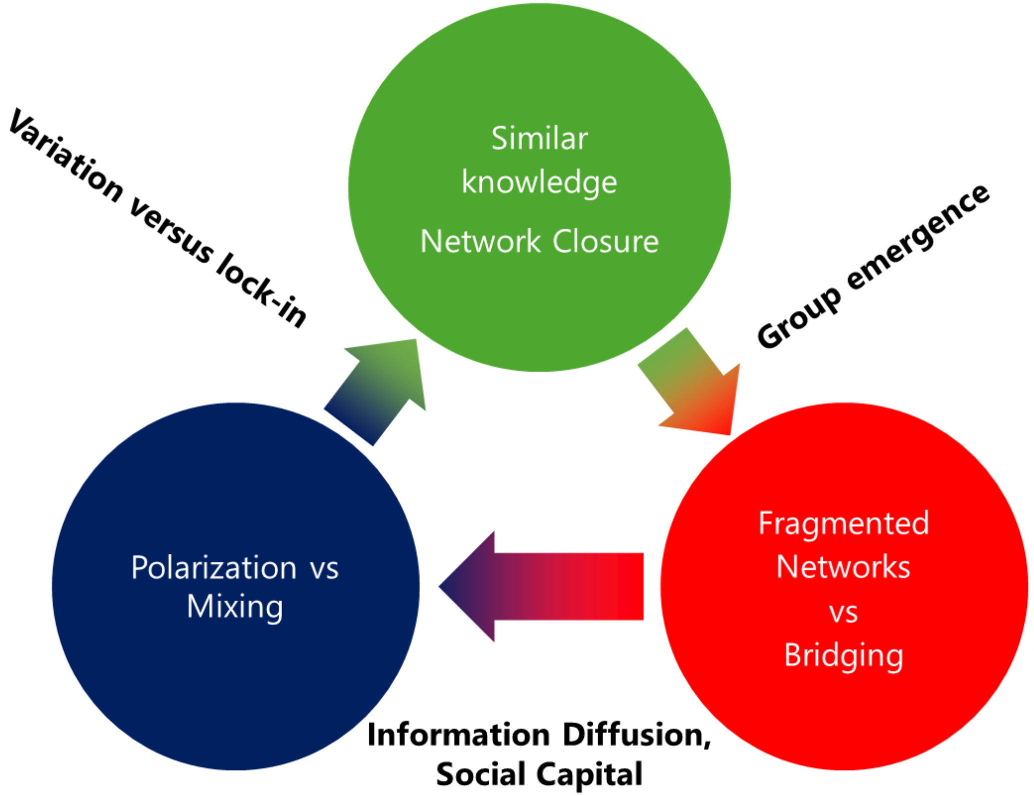
Controls:
1. Average Related Density (van de Wouden and Rigby 2019)
2. Complexity (Hidalgo and Hausmann 2009)
3. Density (Bergé et al., 2018; Breschi and Lenzi, 2016)
4. Isolate (Lobo and Strumsky 2008)
5. Community
6. Population
7. Interregional ties

	Share of atypical patents	Total patent growth rate
SMALLWORLDNESS	0.0115**	0.0336
	(0.0052)	(0.0796)
SPECIALIZATION	0.0558**	1.6982***
	(0.0220)	(0.4397)
SICT	0.1039**	0.5302
	(0.0493)	(0.7598)
SIMILARITY	-0.0231**	-0.3266*
	(0.0094)	(0.1709)
Region FE	YES	YES
Time FE	YES	YES
Observations	1,526	1,526
R^2	0.4172	0.3684

Network dynamics of specialization and radical combination



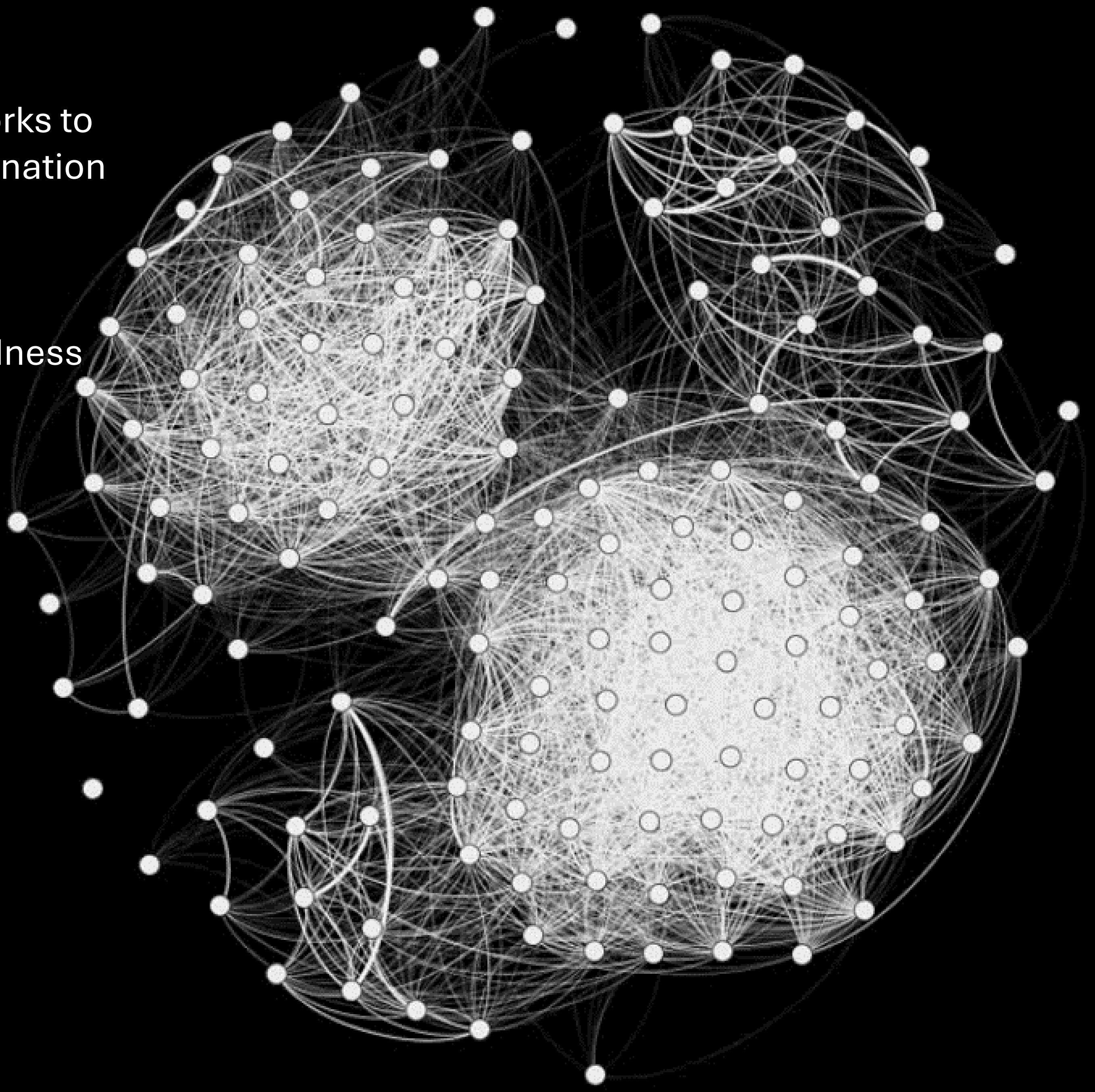




Migliano et al. (2020, Science Advances)

The role of diffusion in networks to understand knowldge combination

- diverse vs cohesive links,
- type of knowledge,
- organizations and firms,
- specialization and relatedness



Aplin et al. (2015) Nature

nature COMMUNICATIONS

ARTICLE

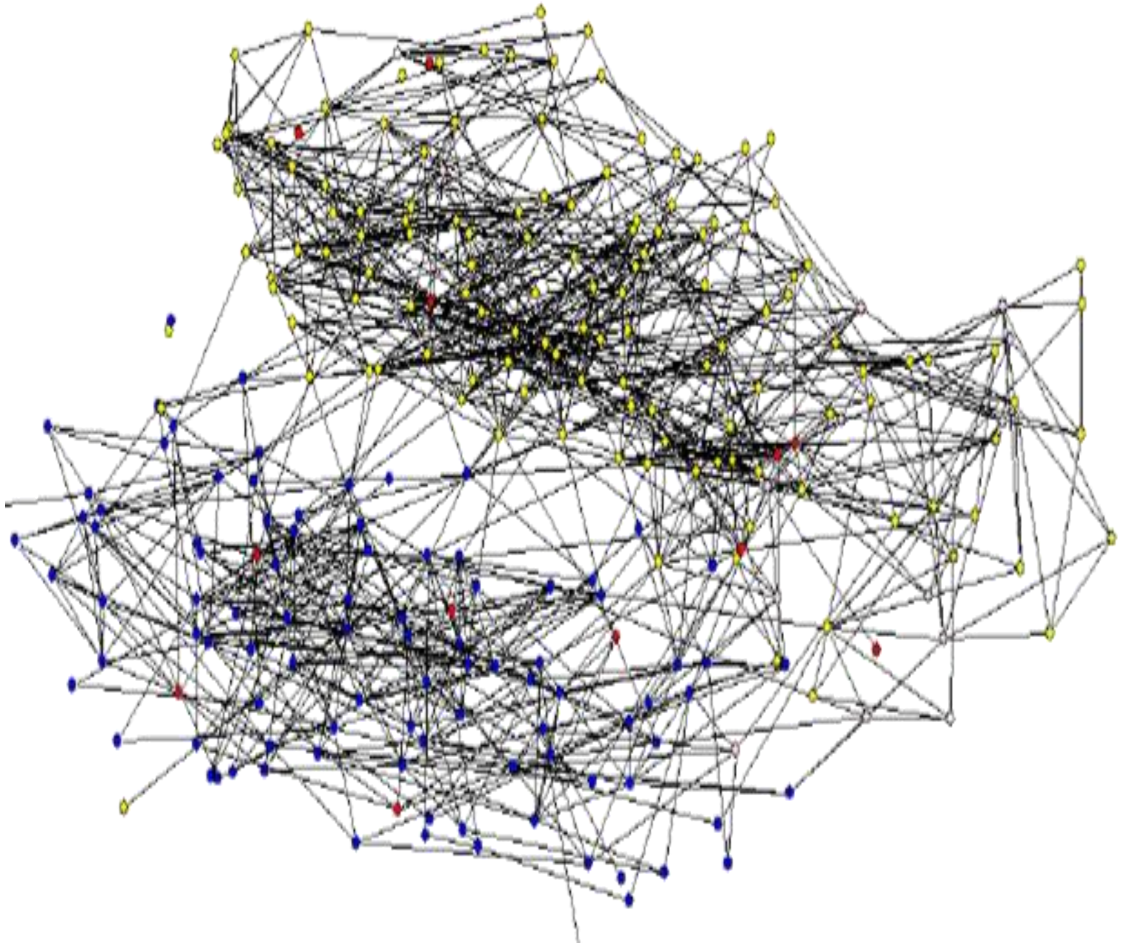
https://doi.org/10.1038/s41467-021-21465-0 OPEN

Inequality is rising where social network segregation interacts with urban topology

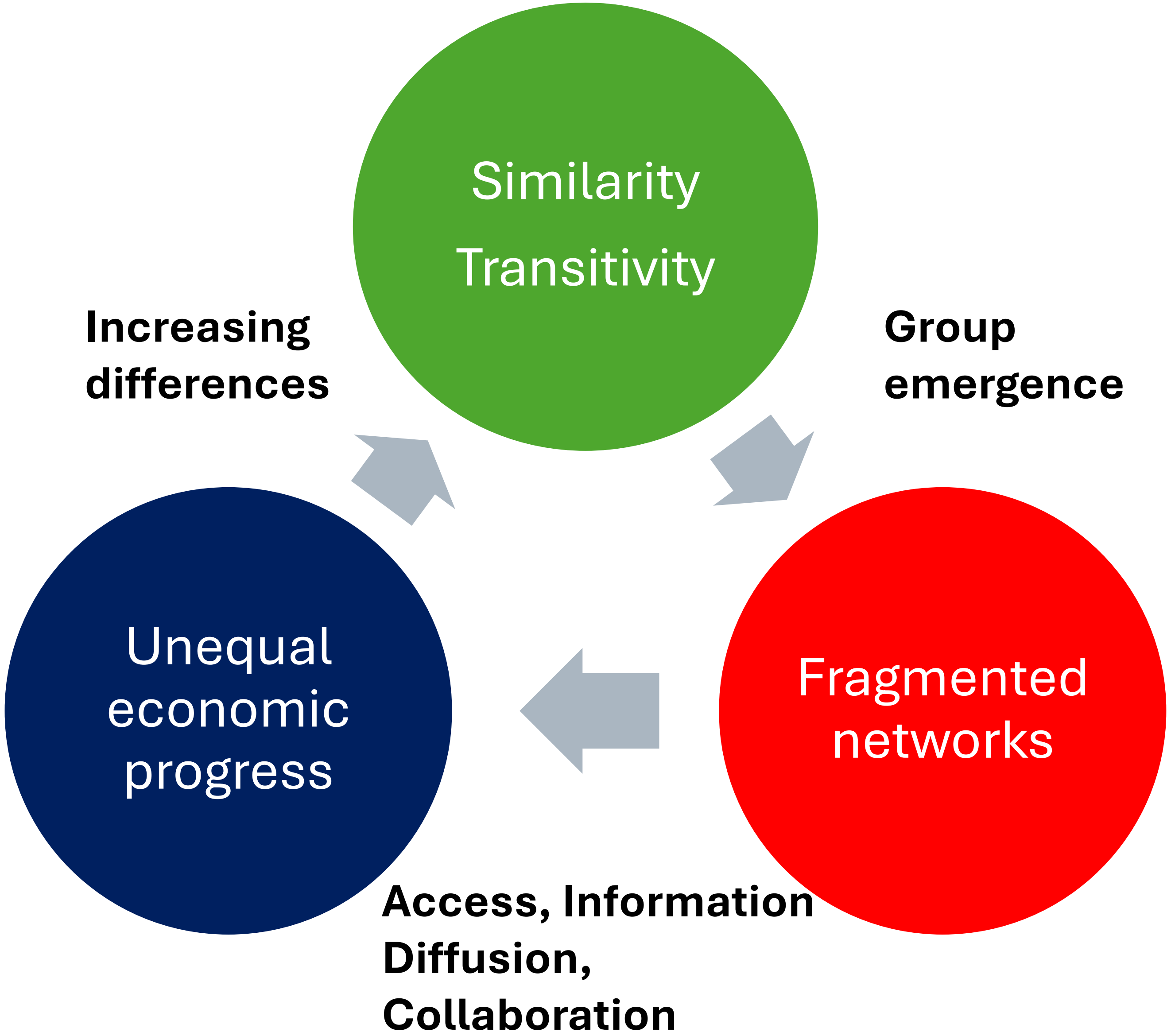
Gergő Tóth[1,2,13], Johannes Wachs[3,4,13], Riccardo Di Clemente[5,6], Ákos Jakobi[7,8], Bence Ságvári[1,9,10], János Kertész[11] & Balázs Lengyel[1,10,12]✉

Social networks and inequalities

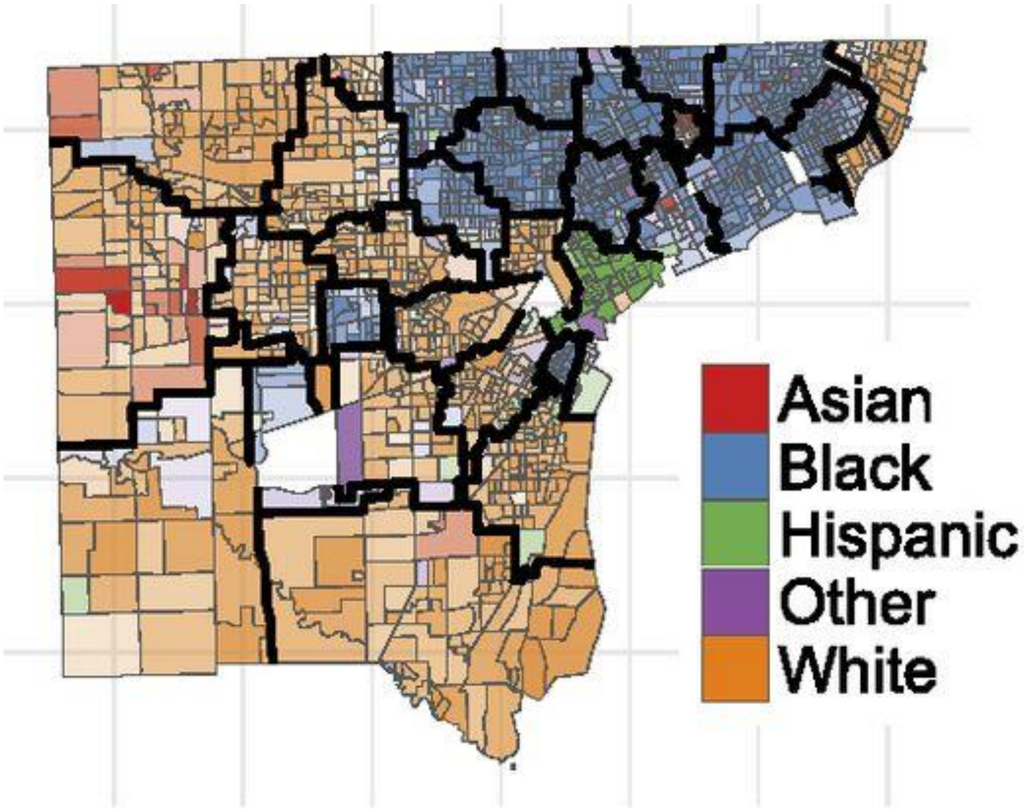
DiMaggio-Garip (2012) Annual Review of Sociology



Currarini-Pin-Jackson (2009) Econometrica



Local inequalities and consequences of spatial segregation



(a) Detroit, $k = 34$


Article | Open access | Published: 01 August 2022

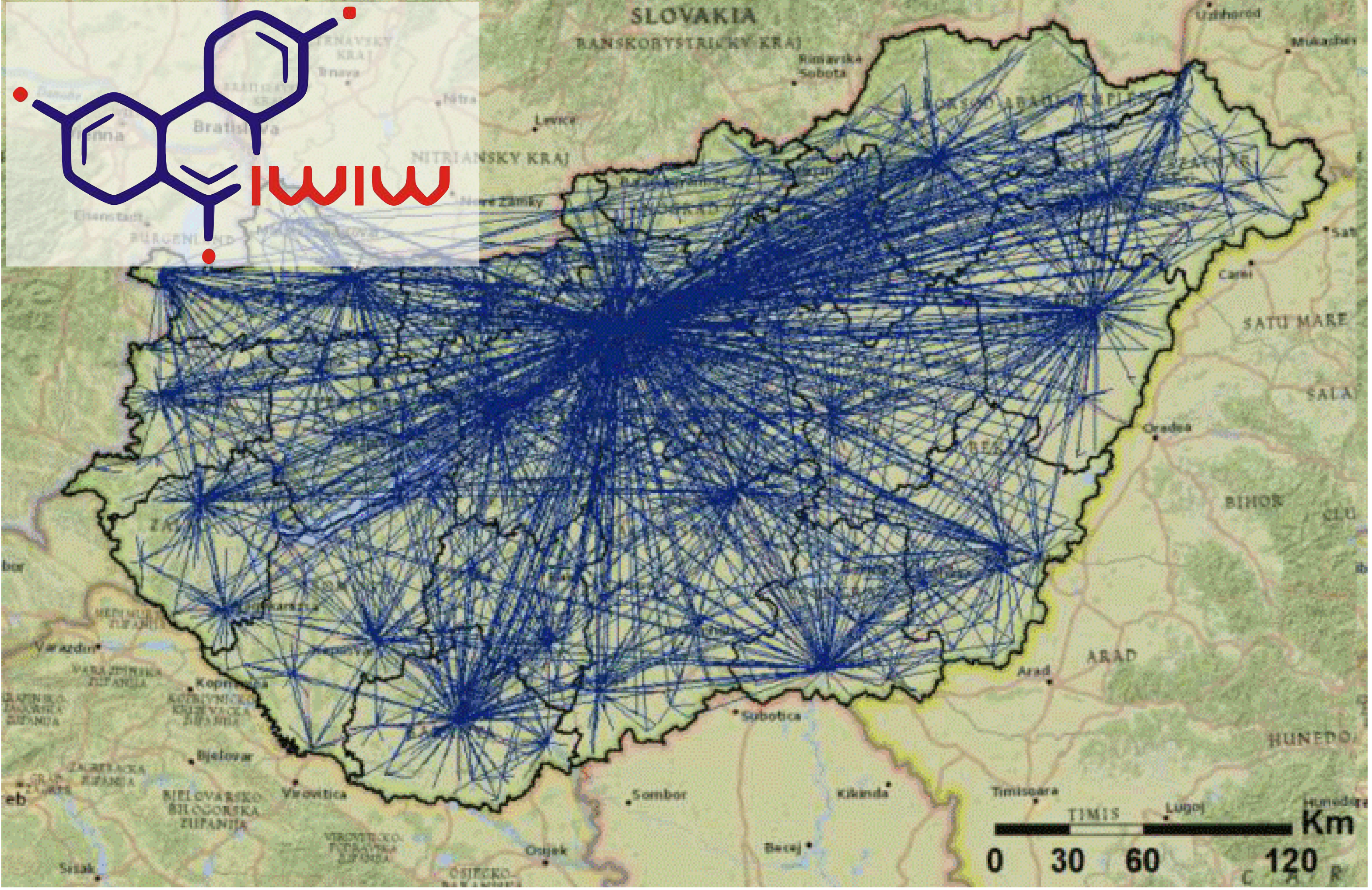
Social capital I: measurement and associations with economic mobility

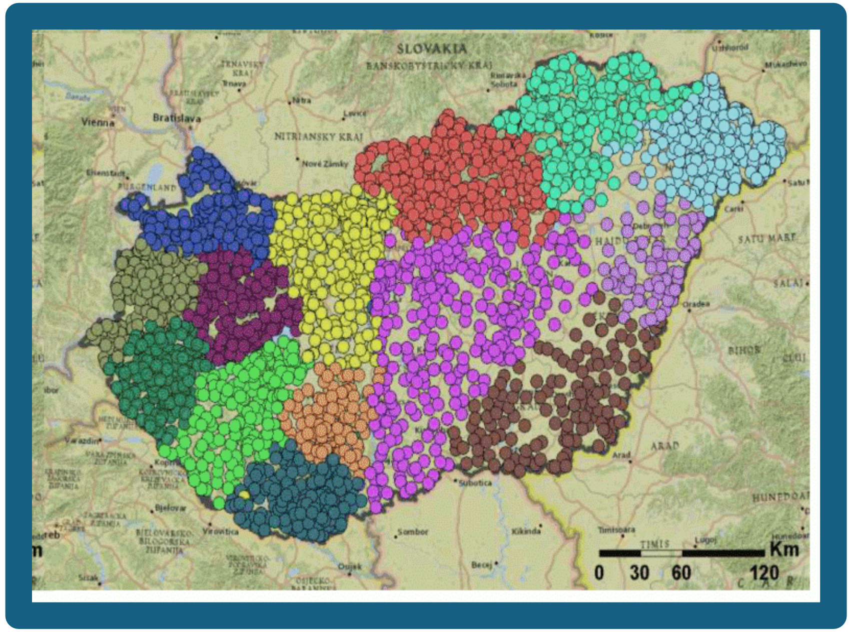
Raj Chetty, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole & Nils Wernerfelt

Nature **608**, 108–121 (2022) | Cite this article

182k Accesses | **151** Citations | **2296** Altmetric | Metrics

Philip S. Chodrow PNAS 2017;114:44:11591-11596





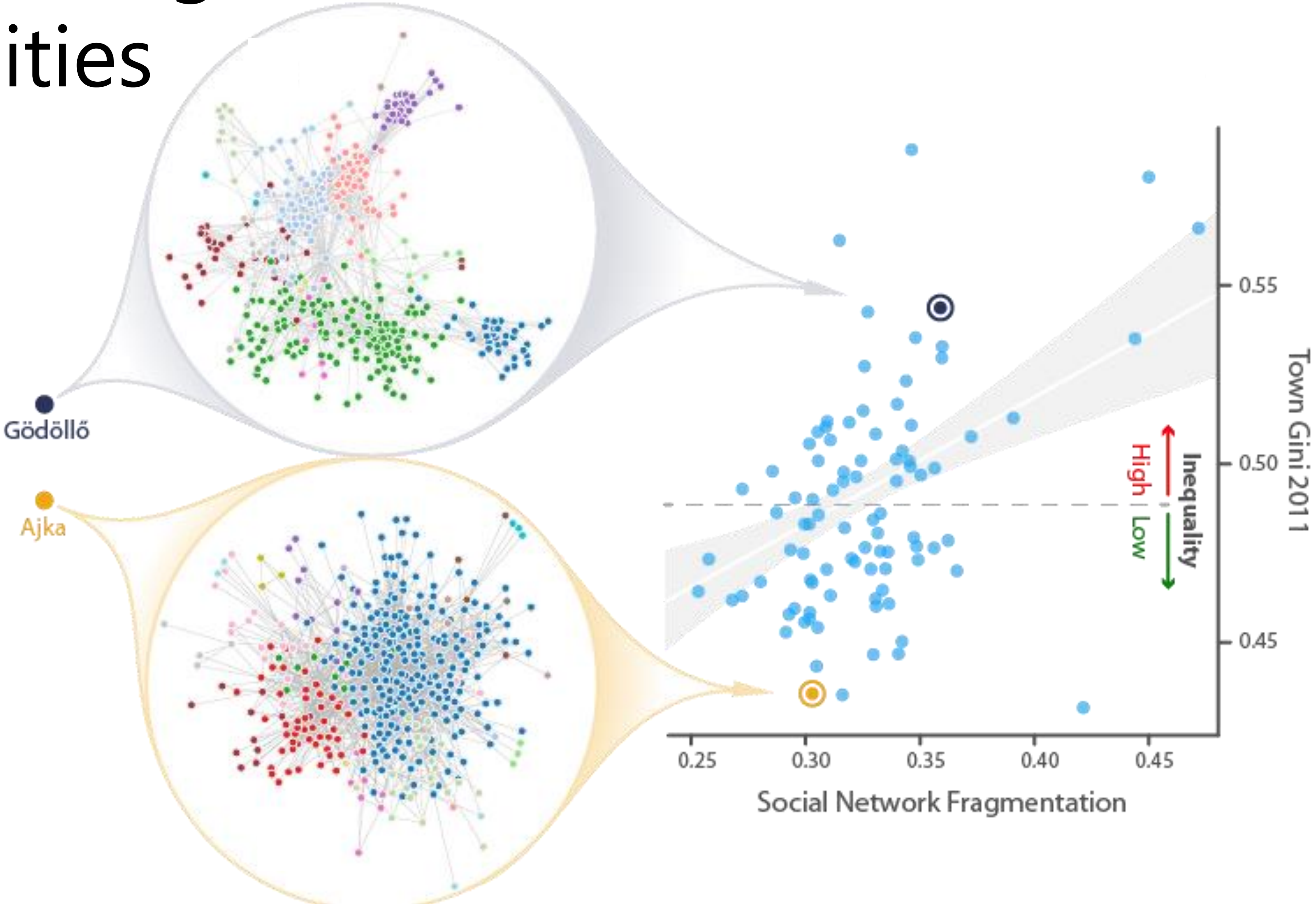
Social media platform

2.8 M Users

- **Location** (self-reported),
- **Date of registration,**
- **Date of last login,**
- **ID of friends,**
- **ID of invitor**

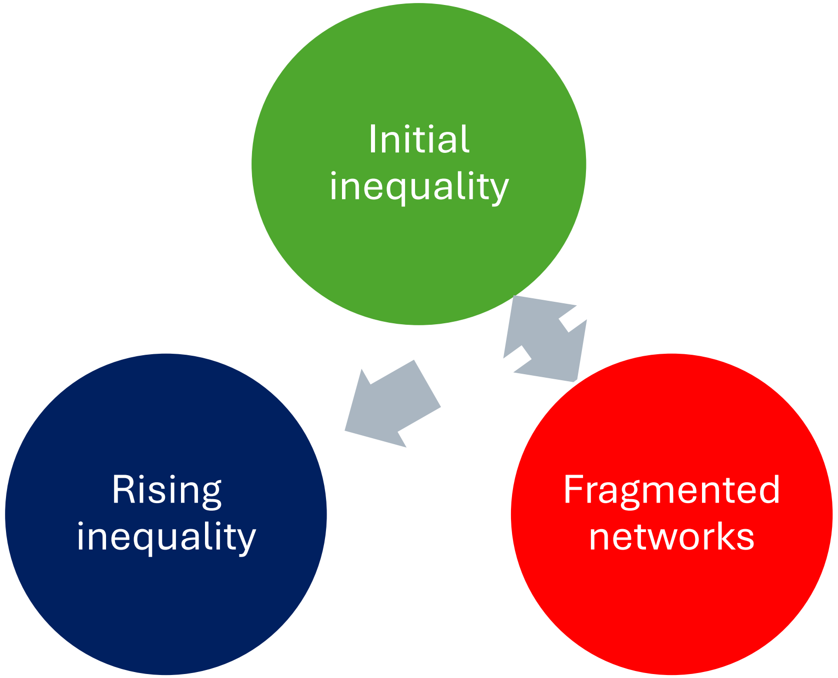
Jakobi and Lengyel (2014) Springer; Lengyel et al. (2015) PLoS ONE

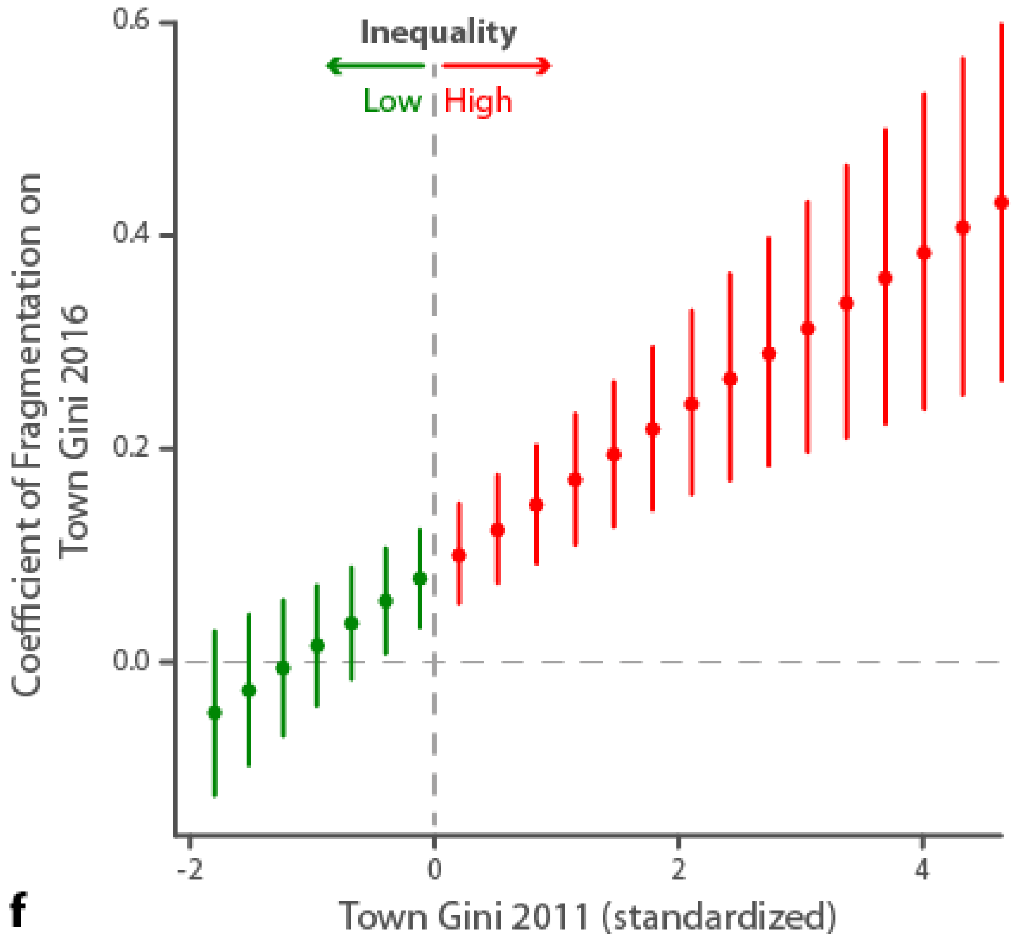
Network fragmentation is correlated with inequalities



Network fragmentation in unequal cities lead to more inequalities

$$G_{i,2016} = \alpha \times G_{i,2011} + \beta \times F_i + \gamma \times (G_{i,2011} \times F_i) + Z_{i,2011} + \epsilon$$



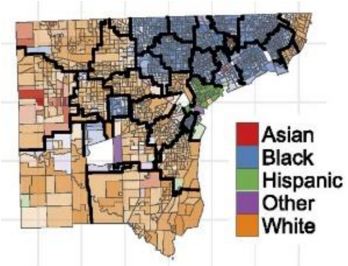


What is the role of urban topology?

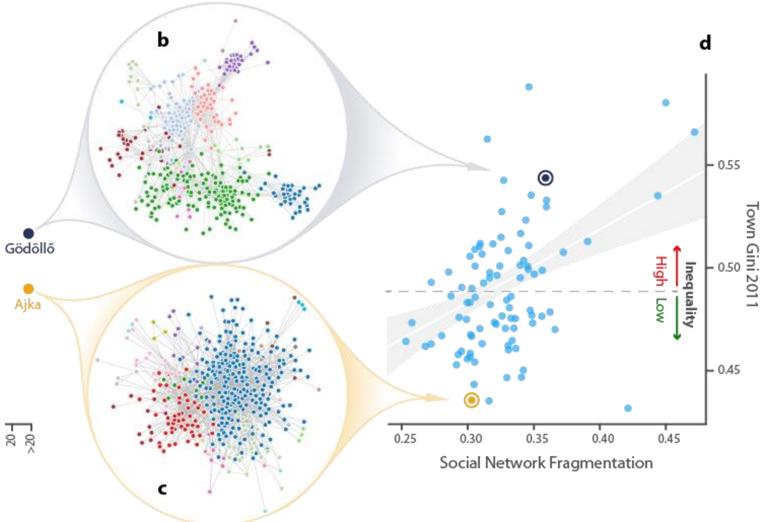
A 2SLS approach

$$F_i = \delta + \gamma IV_i + \delta N_i + e_i$$

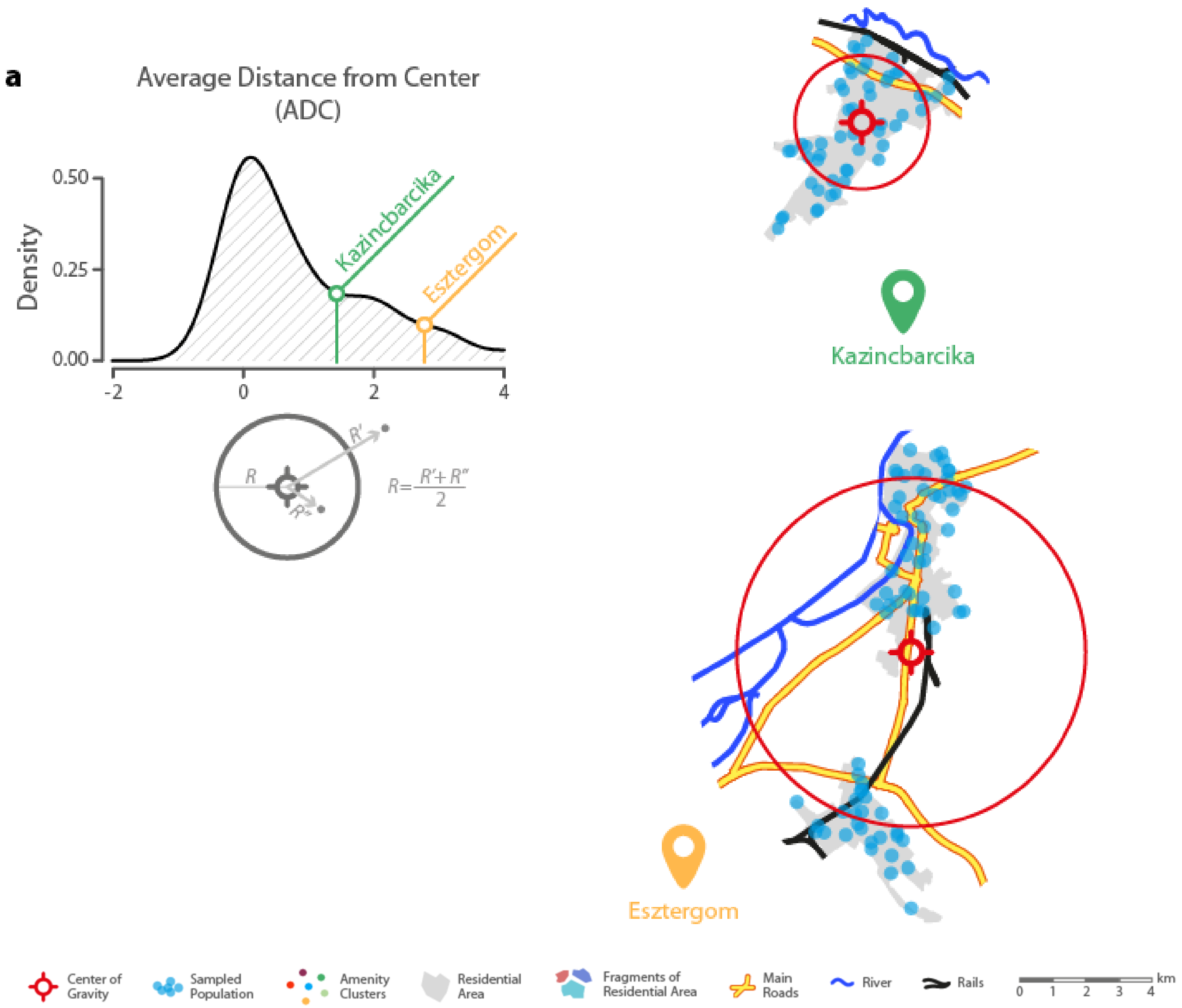
$$G_i = \alpha + \beta_1 \hat{F}_i + \beta_2 X_i + \varphi_k + e_i$$



(a) Detroit, $k = 34$



Urban topology 1: average distance from center



Downtown functions as major hub for social interactions

Urban economics standard models (e.g. von Thünen, Alonso)

$$ADC_i = \frac{\sum_p^P D_{p,c}}{P} / S_i$$

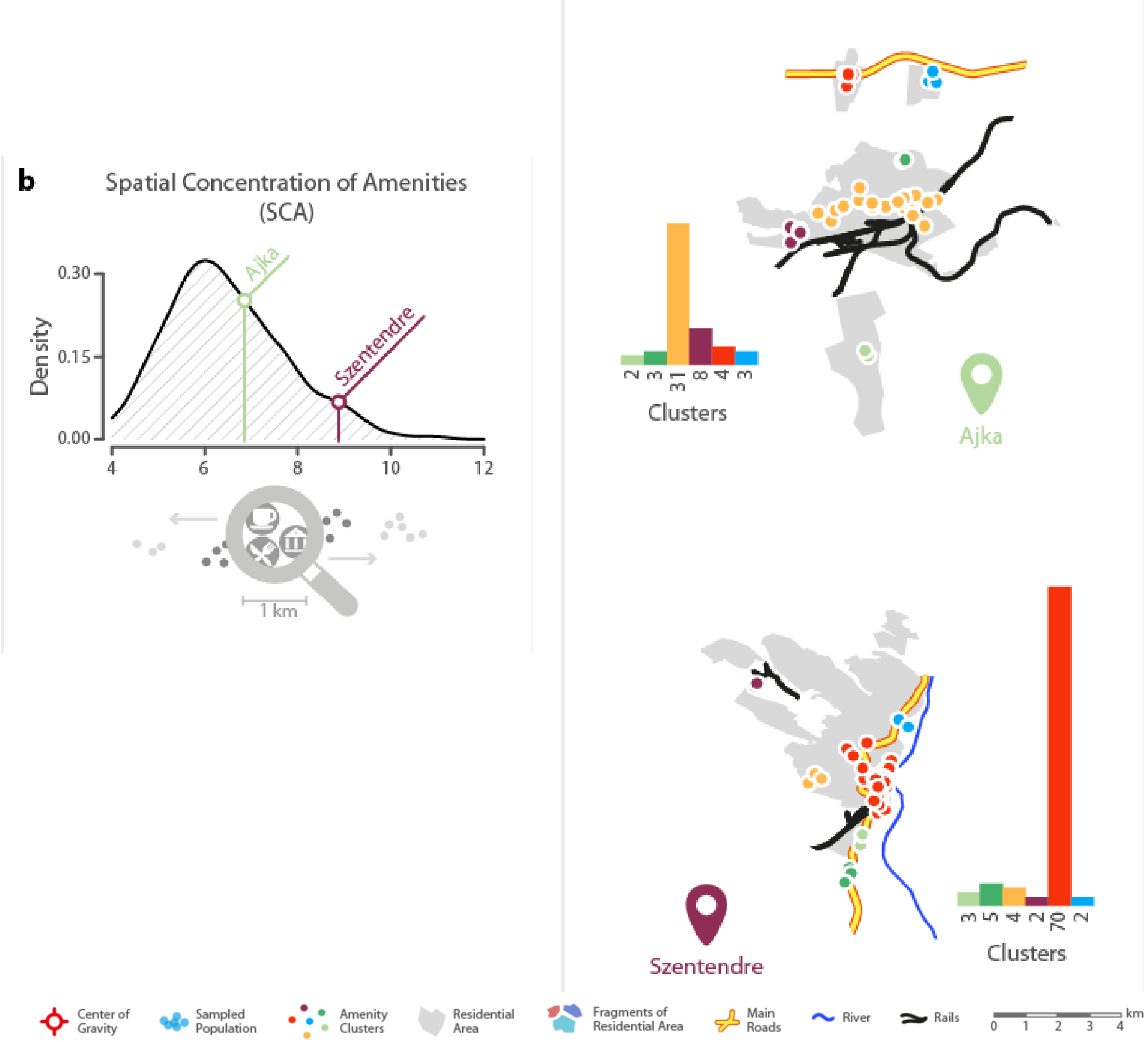
Co-location is important for social interaction and the probability of links decreases as distance grows

-> we expect that fragmentation is higher in towns where ADC is high

Reversed causality is not likely

-> city growth is not likely to be driven by segregated social groups

Urban topology 2: spatial concentration of amenities



H1: spatial concentration of amenities facilitates segregation of the poor living in the periphery
(Brueckner-Thisse-Zenou, 1999, European Economic Review)

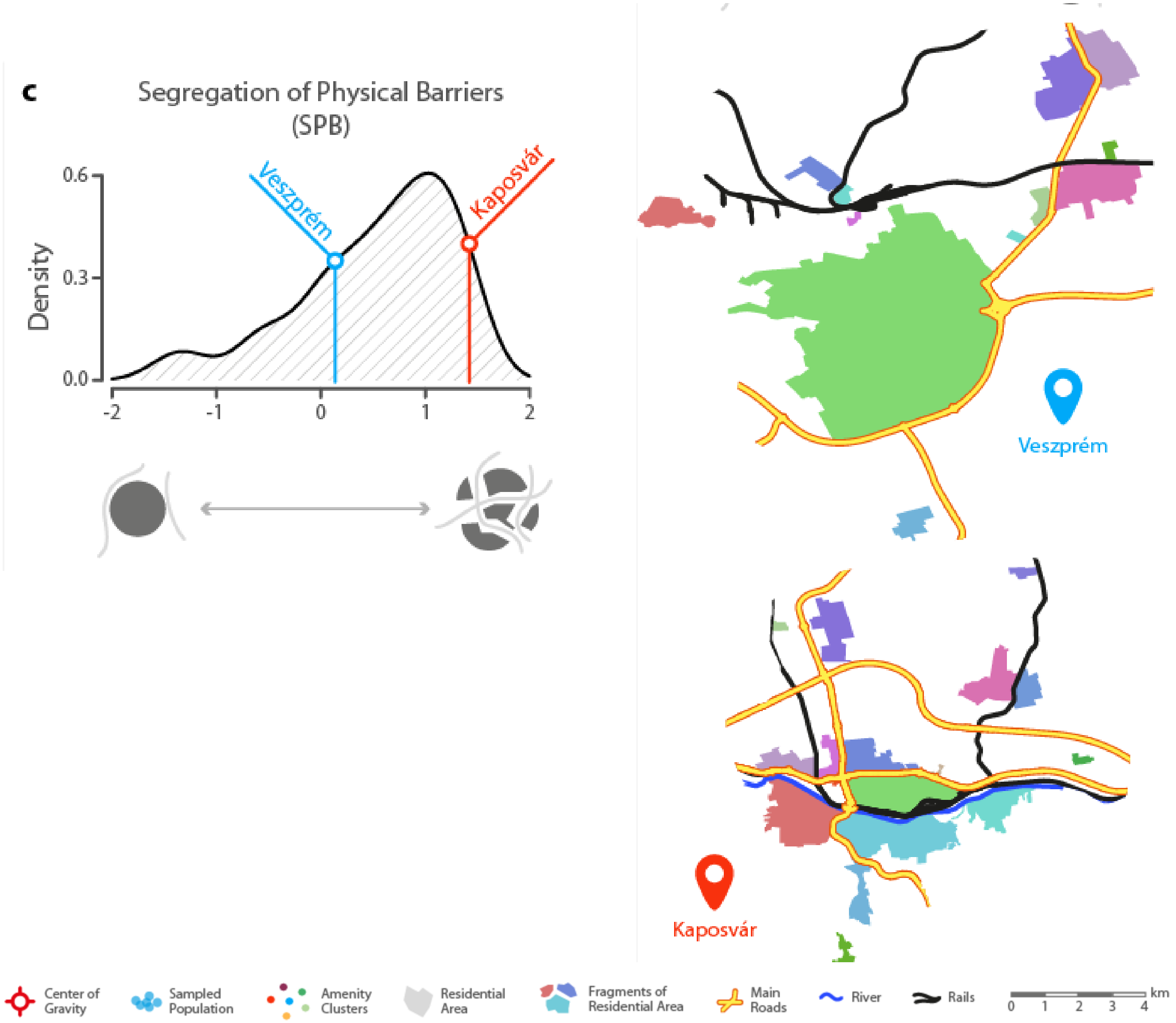
H2: spatial diffuse of amenities provide meeting places in peripheral neighborhoods that facilitates segregation (Zhang-Pryce, 2019, Urban Studies)

$$SCA_i = \frac{\sum_c (p_c \times \log p_c)}{n(c)} / S_i$$

H1 is verified with a positive correlation between the measure and fragmentation.

-> we expect that fragmentation is higher in towns where SCA is high even when controlled for town characteristics.

Urban topology 3: segregation of physical barriers



Built and natural barriers facilitate segregation in cities (Cutler and Glaeser, 1997, QJE)

Used to instrument racial segregation effect on inequality (Ananat, 2011, AEJ: Applied Economics)

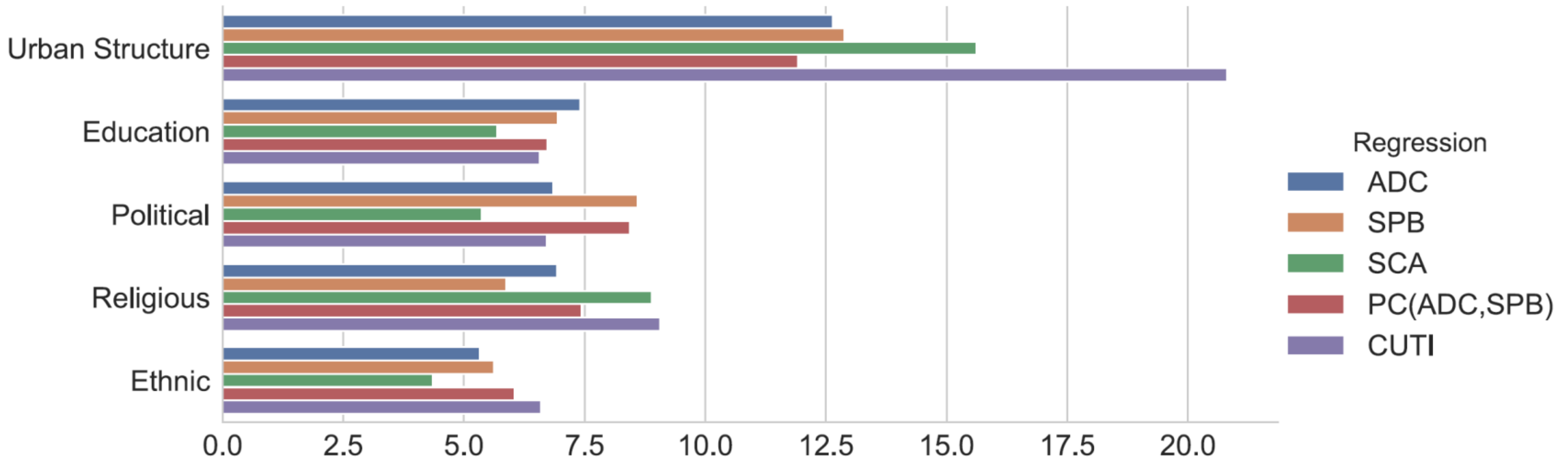
$$SPB_i = 1 - \sum_a (S_a/S_i)^2$$

Barriers decrease the probability of face-to-face interaction (Lengye et al, 2015 PLOS ONE)
Barriers facilitate sorting into neighborhoods (Benton, 2017, Journal of Urban History)
-> we expect that fragmentation is higher in towns where SPB is high

Reversed causality is not likely
-> railroads were not built after the 1940s in Hungary (Körner-Nagy, 2011, Építés-Építészettudomány)

Does urban topology explain network segregation better than other dimensions of social segregation?

Random Forest approach:
- **500 regressions randomly combining the explanatory variables and calculate mean squared error (MSE) from the random sample of these regressions**
- **Randomly vary variables around their mean and recalculate MSE**



1. School inequalities

Coefficient of variance in 6th grade math exam.

2. Political fragmentation

Coefficient of variance in the votes given to Fidesz across voting districts.

3. Religious fragmentation

Entropy of religious distribution.

4. Ethnic fragmentation

Entropy of ethnic distribution.

Does urban topology correlate with network fragmentation? - 1st stage in the 2SLS analysis

$$F_i = \delta + \gamma IV_i + \delta N_i + e_i$$

	Dependent variable: $Fragmentation(F_i)$		
	(1)	(2)	(3)
ADC	0.091**		
	(0.045)		
SCA		0.110**	
		(0.046)	
SPB			0.168***
			(0.044)
User rate	0.367***	0.355***	0.344***
	(0.045)	(0.046)	(0.044)
Constant	−0.000	−0.000	−0.000
	(0.042)	(0.042)	(0.042)
Observations	473	473	473
R^2	0.167	0.170	0.185
Adjusted R^2	0.163	0.166	0.181
Residual Std. Error (df = 470)	0.915	0.913	0.905
F Statistic (df = 2; 470)	47.082***	48.085***	53.259***

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

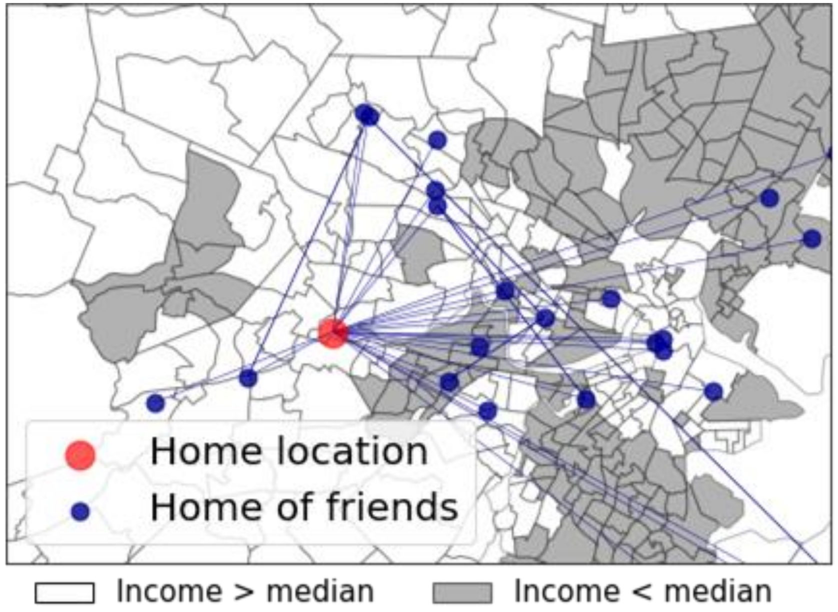
Urban topology <-> network fragmentation <-> inequality 2nd stage in the 2SLS analysis

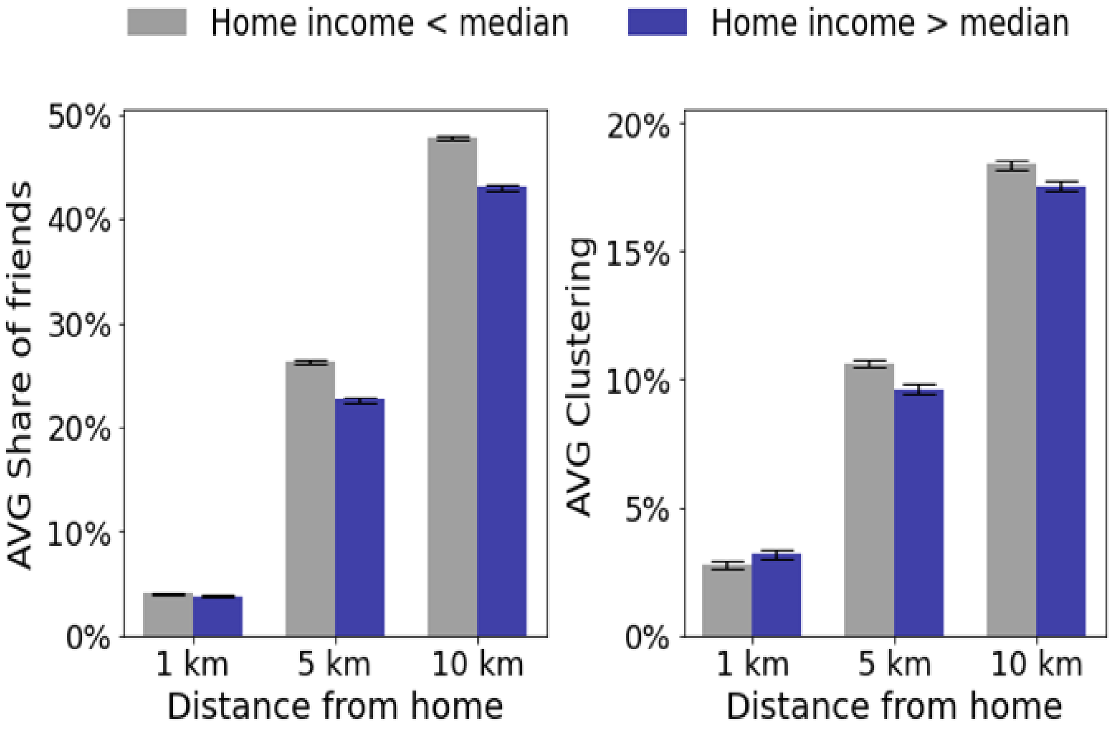
$$G_i = \alpha + \beta_1 \hat{F}_i + \beta_2 X_i + \varphi_k + e_i$$

Control variables:
- Level of FDI
- Change of FDI
- Unemployment rate
- Population density
- Distance to the closest border

	Dependent variable: Gini Coefficient ($G_{i,2016}$)		
	Instrumental Variable		
	ADC	*SCA*	*SPB*
Estimated Fragmentation ($\hat{F}_i$)	0.408***	0.533***	0.288**
	(0.153)	(0.146)	(0.138)
Population density	−0.092*	−0.118**	−0.067
	(0.055)	(0.052)	(0.053)
Distance to border	−0.243***	−0.231***	−0.254***
	(0.059)	(0.062)	(0.058)
Constant	−0.386	−0.392	−0.380
	(0.369)	(0.330)	(0.372)
County FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
First Stage F-test	22.290***	24.009***	26.754***
Wu-Hausman test	1.107	3.729	0.011
Sargan test	0.051	5.349*	1.400
Observations	473	473	473
R^2	0.231	0.192	0.245
Adjusted R^2	0.186	0.145	0.200
Res.St.Err. (df = 446)	0.902	0.924	0.894
Note:	*p<0.1; **p<0.05; ***p<0.01		

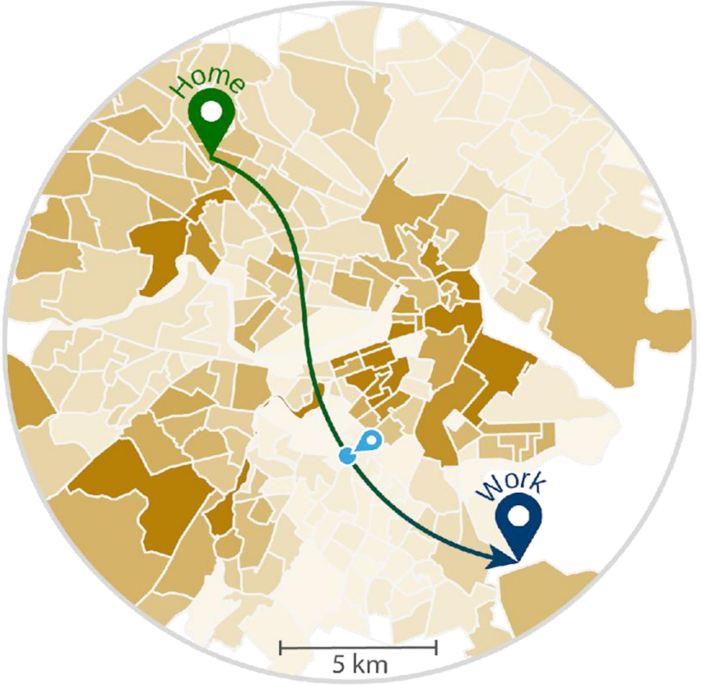
Residents of low income neighbourhoods have spatially concentrated social networks

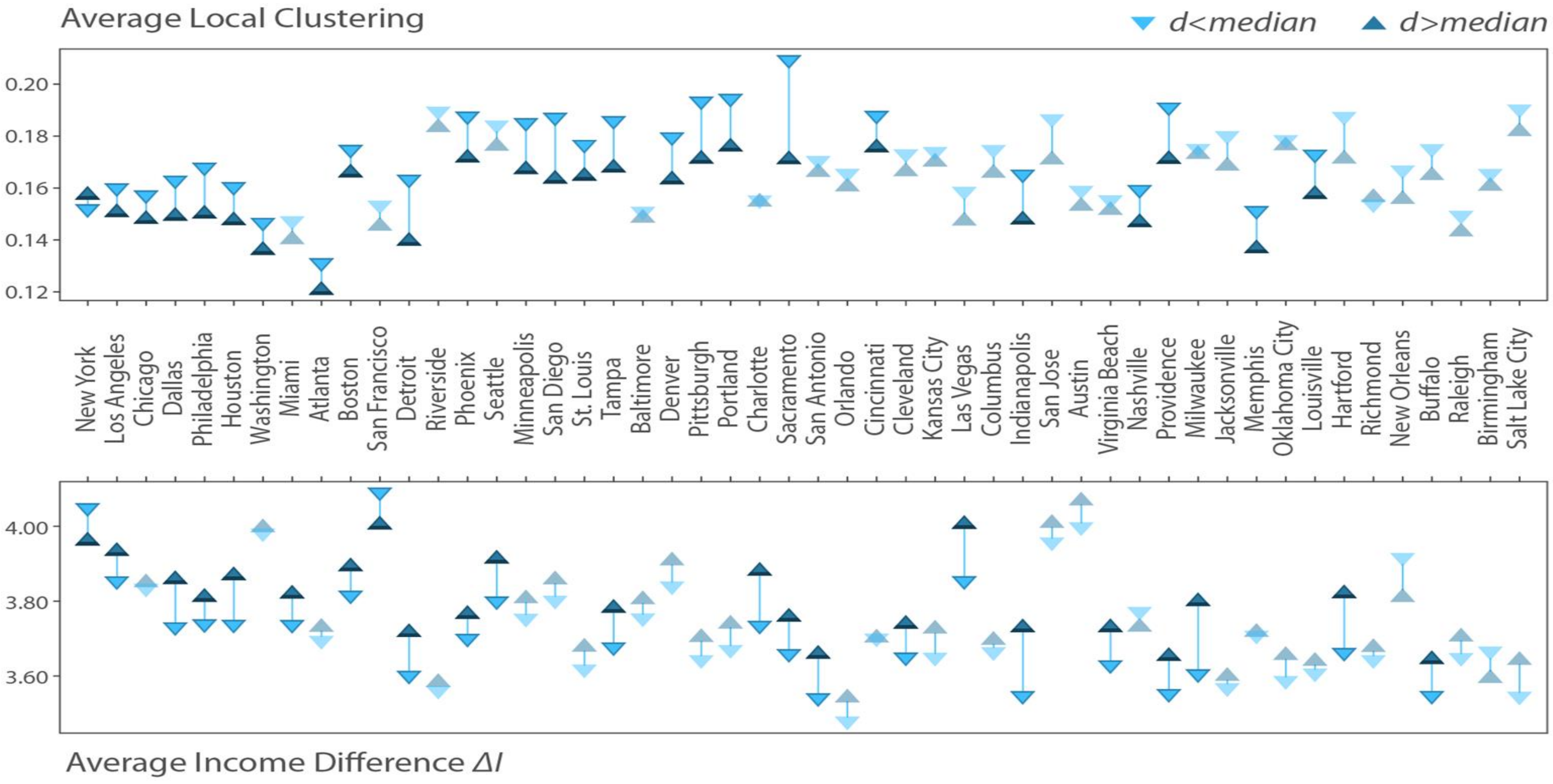




Kovács Á., Juhász S., Bokányi E., Lengyel B.(2023) Environment and Planning B

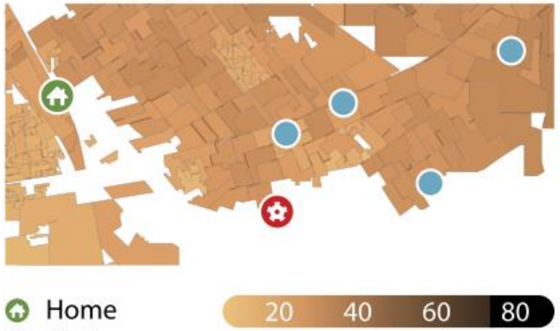
Commuting to work can diversify social networks

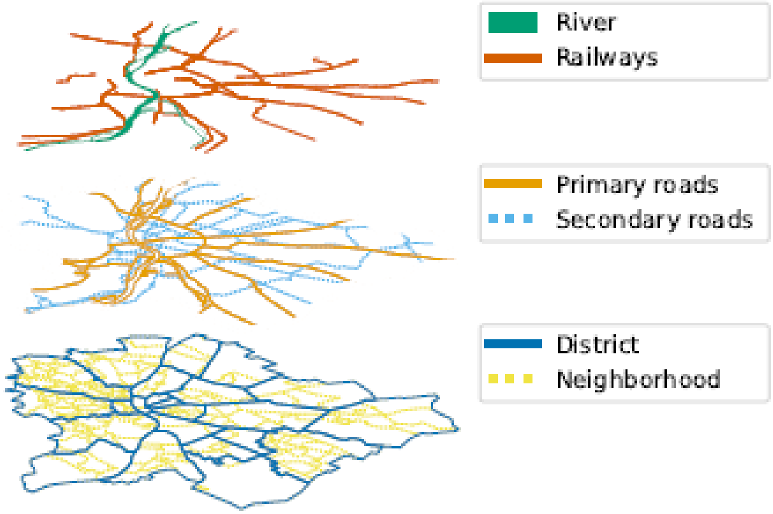


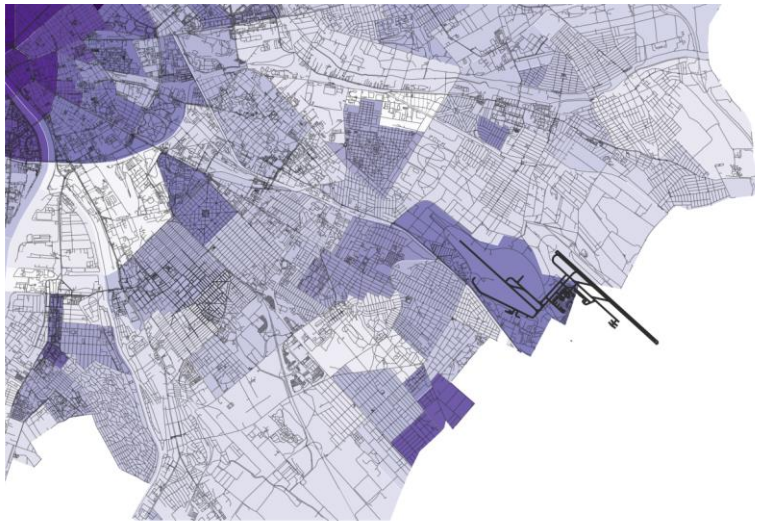


Bokányi E., Juhász S., Karsai M., Lengyel B.(2021) Scientific Reports

Locations with diverse, non-ubiquious services mix socio-economic groups across urban barriers







Juhász S., Pintér G., Kovács Á., Borza E., Mónus G., Lőrincz L., Lengyel B.(2023) EPJ Data Science
Pintér G., Lengyel B. (2024) Barriers of Urban Mobility. arXiv preprint