Network Analysis in Regional Science: New Insights from Old Tools?

Stephan J. Goetz

2019 NARSC Presidential Address

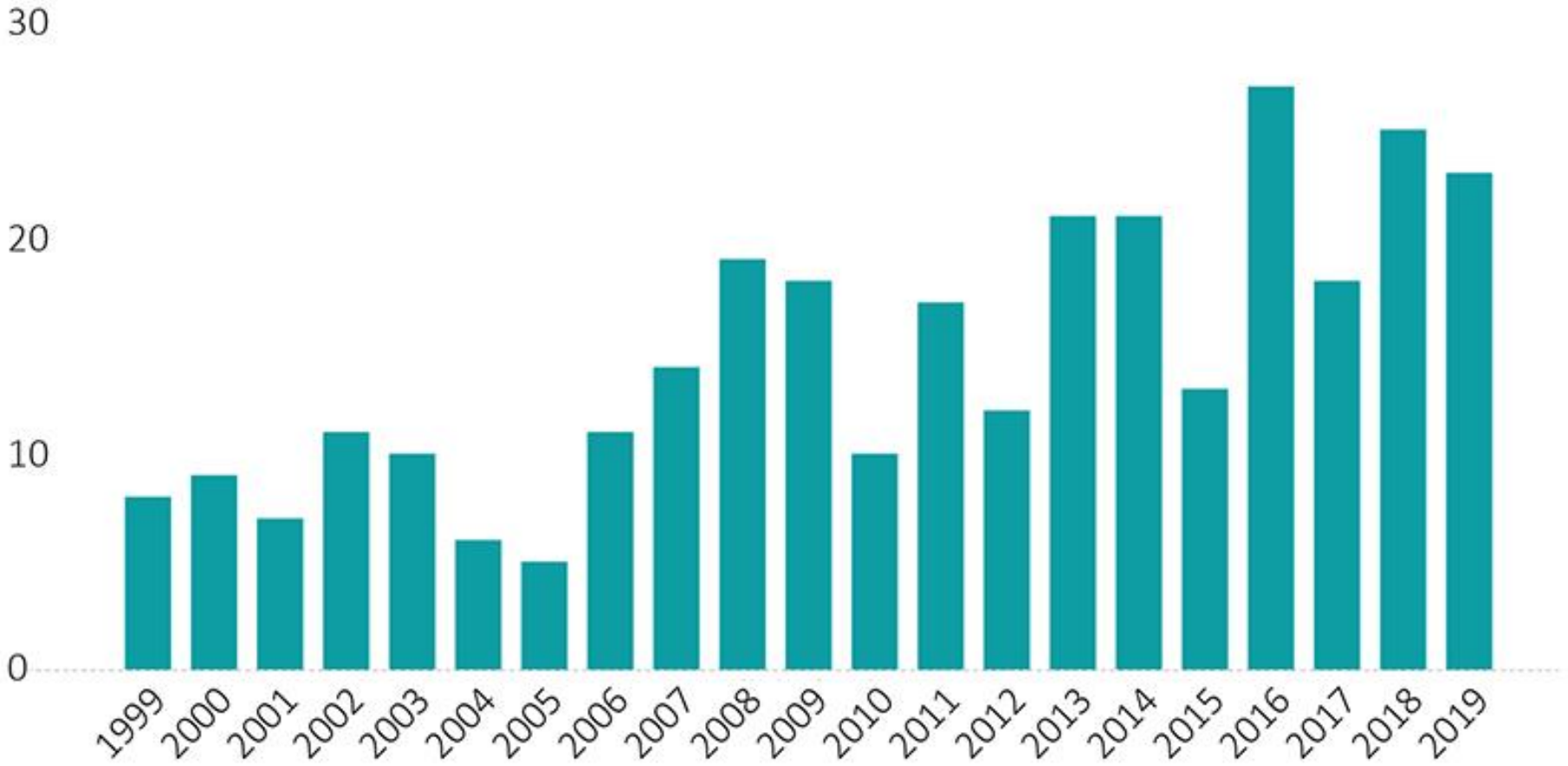
Funding under USDA NIFA Grant no. 2018-51150-28696,
as well as from PSU CAS and NERA, is gratefully acknowledged

PennState
College of Agricultural Sciences

“Networks” in the Regional Science Literature

“Network(s)” in regional science paper titles, by year

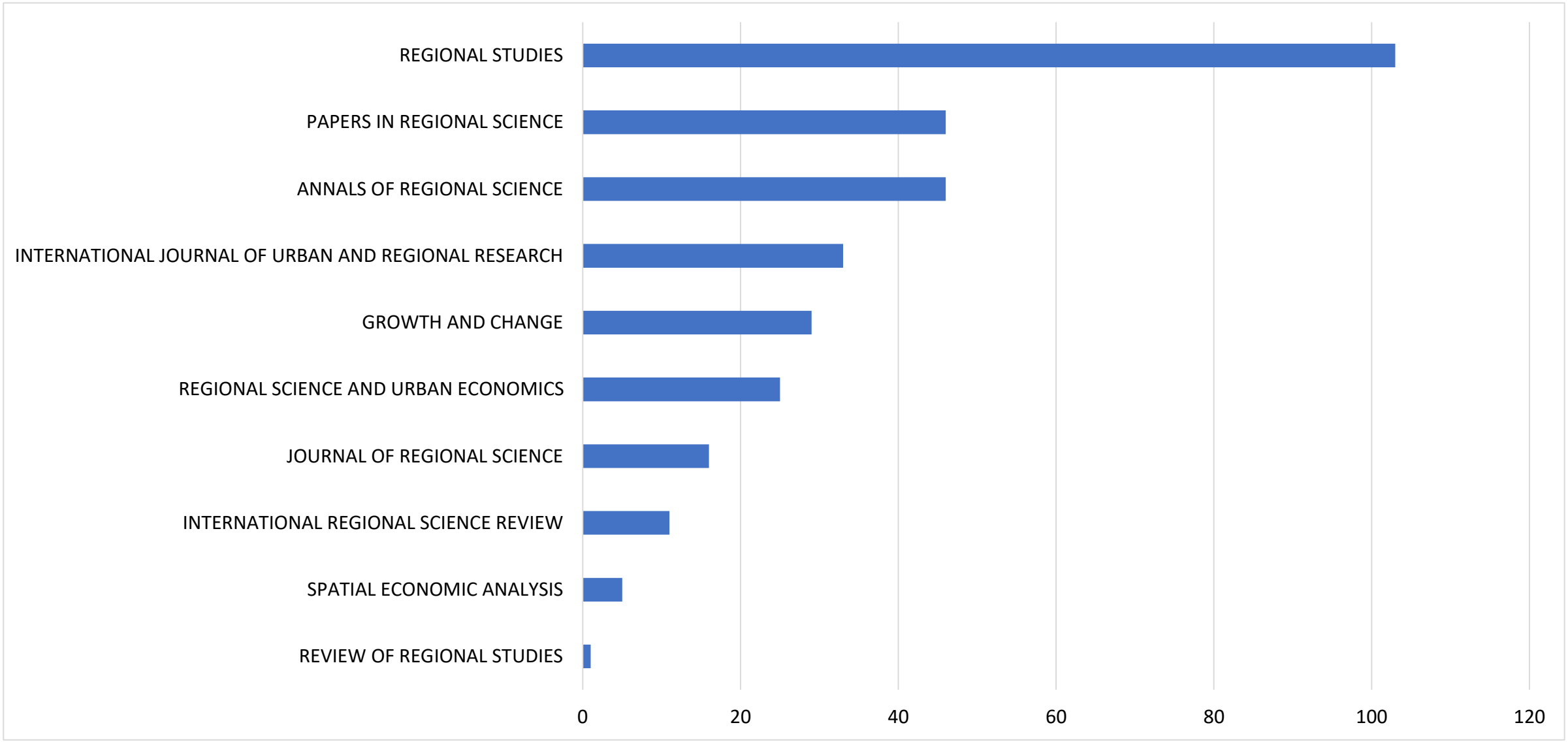
Web of Science Searches



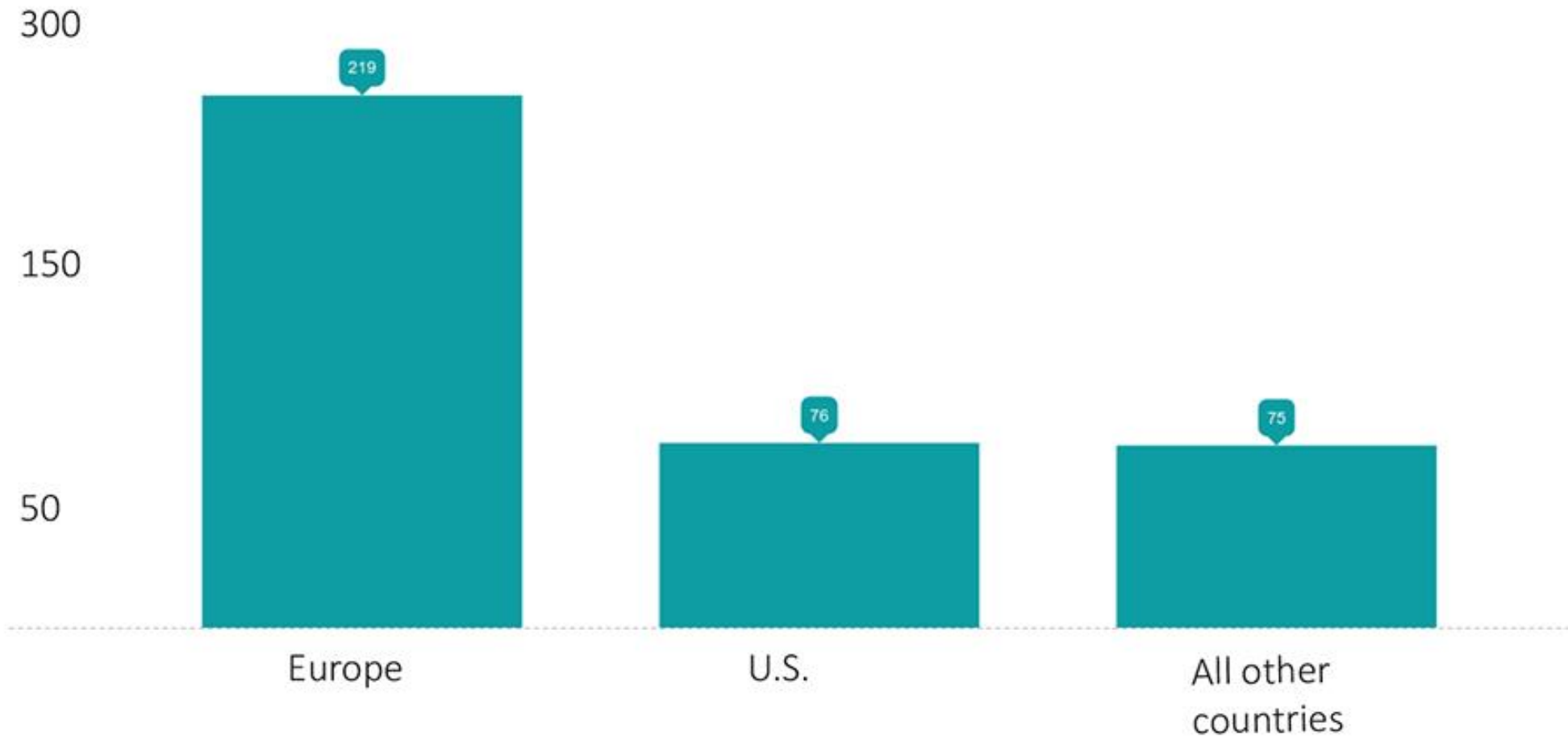
Journal of Regional Science
Regional Science and Urban Economics
Annals of Regional Science
The Review of Regional Studies
Papers in Regional Science
Regional Studies
International Journal of Urban and Regional Research
International Regional Science Review
Growth and Change
Review of Spatial Economics

19 Papers at this conference

No. of Appearances by Journal (not normalized), 1999-present



“Network(s)” in Regional Science Paper Titles, by Country/Region



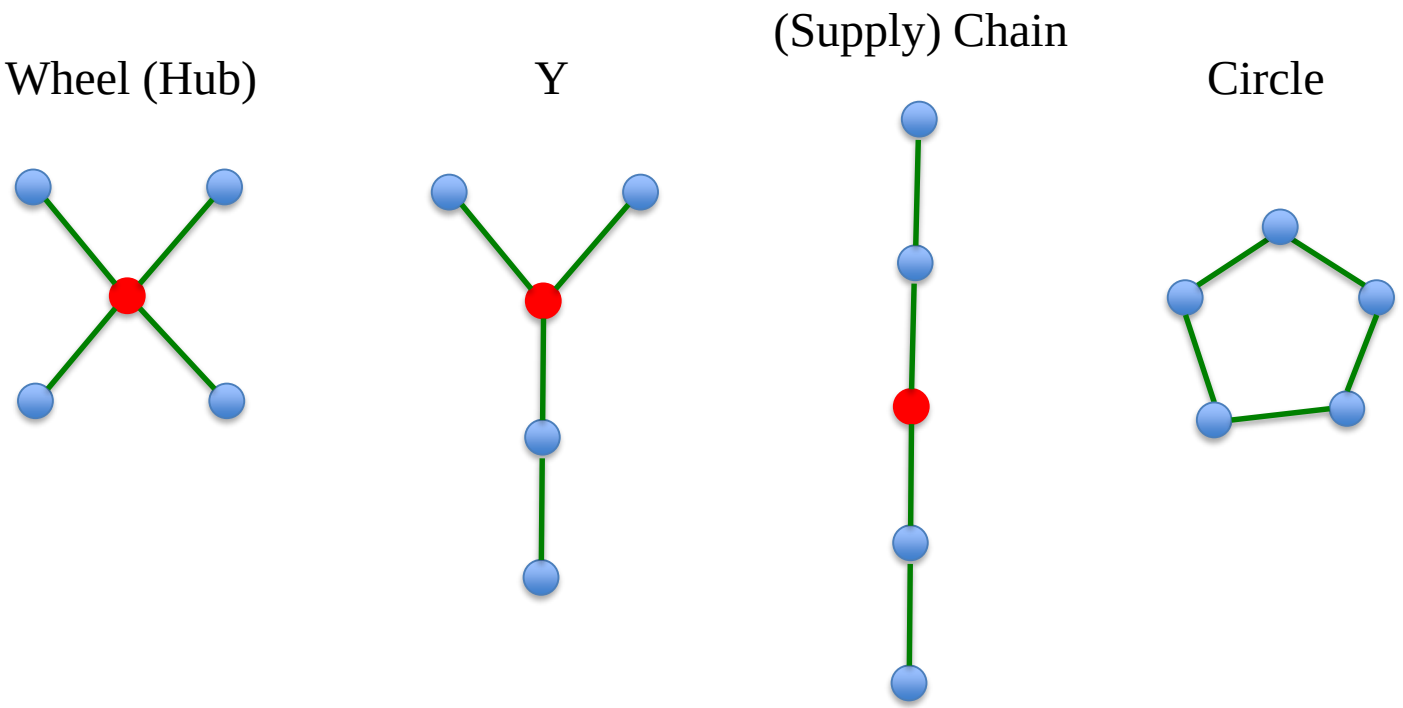
Network Basics

Network Basics

- Networks are about:
  - Nodes (individuals, entities…)
  - Links (edges, connections)
- *The way in which nodes connect makes all the difference*

- Topics/applications
  - Strength of weak ties
  - Random vs. non-random networks
  - Preferential attachment (fit get fitter)
  - Emergence of scaling

Basic Network Structures

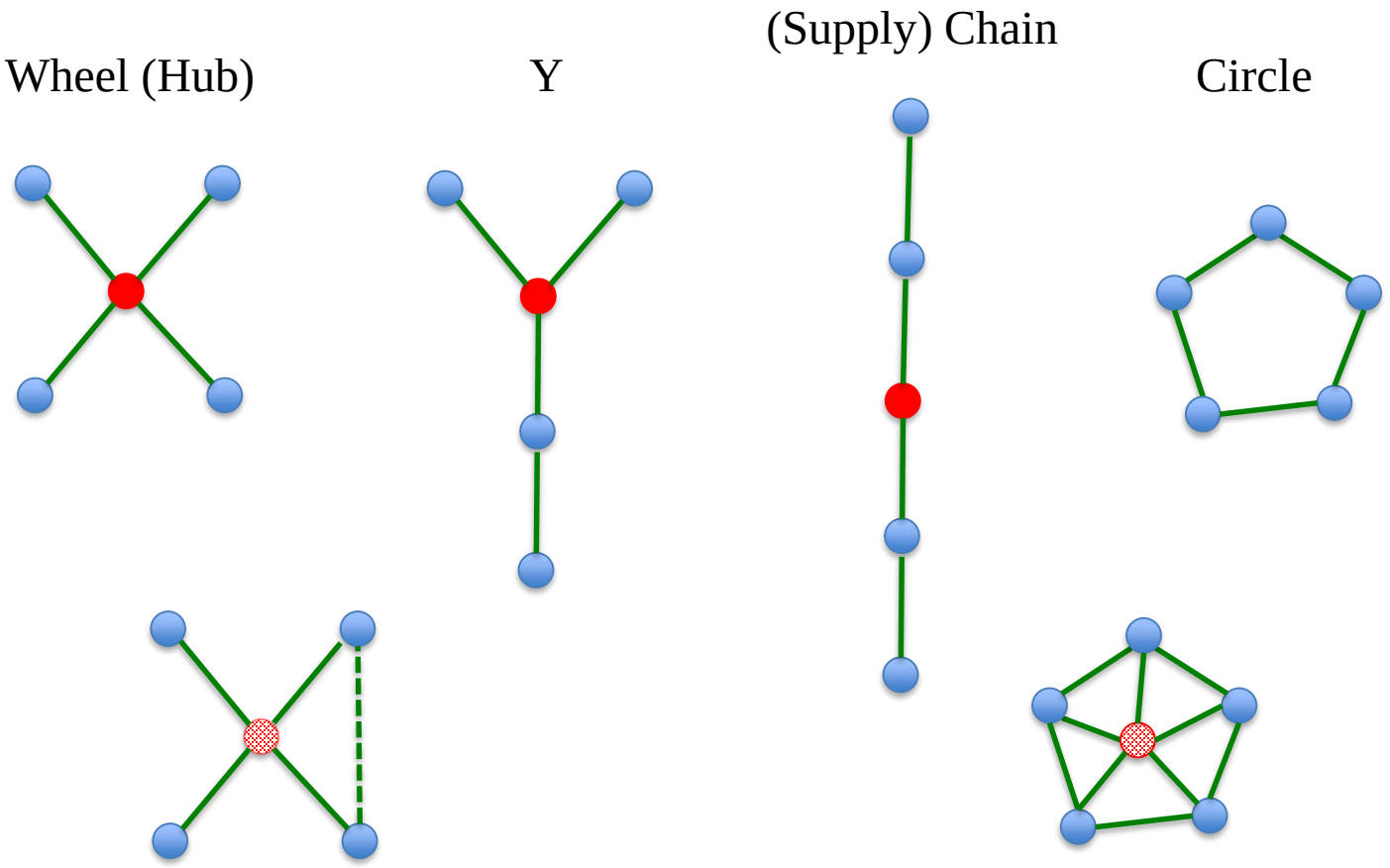


More centralized

More decentralized

Source: Adapted from Borgatti, S.P., et al. (2009), "Network Analysis in the Social Sciences," *Science* 13 Feb. Vol. 323 no. 5916 pp. 892-895

Basic Network Structures

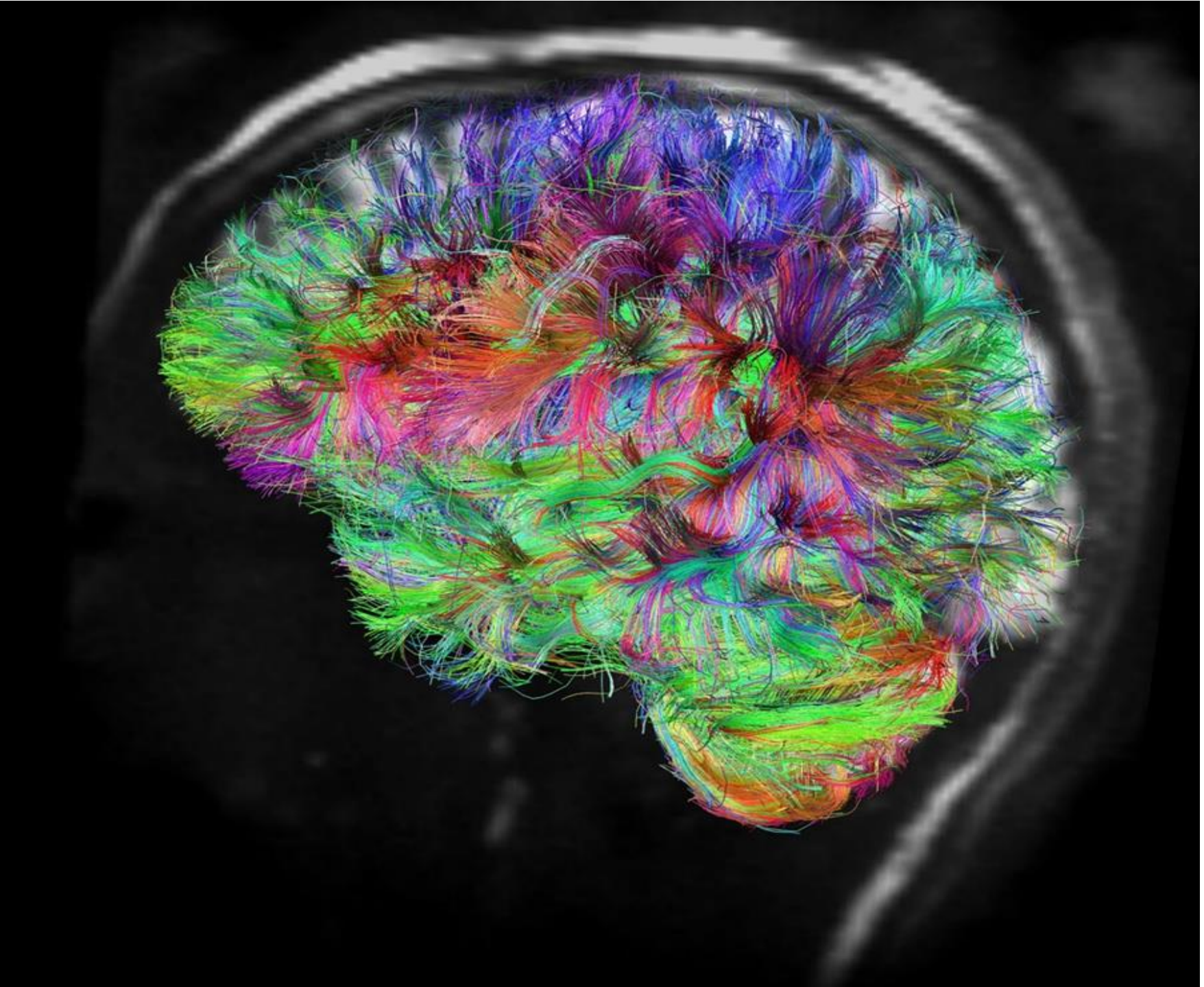


Source: Adapted from Borgatti, S.P., et al. (2009), "Network Analysis in the Social Sciences," *Science* 13 Feb. Vol. 323 no. 5916 pp. 892-895

Why network analysis?

- Additional insights into relationships, causal factors
- Node's position within network adds another potential *dimension* to characteristics explaining outcomes
  - $Y = a + bX$ vs.
  - $Y = a + bX + c\Omega$
  - Ω = the node's position (e.g., centrality) in the network
- *Examples*
  - Mincer-type earnings equation: individual's position in firm's network
  - County economic growth model: county's links to other counties
  - Industry, firm resilience: how central to the local economy

- Networks can be "stacked" (overlayed)

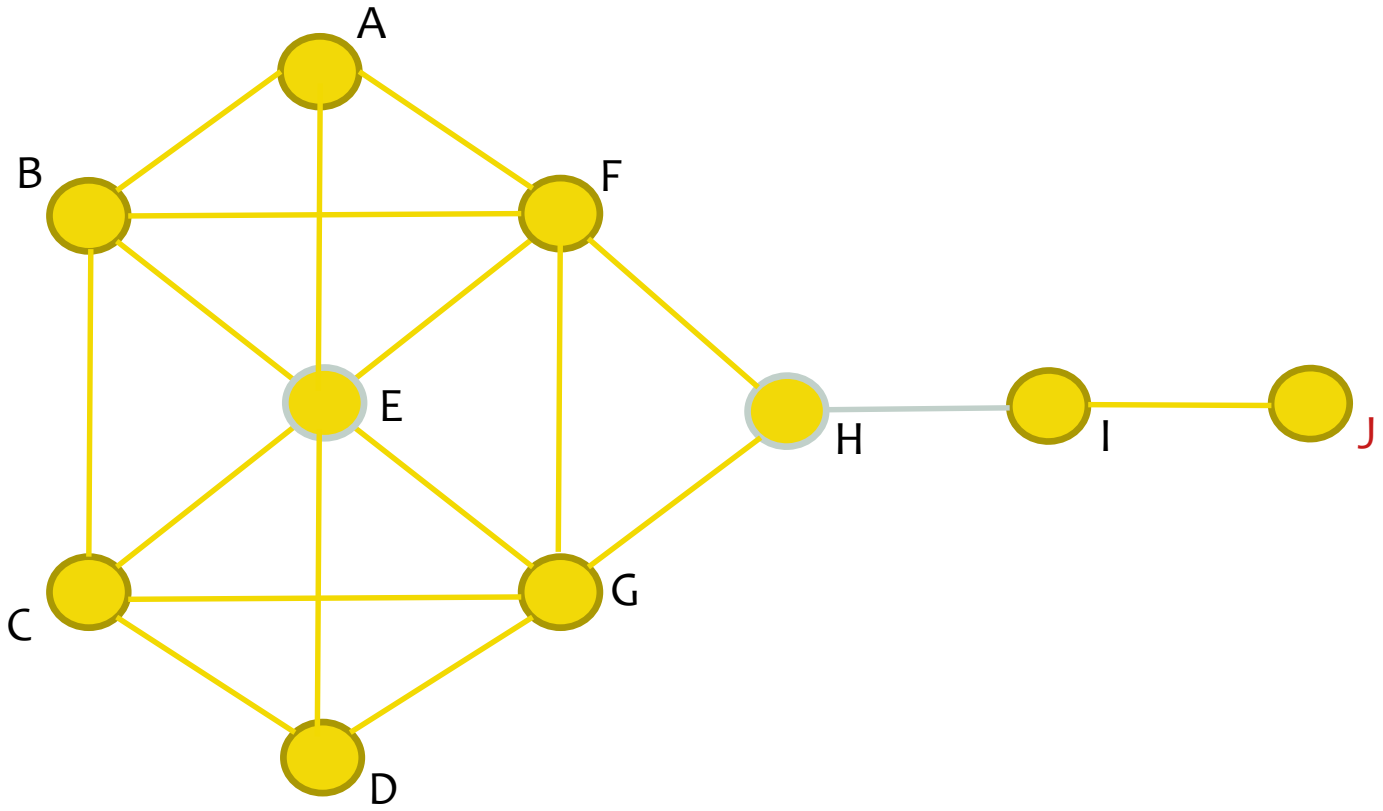


~88,000,000,000 neurons

Each nerve cell may have up to ten thousand connections (***synapse):*** The brain has 100,000,000,000,000 synapses

Source: Dean Drosnes, MD, FASAM; Medical Director, Pennsylvania, Caron Treatment Centers: “America’s Opioid Crisis: Where are we in 2019?”

A Generic Network



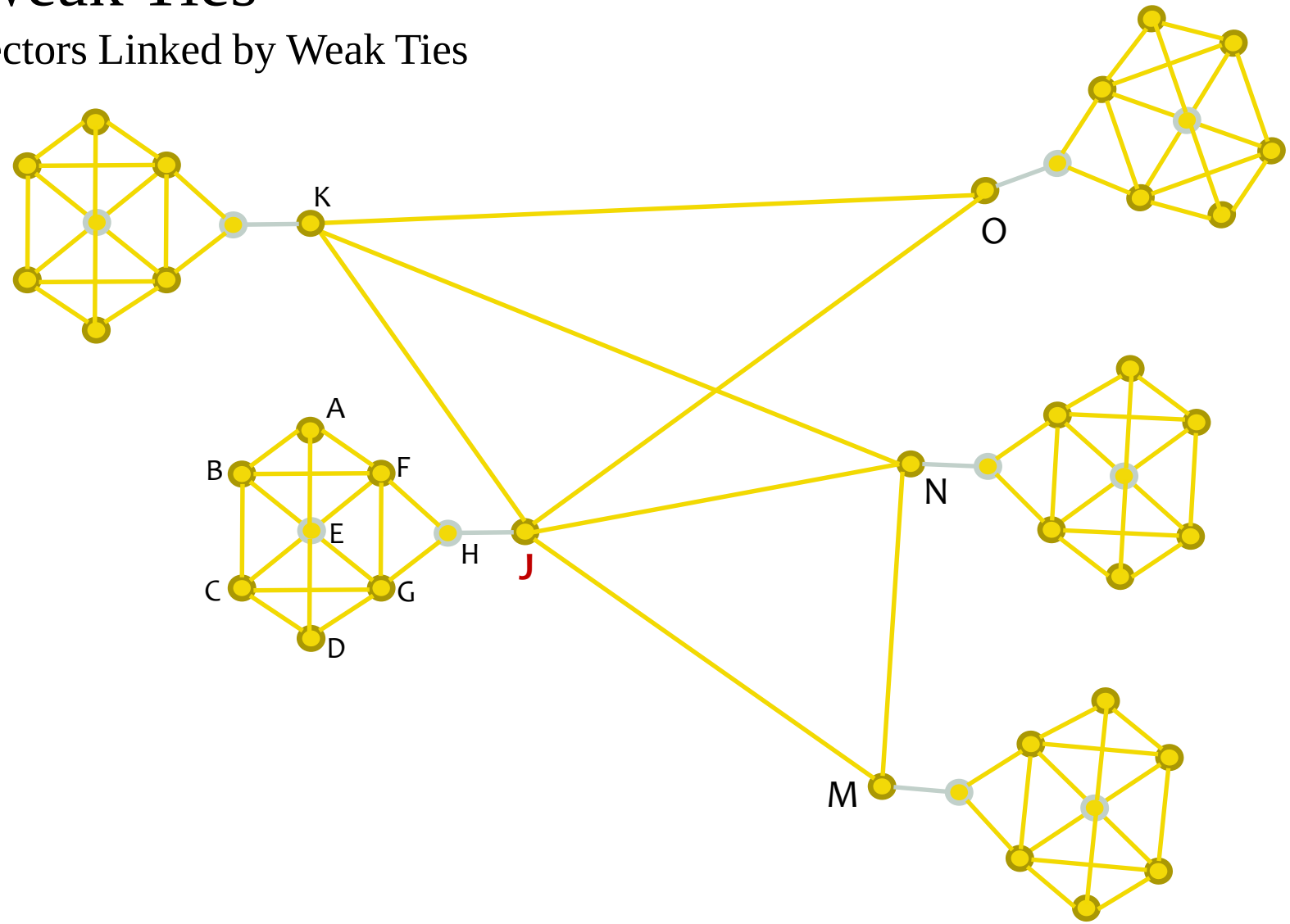
"Centrality" Measures

- Overall connections
  - Highest: E, then F, G
  - Lowest: J
- Betweeness
  - Highest: H, then F, G
  - Lowest: A, D, J
- Closeness
  - Highest: F, G, then E, H
  - Lowest: J

Nodes, individuals, or firms, have different roles in the network, depending on their positions: J may not seem important, but…

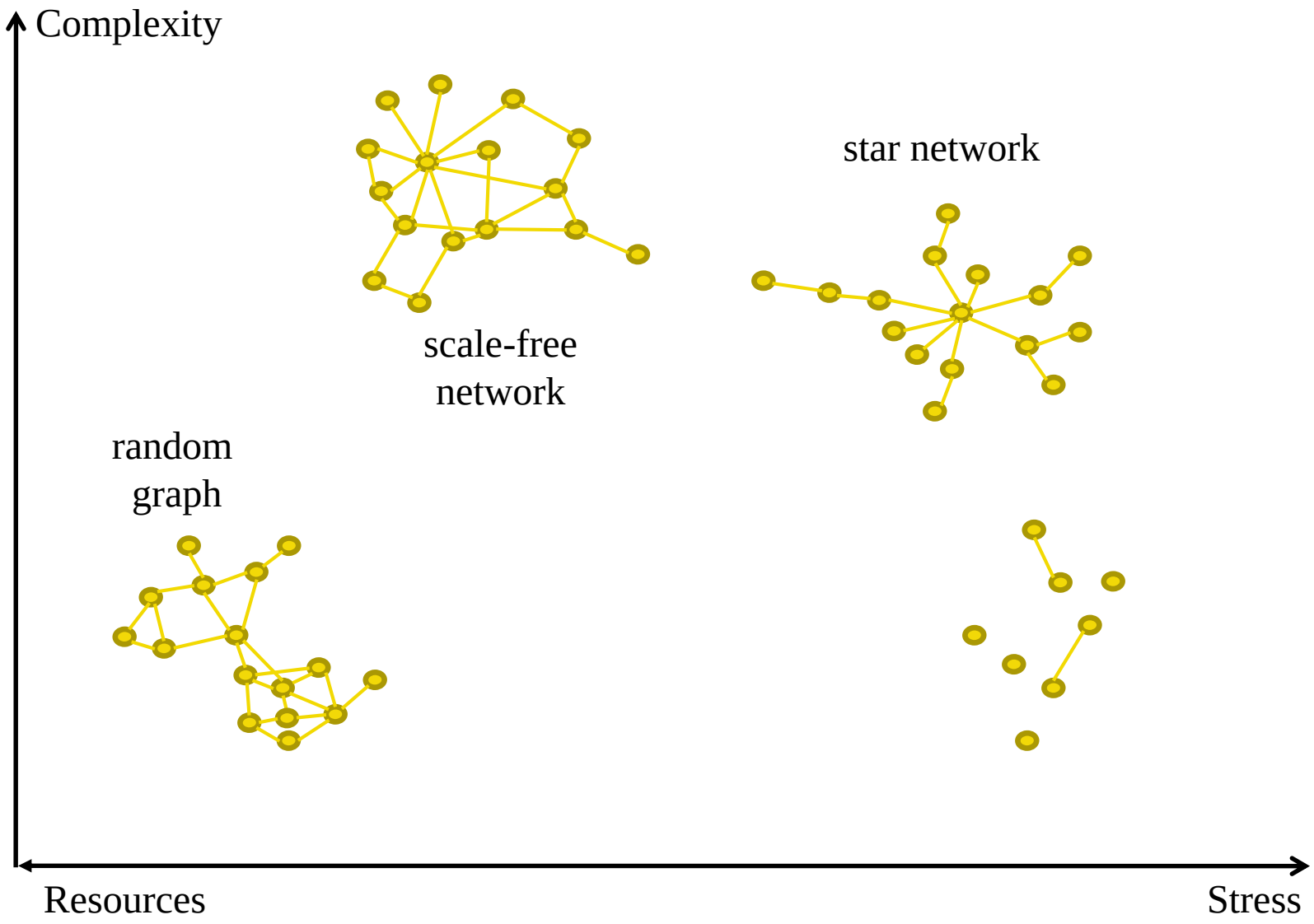
Strength of Weak Ties

A Network of Connectors Linked by Weak Ties



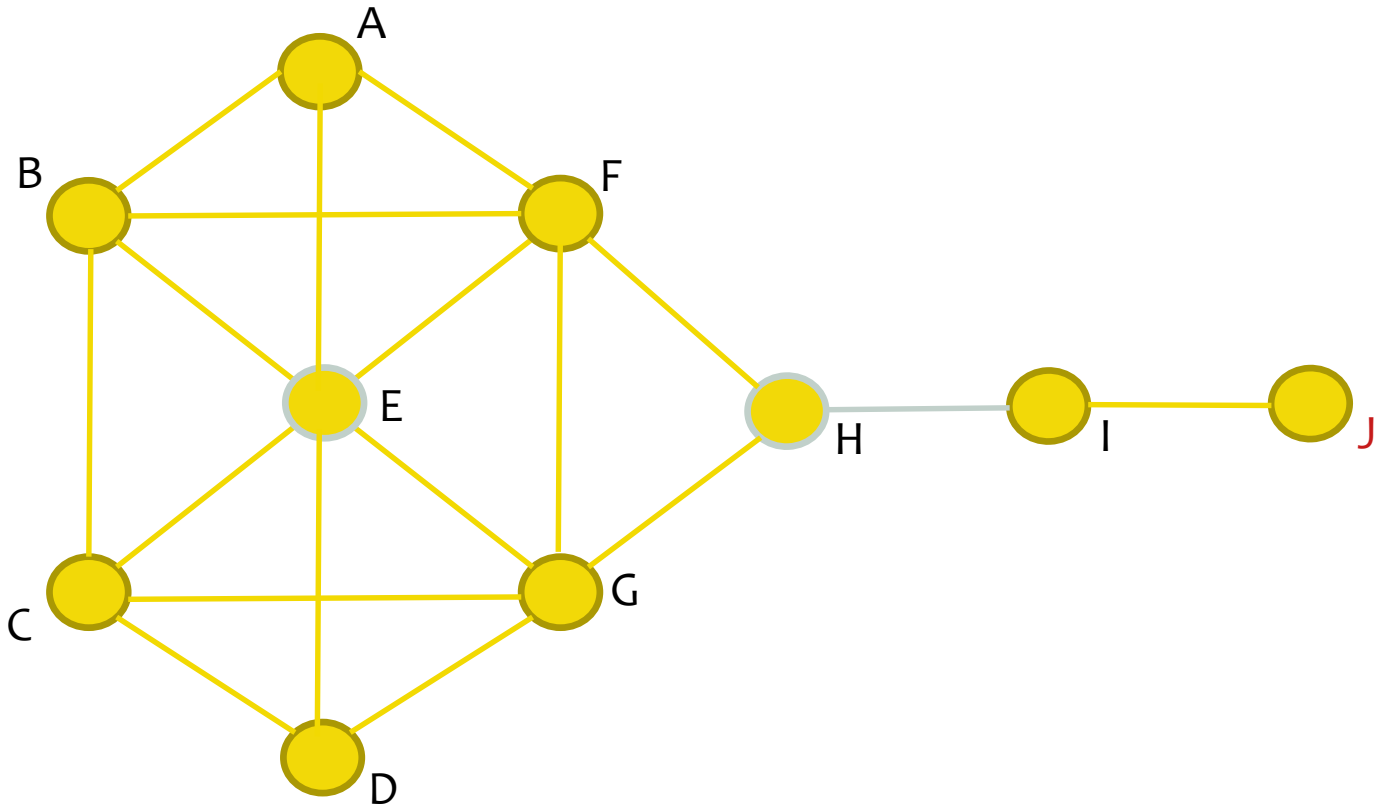
R. Ogle (2007) *Smart World – Breakthrough Creativity and the New Science of Ideas*, Harvard Business School Press, 303*pp.*

How Networks Grow or Fail: Topological Phase Transitions



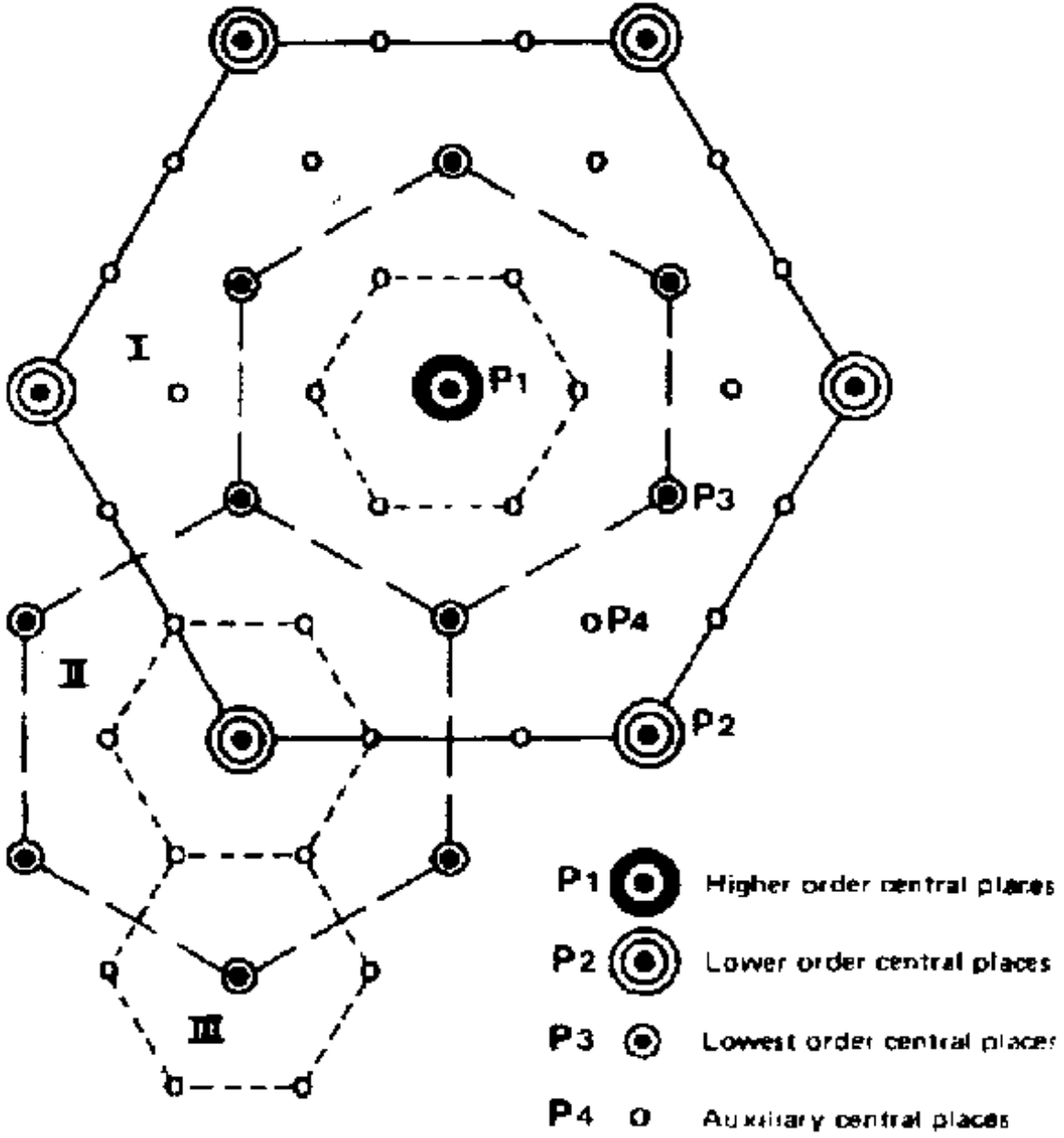
Source: After Csermly, 2009, *Weak Links: The Universal Key to the Stability of Networks and Complex Systems*, Springer

A Generic Network

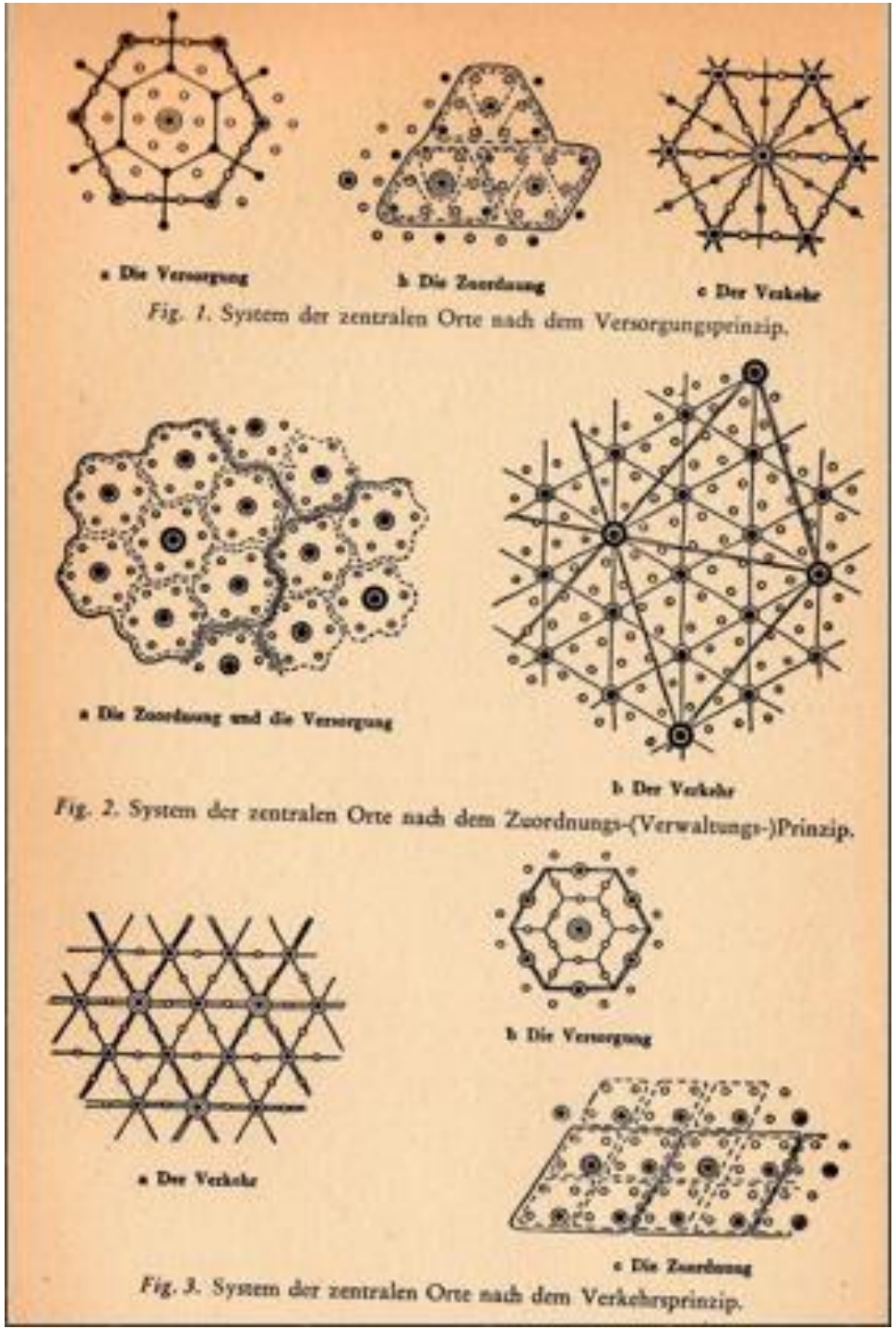


- Individuals
  - Friendship (Facebook)
  - Entrepreneurs (Twitter)
  - Authors
  - Patent holders
  - Artists (co-exhibition)
- Industries
  - Firms
  - Establishments
- Cities, counties
  - Migration
  - Commuting
  - Commodity flows (x, m)
  - Communications (info)
  - Gov websites

Words, planets, flavors, genes, drug interactions…



https://www.pinterest.com/pin/154107618479598158/?nic=1
https://www.pinterest.com/pin/544091198721627749/?nic=1



Input-Output Tables as Networks

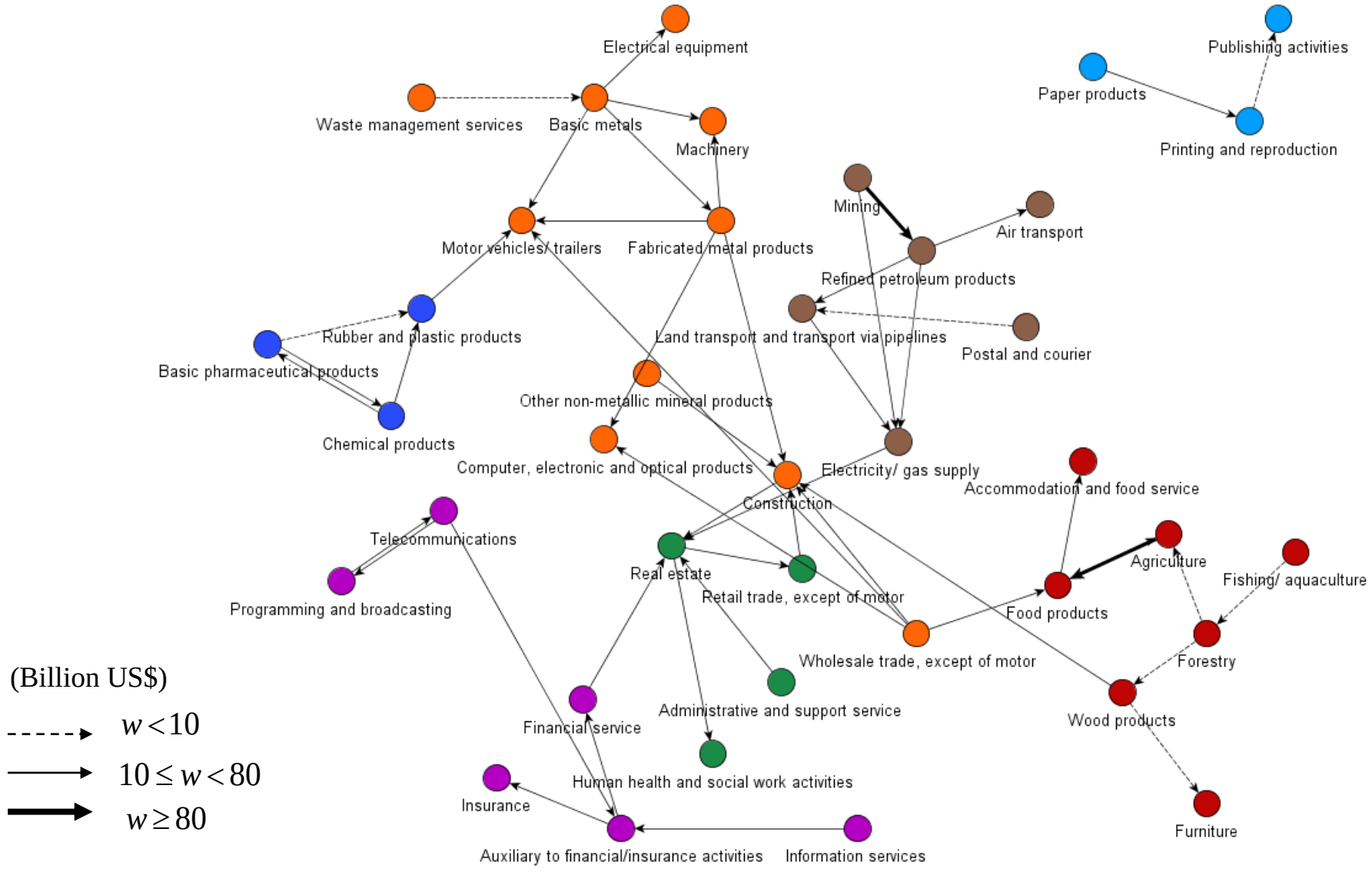
Francois Quesnay, Tableau économique; W. Leontief, Input-output table

Input-Output Tables as Networks

		Industry						Final Demand	Total output
		1	*2*	*3*	*4*	⋯	*t*		
Industry	*1*	x_{11}	x_{12}	x_{13}	x_{14}	⋯	x_{1t}	f_1	X_1
	2	x_{21}	x_{22}	x_{23}	x_{24}	⋯	x_{2t}	f_2	X_2
	3	x_{31}	x_{32}	x_{33}	x_{34}	⋯	x_{3t}	f_3	X_3
	4	x_{41}	x_{42}	x_{43}	x_{44}	⋯	x_{4t}	f_4	X_4
	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
	s	x_{s1}	x_{s2}	x_{s3}	x_{s4}	⋯	x_{st}	f_s	X_s
Value added		v_1	v_2	v_3	v_4	⋯	v_t		
Total output		X_1	X_2	X_3	X_4	⋯	X_t		

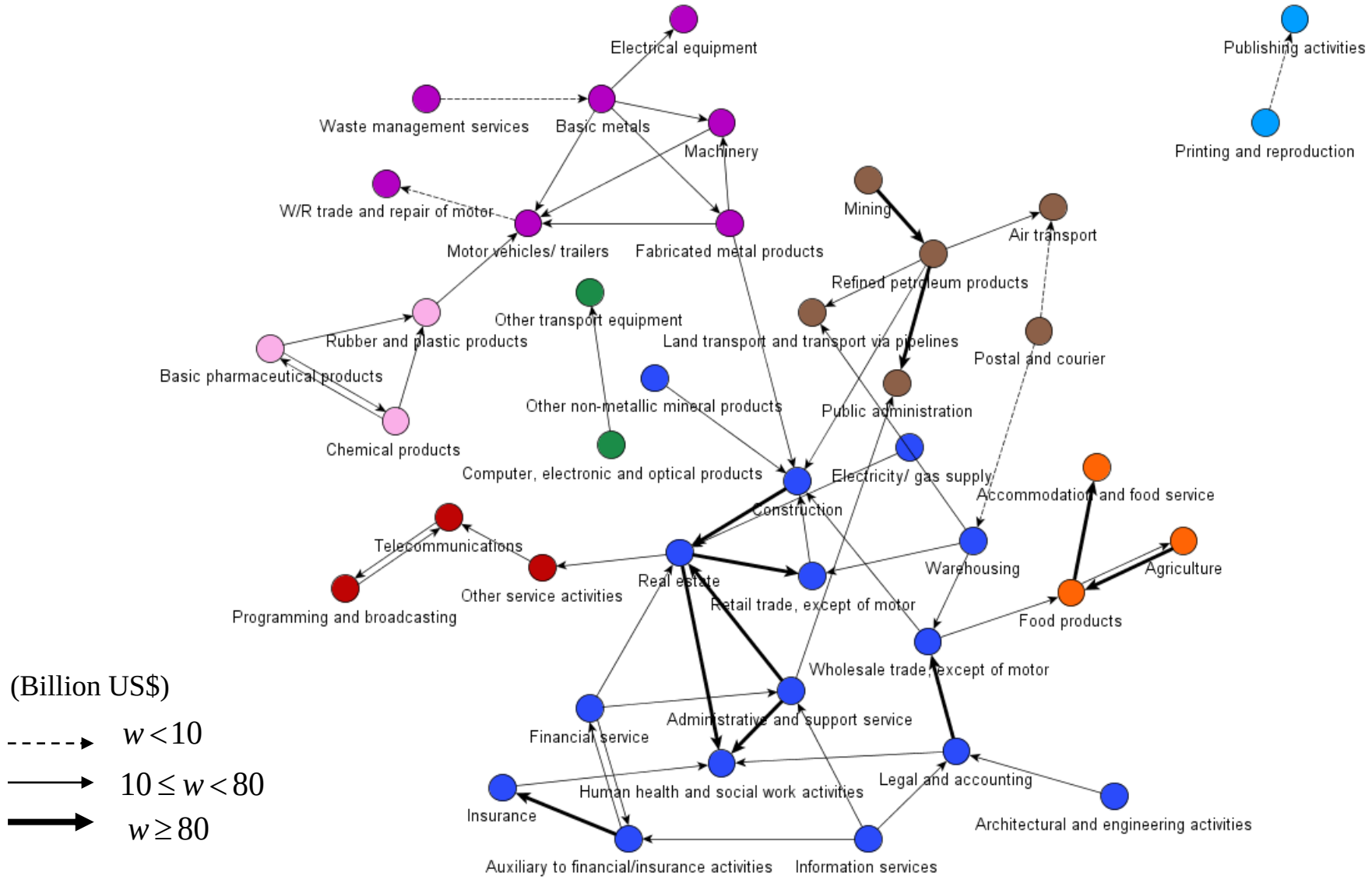
Source: Han Y., Goetz S.J. (2019) "Predicting US county economic resilience from industry input-output accounts" *Applied Economics* 19(51): 2019-2028

US Economy, Backbone, 2000



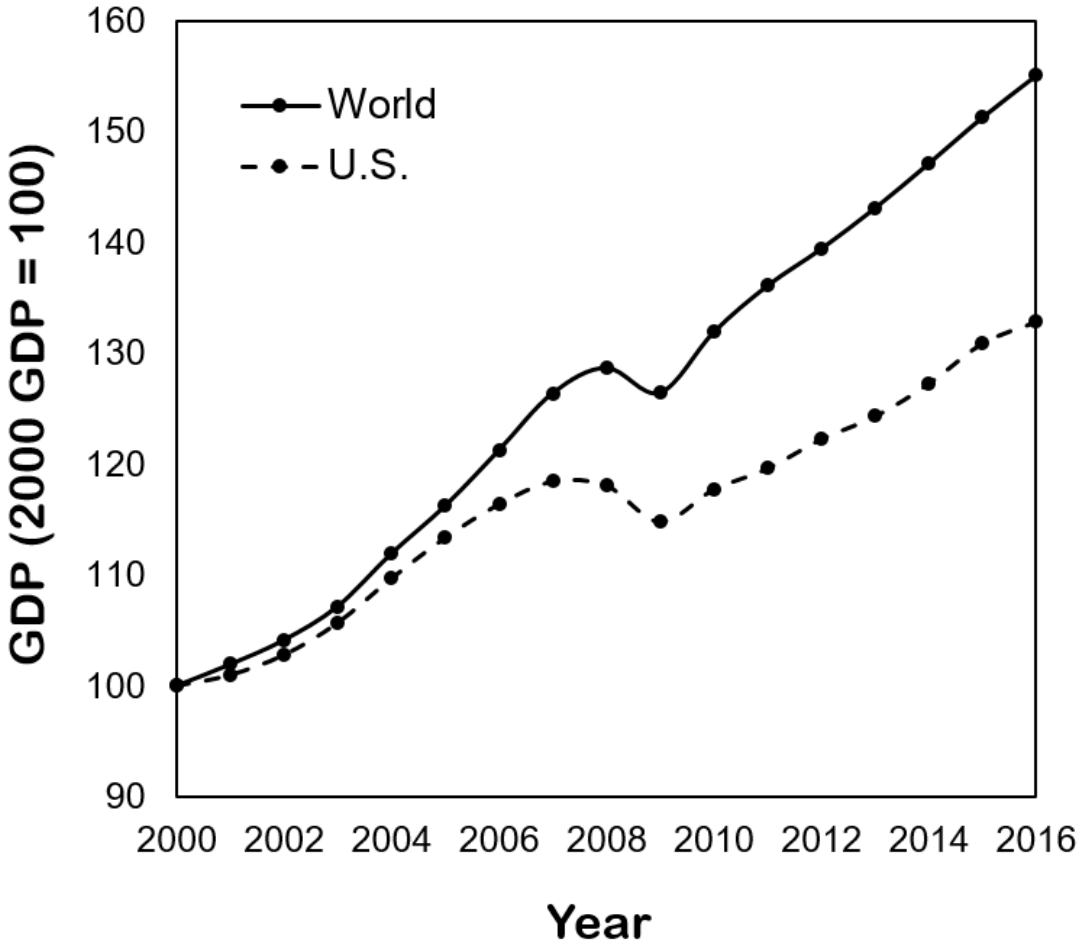
Data source: World Input-Output Table. Backbone detection algorithm of Serrano et al. (2009) for *p*-value 0.02.
Node colors indicate industry communities based on the natural clustering method.

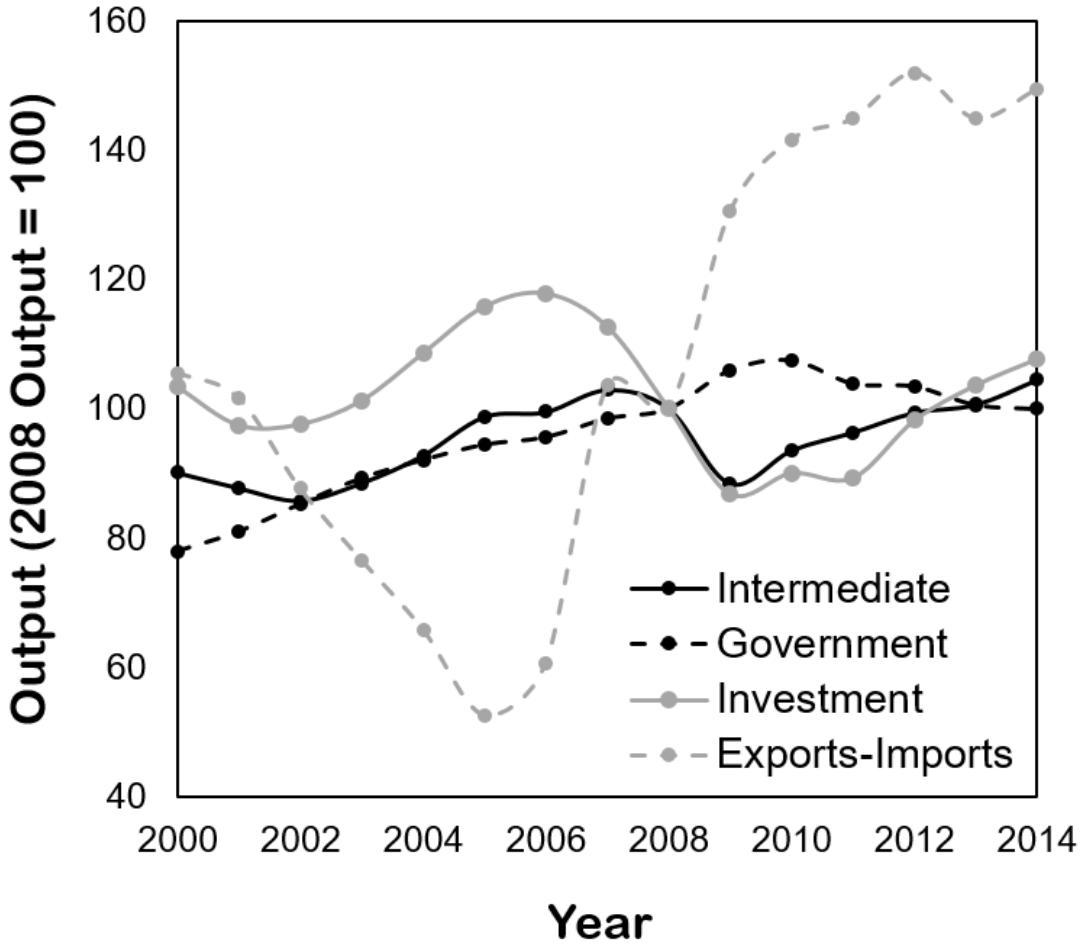
US Economy, Backbone, 2014



Data source: World Input-Output Table. Backbone detection algorithm of Serrano et al. (2009) for *p*-value 0.02. Node colors indicate industry communities based on the natural clustering method.

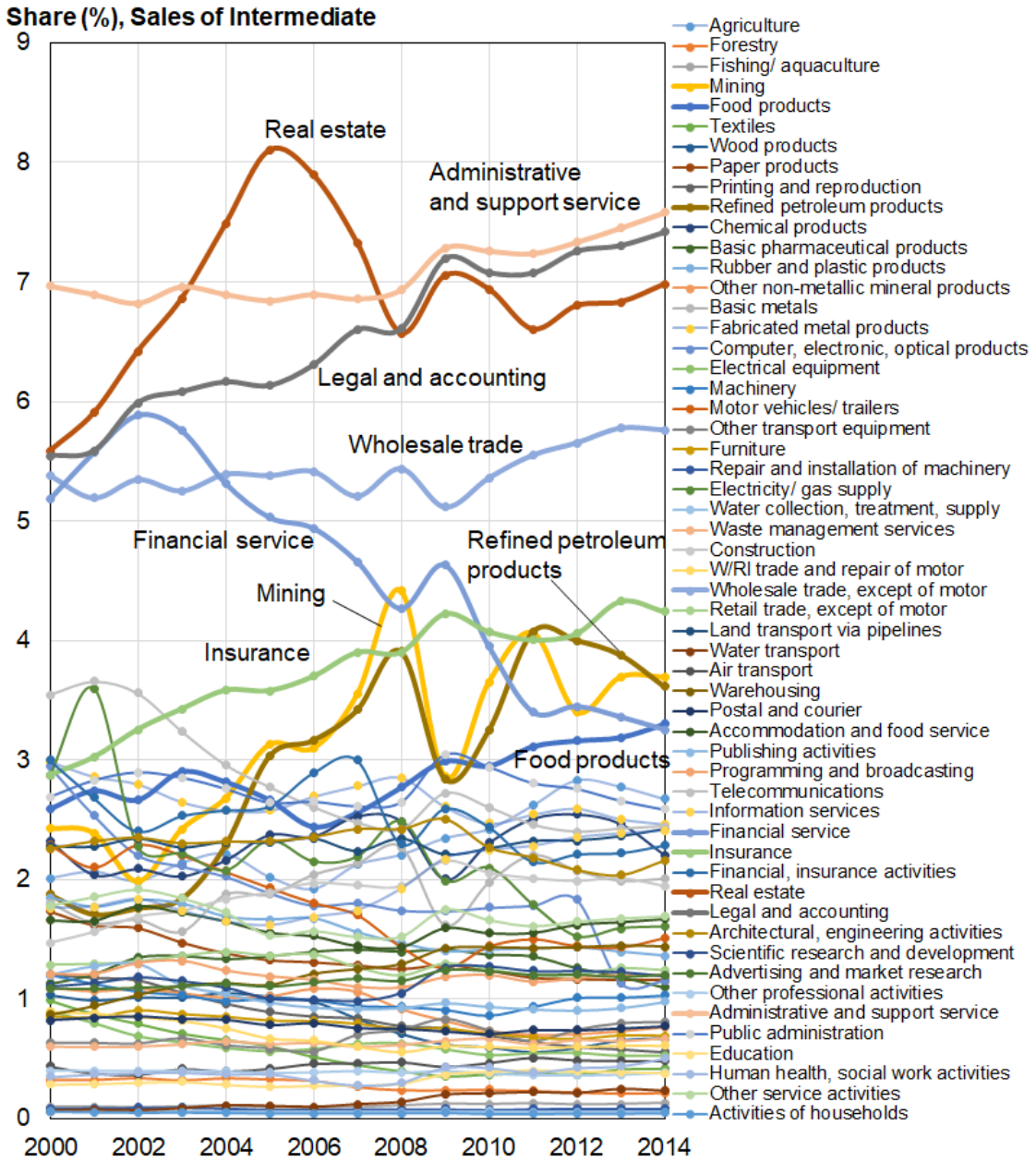
Changes in the U.S. Economy, 2000-2014/16





Note: GDP data are in constant 2014 U.S. dollars. Data source: World Bank and Authors

Changes in Industry Sales as % of total, 2000=2014



Data source:
World Input-Output Table
55 intermediate industries

Highest Sales Shares (top 10), as measure of centrality

US Economy, sales	**2000**	**2014**
Administrative and support service	6.97	7.58
Legal and accounting	5.54	7.42
*Real estate	5.59	6.98
Wholesale trade, except of motor	5.38	5.76
Insurance	2.87	4.24
Mining	2.43	3.69
Refined petroleum products	1.88	3.61
Food products	2.59	3.30
Financial service	5.19	3.25
Agriculture	2.01	2.68

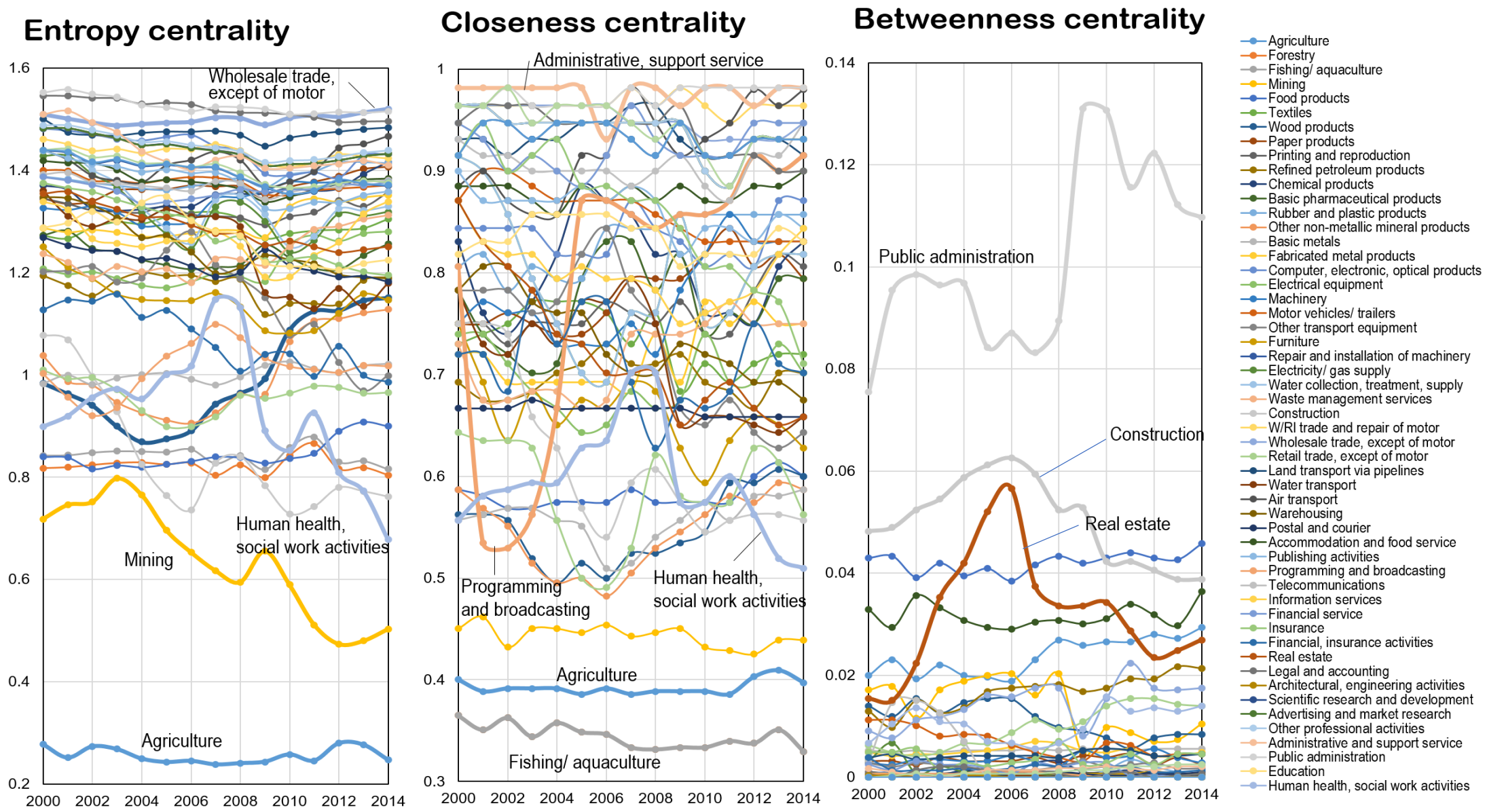
Highest Betweeness Centralities, U.S. industries

Betweenness	**2000**	**2014**
*Public administration	0.0755	0.1097
Food products	0.0430	0.0458
Construction	0.0482	0.0388
Accommodation and food service	0.0328	0.0363
Agriculture	0.0199	0.0294
Real estate	0.0154	0.0269
Refined petroleum products	0.0129	0.0213
Wholesale trade, except of motor	0.0091	0.0175
Retail trade, except of motor	0.0063	0.0140
Human health and social work activities	0.0066	0.0140

Highest Closeness Centrality, U.S. industries

Closeness	**2000**	**2014**
Administrative and support service	0.982	0.982
Public administration	0.964	0.982
*Air transport	0.885	0.980
W/R trade and repair of motor	0.964	0.964
Financial service	0.947	0.947
Telecommunications	0.931	0.947
Wholesale trade, except of motor	0.964	0.947
Activities of households as employers	0.915	0.931
Advertising and market research	0.964	0.931
Architectural and engineering activities	0.964	0.931

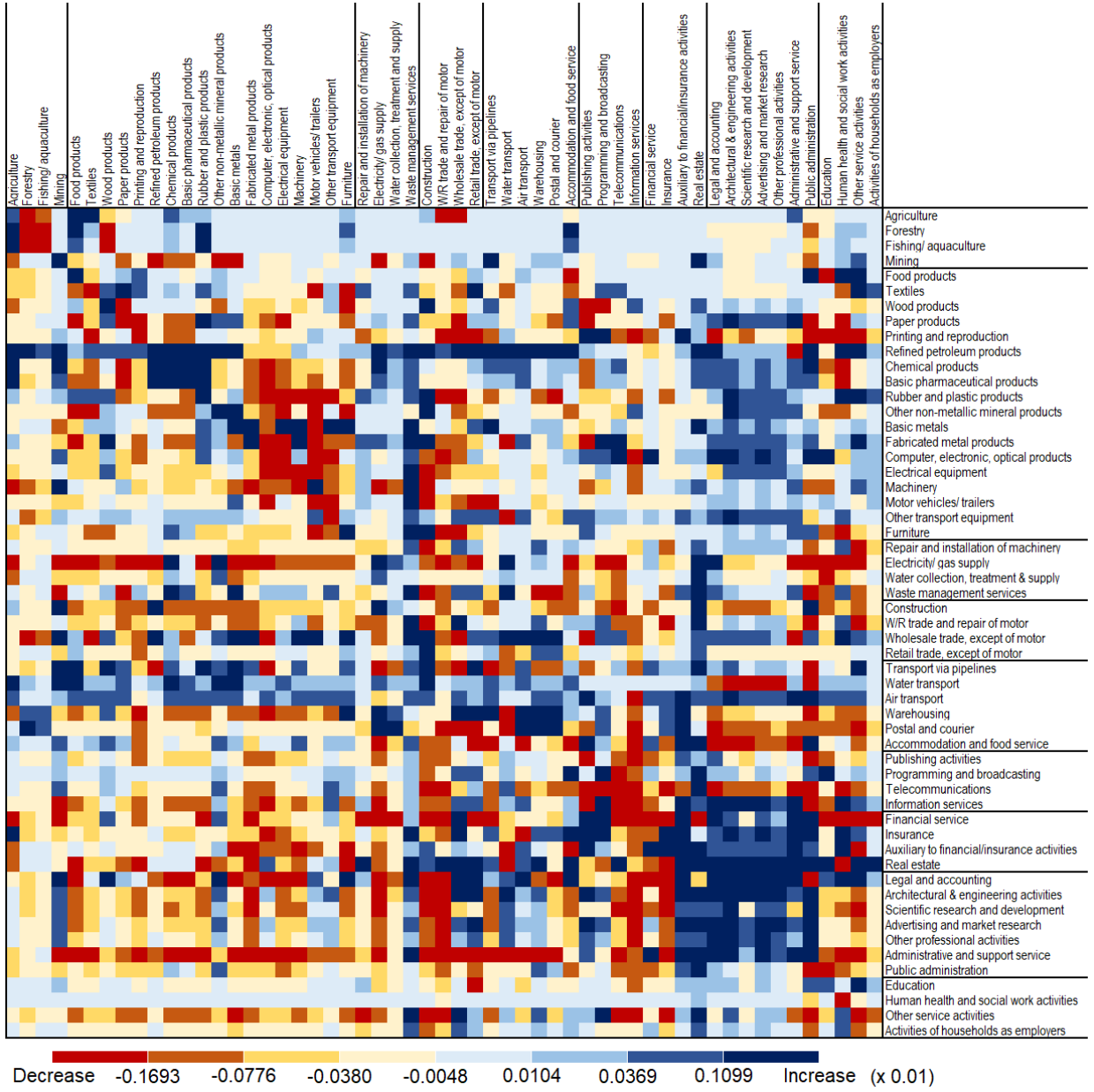
Changes in Network Centralities, U.S. Industries, 2000-2014



Data source: World Input-Output Table. 5 intermediate industries

Changes in the Direct Flow Matrix, 2002-2006

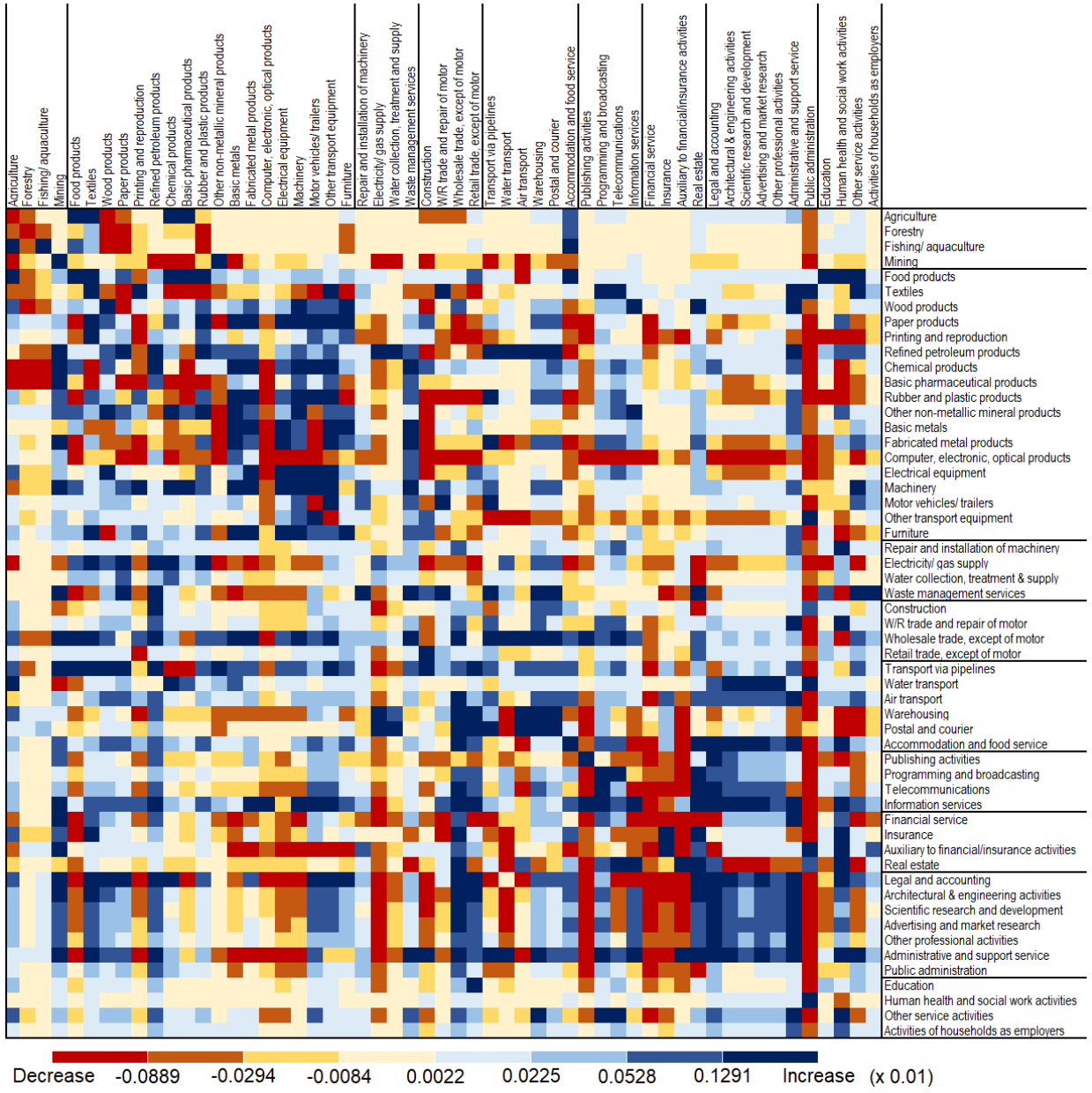
C(pre-recession) =A(2006)-A(2002)



Data source:
World Input-Output Table
55 intermediate industries

Changes in the Direct Flow Matrix, 2010-2014

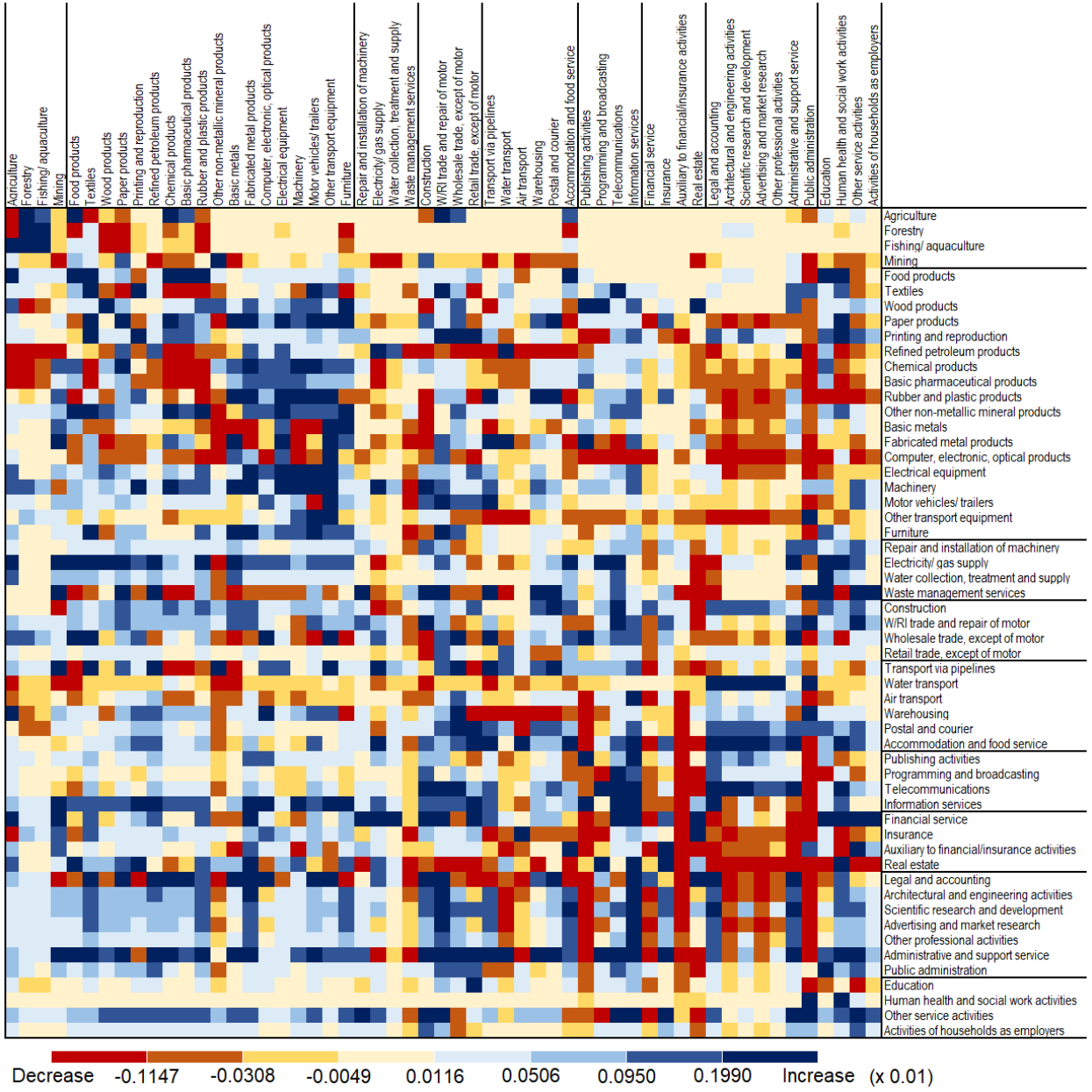
C(post-recession) =A(2014)-A(2010)



Data source:
World Input-Output Table
55 intermediate industries

Rewiring of the Direct Flow Matrix

Rewiring=
C(post-recession)
-C(pre-recession)

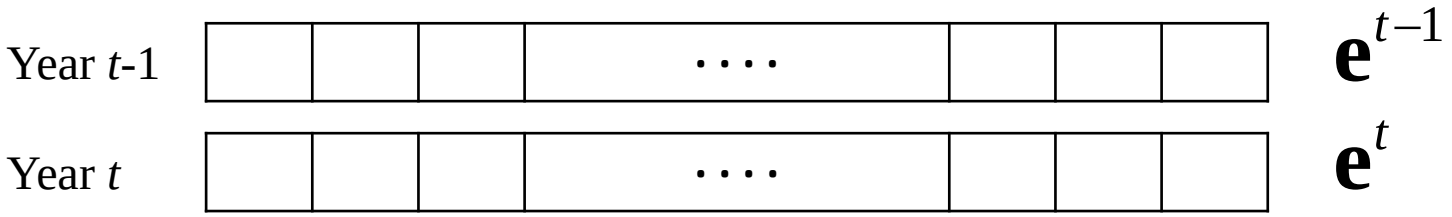


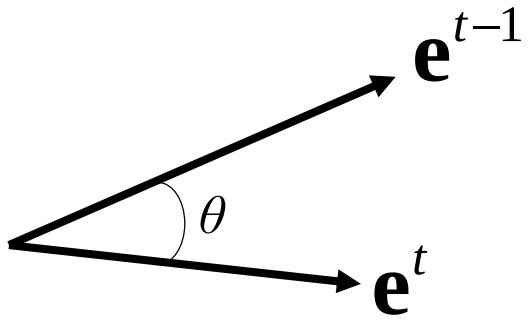
Data source:
World Input-Output Table
55 intermediate industries

Measuring Change in Networks (Rewiring) using cosine similarity or distance

Cosine similarity: Measuring network change (or distance)

Vectors of Inter-Industry Transactions (U.S.); or Employment at 6-digit NAICS level, U.S. counties



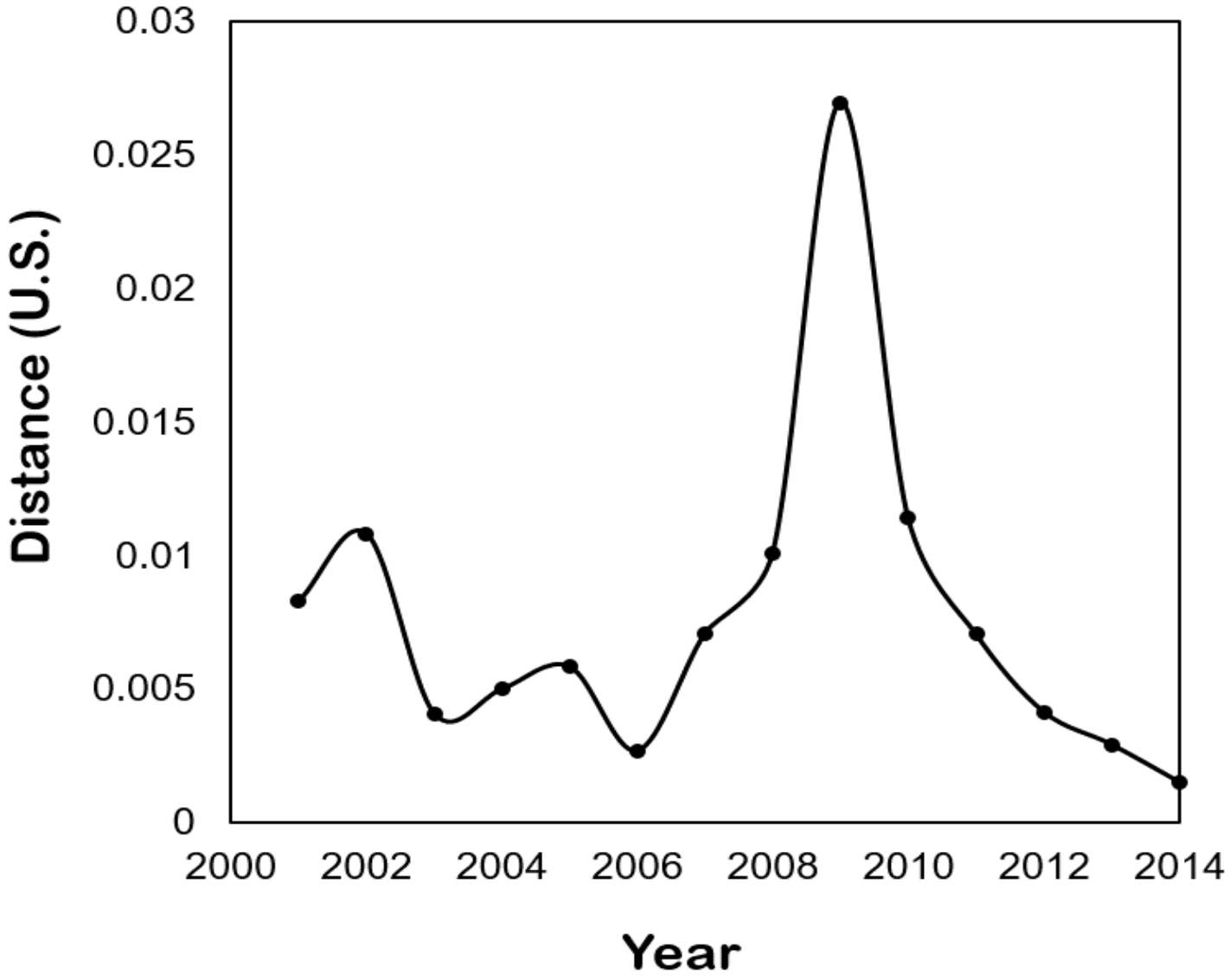


$$\text{distance}(\mathbf{e}^{t-1},\mathbf{e}^{t}) = 1 - \frac{\mathbf{e}^{t-1}\cdot\mathbf{e}^{t}}{\|\mathbf{e}^{t-1}\|\|\mathbf{e}^{t}\|} = 1 - \frac{\sum_j e_j^{t-1}\cdot e_j^{t}}{\sqrt{\sum_j (e_j^{t-1})^2}\sqrt{\sum_j (e_j^{t})^2}} = 1 - \cos\theta^{t}$$

e_j^t is the number of workers employed in industry j at time t

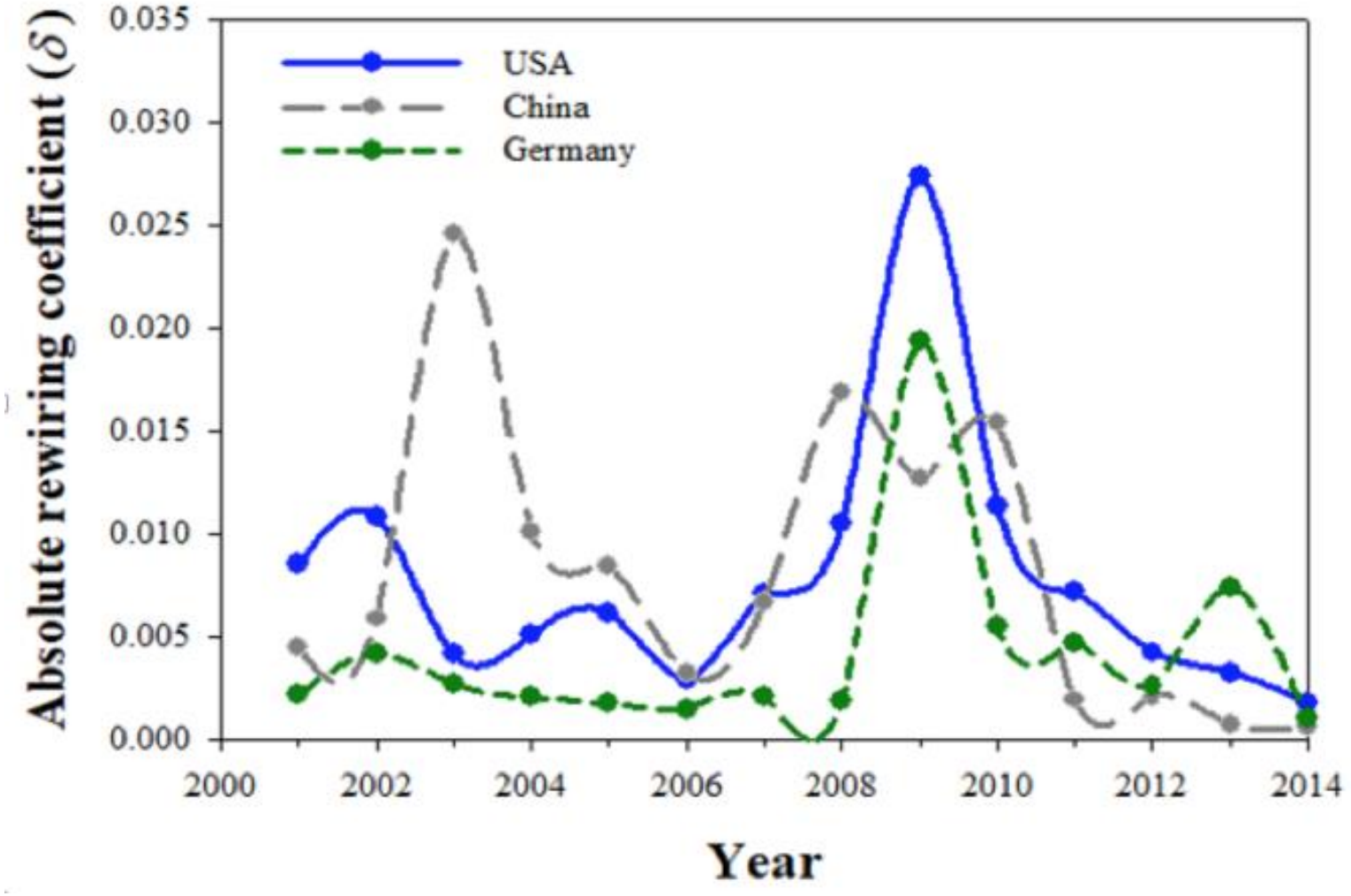
Example 1: National Economies

Year-to-Year “distance” in the U.S. Economy (shows amount of rewiring)



Data source: World Input-Output Table. 55 intermediate industries; from: Goetz and Han (2018) IAES conference, London UK

Rewiring during the GR (cosine similarity): USA, Germany and China



Source: Han, Y. and Goetz S.J. (2019) "Measuring Network Rewiring over Time" *PLOS ONE*

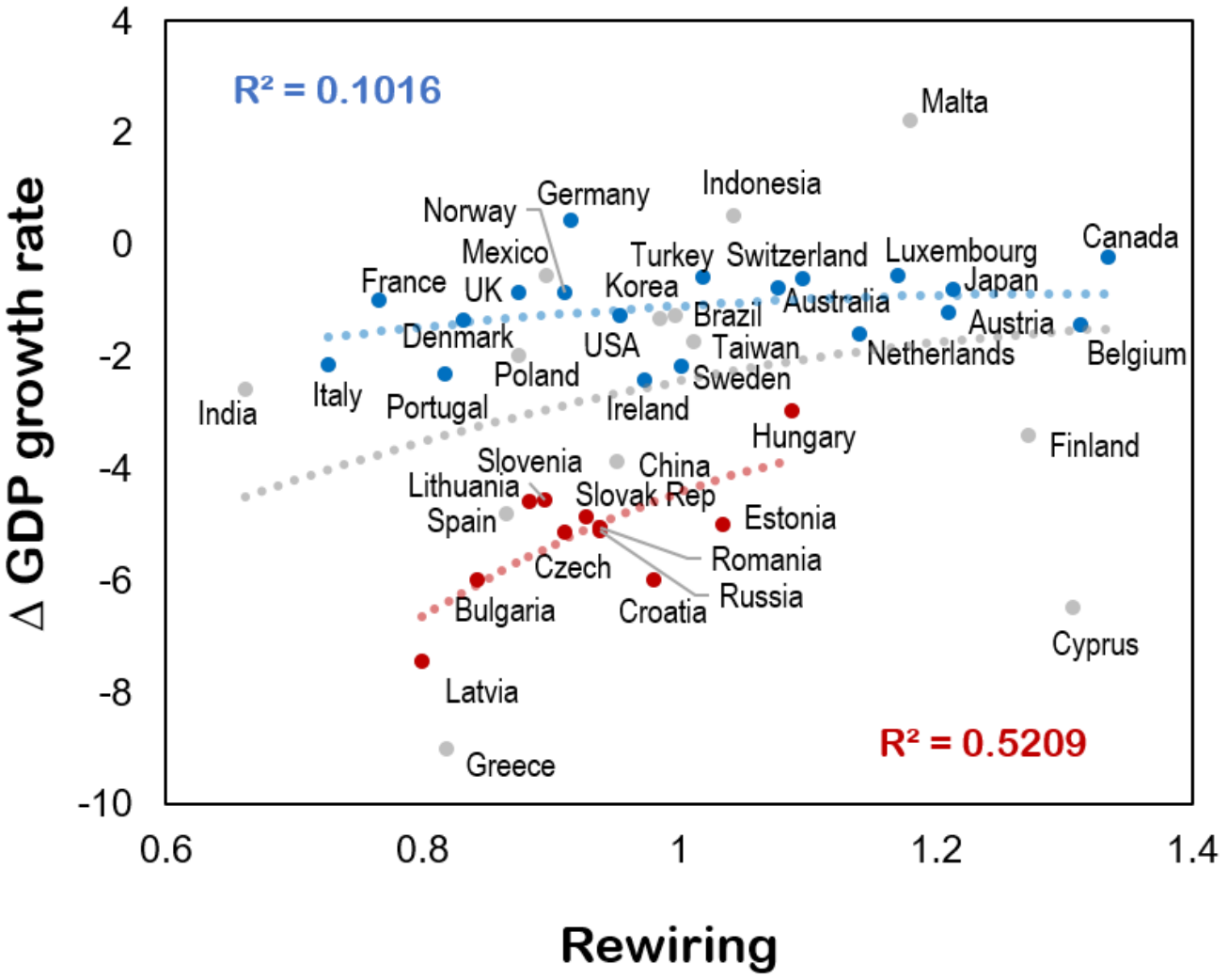
- https://doi.org/10.1371/journal.pone.0220295

Changes in GDP Growth Rates

43 All countries

19 Early OECD
(except Spain and Greece)

11 Eastern Europe
(except Poland)

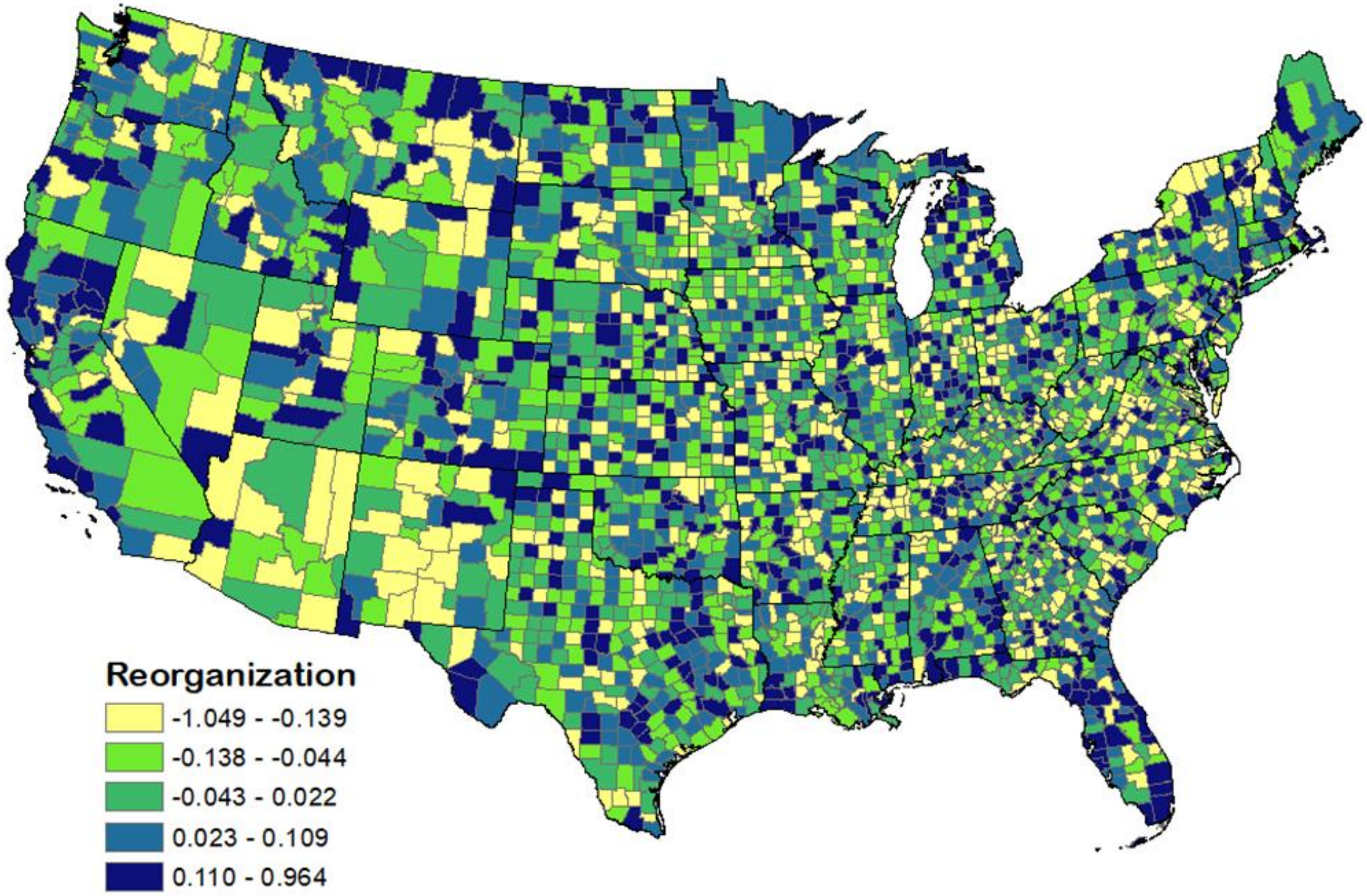


Δ GDP growth rate = GDP growth rate (2010-2014) - GDP growth rate (2002-2006).
Data source: Word Input-Output Database. World Bank, GDP data are in constant 2010 U.S. dollars

Example 2: County Economy Employment Transition between Sectors

Map of employment reorganization during the Great Recession

Cosine similarity applied to employment by industry vectors, 2008-2009



Regression results for changes in the county-level employment growth rate, 2003-2007 and 2011-2015 (using cosine similarity)

	Basic	Model 1	Model 2	Model 3	Model 4	Model 5
Ind_diversity	-0.130*	-0.113*	-0.127*	-0.197***	-0.075	-0.076
JobFlows		-0.021				
Scale of reorg. (s)			-0.129***	0.425***		
Scale_square (s^2)				-0.571***		
Direction of reorg. (d)					0.286***	
Reorganization (s x d)						0.294***
Const.	**	*		***	**	**
Adj. R-square	0.0310	0.0309	0.0469	0.0626	0.1094	0.1143

Dependent variable: Changes in employment growth rate. N=3,108 U.S. counties

Robust standard errors and state-fixed effect, *t* statistic in parentheses. Significance levels: different from zero at *10%, **5%, and ***1% or lower.

JobFlows data from Partridge and Tsvetkova (2018). Other controls: Popdensity, log; Landarea, log; %Edu_High-; %Edu_Bachelor+; Social_Capital

Economic Resilience Based on Input-Output Tables

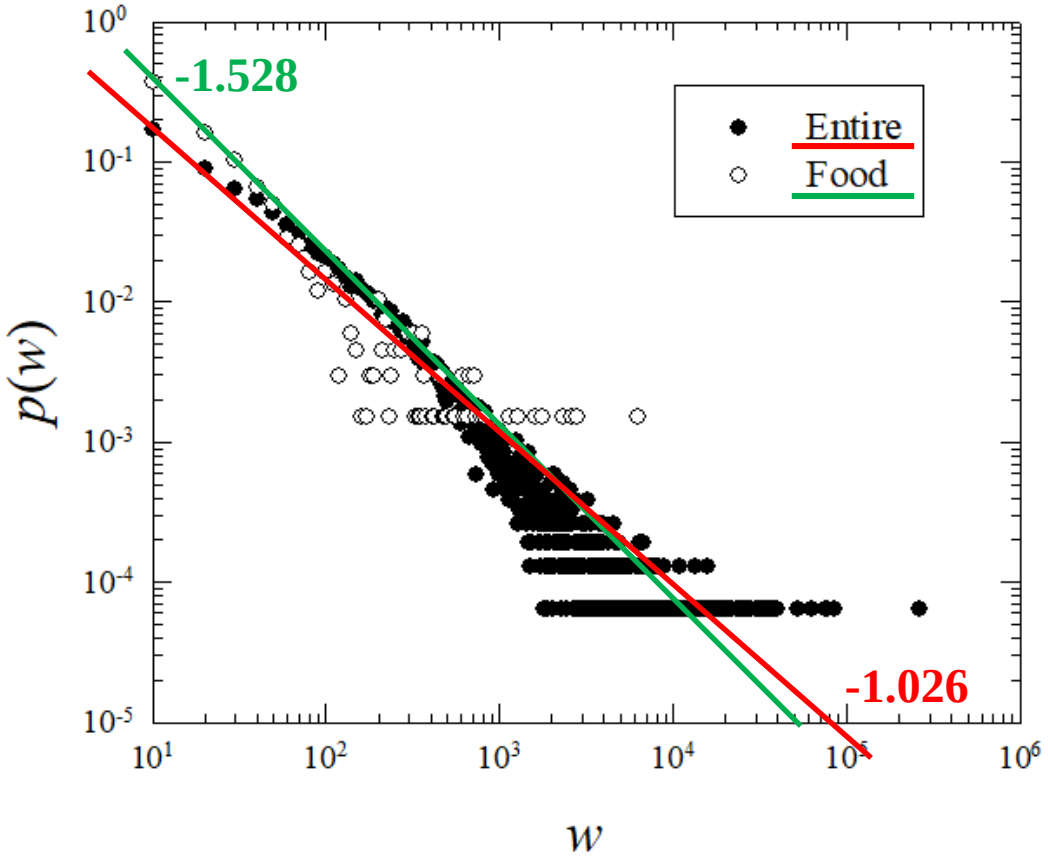
Based on Goetz, S.J. and Han Y. (2013), “Predicting the economic resilience of US counties from industry input-output accounts,” paper pres. at the SRSA meeting, Washington DC.

Chopra, S.S., V. Khanna (2015)“Interconnectedness and interdependencies of critical infrastructures in the US economy: Implications for resilience,” *Physica A: Statistical Mechanics and its Applications*, Vol. 436, 15 Oct., Pages 865-877

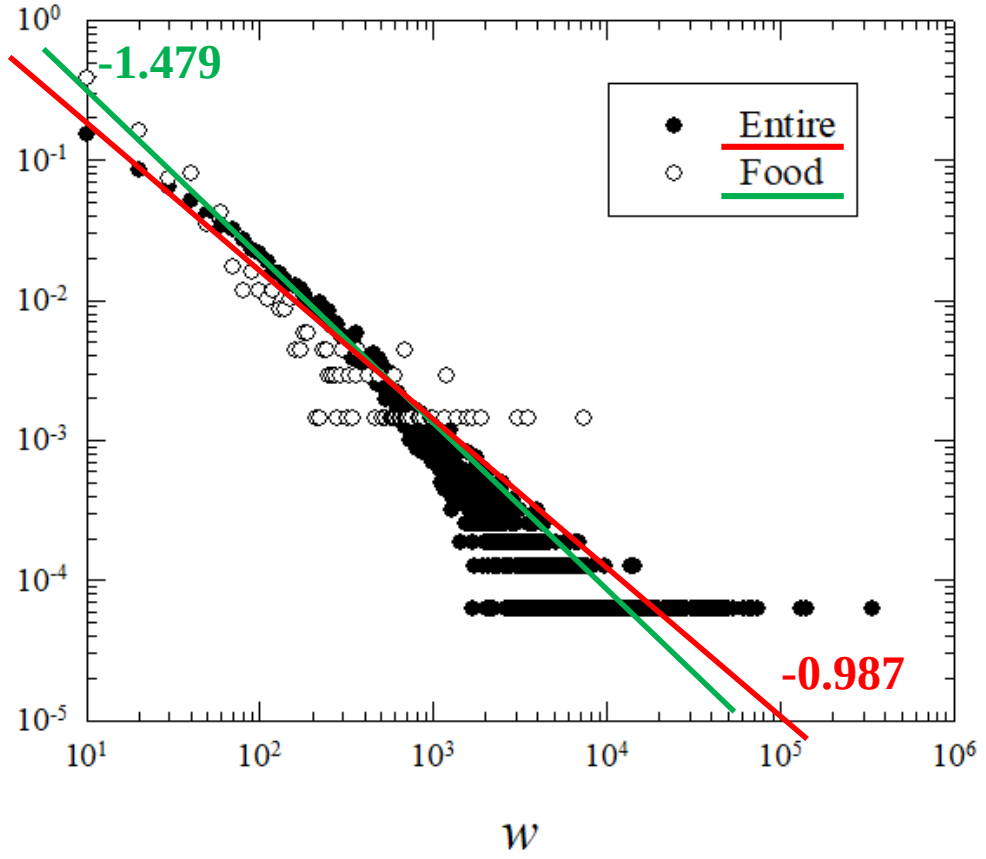
Degree Distributions (and Resilience) of the US Economy and the Food System

Economy-wide resilience based on the input-output table

1997



2014

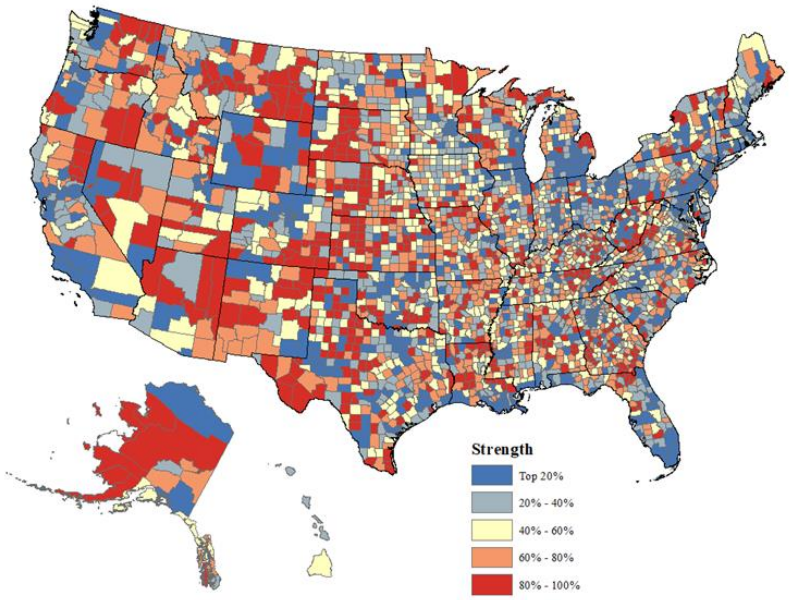


Both systems have become more concentrated (flatter slope, larger or less negative gamma coefficient) and thus more vulnerable to a targeted attack or failure, but less vulnerable to a random attack.

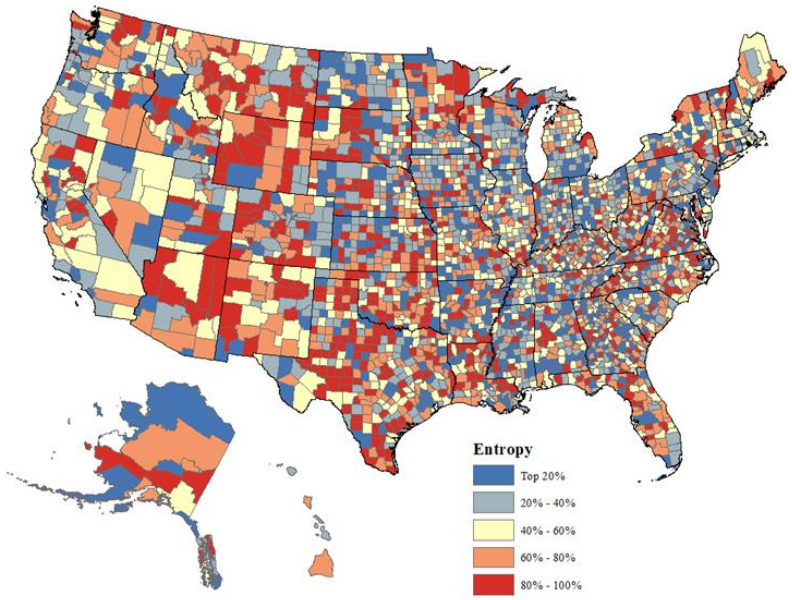
County centrality scores stepped down to the county level

Centrality scores based on the input output table

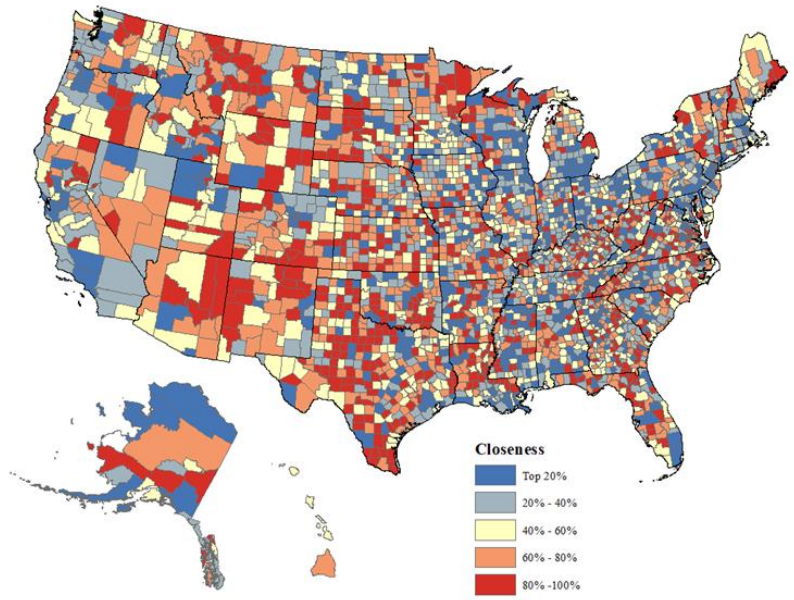
Strength Centrality



Entropy Centrality



Betweeness Centrality



Source: Han Y., Goetz S.J. (2019) "Predicting US county economic resilience from industry input-output accounts" *Applied Economics* 19(51): 2019-2028

Impacts of Centralities on *Resilience* (=dep. variable)

Source: Han Y., Goetz S.J.: (2015) *Review of Regional Studies and* (2019) *Applied Economics*

Beta Coefficients	Strength	Entropy	Closeness
%Foreign born	0.121***	0.121***	0.120***
In-migration	0.084***	0.085***	0.085***
Commuting	0.180***	0.186***	0.182***
Strength	0.043*		
Entropy		0.038**	
Closeness			0.059***
Emp density, log	0.212***	0.226***	0.211***
Land area, log	0.200***	0.207***	0.200***
%Age_2444	0.096***	0.092***	0.09***
Income, log	-0.016	-0.010	-0.016
Education	0.168***	0.166***	0.168***
_cons	*		*
	(0.161)	(0.158)	(0.159)

N=3,087 counties, SFE included.

Spatial Supply Chains based on Input-Output Tables

Local and Regional Food Supply Chain Elements & Locations!

Production: F&V, Meat (NAICS 11)

Livestock Merchant Wholesalers: 424520

Farm product warehousing and storage (except refrigerated) 493130

Specialized Freight Trucking, Local, including refrigerated: 484220

Short Line Railroads: 482112

311

Frozen Fruit, Juice, &Vegetable Mfg.: 411

Fruit and Vegetable Canning: 421

Dried & Dehydrated Food Mfg: 423

Flour Milling: 211

Soybean & Other Oilseed Processing: 221

Perishable Prepared Food Mfg: 991

Refrigerated Warehousing and Storage: 493120

Animal (except Poultry) Slaughtering: 611

Meat Processed from Carcasses: 612

311

Poultry Processing: 615

Rendering and Meat Byproduct Processing: 613

Seafood Product Preparation and Packaging: 311710

Wholesaler level: 424

Fresh Fruit and Vegetable Merchant Wholesalers: 480

Frozen Vegetable Merchant Wholesalers: 420

Poultry and Poultry Product Merchant Wholesalers: 440

Meat and Meat Product Merchant Wholesalers: 470

Fish and Seafood Merchant Wholesalers: 460

Other Grocery and Related Products Merchant wholesalers : 490

Retailer level: 445

Fruit and Vegetable Markets: 210

Meat Markets: 210

Fish and Seafood Markets: 220

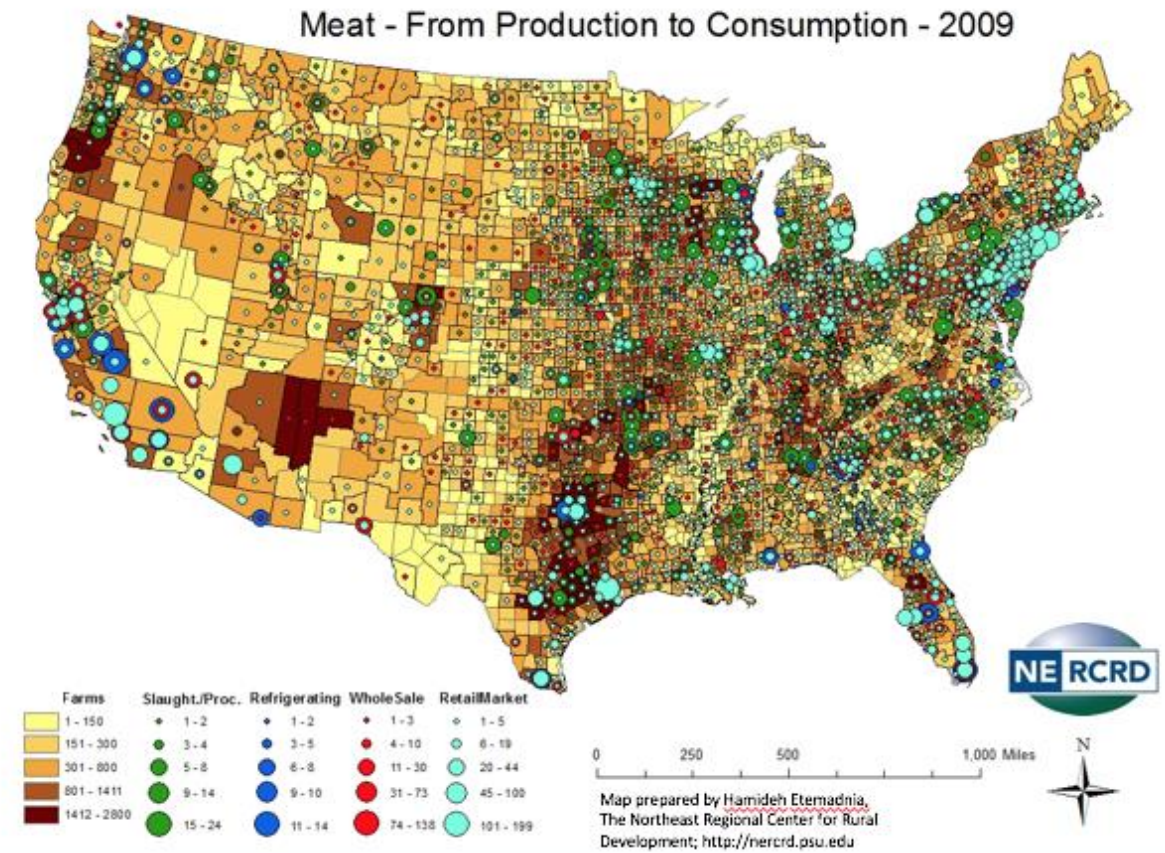
Supermarkets and Other Grocery (except Convenience) Stores: 110

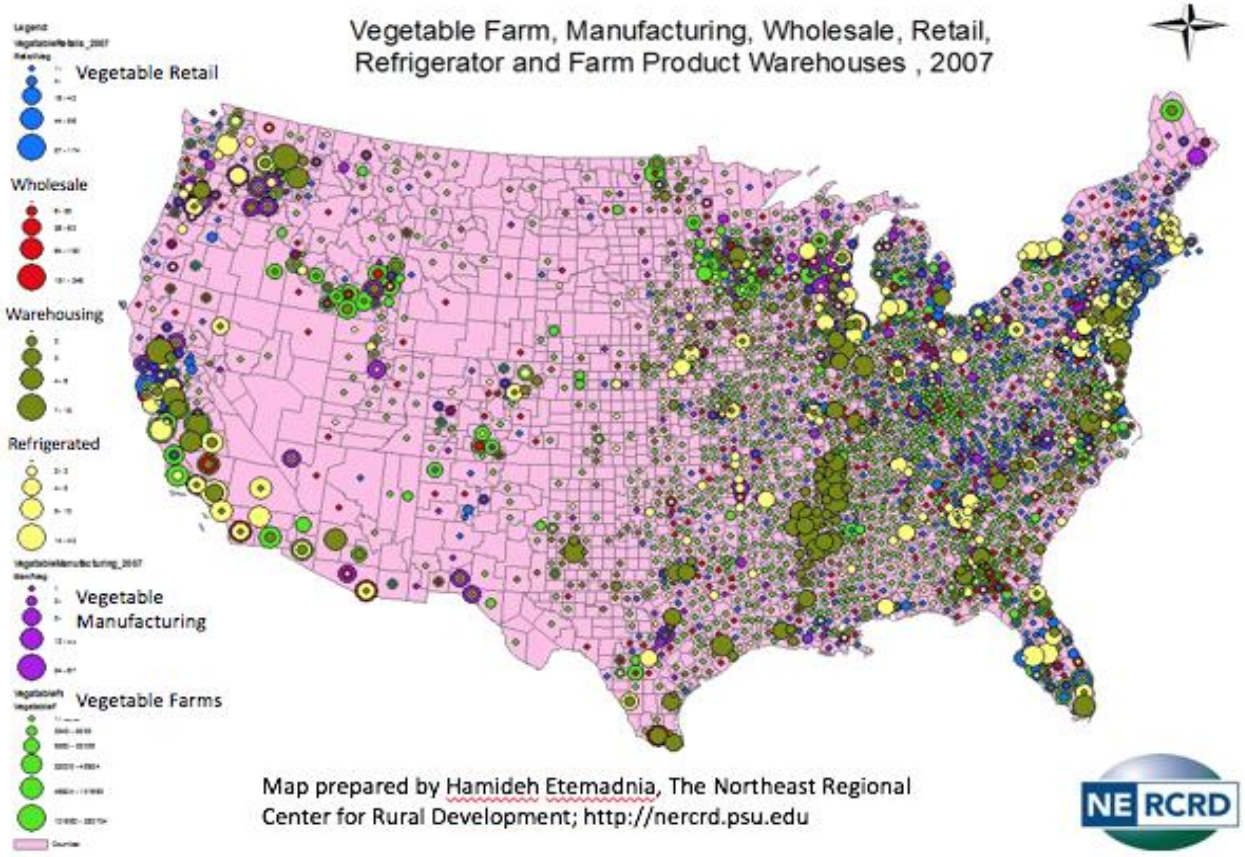
Convenience Stores: 120

Warehouse and Supercenters: 452910

Consumption

NAICS codes used in County Business Patterns





Based on: H. Etemadnia, S.J. Goetz, P. Canning, M. Sadegh Tavallali (2015) "Optimal wholesale facilities location within the fruit and vegetables supply chain with bimodal transportation options: An LP-MIP heuristic approach," *European Journal of Operational Research*, 244 (2): 648-661

Background and Purpose

Economic flows in local economies

- Local spatial linkages of goods, services, and information
- Input-Output table (IO table) captures economic flows (transactions) between industries
- But not flows between local economies

Modeling flows in a local IO network

- Sectors in local economies → Nodes
- Transactions between local sectors → Edges
- We estimate the local IO network
  - using the national IO table and industry structures in local economies
  - based on the idea that complex networks connect themselves
- Then, we analyze spatial linkages in the local IO networks

Source: Han Y., Goetz S.J., and Schmidt C.; Pres. at the 59th ERSA Congress (2019), Lyon, France

Method

Assumptions

- Transactions between local sectors follow the gravity rule and the preferential attachment rule in complex networks (Barabási and Albert, 1999)
- Production and consumption of a local sector are proportional to employment (or population) in the local sector

Linkages

- Preferential attachment: A node prefers to attach to a node that already has a large number of connections
- Gravity effect: Amount of transactions between local industries are inversely proportional to distance
- Total production and consumption of local industries, and transactions between industries are fixed

Barabási, A.-L., and R. Albert (1999) Emergence of Scaling in Random Networks. *Science* 286(5439): 509-512.

Model (balance production – consumption over space):

Uses IO table to examine each segment of the supply chain

Balance Equation of industry *S* in IO table

$$X^s = \sum_t^m x^{st} + \sum_u^n f^{su} = \sum_i^N \sum_j^N \left(\sum_t^m x_{ij}^{st} + \sum_u^n f_{ij}^{su} \right)$$

x^{st} : flows from industry s to industry t

f^{su} : flows from industry s to final demand u

x_{ij}^{st} : flows from industry s in place i to industry t in place j

f_{ij}^{su} : final demand flows from industry s in place i to final demand u in place j

Production, consumption

$$p_i^s = X^s \frac{m_i^s}{\sum_k m_k^s}, \quad c_j^t = Y^t \frac{m_i^t}{\sum_k m_k^t}, \quad f_j^u = Y^u \frac{m_i^u}{\sum_k m_k^u}$$

p_i^s: total production of industry s in place i

c_j^t: total consumption of industry t in county j

f_j^u: total consumption of final demand u in county j

m_i^s: employment in industry s in county i

X^s: Total output of industry s

Y^t: Total intermediate consumption of industry t

Transactions

$$x_{ij}^{st} = k_{ij}^{st} \cdot p_i^s \cdot \Pi_{ij}^{st}, \quad \Pi_{ij}^{st} = \frac{c_j^t \cdot x^{st} \cdot d_{ij}^{-\gamma}}{\sum_k \sum_m c_k^m \cdot x^{sm} \cdot d_{ik}^{-\gamma}}$$

$$f_{ij}^{su} = k_{ij}^{su} \cdot p_i^s \cdot \Pi_{ij}^{su}, \quad \Pi_{ij}^{su} = \frac{f_j^u \cdot x^{su} \cdot d_{ij}^{-\gamma}}{\sum_k \sum_n f_k^n \cdot x^{sn} \cdot d_{ik}^{-\gamma}}$$

k_{ij}^{st}: a proportionality constant

d_{ij}: distance from county i to county j

γ: exponential coefficient for distance effect

Π_{ij}^{st}: probability that a unit commodity s produced in place i is consumed in industry t in place j

Conditional equation

$$p_i^s = \sum_j \left(\sum_t x_{ij}^{st} + \sum_u f_{ij}^{su} \right), \quad c_j^t = \sum_i \sum_t x_{ij}^{st}, \quad f_j^u = \sum_i \sum_u f_{ij}^{su}, \quad x^{st} = \sum_i \sum_j x_{ij}^{st}, \quad f^{su} = \sum_i \sum_j f_{ij}^{su}$$

Data

IO table and sectors

- Total 416 sectors from the 2012 Benchmark IO table (BEA)

Local industry structure

- Employment based on NAICS in 3,148 counties
- Farming sectors: commodity products from AG Census
- Some sectors in final demand: Population from US Census

Distance between counties

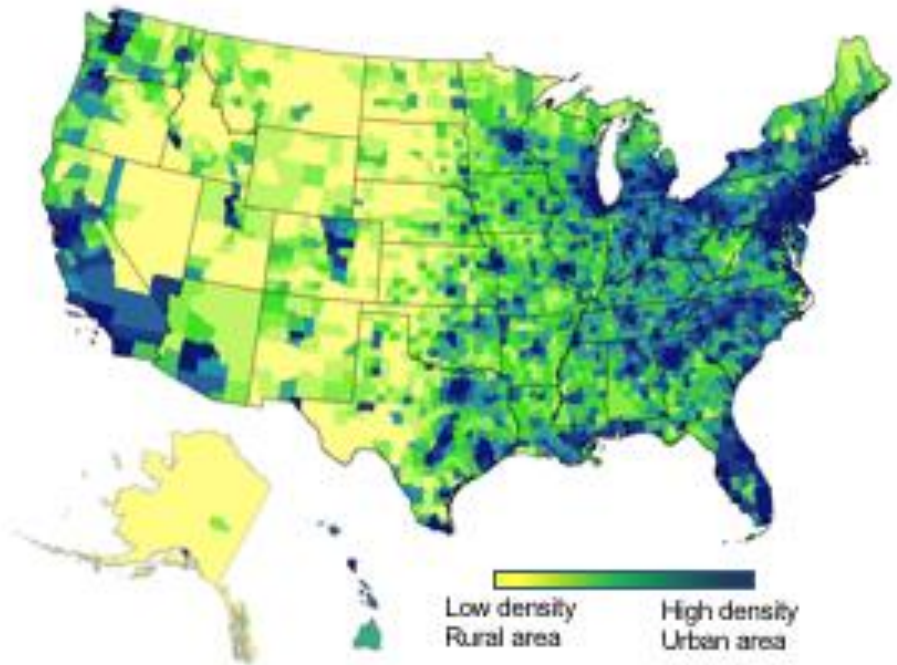
- Network impedance from Oak Ridge Natl. Lab.
- Transportation cost depends on the geographical environment (mountains, river, etc.) and road types (highway, railroad, ship, etc.)
- 1 mile distance (highway), 1/3.3 (rail), 1/5.0 (inland barge), 1/5.8 (Great Lakes), 1/6.5 (marine shipping)

Rural-urban linkages

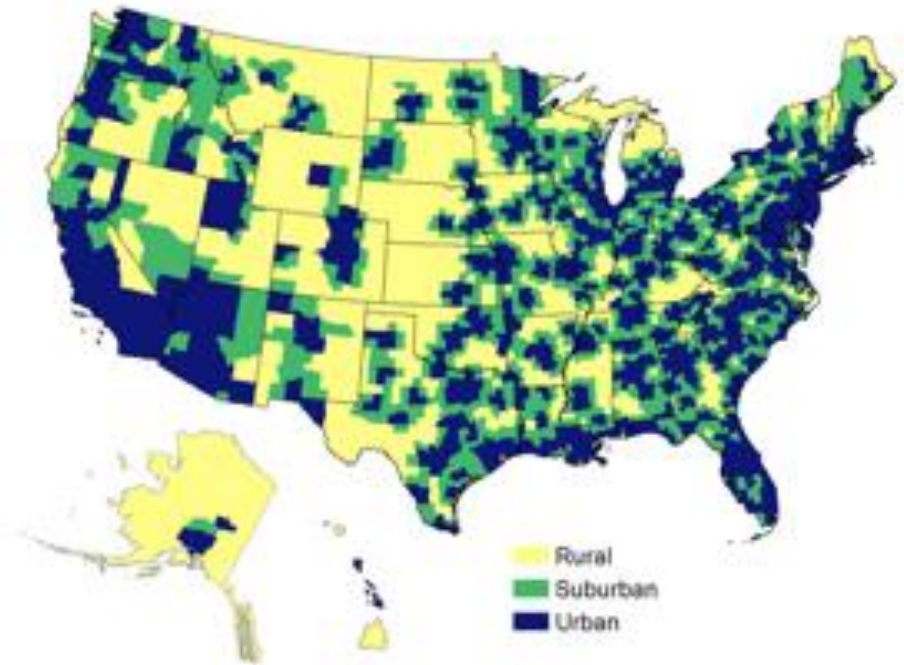
- Population density (per square miles)
  - High population density: urban
  - Low population density: rural
- Rural-urban continuum code (RUCC)

Code	Description	Type
1	Counties in metro areas of 1 million population or more	Urban
2	Counties in metro areas of 250,000 to 1 million population	Urban
3	Counties in metro areas of fewer than 250,000 population	Urban
4	Urban population of 20,000 or more, adjacent to a metro area	Suburban
5	Urban population of 20,000 or more, not adjacent to a metro area	Rural
6	Urban population of 2,500 to 19,999, adjacent to a metro area	Suburban
7	Urban population of 2,500 to 19,999, not adjacent to a metro area	Rural
8	Completely rural or less than 2,500 urban pop., adjacent to a metro area	Suburban
9	Completely rural or less than 2,500 urban pop., not adjacent to a metro area	Rural

Population density, 2016



Rural-urban continuum code (RUCC)



Examining Rural-Urban linkages

Examine three industries

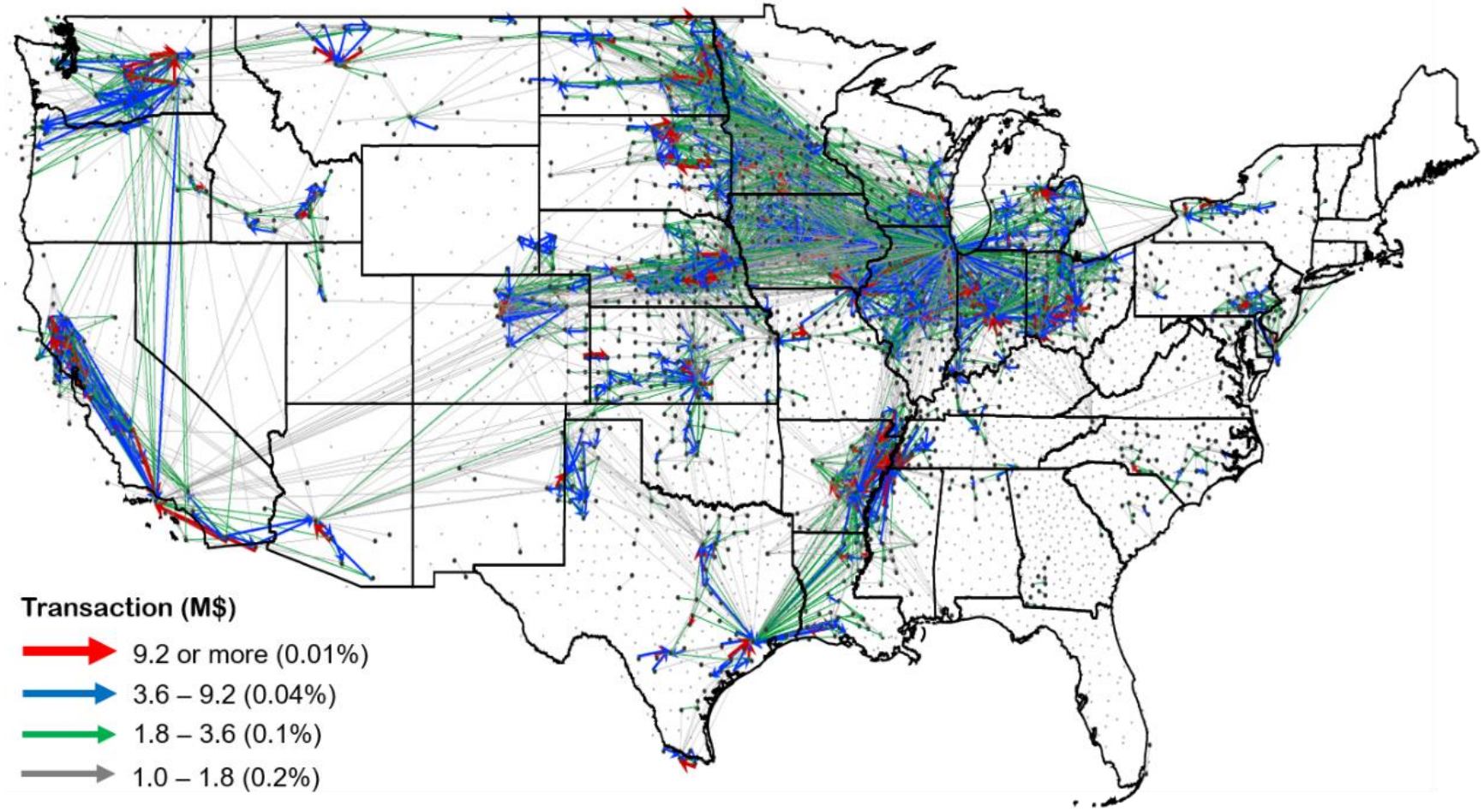
- Agriculture: Grain farming (primary)
- Manufacturing: Motor vehicle gasoline engine and engine parts (secondary)
- Information: Data processing, hosting, and related services (tertiary)

Rural-urban linkages

- Discrepancies between production and consumption
- Rural-urban linkages sorted by population density and RUCC
- Generate a network map based on linkages
- Show rural-urban linkages in the two largest downstream industries

Spatial supply chain

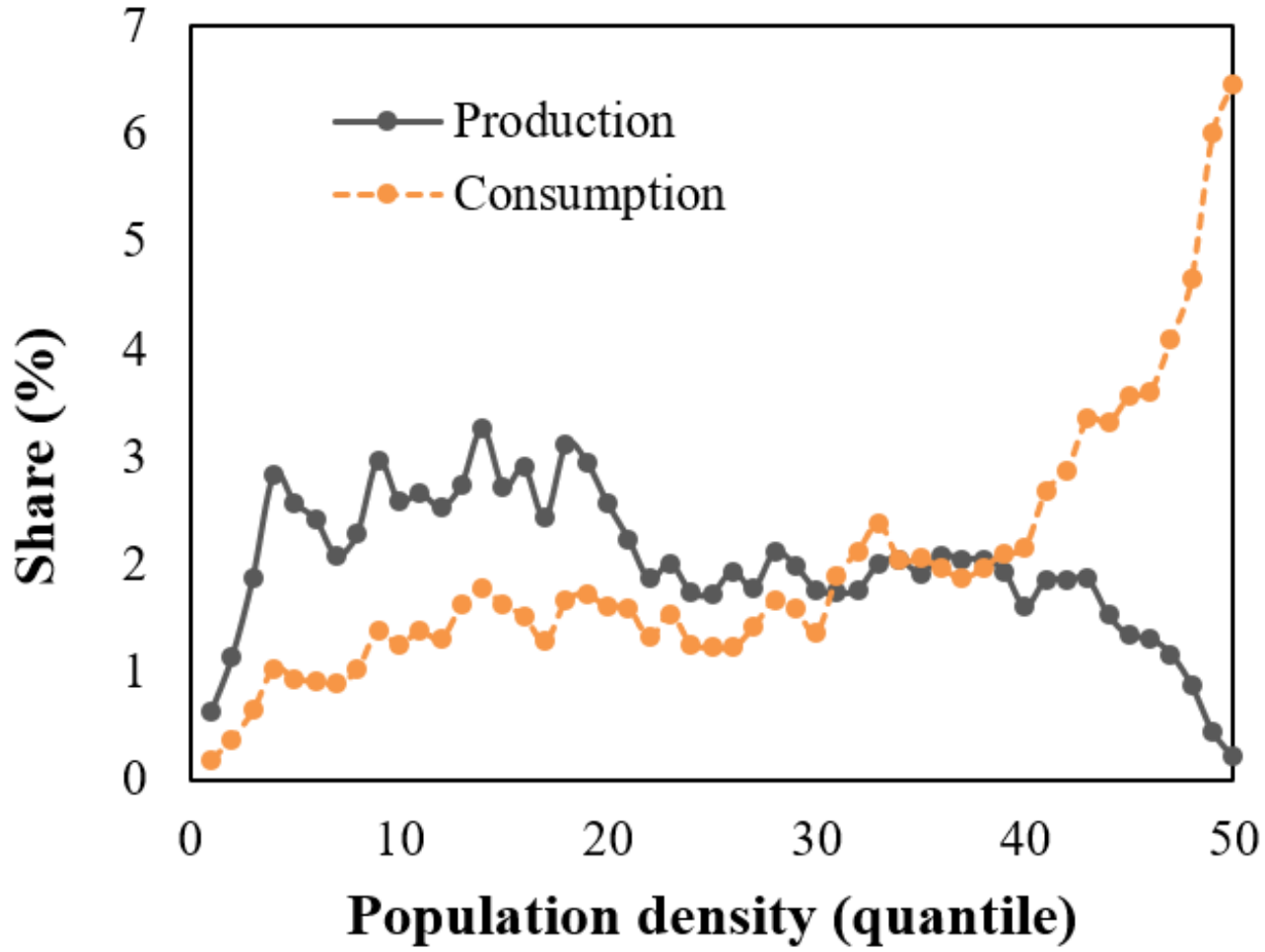
Grain farming



Source: Han Y., Goetz S.J., and Schmidt C.; Pres. at the 59th ERSA Congress, 2019, Lyon, France

Production-consumption

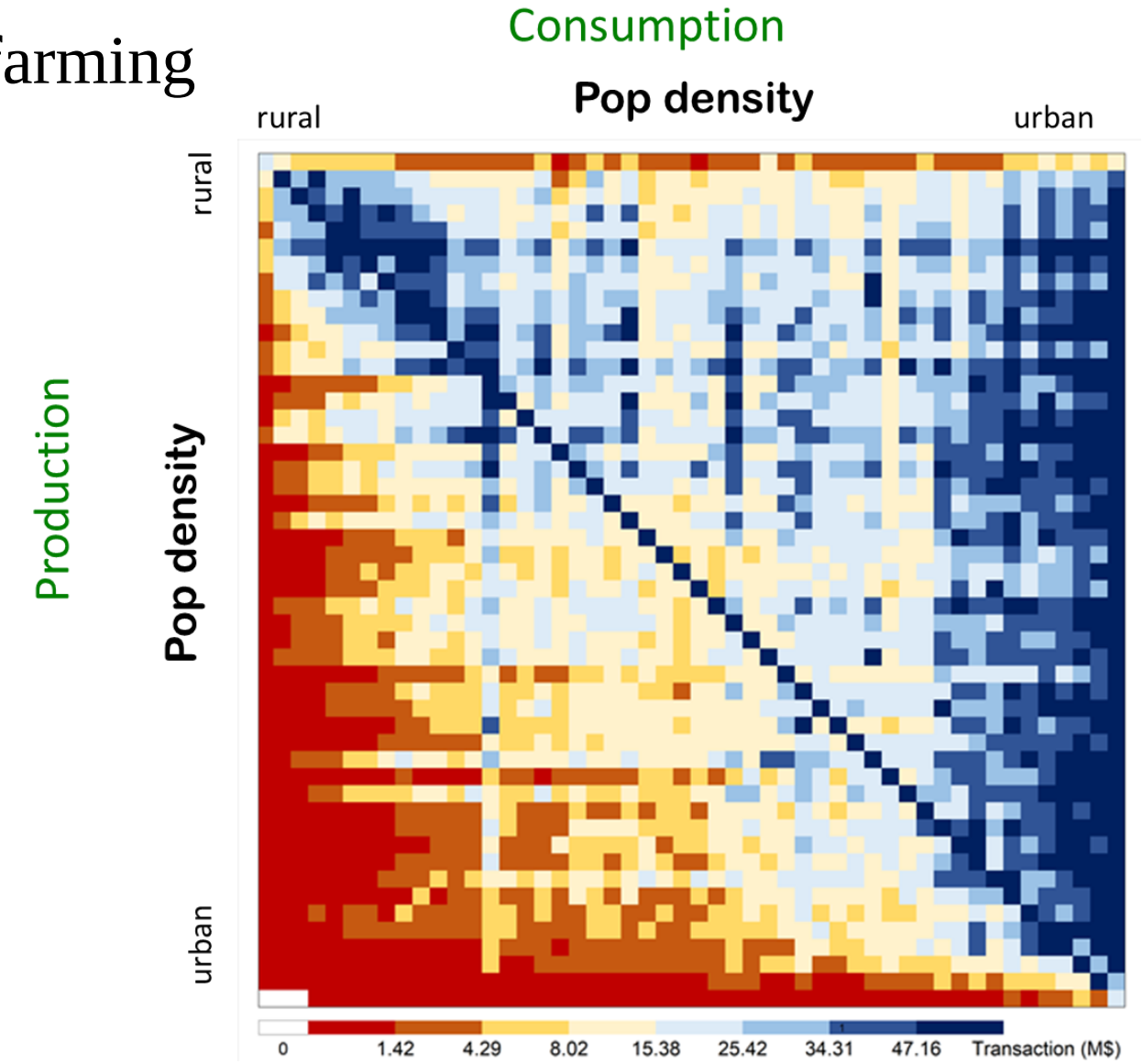
Grain farming



Discrepancies between production and consumption create rural-urban linkages

Rural-urban linkages

Grain farming



Rural-urban linkages

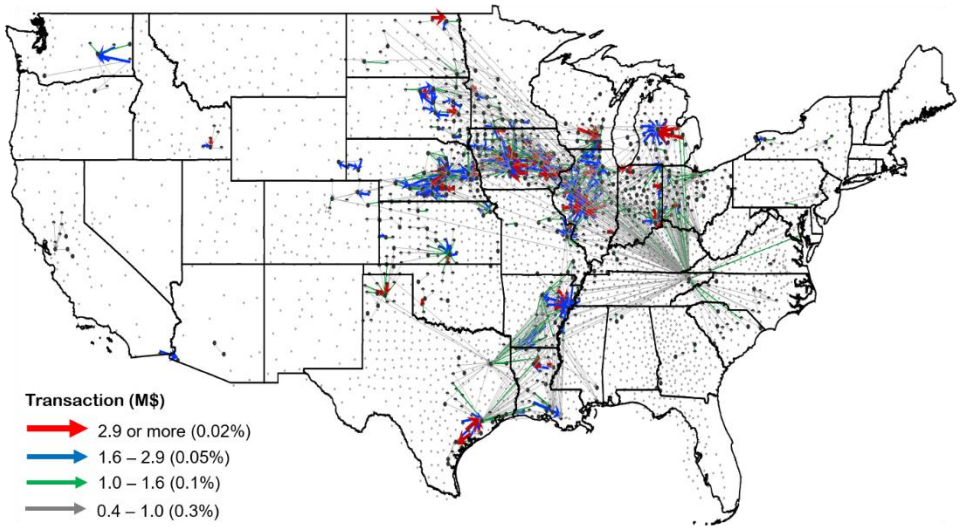
Grain farming

%	Rural	Suburban	Urban	Production
Rural	11.6	7.4	15.8	34.8
Suburban	5.8	9.6	20.3	35.6
Urban	3.5	6	20.1	29.6
Consumption	20.8	23	56.2	100

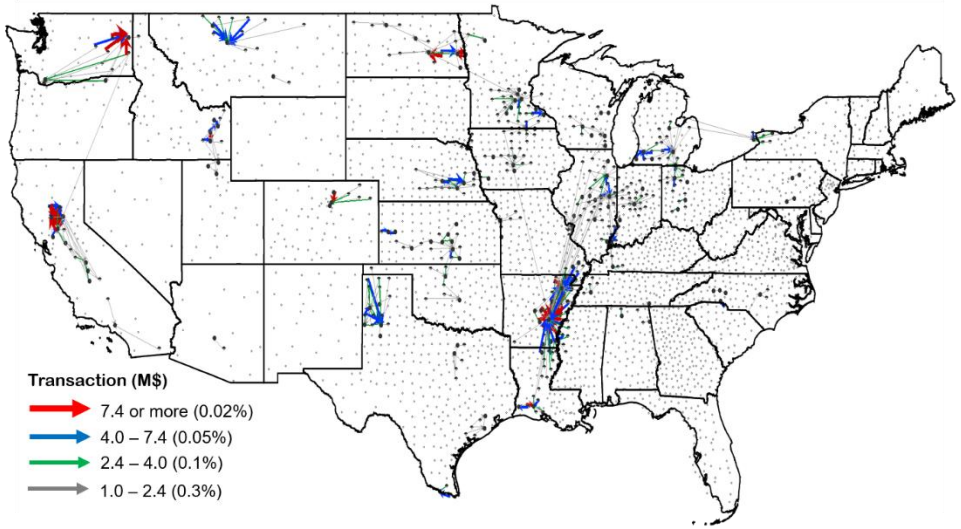
Spatial supply chain

From Grain farming to…

Other basic organic chemical manufacturing



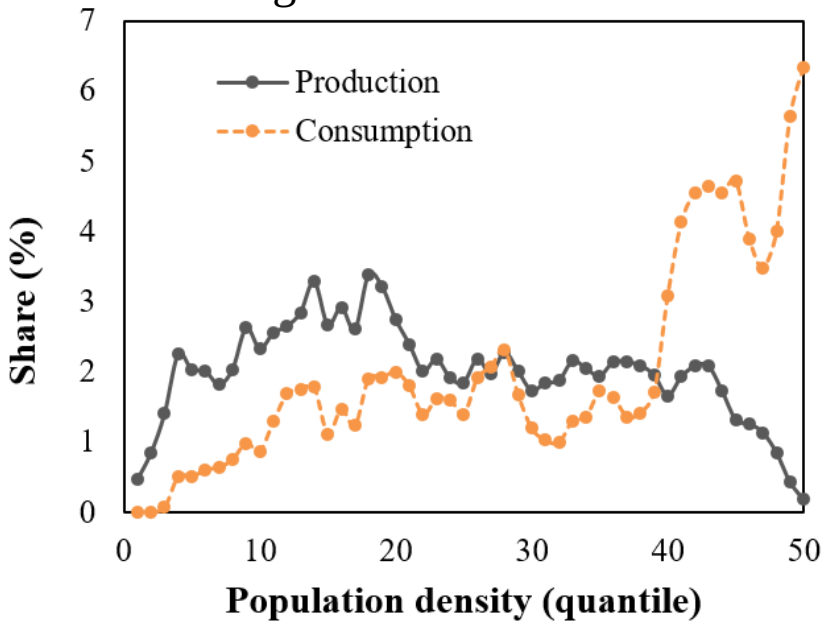
Flour milling and malt manufacturing



Spatial supply chain

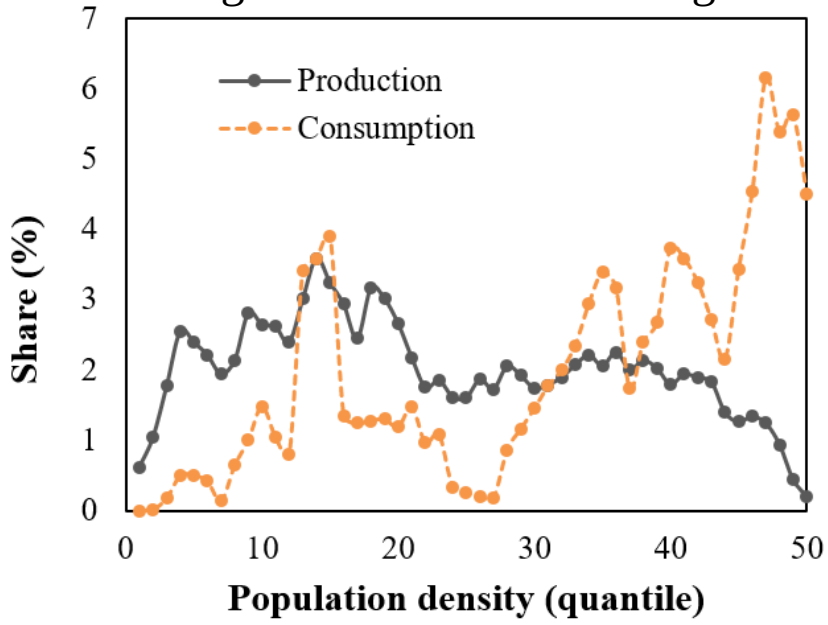
From Grain farming to…

Other basic organic chemical manufacturing



%	Rural	Suburban	Urban	Production
Rural	11.6	5.9	16.1	33.7
Suburban	6.1	7.5	22.9	36.6
Urban	3.8	4.9	21.1	29.7
Consumption	21.5	18.4	60.1	100

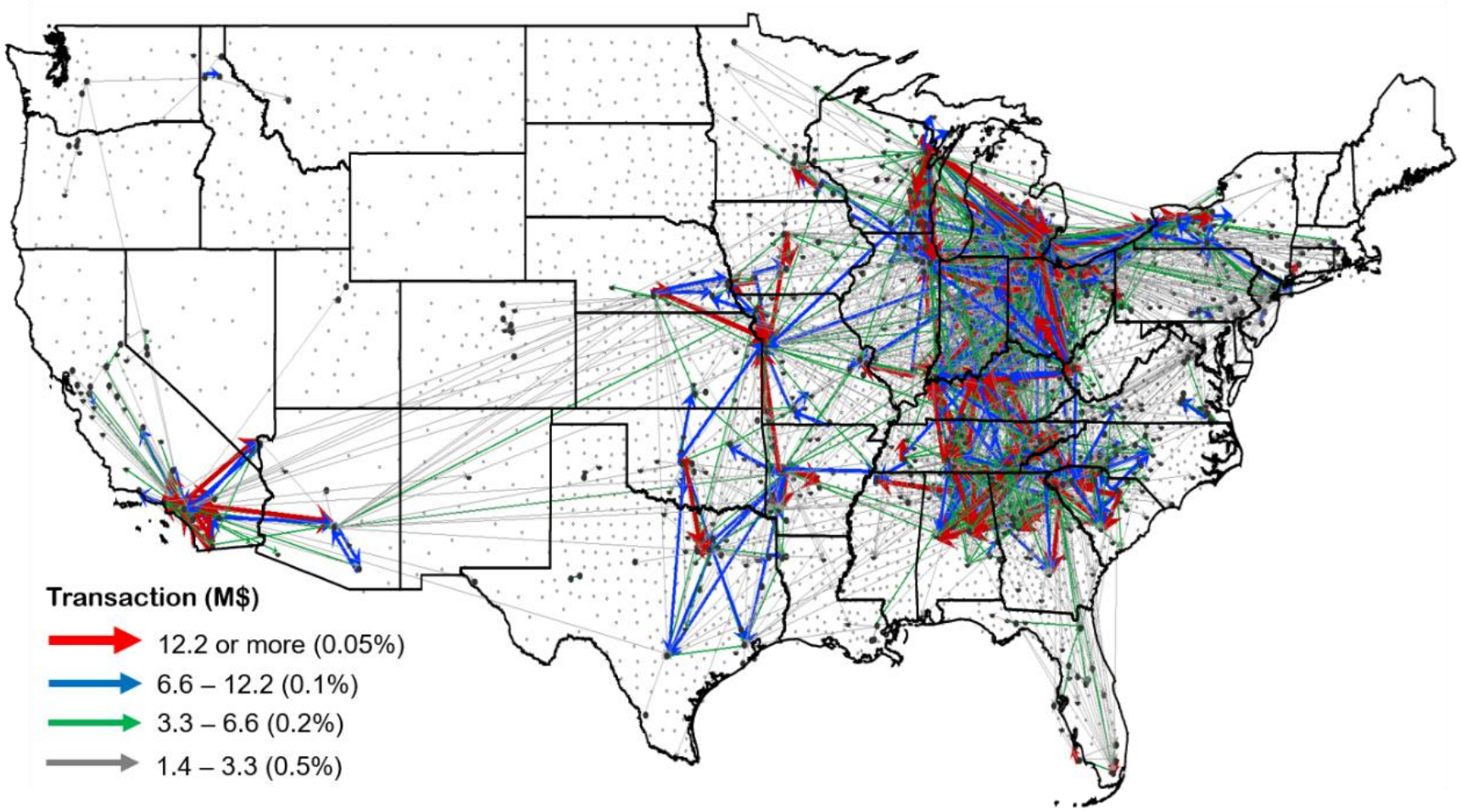
Flour milling and malt manufacturing



%	Rural	Suburban	Urban	Production
Rural	5.2	8.6	18.1	31.8
Suburban	2.8	10.8	23.6	37.3
Urban	1.7	6.5	22.7	30.9
Consumption	9.7	25.9	64.4	100

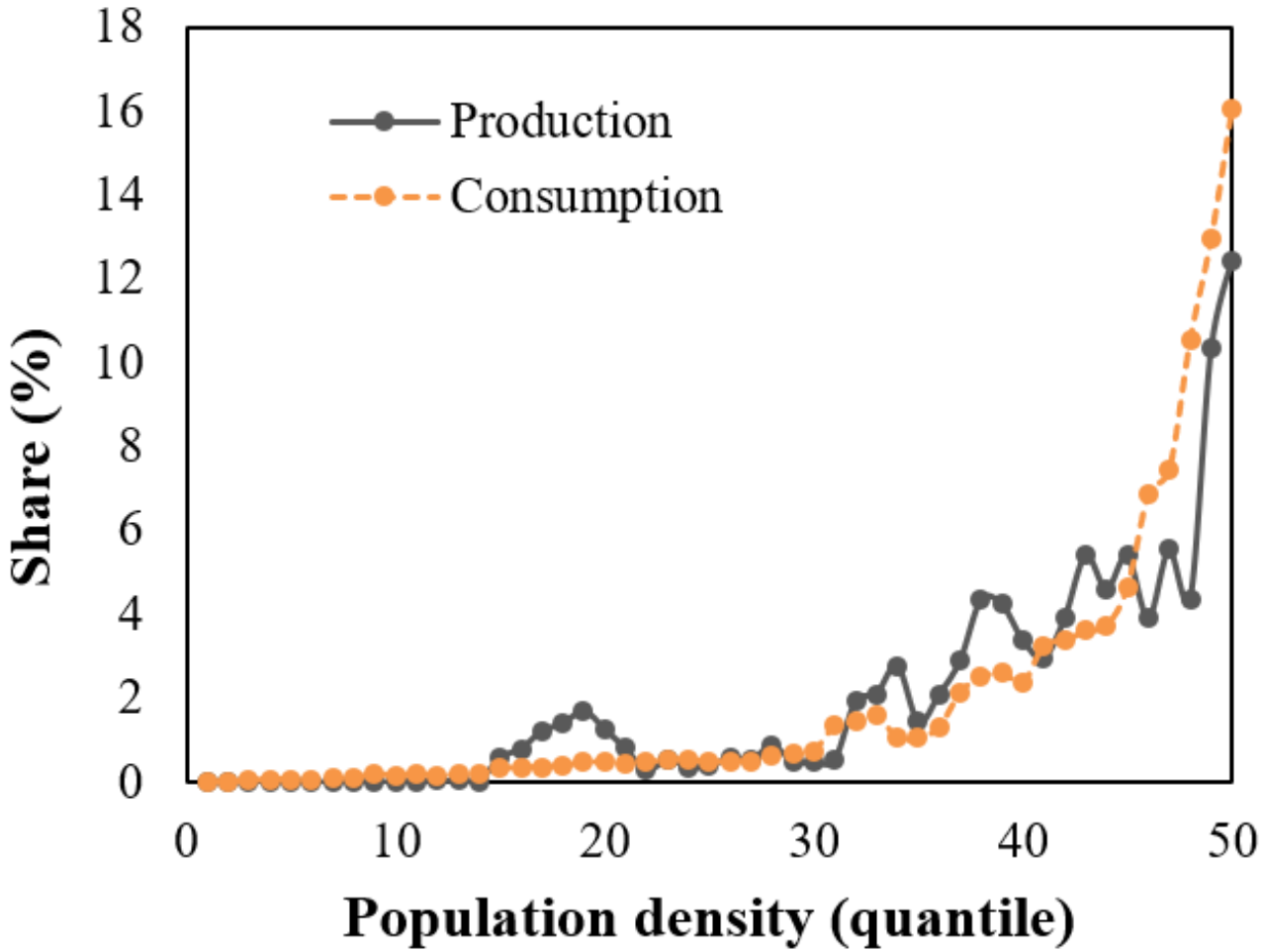
Spatial supply chain

Motor vehicle gasoline engine and engine parts mfg



Production-consumption

Motor vehicle gasoline engine and engine parts mfg

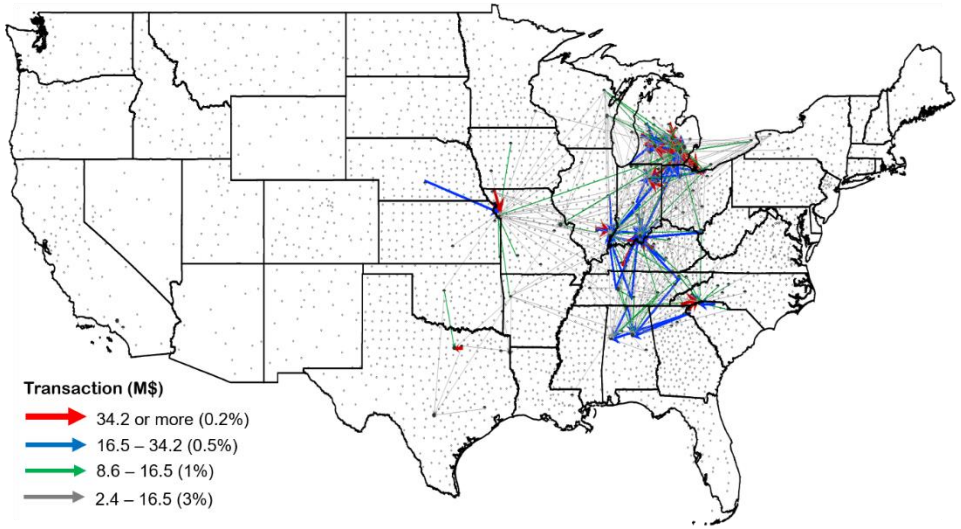


Discrepancies between production and consumption create rural-urban linkages

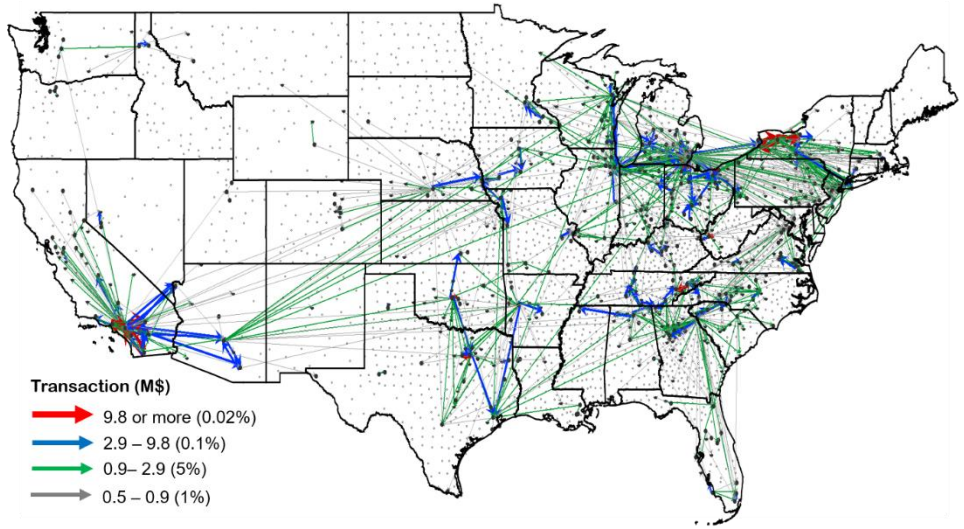
Spatial supply chain

From Motor vehicle gasoline engine to

Light truck and utility vehicle manufacturing



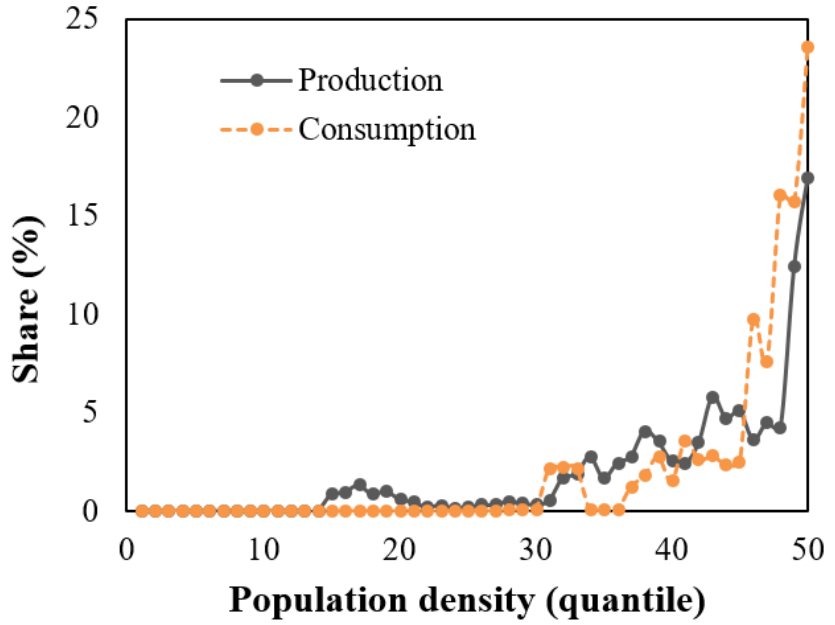
Automotive repair and maintenance



Spatial supply chain

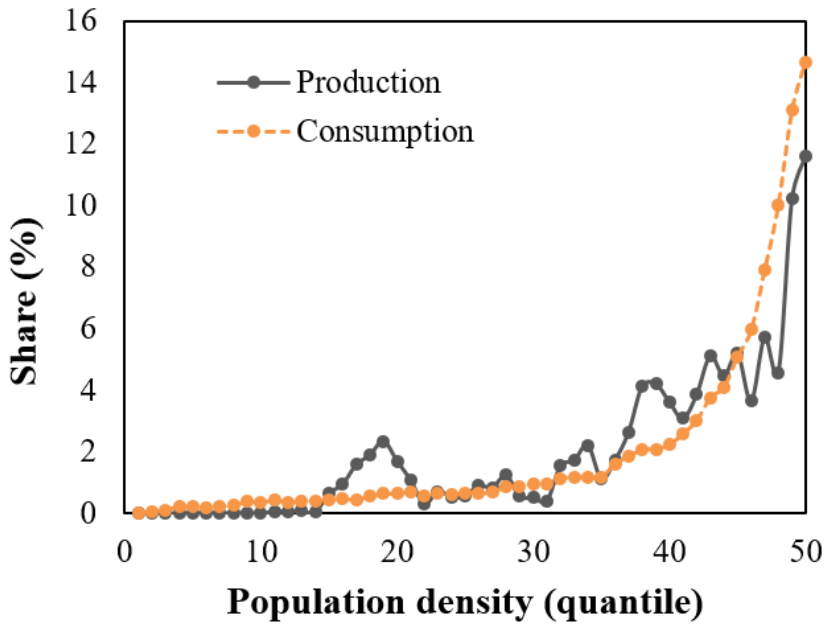
From Motor vehicle gasoline engine to

Light truck and utility vehicle manufacturing



%	Rural	Suburban	Urban	Production
Rural	0.001	1.8	2.8	4.6
Suburban	0.004	3.3	19.0	22.3
Urban	0.003	5.3	67.8	73.1
Consumption	0.008	10.4	89.6	100

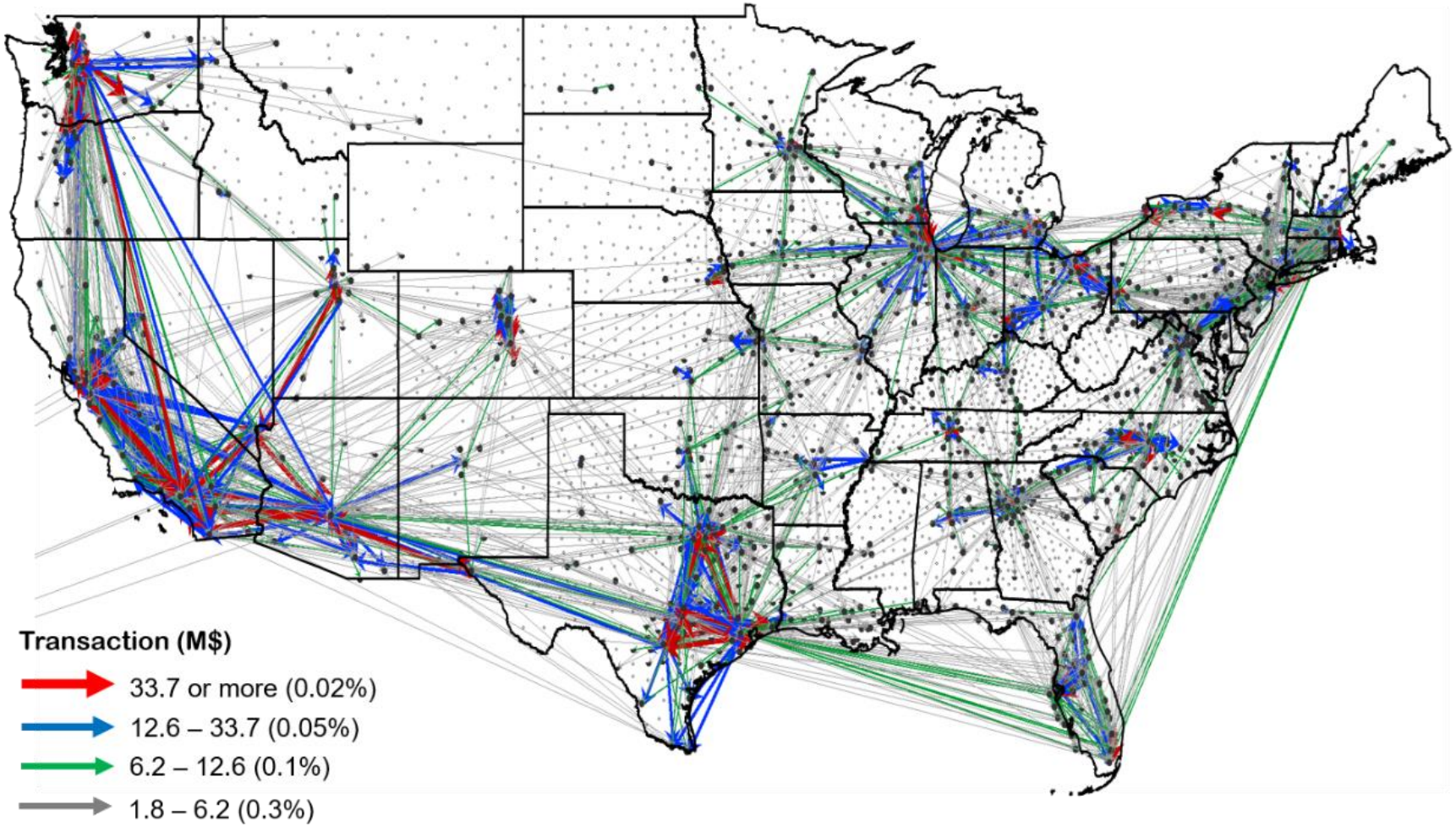
Automotive repair and maintenance



%	Rural	Suburban	Urban	Production
Rural	0.8	0.7	3.3	4.7
Suburban	3.4	4.4	19.9	27.7
Urban	3.3	6.6	57.7	67.6
Consumption	7.5	11.6	80.8	100

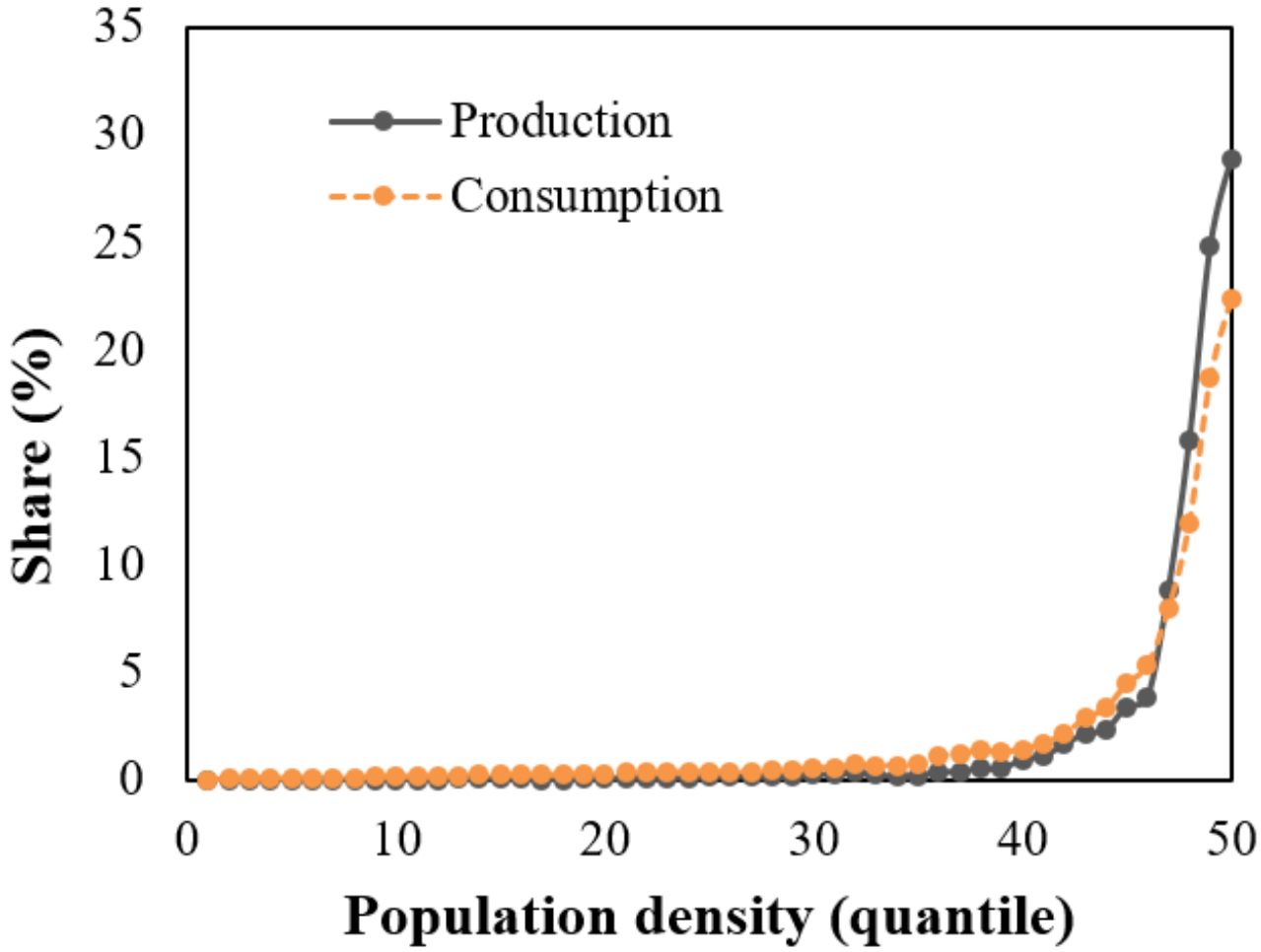
Spatial supply chain

Data processing, hosting, and related services



Production-consumption

Data processing, hosting, and related services

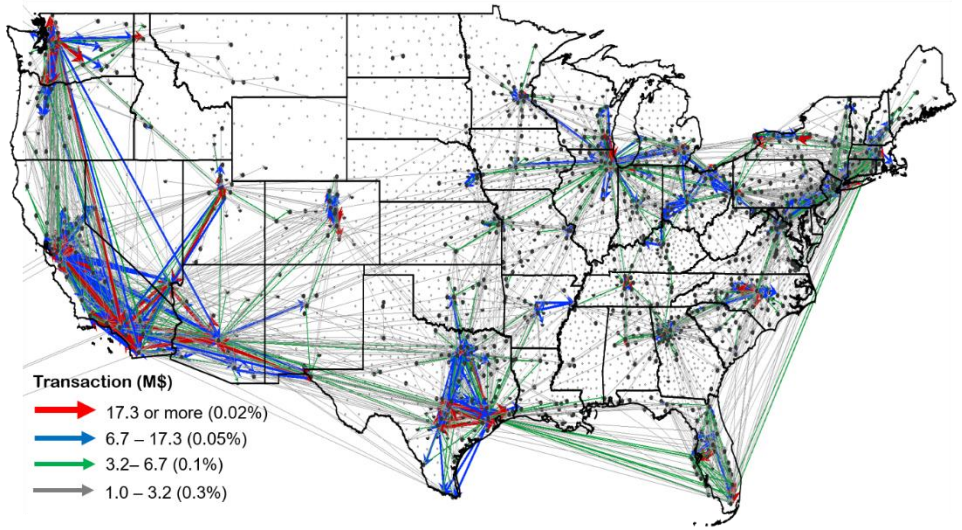


Discrepancies between production and consumption create rural-urban linkages

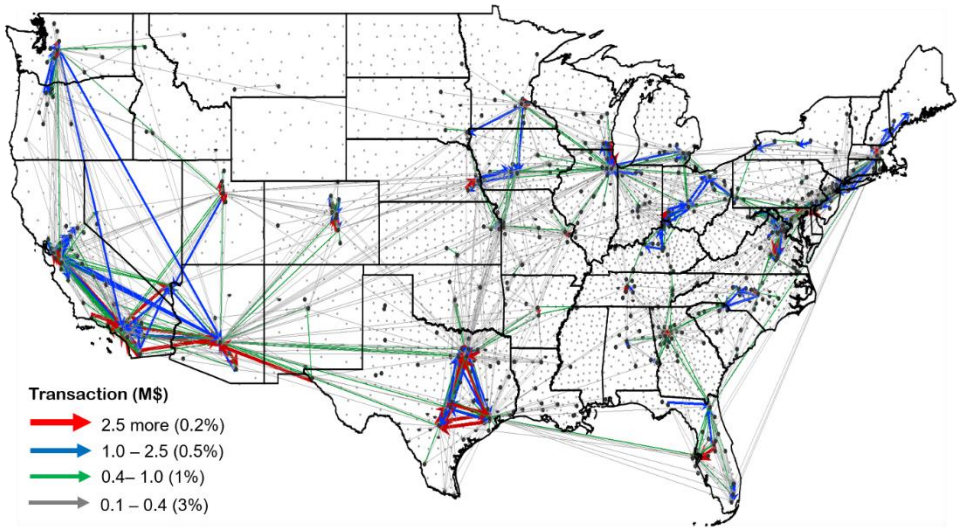
Spatial supply chain

From Data processing, hosting, and related services to

Personal consumption expenditures



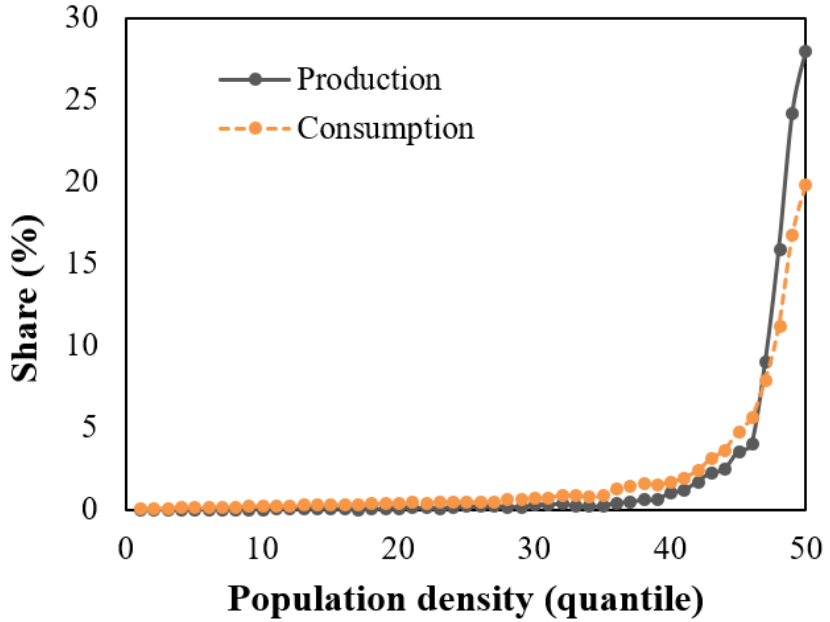
Nondepository credit intermediation activities



Spatial supply chain

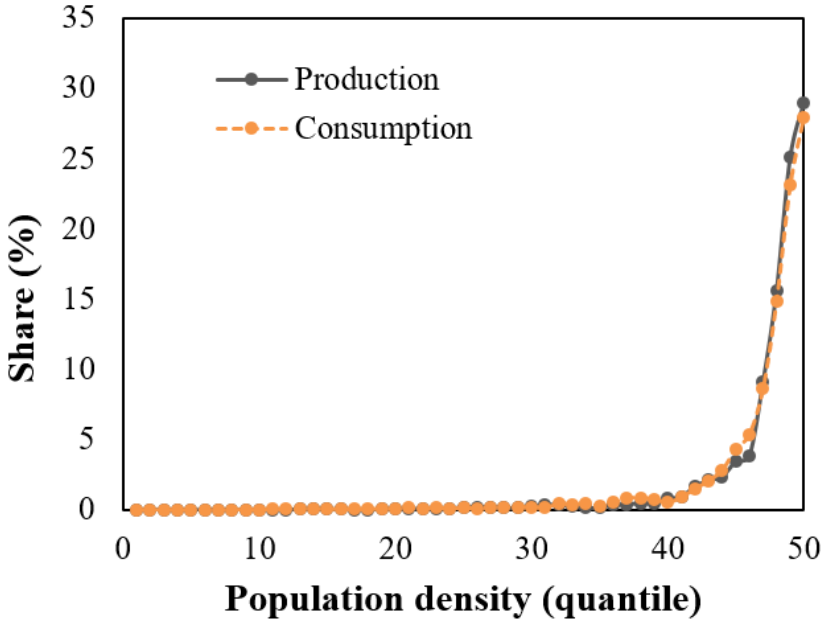
From Data processing, hosting, and related services to

Personal consumption expenditures



%	Rural	Suburban	Urban	Production
Rural	0.3	0.3	1.1	1.6
Suburban	0.2	0.4	1.4	1.9
Urban	3.5	7.4	85.5	96.5
Consumption	4.0	8.0	87.9	100

Nondepository credit intermediation activities



%	Rural	Suburban	Urban	Production
Rural	0.2	0.0	1.0	1.2
Suburban	0.0	0.1	1.4	1.4
Urban	0.9	1.8	94.7	97.4
Consumption	1.0	1.9	97.1	100

County Commuting and Migration Networks

Current definitions of LMAs

LMAs are mutually exclusive.
A county can belong to only one LMA.

Large cities (DC, NYC) may belong to multiple LMAs.

But: A county in New Jersey may send commuters to DC, NYC *and* Philadelphia

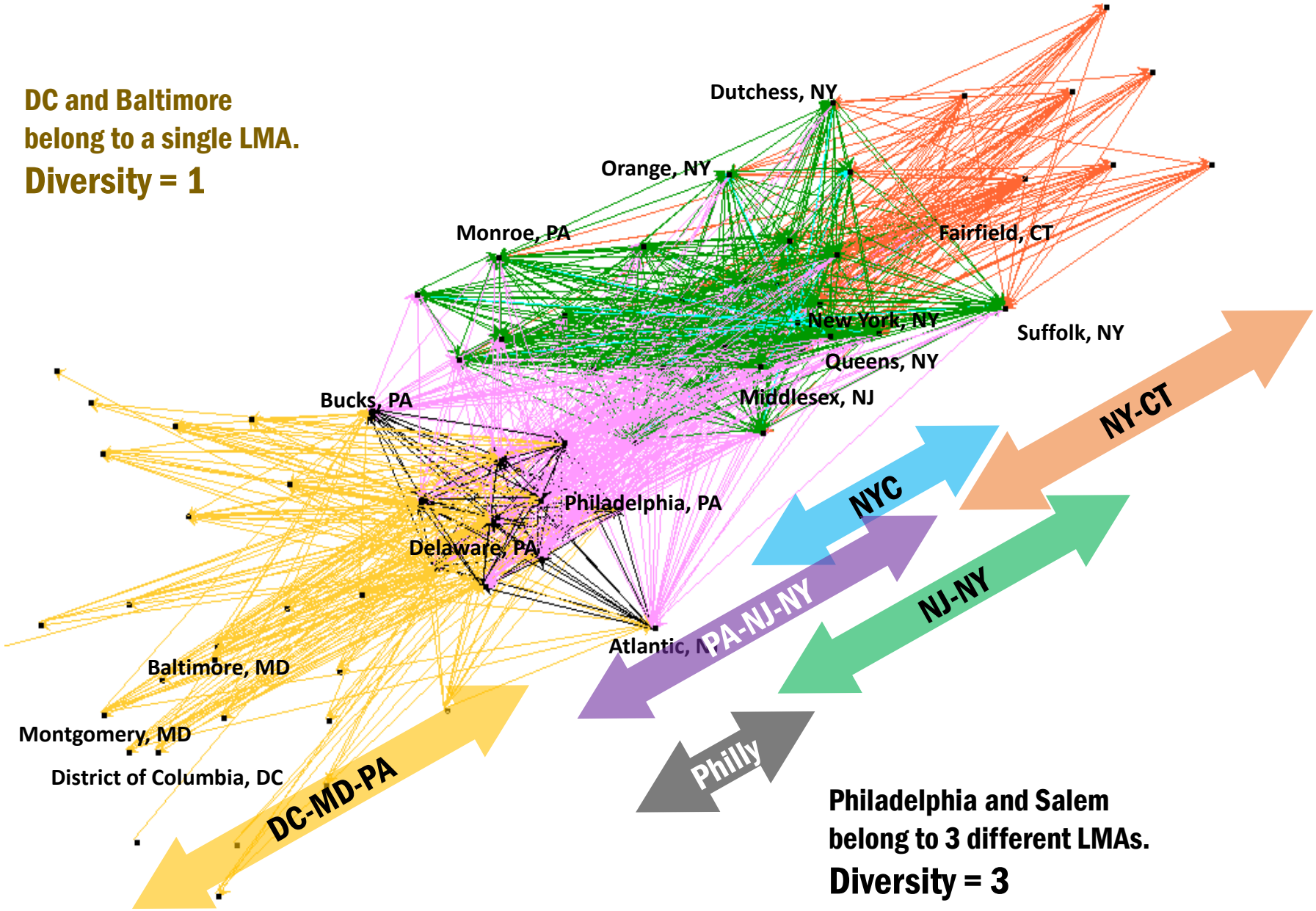
Labor market areas for the US, 1990.
USDA ERS Commuting Zones and Labor Market Areas

LMA Diversity

- Consider that labor market areas may overlap
- We calculate the number of different LMAs to which a county may belong, using commuting data and the link community method of Ahn et al. (2010)
  - Membership in more LMAs would provide more diverse economic opportunities, more stability?
  - Diversity = the number of LMAs to which a county belongs
  - Vector-similarity (or diversity) calculated from the 3,151 x 3,151 US county matrix

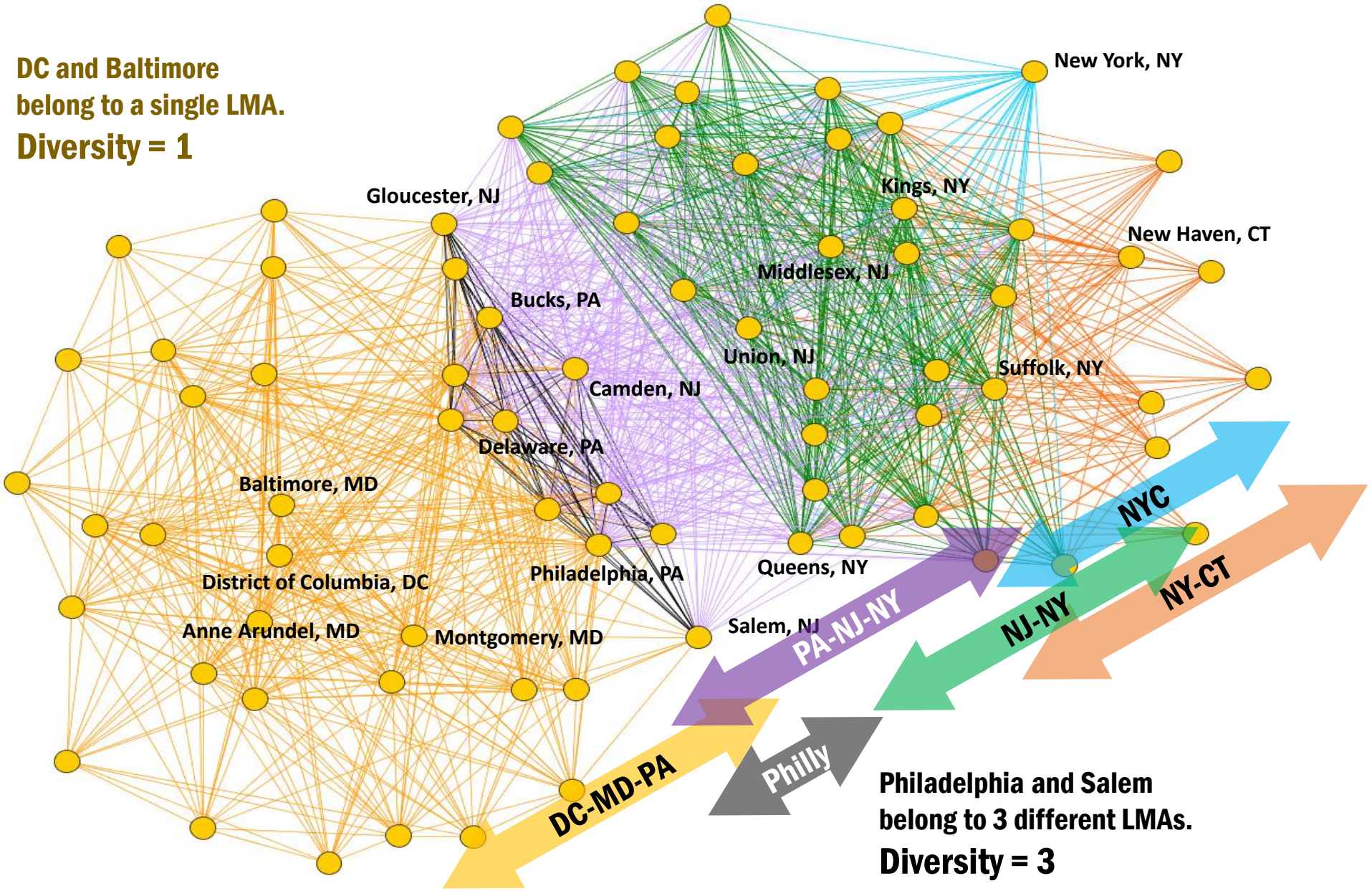
Y.-Y. Ahn, J.P. Bagrow, S. Lehmann. 2010. Link Community reveal multi-scale complexity in networks. *Nature* 446: 761-764 Supp. p.

Overlapping commuting zones, NE (geographical representation)



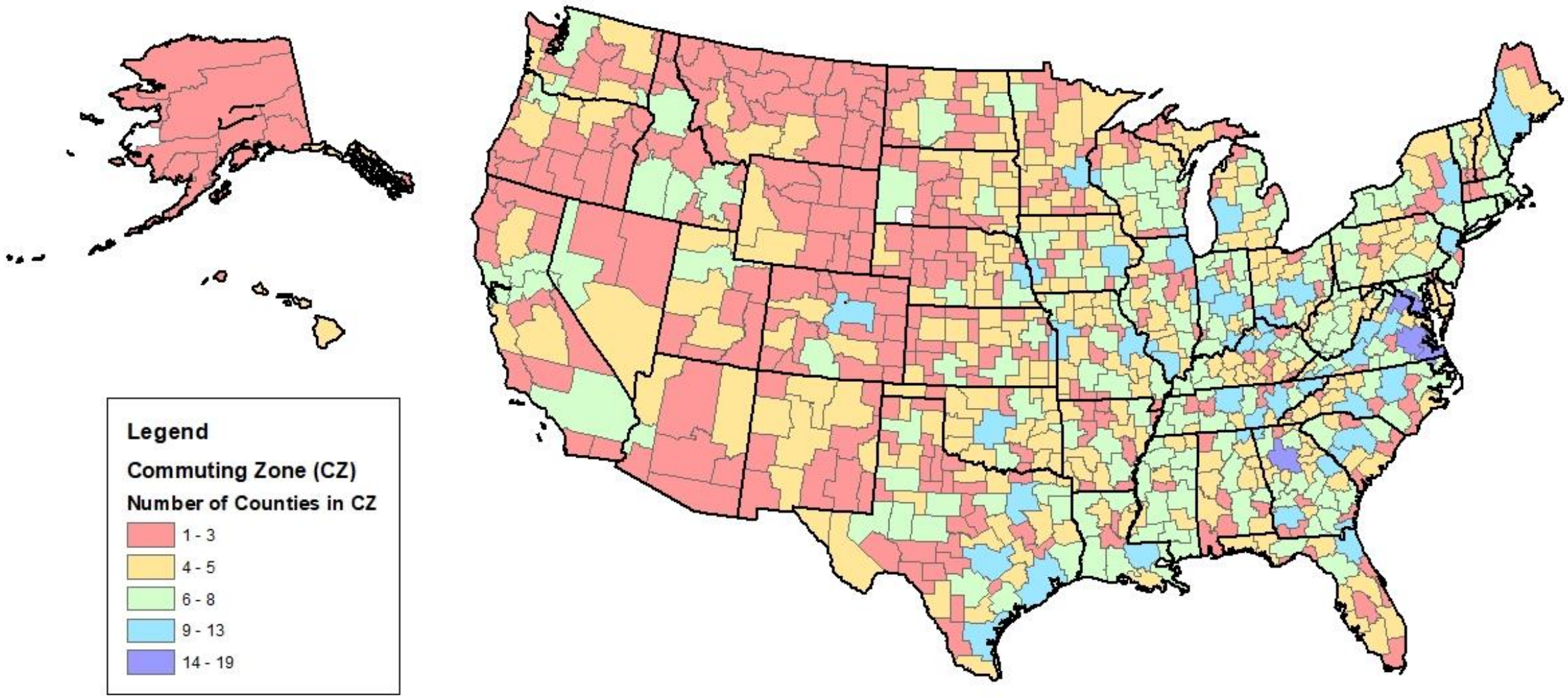
Y. Han, S.J. Goetz (2019) "Overlapping labour market areas based on link communities" *Papers in Regional Science* 98 (1), 539-553

Overlapping commuting zones, NE (topological representation)



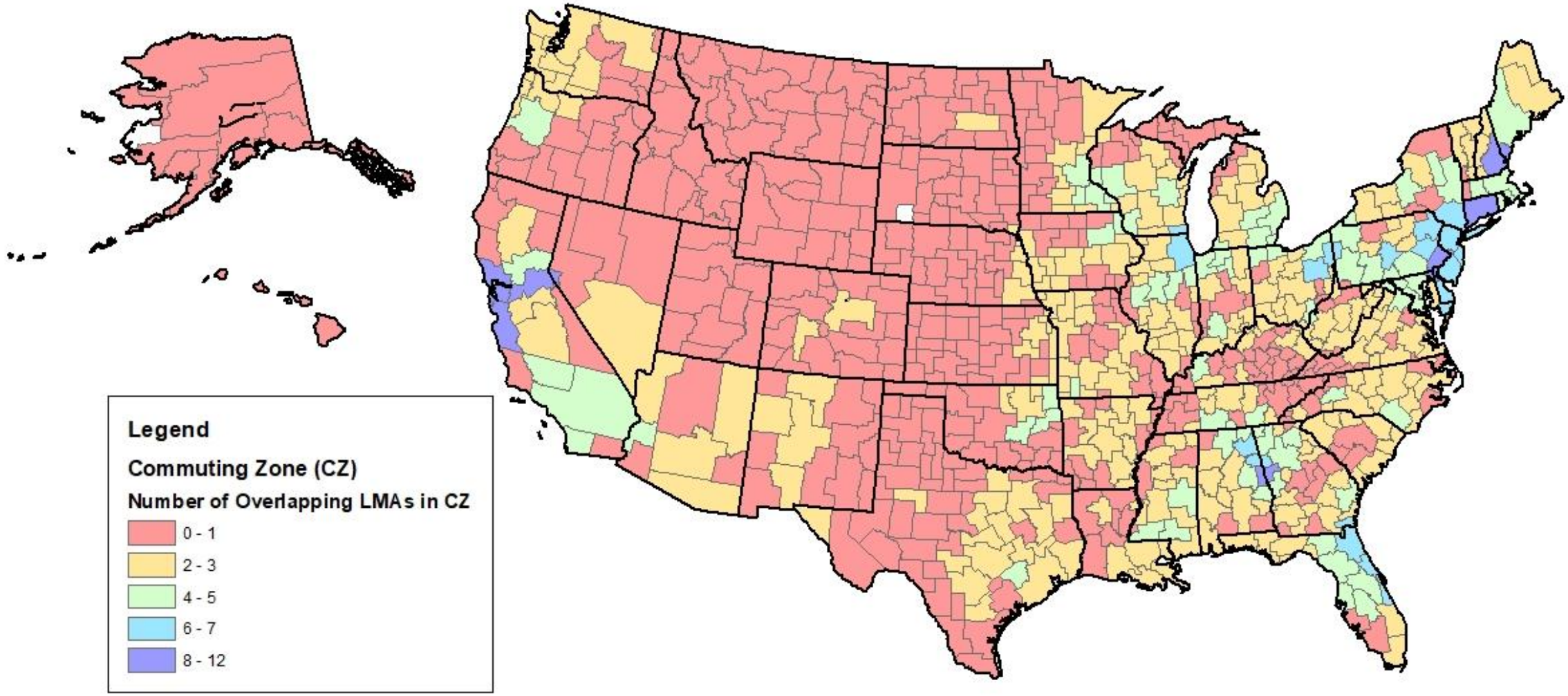
Y. Han, S.J. Goetz (2019) "Overlapping labour market areas based on link communities" *Papers in Regional Science* 98 (1), 539-553

Number of counties per commuting zone (Han & Goetz, 2019, *PIRS*)



Map: D. Meadowcroft

Number of Overlapping LMA's in each CZ (Han & Goetz, 2019, *PIRS*)



Map: D. Meadowcroft

CZ-level Models, Innovation Index as Outcome Variable

Dependent variable: Innovation Index from:

B.E. Whitacre, D. Meadowcroft & R. Gallardo (2019) *"Firm and regional economic outcomes associated with a new, broad measure of business innovation,"* *& Regional Development* Vol. 31, Issue 9-10: *930-952*

	(1)	(2)	(3)	(4)
Number of Overlapping LMA's	0.004***	0.004***		0.002**
Total Land Area		7.490e-5**	4.530e-5	5.800e-5*
Population Density		5.150e-5***	3.320e-5***	2.780e-5***
Perc. of 20 to 40 Year Old's			-0.108	-0.127
Perc. of 40 to 60 Year Old's			-0.034	-0.090
Percentage White			0.018	0.021*
Perc. with Bachelor's Degree			-0.104	-0.069
Perc. with Graduate Degree			0.299**	0.284*
Poverty Rate			-0.120**	-0.116**
Perc. Creative Class			0.154	0.114
Constant	0.378***	0.375***	0.387***	0.403***
Number of Observations	586	586	586	586
R^2	0.027	0.041	0.119	0.125

*Note: *, **, and *** represents statistical significance at the 10%, 5%, and 1% levels, respectively*

Regression result: Devon Meadowcroft

Overlapping Migration and Commuting Networks

- **Nodes:** Counties
- **Links:** Flows of people
- *Motivation* differs: migrate vs. commute
  - Migration: employment vs. retirement vs. other
  - Commuting: access employment
- Assume: migrants adopt same commuting behavior as residents

Networks can be *overlapped*

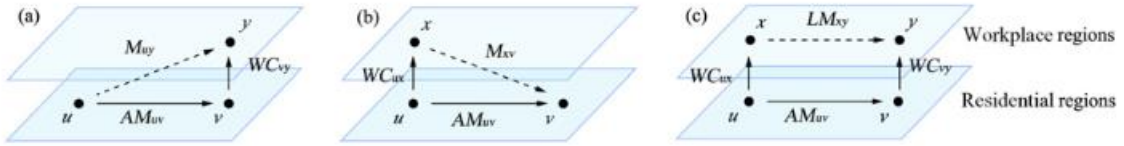


FIGURE 2. CONCEPT OF LATENT MIGRATION.
Note: Letters *u* and *v* refer to places of residence, and *x* and *y* are places of work.

Han, Y., Goetz S.J., Kim, T. and Lee J. (2013) "Estimating Employment-Related Migration from Overlapping Migration and Commuting Networks" *Growth and Change* (44) 3: 474–493

Overlapping Migration and Commuting Networks

We assume in-migrants have the same commuting patterns as existing residents

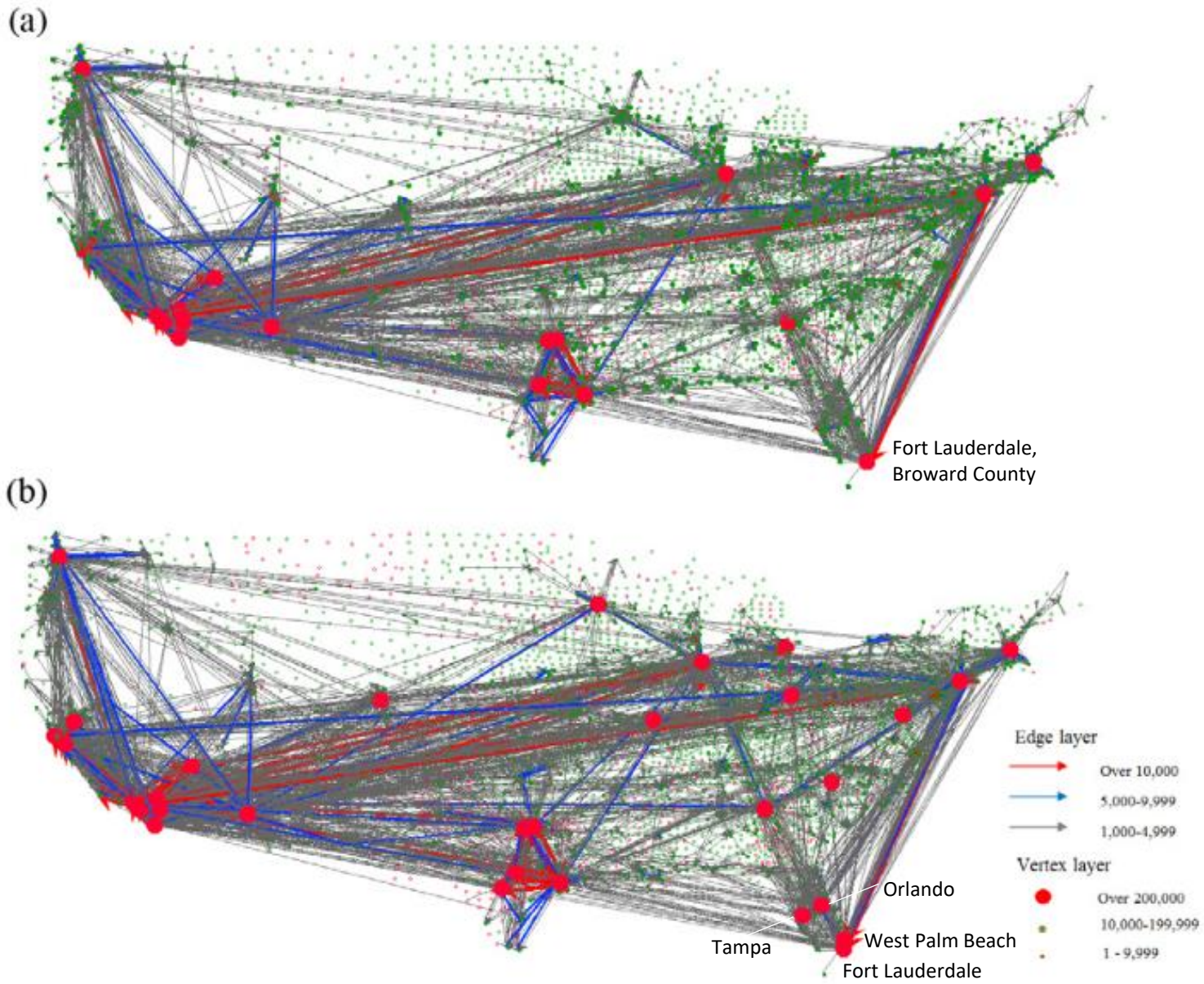


FIGURE 3. ILLUSTRATION OF THE U.S. (A) ACTUAL AND (B) LATENT MIGRATION NETWORK IN 2000 (BEST VIEWED IN COLOR).

Han, Y., Goetz S.J., Kim, T. and Lee J. (2013) "Estimating Employment-Related Migration from Overlapping Migration and Commuting Networks" *Growth and Change* (44) 3: 474–493

Spatial Networks Based on Communications Links

Network Diversity and Economic Development

Nathan Eagle, *et al.* *Science* **328**, 1029 (2010)

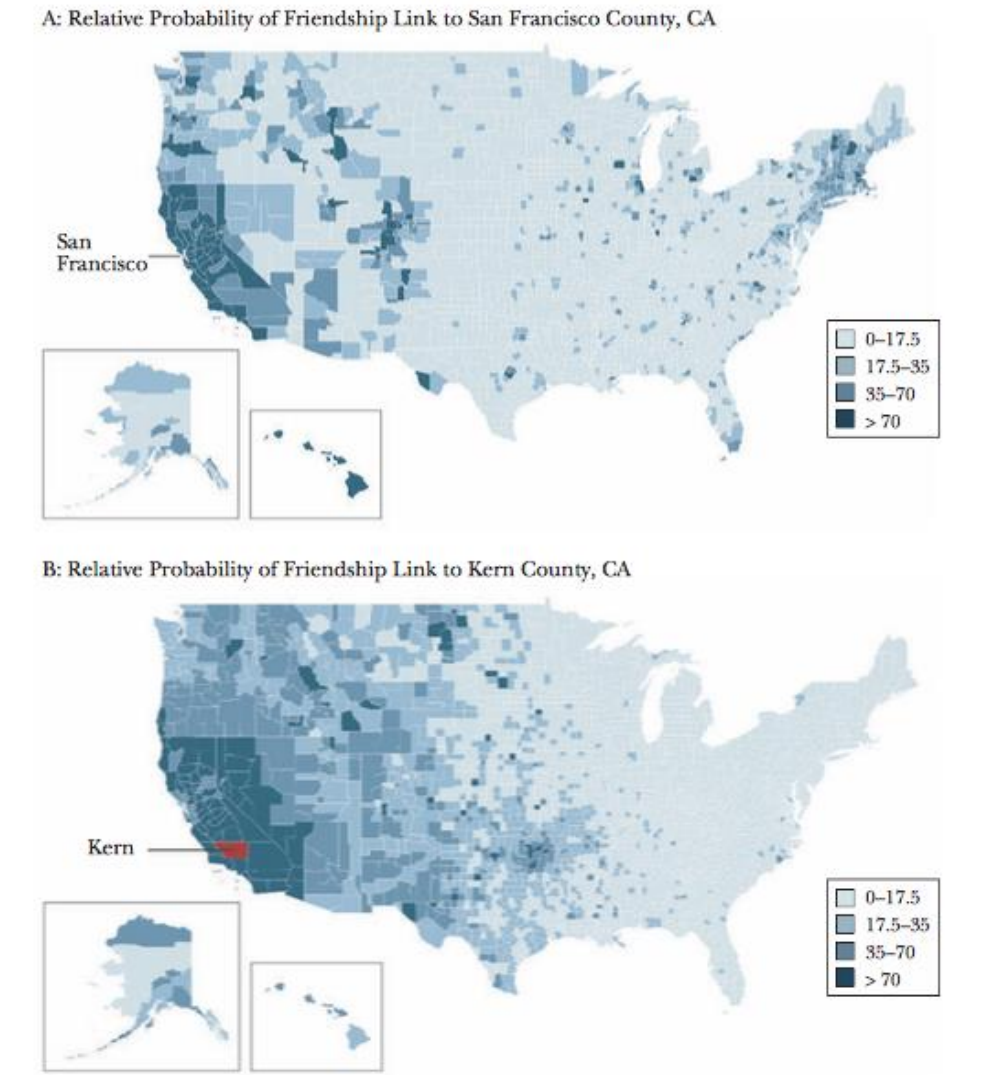
Fig. 1. An image of regional communication diversity and socioeconomic ranking for the UK. We find that communities with diverse communication patterns tend to rank higher (represented from light blue to dark blue) than the regions with more insular communication. This result implies that communication diversity is a key indicator of an economically healthy community. [(*29*) Crown copyright material is reproduced with the permission of the Controller of Her Majesty's Stationery Office]

Facebook-based networks

Bailey* et al. (2018), "Social Connectedness: Measurement, Determinants and Effects," *J. Econ. Perspectives*

*Senior Economic Research Scientist at Facebook, Menlo Park, CA

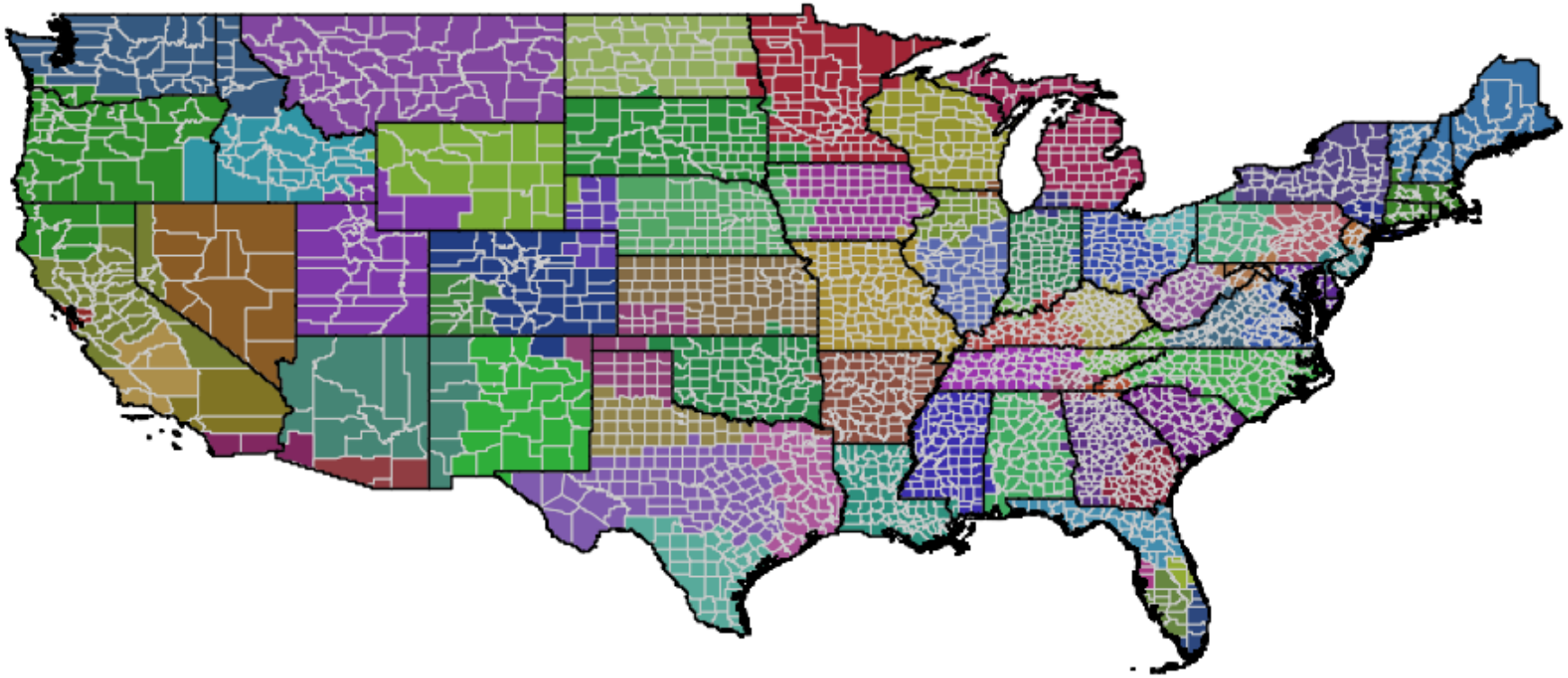
Figure 1
County-Level Friendship Maps



Note: The heat maps show the relative probability that a Facebook user in each county *j* has a friendship link to San Francisco County, CA (Panel A) and Kern County, CA (Panel B). Darker colors correspond to counties in which there is a higher probability of a friendship link between a person in home county *i* (San Francisco or Kern) and county *j*. The "relative probability of friendship" is constructed by taking the Social Connectedness Index between counties *i* and *j* and dividing it by the product of the number of Facebook users in the two counties.

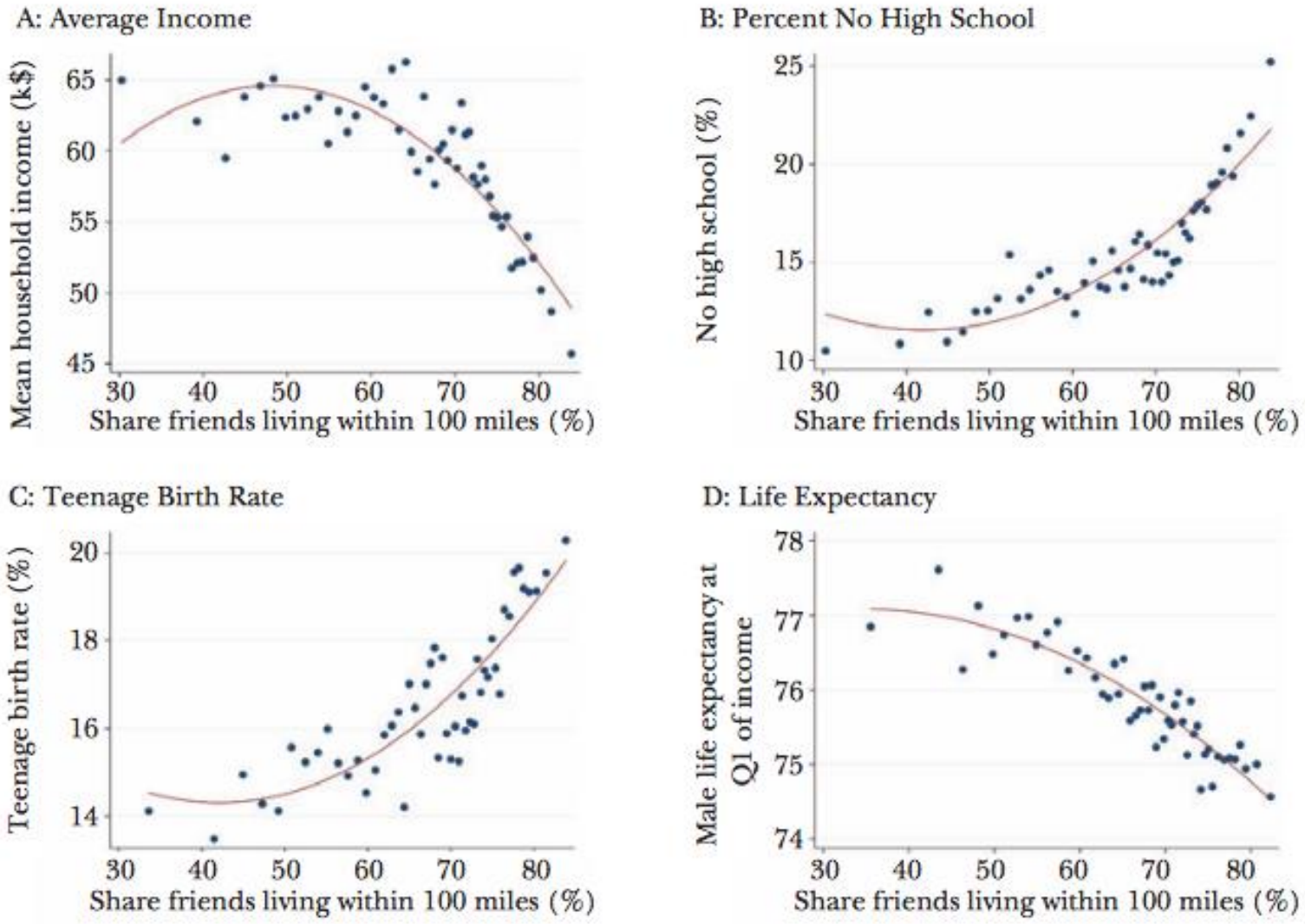
75 Regional County Clusters, the U.S.

Figure A5: Connected Communities within the United States - 75 Distinct Units



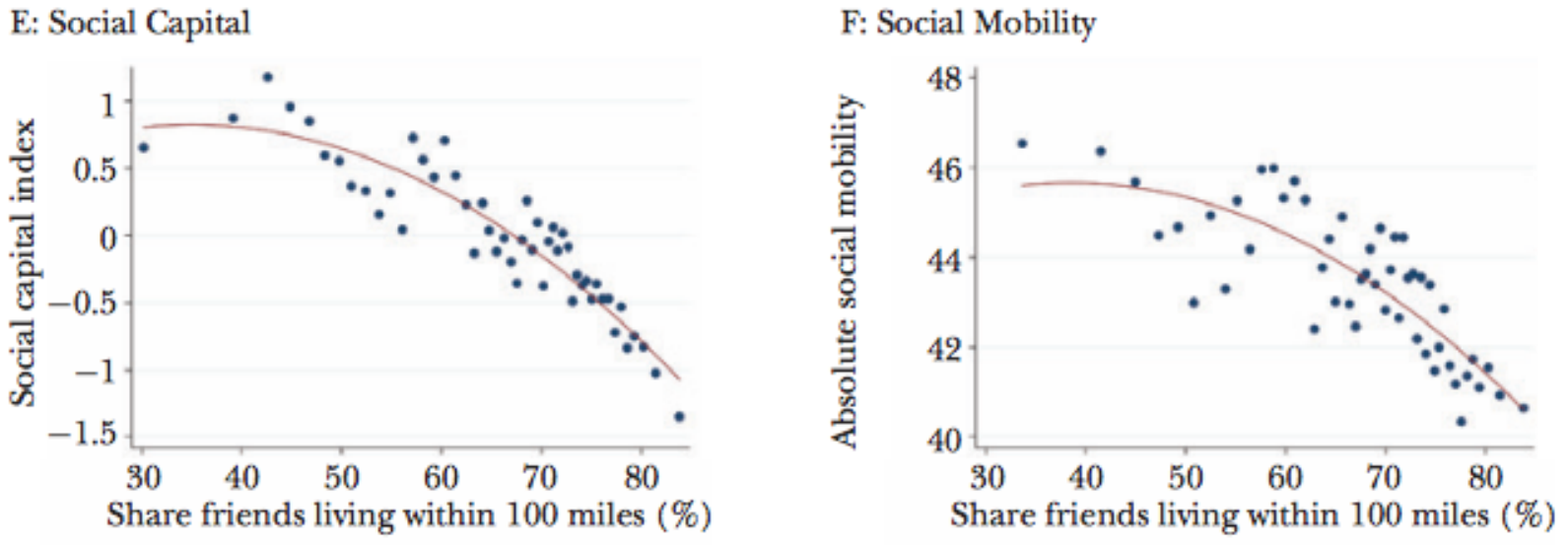
Note: Figure shows US counties grouped together when we use hierarchical agglomerative linkage clustering to create 75 distinct groups of counties. The algorithm assigns both Hawaii and Alaska, not pictured, to two distinct clusters including only the respective state.

About the Strength of Weak Ties…



"Social Connectedness: Measurement, Determinants and Effects," Bailey* et al. (2018), *J. Econ. Perspectives*

*Senior Economic Research Scientist at Facebook, Menlo Park, CA

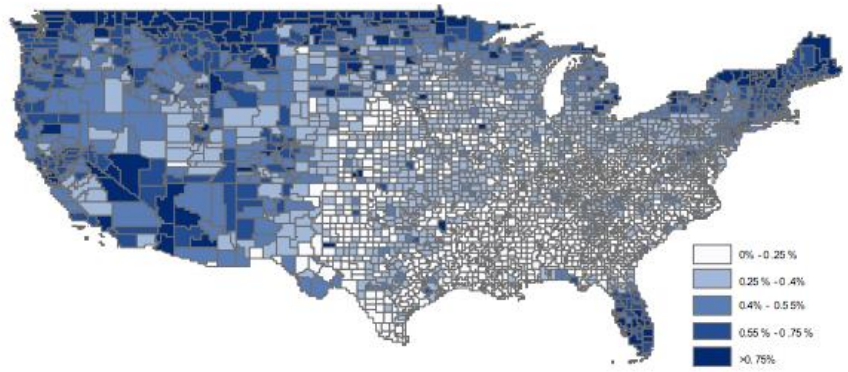


Notes: Panels show binned scatterplots with counties as the unit of observation. To generate each binned scatterplot, we group the x-axis variable into 50 equal-sized bins. We then compute the mean of the x-axis and y-axis variables within each bin and create a scatterplot of these 50 data points. The horizontal axes measure the share of friends of the county population that live within 100 miles. On the vertical axes are a number of county-level measures of socioeconomic outcomes: the mean county income in Panel A; the share of the population with no high school degree in Panel B; the teenage birth rate as provided by Chetty, Hendren, Kline, and Saez (2014) in Panel C; the life expectancy of males in the first quarter of the national income distribution from Chetty et al. (2016) in Panel D; the measure of social capital in 2009 as defined by Rupasingha, Goetz, and Freshwater (2006) in Panel E; and the absolute measure of social mobility from Chetty et al. (2014) in Panel F. The red line shows the fit of a quadratic regression. The online Appendix (http://e-jep.org) provides more details.

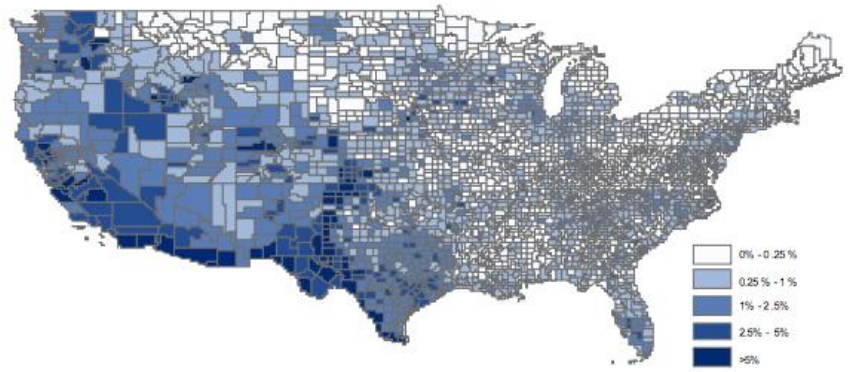
"Social Connectedness: Measurement, Determinants and Effects," Bailey* et al. (2018), *J. Econ. Perspectives*

Figure A31: International Social Connectedness

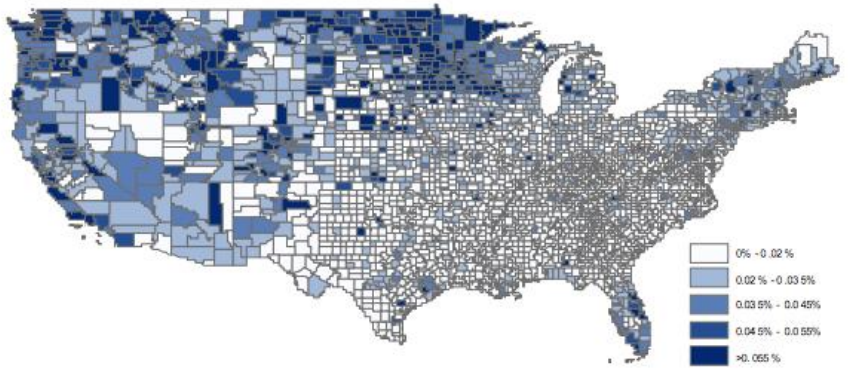
(A) Canada



(B) Mexico



(C) Norway



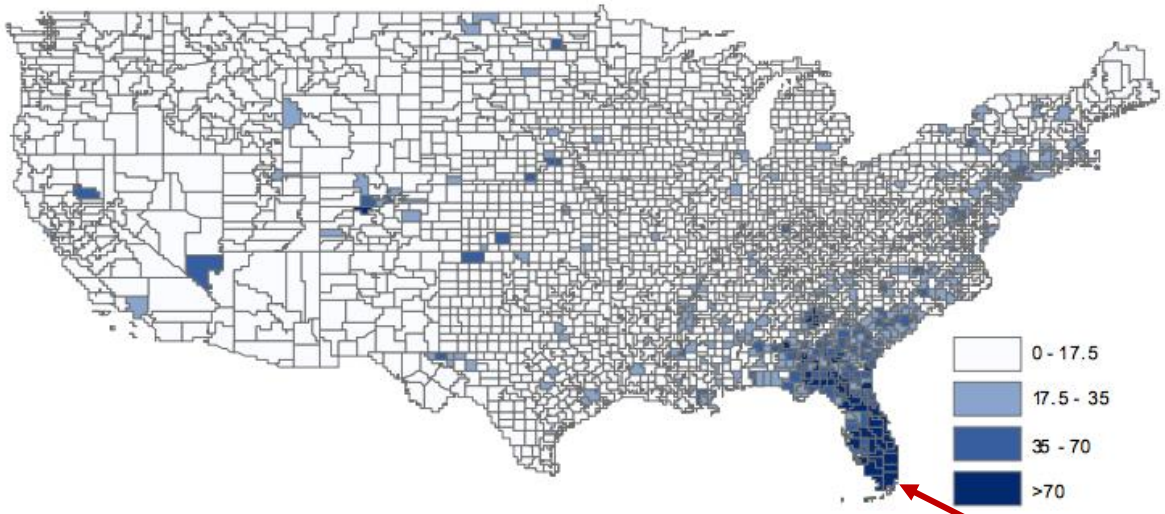
Bailey* et al. (2018), "Social Connectedness: Measurement, Determinants and Effects," *J. Econ. Perspectives*

RE: Overlapping Migration and Commuting Networks

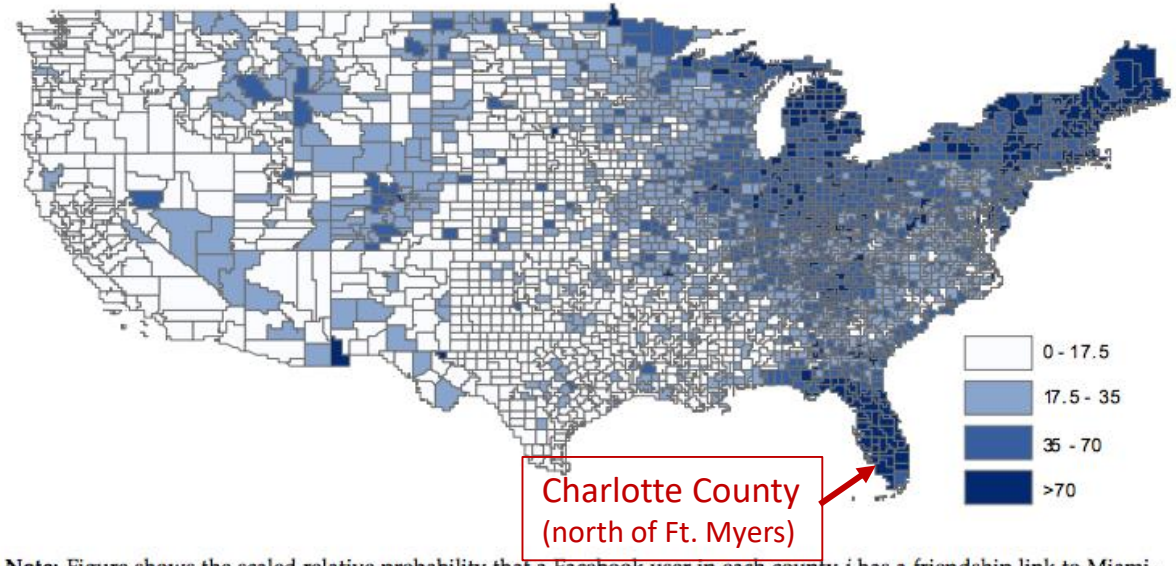
Bailey* et al. (2018), "Social Connectedness: Measurement, Determinants and Effects," *J. Econ. Perspectives*

Figure A16: Florida Retirement Communities and Friendship Network Distribution

(A) Relative Probability of Friendship Link to Miami-Dade County, FL ($RelativeProbFriendship_{i,j}$)



(B) Relative Probability of Friendship Link to Charlotte County, FL ($RelativeProbFriendship_{i,j}$)



Note: Figure shows the scaled relative probability that a Facebook user in each county j has a friendship link to Miami-Dade County, FL in Panel A, and Charlotte County, FL in Panel B. It is constructed as in equation A1. Darker colors correspond to counties in which there is a higher probability of a friendship link between a person in home county i (Miami-Dade or Charlotte) and county j.

In Conclusion…

In Conclusion…

THANK YOU

sgoetz@psu.edu

Conclusion

Spatial linkages

- Inconsistency between production and consumption areas

Modeling spatial linkages between local industries

- Using production and consumption in local economies
- Attachment method in networks and gravity rule
- Apply the model to the U.S. economy

Contribution

- Basic data for understanding the National economic system and the environment of local economies
- We can apply this model to estimate the economic innovation, diversity, resilience, vulnerability
- and to predict the change in local economies from shocks

Outline; acknowledgements!

- Why network analysis?
- Network (very) basics
- “Network*” in Regional Science Journals paper titles
- Old tool, new insights: input-output table as a network
  - Economic rewiring (*PLoS One*)
  - Resilience (*Appl. Econ.*)
  - Latent Local Innovation
  - Visualizing Spatial Supply Chains (ERSA presentation)
- Labor market adjustments
- Commuting, migration networks
  - Employment-based migration
  - Overlapping
- Social media-based spatial networks
  - Twitter, entrepreneurs
  - Facebook, entrepreneurs: value of county connectedness; *JEP* paper

CZs/LMAs in the Northeast region

Commuting data of n=75 Northeast counties are used;
we detect 6 LMAs at the maximum density (=threshold)