Nation-scale social networks

..

KOR data presentation - May 20, 2025

Andreas Bjerre-Nielsen

DANMARKS STATISTIK

UNIVERSITY OF COPENHAGEN

DTU

Social relations @ scale

- State-of-the-art approaches
  - Surveys
    - Detailed but expensive and likely biased
  - Social media
    - Allows new way of tracking social relations at scale, e.g., by Meta on their platforms

- Problems with SoME plaforms
  - Participation is selective
  - Recent technology and we cannot track long changes in social fabric
  - Data is very shallow – have to predict/infer outcomes, e.g., income
  - Monopoly of data – cannot share data openly, and likely to pursue positive reputation projects

(generated using ChatGPT)

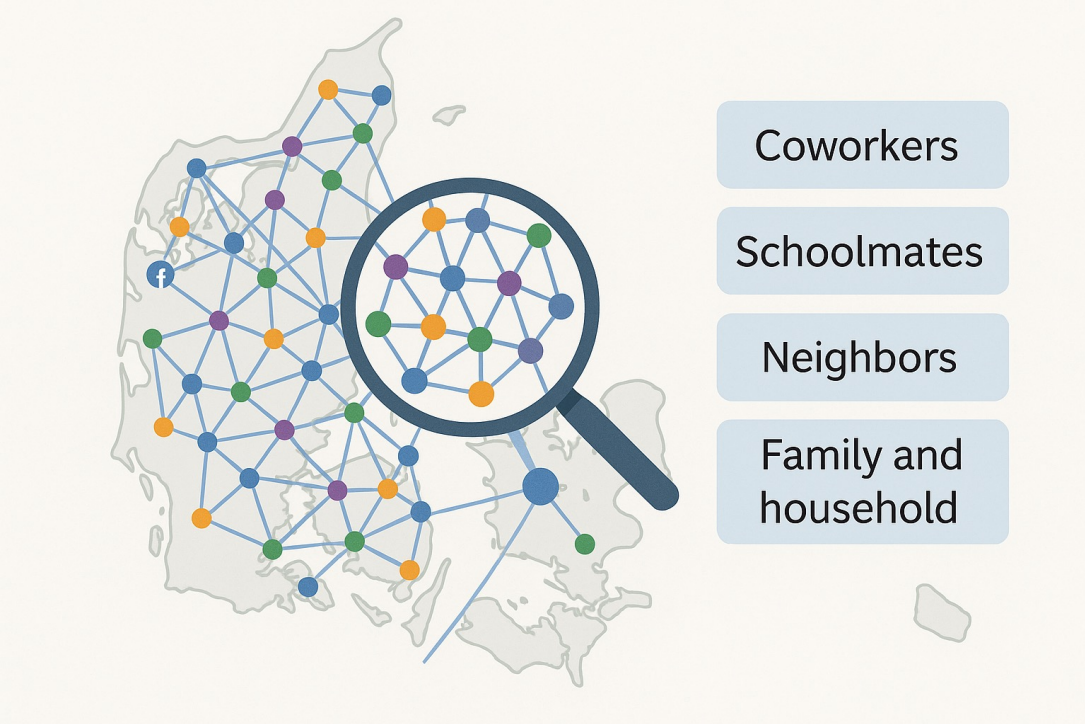
Unveiling the Social Fabric Through a Temporal, Nation-Scale Social Network and its Characteristics

Jolien Cremers[1,+], Benjamin Kohler[1,2,7,+], Benjamin Frank Maier[1,3,+], Stine Nymann Eriksen[1], Johanna Einsiedler[1,2], Frederik Kølby Christensen[1], Sune Lehmann[2,3], David Dreyer Lassen[2,4], Laust Hvas Mortensen[1,5,6], and Andreas Bjerre-Nielsen[2,4,*]

[1]Methods and Analysis, Statistics Denmark, Copenhagen, 2100, Denmark
[2]Center for Social Data Science (SODAS), University of Copenhagen, Copenhagen, 1353, Denmark
[3]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, 2800, Denmark
[4]Department of Economics, University of Copenhagen, Copenhagen, 1353, Denmark
[5]Department of Public Health, University of Copenhagen, Copenhagen, 1353, Denmark
[6]ROCKWOOL Foundation, Copenhagen, 1472, Denmark
[7]Center for Law and Economics, ETH Zurich, Haldeneggsteig 4, 8006, Switzerland
[*]abn@sodas.ku.dk
[+]These authors contributed equally to this work

A nation-scale social network

- Following Van der Laan et al. (2023)
  - Construct a **combined network** from registry data
  - Construct network for **each year** by **merging layers**



(generated using ChatGPT)

Danish registry network

- Neighbors
  - Sample of 10 households within 50 meters (if there are any)

- Colleague
  - Work in same place – using Danish population in November
  - For large firms - random subsample of 100

- School and classmate layer
  - Enrolled in same educ. institution

Structure?

- Family layer
  - first and second degree family relations (e.g., cousin, grand-parent)
  - basis of legal child-parent relationships (biological and adoptive)

- Household layer
  - same address (Stat. Denmark designation)

Dual representation

- ***Individual-centered***: relations between individuals
- ***Bipartite***: relationer via container nodes (e.g. schools)

  - *Employer-employee* <> *coworkers*
  - *School-student* <> *schoolmates* (*and classmates*)
  - *Neighborhood-residents* <> *neighbors*

Our contributions

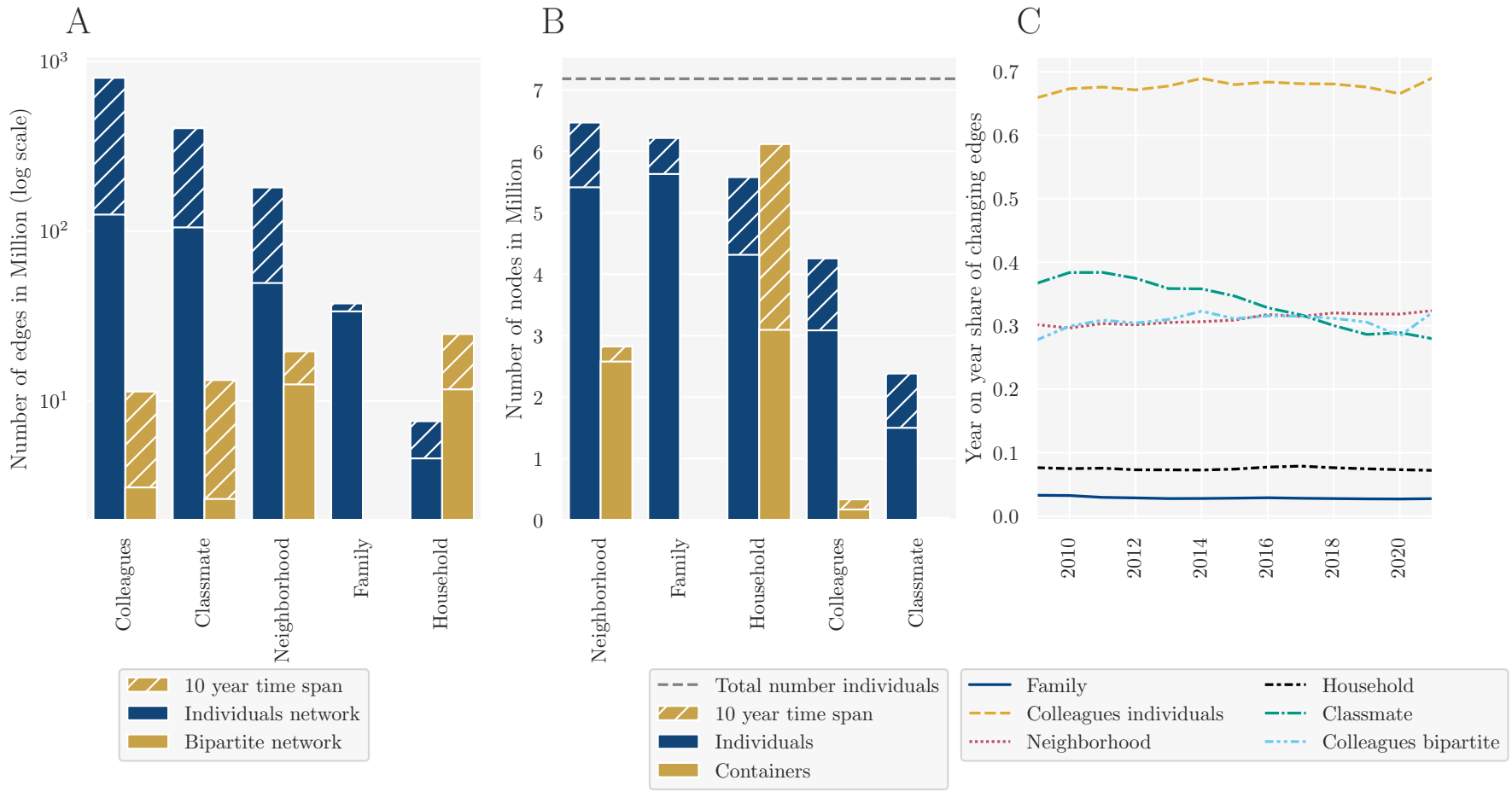
- **Leverage dual representation to**
  - Obtain computational improvements
  - Heterogeneous edge weights - more realistic networks

- **Time dimension**
  - Aggregate layers across time – allow for connectivity to former schoolmates

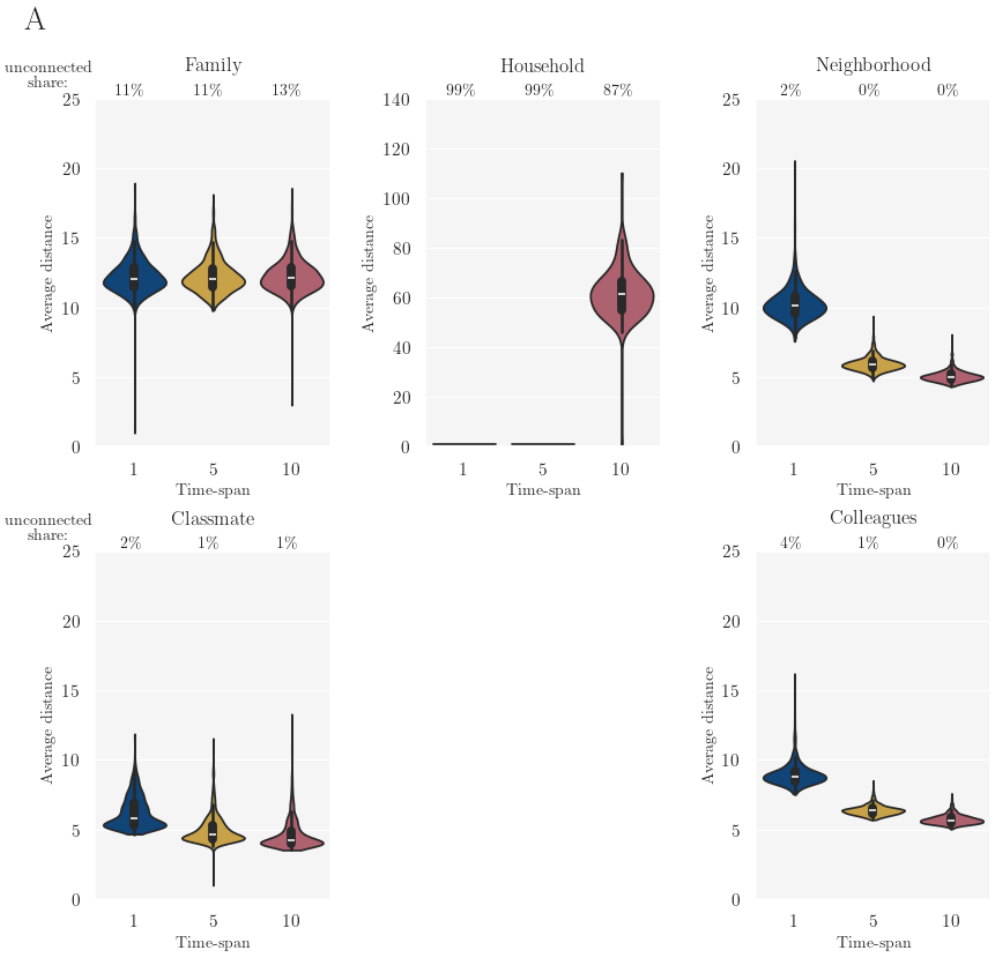
- **Patterns uncovered**
  - Life cycle and heterogeneous connectivity
  - Interlayer spillover
  - Realistic small world properties

Basic network properties

Linking statistics



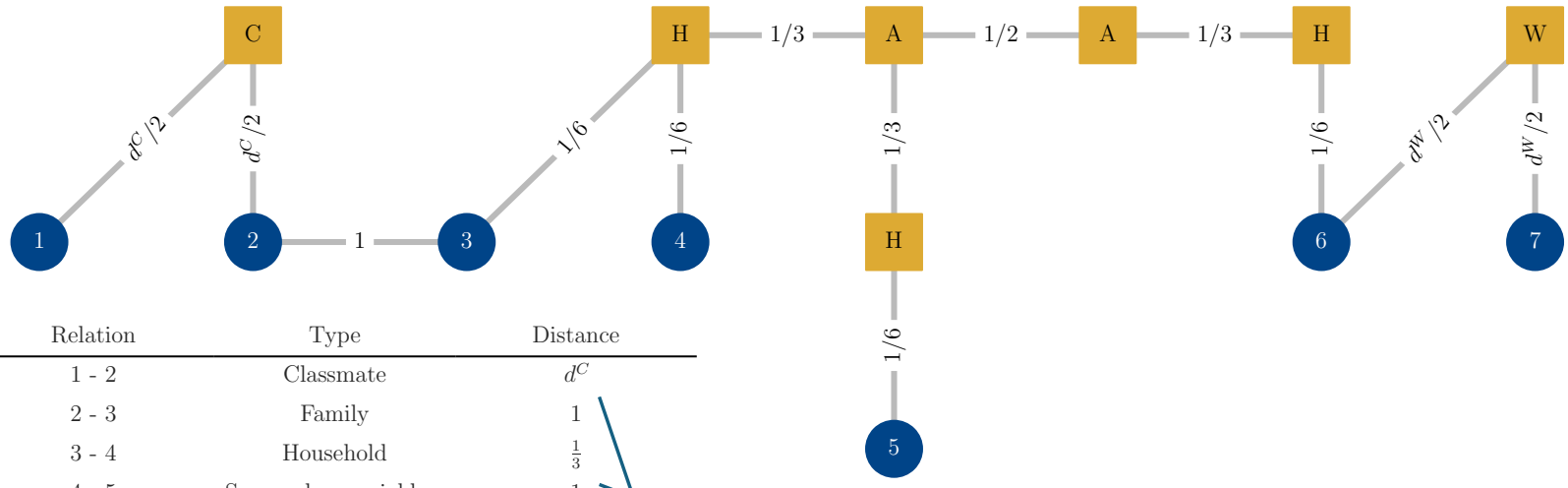
Connectivity by layers



New network properties

How to combine layers?

Flexible weights according to priors



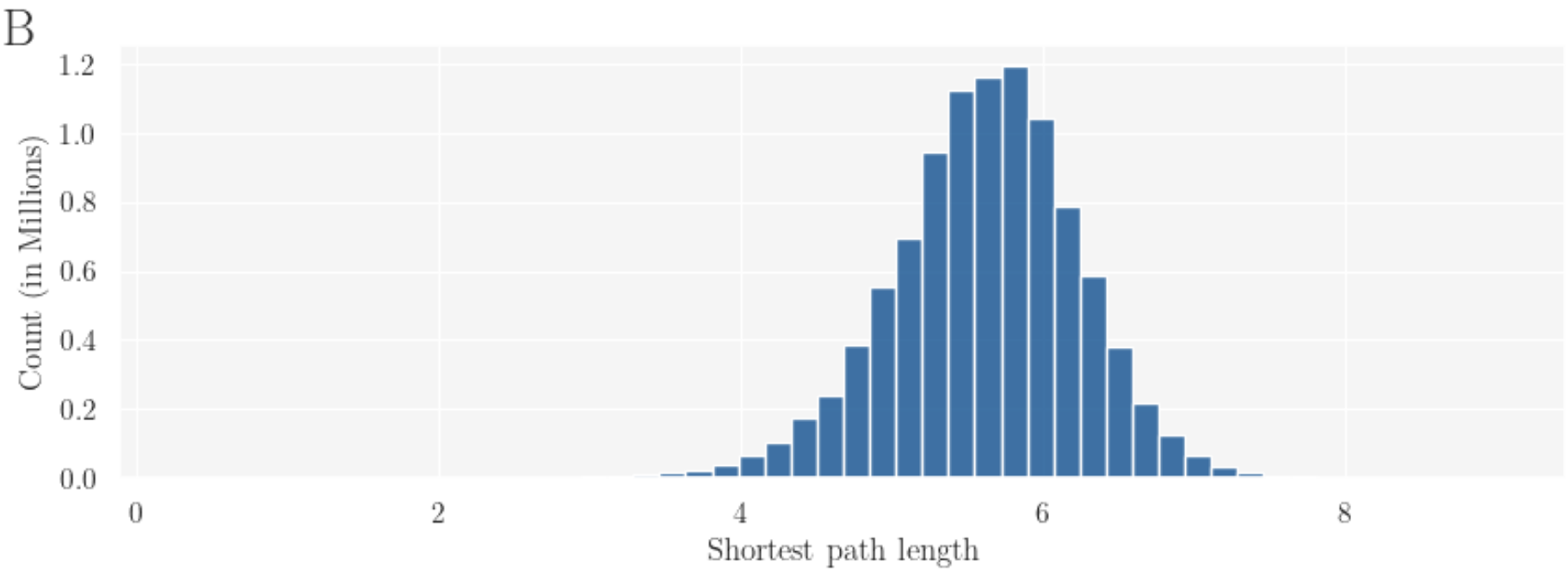
Relation	Type	Distance
1 - 2	Classmate	d^C
2 - 3	Family	1
3 - 4	Household	$\frac{1}{3}$
4 - 5	Same adress neighbor	1
5 - 6	Other adress neighbor	$\frac{3}{2}$
6 - 7	Colleague	d^W

Likelihoods of connecting

Grows with container node size

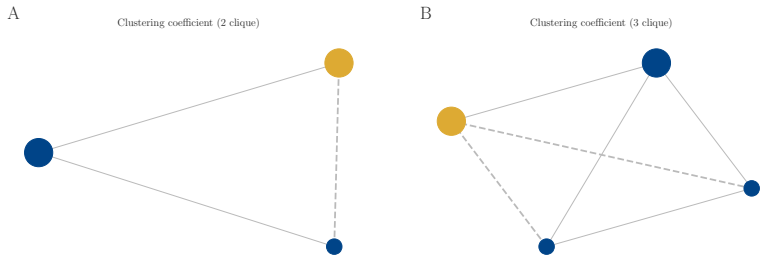
Connectivity overall

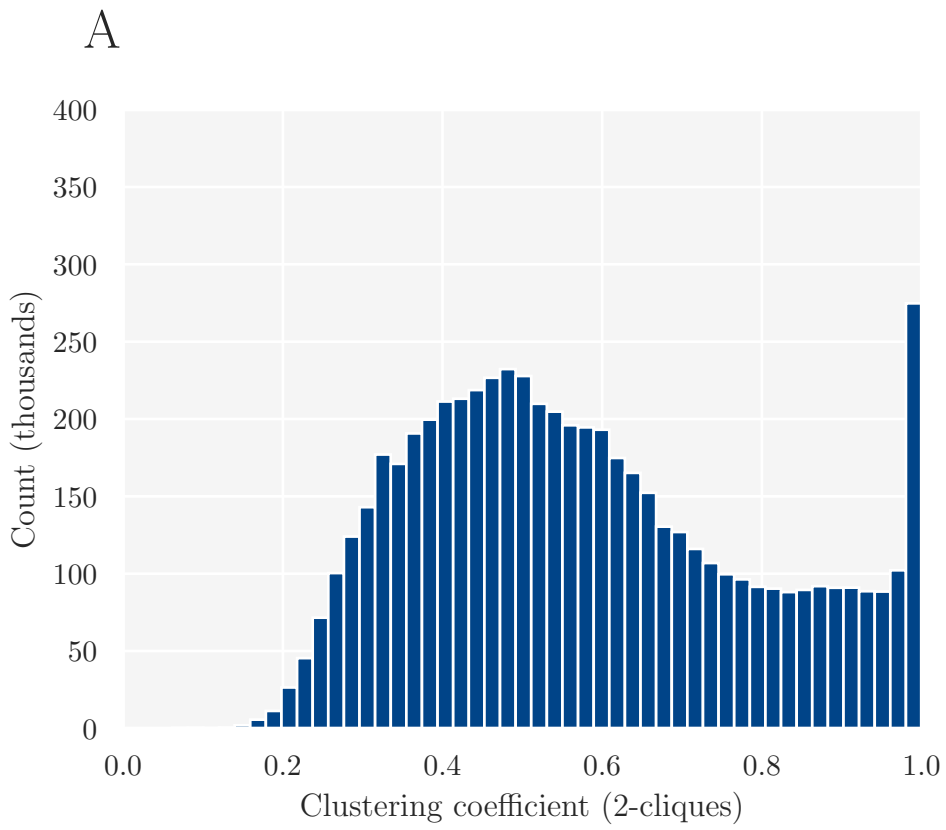
Are people six degrees of apart?

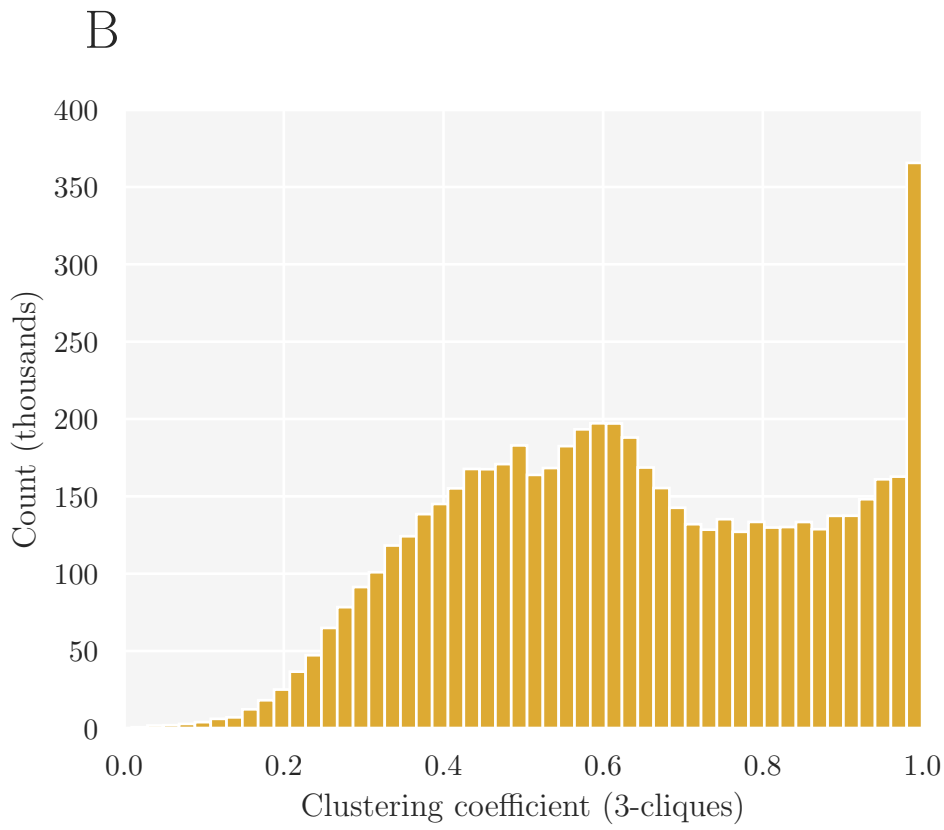


Connection overlap:

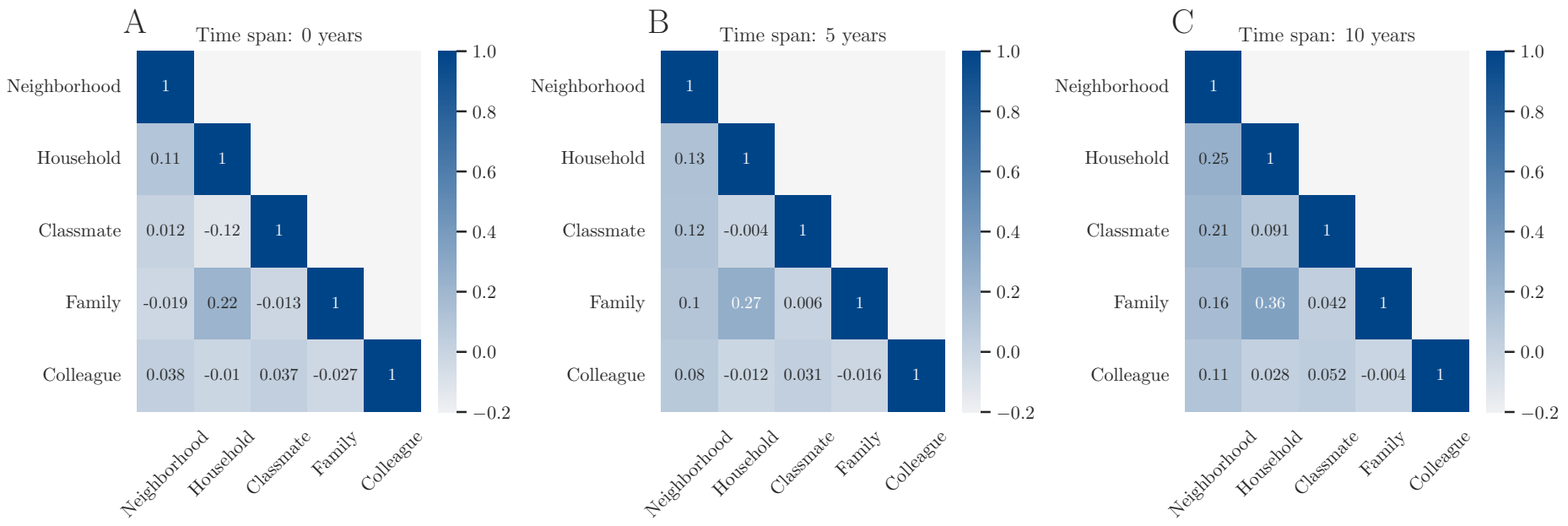
are my links are also mutually linked?



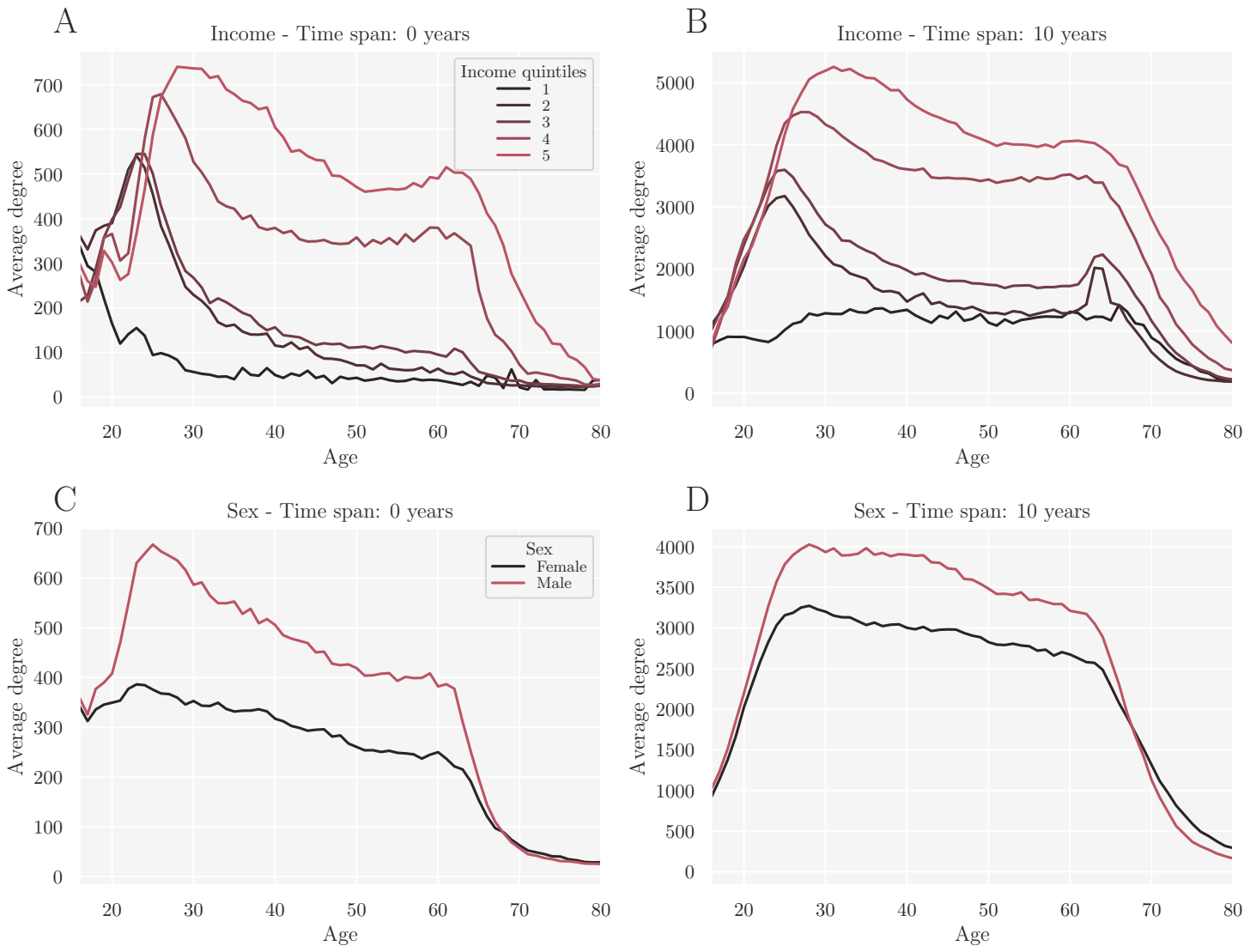




Do network layers interact?



Who have more links?



Outro

Further use and applications

- Data access and software
  - **regnet**: A python package for analysis and network construction
  - Data available through researcher access at Statistics Denmark

- More applications of data
  - Social cohesiveness, segregation and access to social capital
  - Social and match factors
    - Institutional effects (schools, workplaces)
    - Combination with exogenous shocks > peer effects
  - Epidemiology and disease transmission
  - Enrich network – survey of connections within layers

Thank you for listening!

- Contact: abn@sodas.ku.dk

- Please ask any questions, and thanks for your time!