

# Applications of social network analysis on historical data

Dave Griffiths & Paul Lambert

University of Stirling

15<sup>th</sup> May 2013

# What is Social Network Analysis

- SNA is a set of tools for understanding how social interactions influence the actors involved
- It concentrates on relationships between Actor  $i$ , Actor  $j$ , and the implications of their association
- This can focus on the effects for the individual, their whole network or even wider society
- SNA is about the structure of interactions.

# Members of the committees of the UK Social Network Association

<b>Dimitris Christopoulos</b>	<b>Politics</b>	Relational attributes of political entrepreneurs: a network perspective
<b>Bruce Cronin</b>	<b>Business</b>	Director networks and UK firm performance
<b>Martin Everett</b>	<b>Sociology</b>	The human factor: Re-organisations in public health policy
<b>Tom Alcott</b>	<b>(Practitioner)</b>	
<b>Riccardo De Vita</b>	<b>International business and economics</b>	Managing resources for corporate entrepreneurship: the case of Naturis
<b>Bernie Hogan</b>	<b>Communications</b>	Collecting social network data to study social activity-travel behaviour: an egocentered approach
<b>Paola Tubaro</b>	<b>Economic sociology</b>	Norms, status and the dynamics of advice networks: a case study
<b>Federico Varese</b>	<b>Criminology</b>	Mafias on the Move: How Organized Crime Conquers New Territories
<b>Pietro Panzarasa</b>	<b>Business and management</b>	Community structure and patterns of scientific collaboration in Business and Management
<b>Elisa Bellotti</b>	<b>Sociology</b>	What are friends for? Elective communities of single people

- Social Network Analysis (SNA) examines connections and inter-connectivity to understand underlying social structures
- Interested in patterns of interactions between actors, rather than just their attributes
- This involves
  - Actors: a set of individuals, companies, countries, etc
  - Ties: connections joining two actors together (regardless of nature of connection or if all actors can connect)
  - Direction of ties: Directed (A likes B; A manages B) or undirected (A works with B)
  - Strength of ties: Might ('No. of trades') or might not ('do they trade?') be useful
- What are essential are clearly defined and observable rules determining whom is in the network and how connectivity will be measured

- Assumption of interdependence between actors
  - cf. other statistical analyses treating actors as independent
- Educational attainment example
  - Standard regression: grades as if independent of other pupils
  - Hierarchical modelling: grades can depend upon their school, therefore pupils are clustered by school to retain independence
  - Social network approach: grades potentially dependent on the performance of who pupils befriend; “birds of a feather flock together”, “getting in with the wrong crowd”, “pushing each other along”
- Note: dependencies in other approaches are usually one way (i.e., parents influence children, but children cannot influence parental attainment)

# What are social networks?

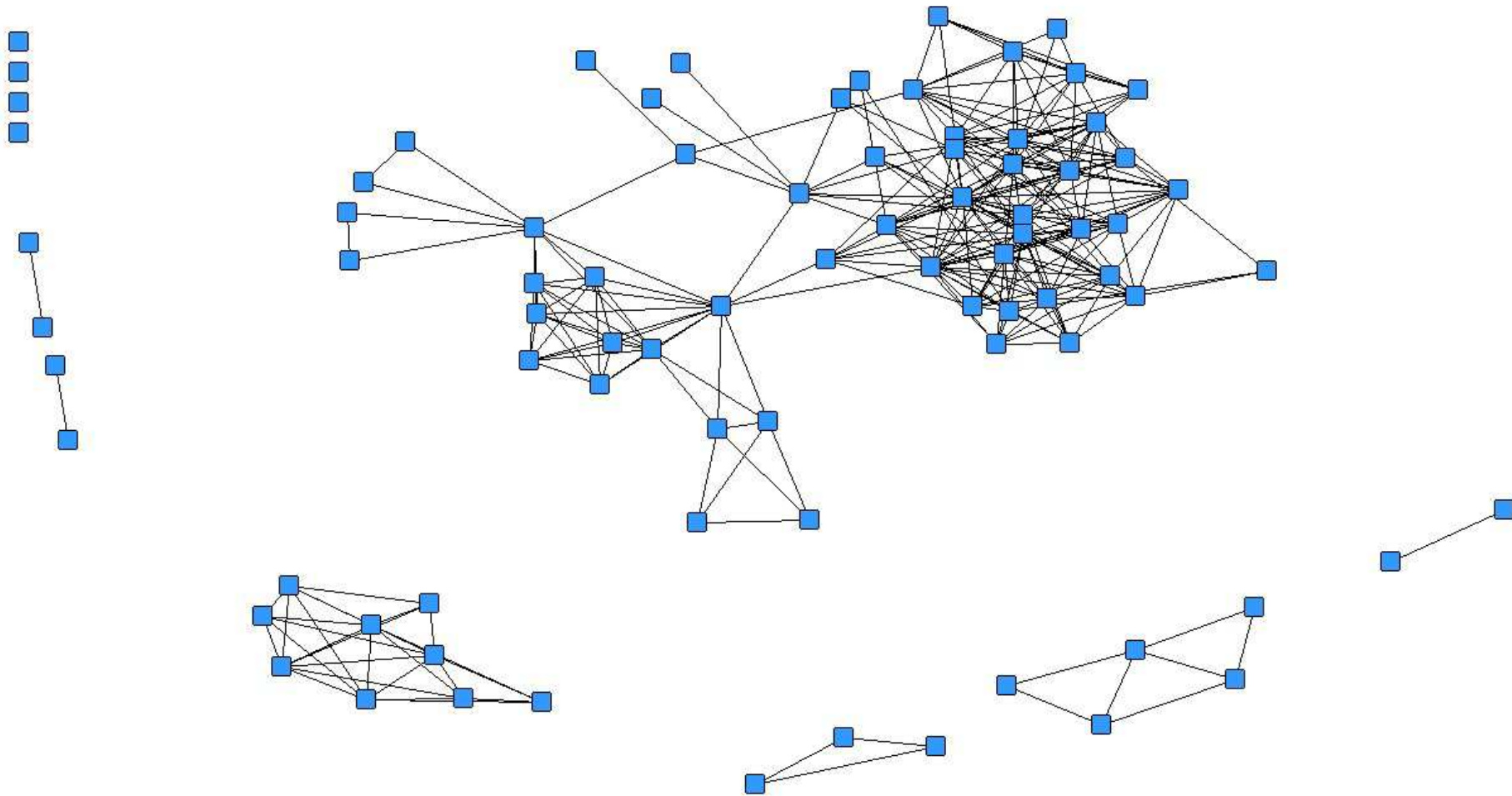
- Networks involving social structure or social process
- This can be formal/conscious networks, such as friendships, who you give birthday cards to, who you invite to parties
- They can be informal/unconscious networks, such as people who attend the same concerts, shop in the same record shops or download music from similar bands
- They can involve people who are on first name terms, or even people who have never met
- They can involve entities, such as countries which trade, companies which share directors, TV series which share actors
- They can involve animals, such as baboons grooming each other
- They can involve concepts, such as linking words or variables

# What isn't Social Network Analysis



Social network analysis = analysis of networks existing in the social world

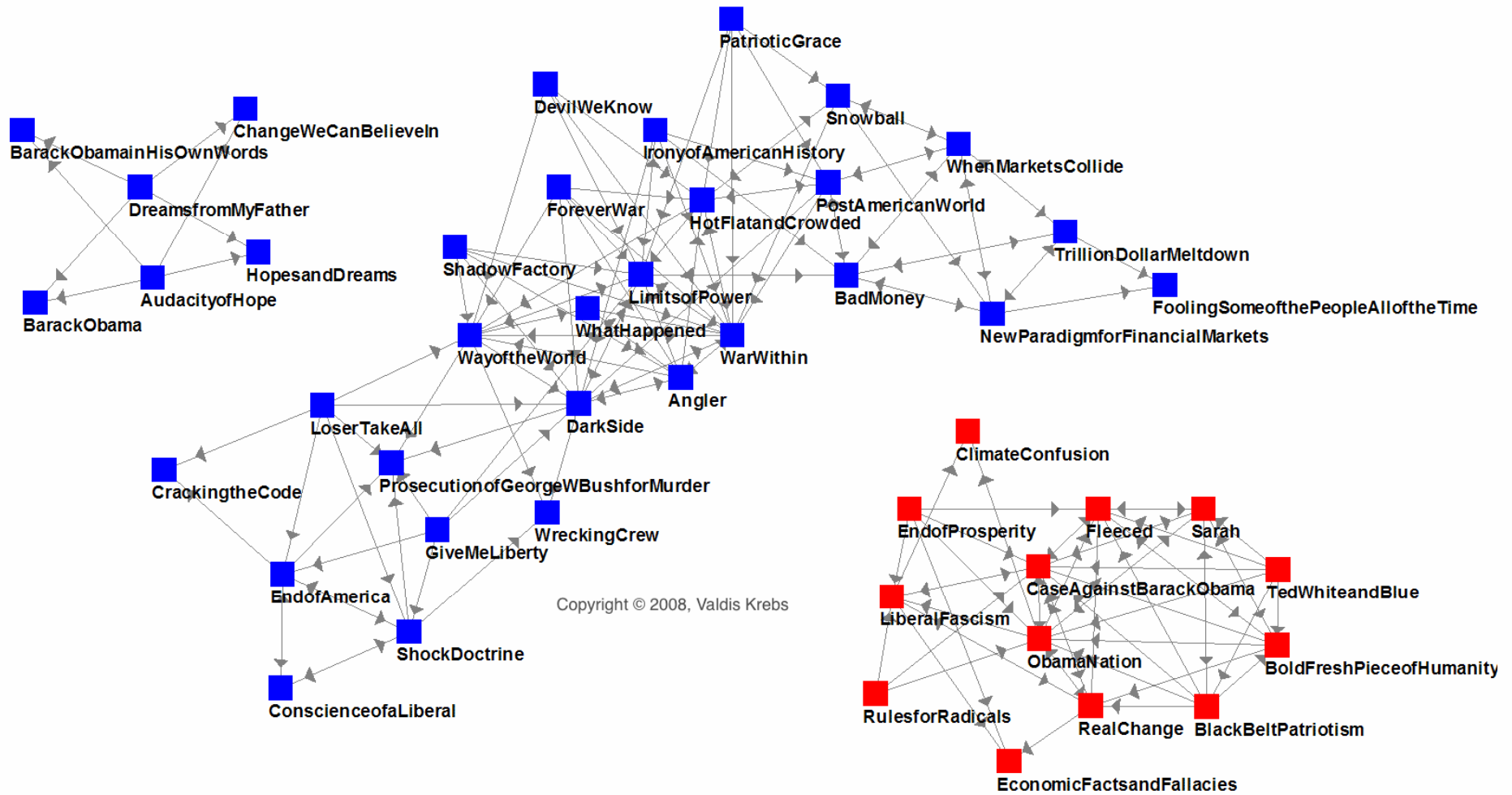
Social networking sites = sites for people to network socially



## Visualisation of my personal Facebook network (August 2012)

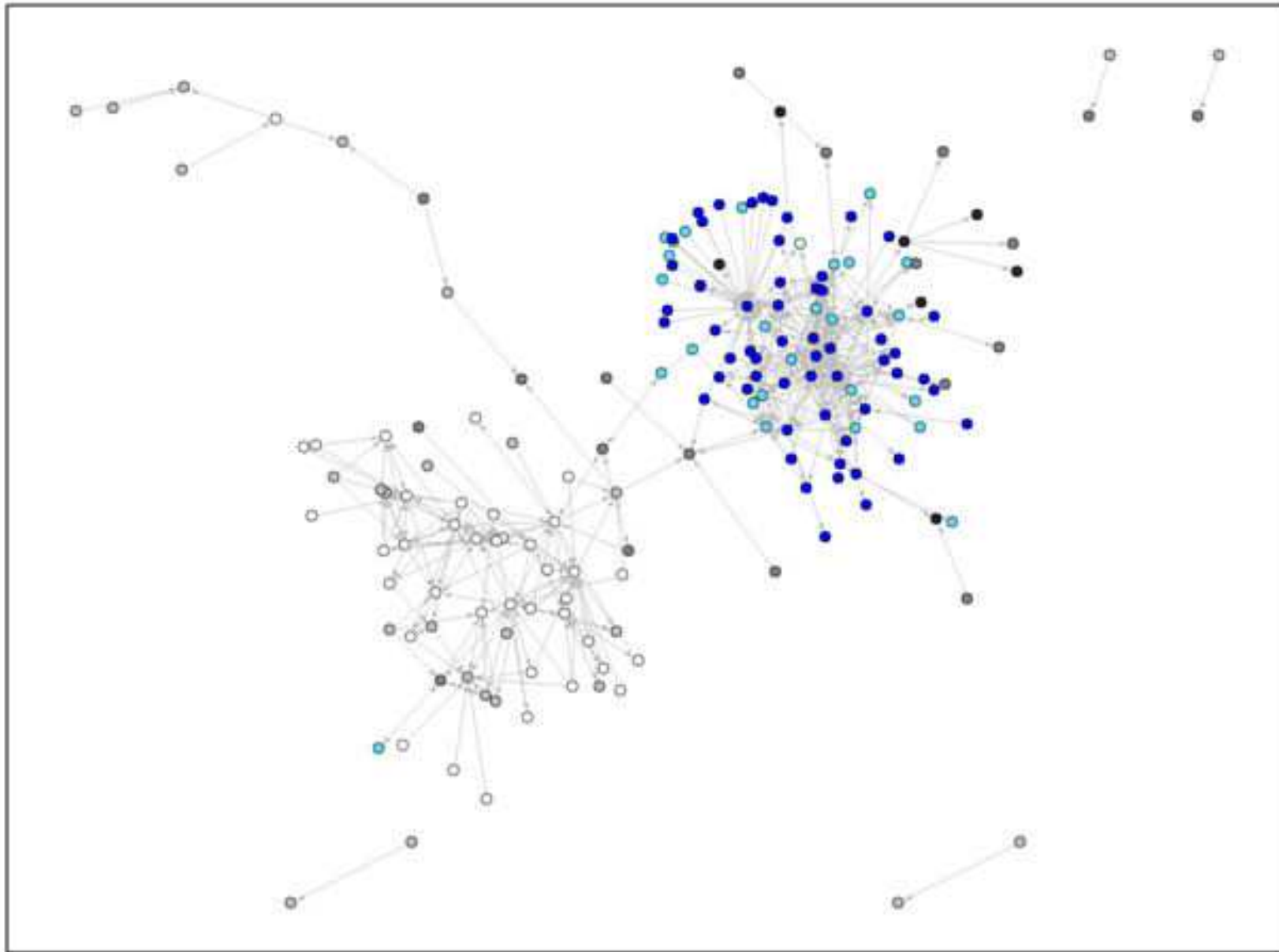
Generated using: <https://apps.facebook.com/namegenweb>





Amazon.com political book recommendations, August 2008

<http://www.orgnet.com/divided.html>



Network of occupational connectivity by marriage and educational level, USA 2000.

Griffiths D., and Lambert, P.S. (2012) [Dimensions and Boundaries: Comparative Analysis of Occupational Structures Using Social Network and Social Interaction Distance Analysis](http://www.socresonline.org.uk/17/2/5.html), *Sociological Research Online*, 17(2), 5, <http://www.socresonline.org.uk/17/2/5.html>

# Practical uses of SNA

- **Power** (how can companies utilise privileged positions)
- **Influence** (which individuals can control decision-making)
- **Isolation** (who is excluded from certain situations)
- **Knowledge transfer** (who can best receive and send information)
- **Efficiency** (do resources flow through a network well)
- **Variance** (does network position affect social position)

Elections and contracts are won or lost through network performance. Networks help us understand not merely how structures operate, but how we can improve and mobilise them.

# Growing interest in historical networks

- Access to data on social connections in the past
- (e.g. recent microdata access projects; digitisation of records)
- Examples of papers in Historical Networks streams of INSNA Sunbelt Conference (Hamburg, late May 2013)
  - Luca De Benedictis & Silvia Nenci – “A Network Analysis of Preferential Trade Agreements: 1815-1914”
  - Christine Fertig – “Personal Networks and Social Classes in Rural Society: A Microstudy of Two Parishes in 18<sup>th</sup> and 19<sup>th</sup> Century Westphalia”
  - Hilde Bras, Alice Kasakoff, Diansheng Guo & Cuglar Koylu – “Visualizing Historical Kinship Networks using Data from Marriage Registers: The Netherlands, 1830-1950”
  - Martin Stark – “Networks of Creditors and Debtors: A Rural Credit Market in 19<sup>th</sup> Century Germany”
  - <http://hamburg-sunbelt2013.org/>
- Marten During runs a useful website collating historical network studies
- <http://historicalnetworkresearch.org/>

# *Studying social connections?*

- Many research methods have been ‘individualist’
  - In statistical analysis & explanatory frameworks
- To study empirical data on social connections...
  - Individualist approach: Use data about the alter(s) to inform analysis of the individual
  - Structural approach: Use data about the connections to inform understanding of the structure
- In social history...
  - Data on social connections is one of few forms of readily available large scale microdata, and is increasingly accessible
  - Social connections are central to interesting social trends, e.g. in social mobility; homogamy; industrialisation; etc

1991



Exemplar  
Geller households  
from TV series  
*Friends* (1991-97)

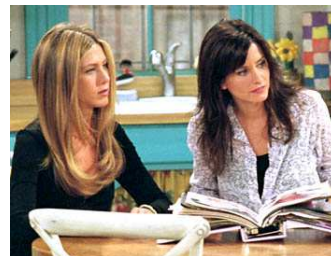
1993



1995

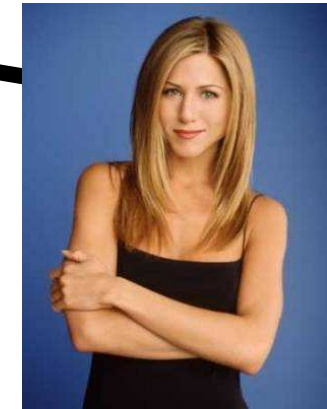


1997

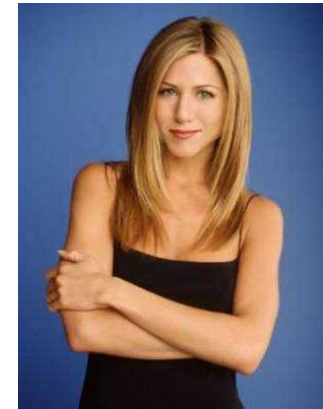




Grouped by  
cohabitation  
networks



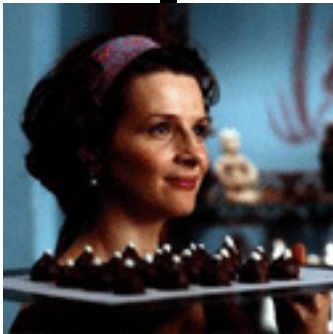
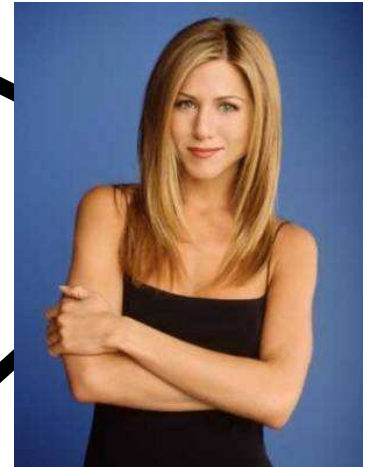
Grouped by family ties

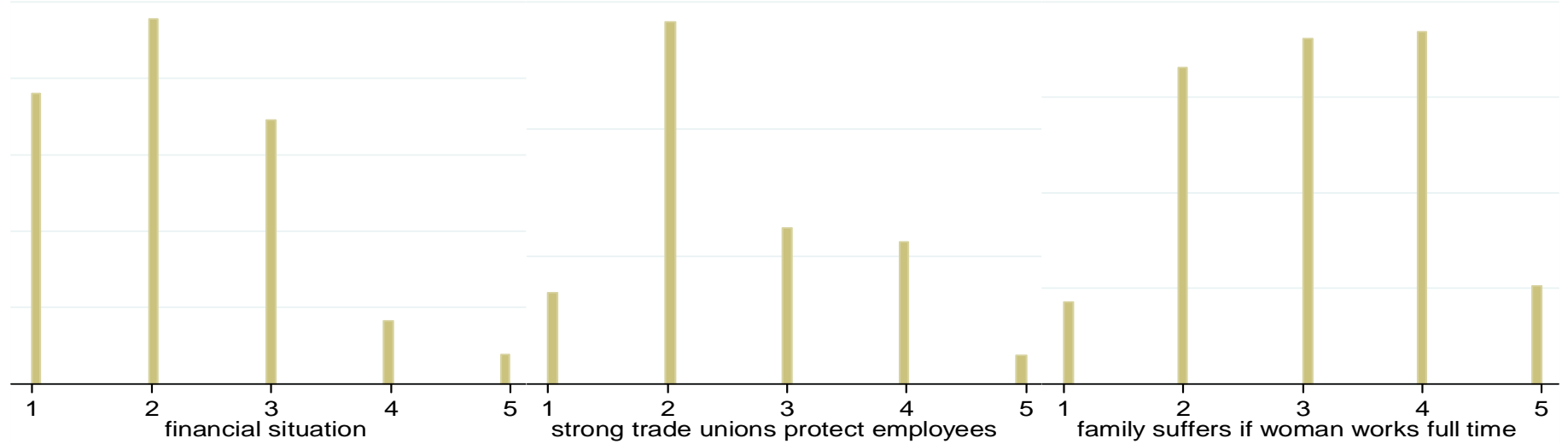
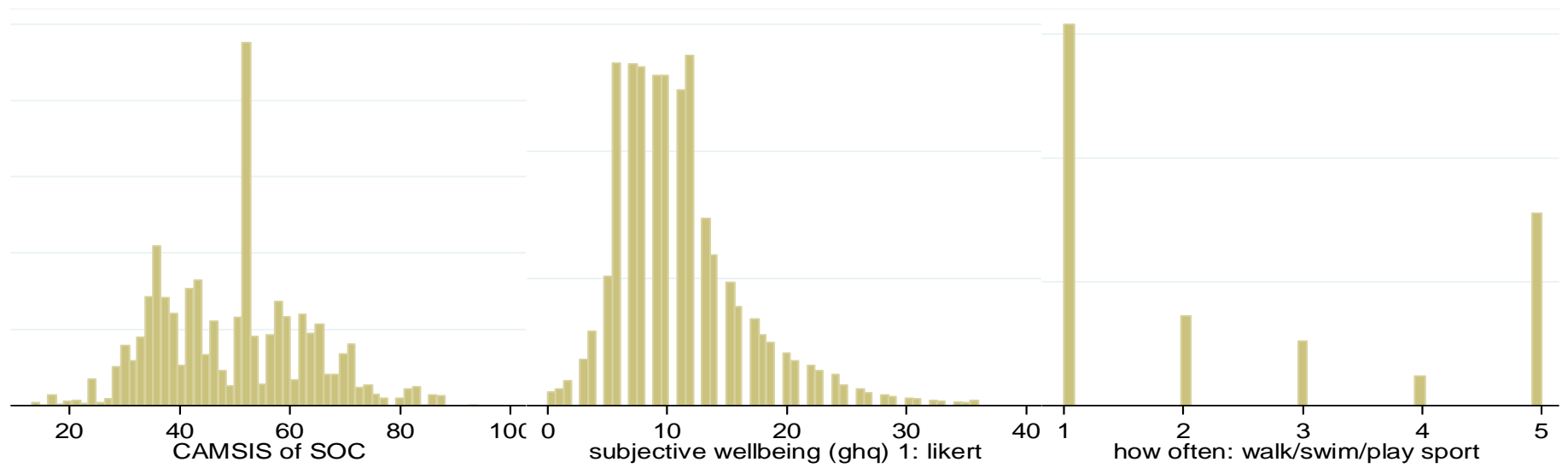






Grouped by  
occupation



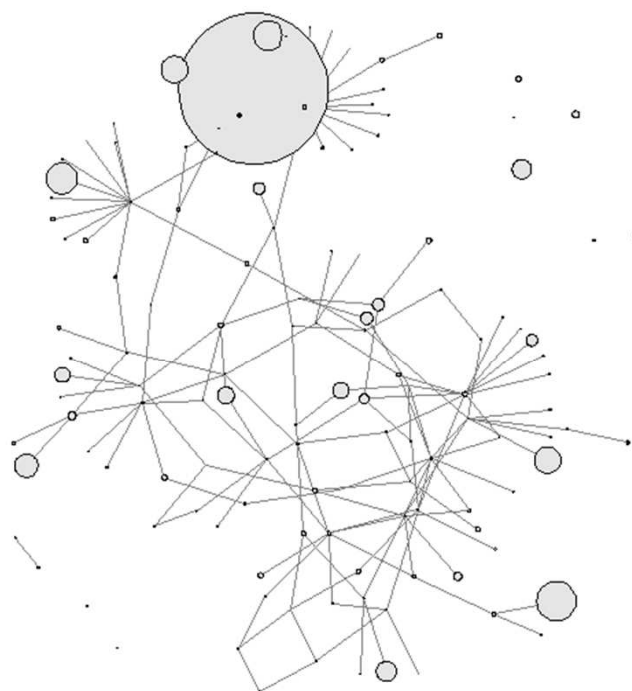


- CAMSIS score of occupational advantage
- Self-rated health
- Participation in exercise
- Feeling financial secure
- Attitudes towards trade unionism
- Attitudes towards motherhood and employment

Outcome 3: Scale ranking for self-rated sports participation level (scale from 1 to 5, 1=very active, modelled as linear scale)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	13.5*	12.8*	13.1*	13.4*	12.9*	<b>12.9*</b>	12.8*
Female	0.86*	0.82*	0.80*	1.12*	0.81*	<b>1.02*</b>	1.14*
(Age – 40)/10	0.42*	0.63*	0.59*	0.42*	0.61*	<b>0.61*</b>	0.63*
(CAMSIS -50)/10	-0.08	0.05	0.13	-0.11	0.08	<b>0.06</b>	-0.02
(Age*CAMSIS)/10	-0.24*	-0.25*	-0.27*	-0.23*	-0.26*	<b>-0.25*</b>	-0.25*
Deviance		154306	154393	155459	154356	<b>154255</b>	154310
AIC	155522	154320	154407	155473	154415	<b>154273</b>	154419
ID variance ICC	100%	71.6%	74.2%	99.1%	71.4%	<b>71.2%</b>	70.9%
FID variance ICC		28.4%			19.8%	<b>19.3%</b>	19.9%
NID variance ICC			25.8%		8.7%	<b>8.9%</b>	8.3%
SOC variance ICC				0.9%		<b>0.6%</b>	1.0%
Fem   soc variance							0.3%
Notes: For model (7), the ICC estimates refer to variance proportions for males at the intercept (due to the 'random coefficients' formulation of that model).							

	CAMSIS	Health	Sports	Financial security	Working mothers	Trade unions
ID variance ICC	<b>71.3%</b>	<b>89.1%</b>	<b>71.2%</b>	<b>74.5%</b>	<b>83.2%</b>	<b>77.3%</b>
FID variance ICC	<b>7.9%</b>	<b>9.3%</b>	<b>19.3%</b>	<b>19.8%</b>	<b>11.6%</b>	<b>7.0%</b>
NID variance ICC	<b>20.8%</b>	<b>1.3%</b>	<b>8.9%</b>	<b>4.6%</b>	<b>4.7%</b>	<b>10.9%</b>
SOC variance ICC		<b>0.2%</b>	<b>0.6%</b>	<b>1.0%</b>	<b>0.5%</b>	<b>4.3%</b>
Fem   soc variance				<b>0.1%</b>		<b>0.5%</b>

**Figure 3.3: Marriage links between occupations in Scotland, 1881, using 'Threshold method'**  
*(link occurs two or more times as often as would be expected by chance, and for at least 5 couples)*



Male job	Female job	# ties	Male job	Female job	# ties
...	...		...	...	
344. Coal Merchant	399. General Shopkeeper, Dealer	6	6. Municipal, Parish, Union, District, Officer	6. Municipal, Parish, Union, District, ..	12
348. Stone Quarrier	236. Grocer. Tea, Coffee, ...	8	6. Municipal, Parish, Union, District, Officer	34. School Service, & others connected	6
348. Stone Quarrier	270. Dyer, Printer, Scourer, Bleacher, ...	16	32. Schoolmaster	32. Schoolmaster	100
348. Stone Quarrier	285. Shirt Maker, Seamstress	9	47. Musician, Music Master	47. Musician, Music Master	35
348. Stone Quarrier	348. Stone Quarrier	6	50. Actor	47. Musician, Music Master	6
360. Road Labourer	82. Toll Collector, Turnpike Gate Keeper	10	50. Actor	50. Actor	59
360. Road Labourer	404. General Labourer	7	...	...	

# HIS-CAM (Historical CAMSIS) scales

- **Summary measure of HISCO occupational positions**
  - Differentiates finer occupational details
    - *Typically 300+ occupational units assigned different scores*
  - Emphasises a hierarchical structure of inequality
  - ***An instrumental measure (of the relative advantage typically associated with incumbents of an occupational position)***
- **Explorative device for understanding occupations**
  - Measure multiple relative structures of stratification between countries, time periods, gender based groups..?

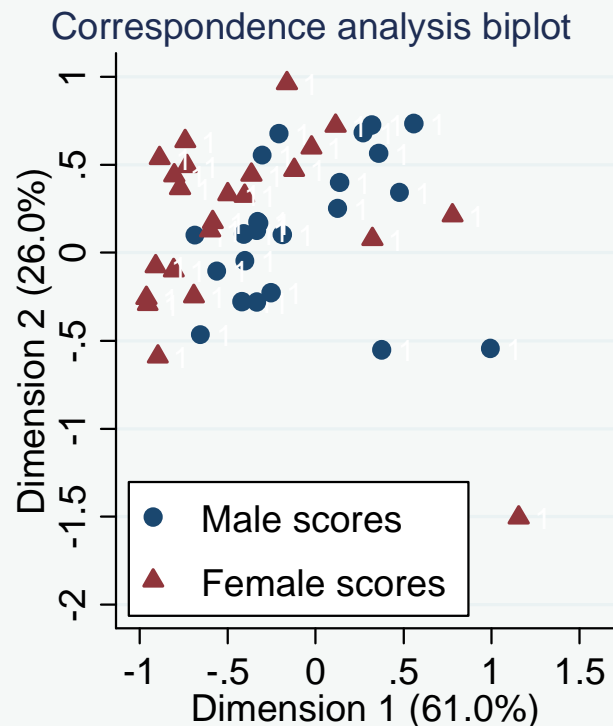
Lambert, P. S., Zijdeman, R. L., Maas, I., van Leeuwen, M. H. D., & Prandy, K. (2013). The construction of HISCAM: A stratification scale based on social interactions for historical research. *Historical Methods*, 46(2), 77-89.

## Husband's Job Units

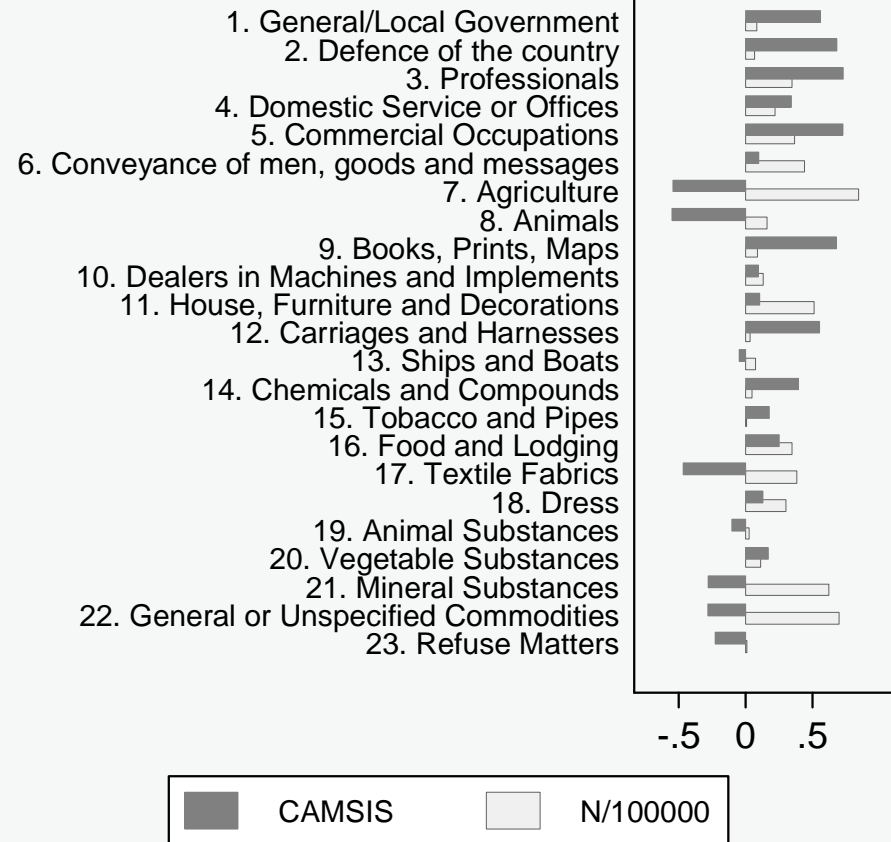
		Husband's Job Units				
Occ Units ↓ →		1	2	..	407	
Derived scores ↓ →		75.0	70.0	..	10.0	
Wife's Job Units	1	72.0	<b>30</b>	<b>15</b>	..	<b>0</b>
	2	72.5	<b>13</b>	<b>170</b>	..	<b>1</b>
	..	..	..	..	..	..
	407	11.0	<b>0</b>	<b>2</b>	..	<b>80</b>

- *Derived scores predict frequency of interactions (#cases per cell)*
- The scales describe one or more dimensions of a **structure of social interaction...**
  - ...this turns out to also represent a **structure of social stratification...**
  - ...resulting in scale scores which measure an occupation's relative position within the structure of stratification.

In the SONOCS project ([www.camsis.stir.ac.uk/sonocs](http://www.camsis.stir.ac.uk/sonocs)), we used SID analyses on large-scale historical datasets



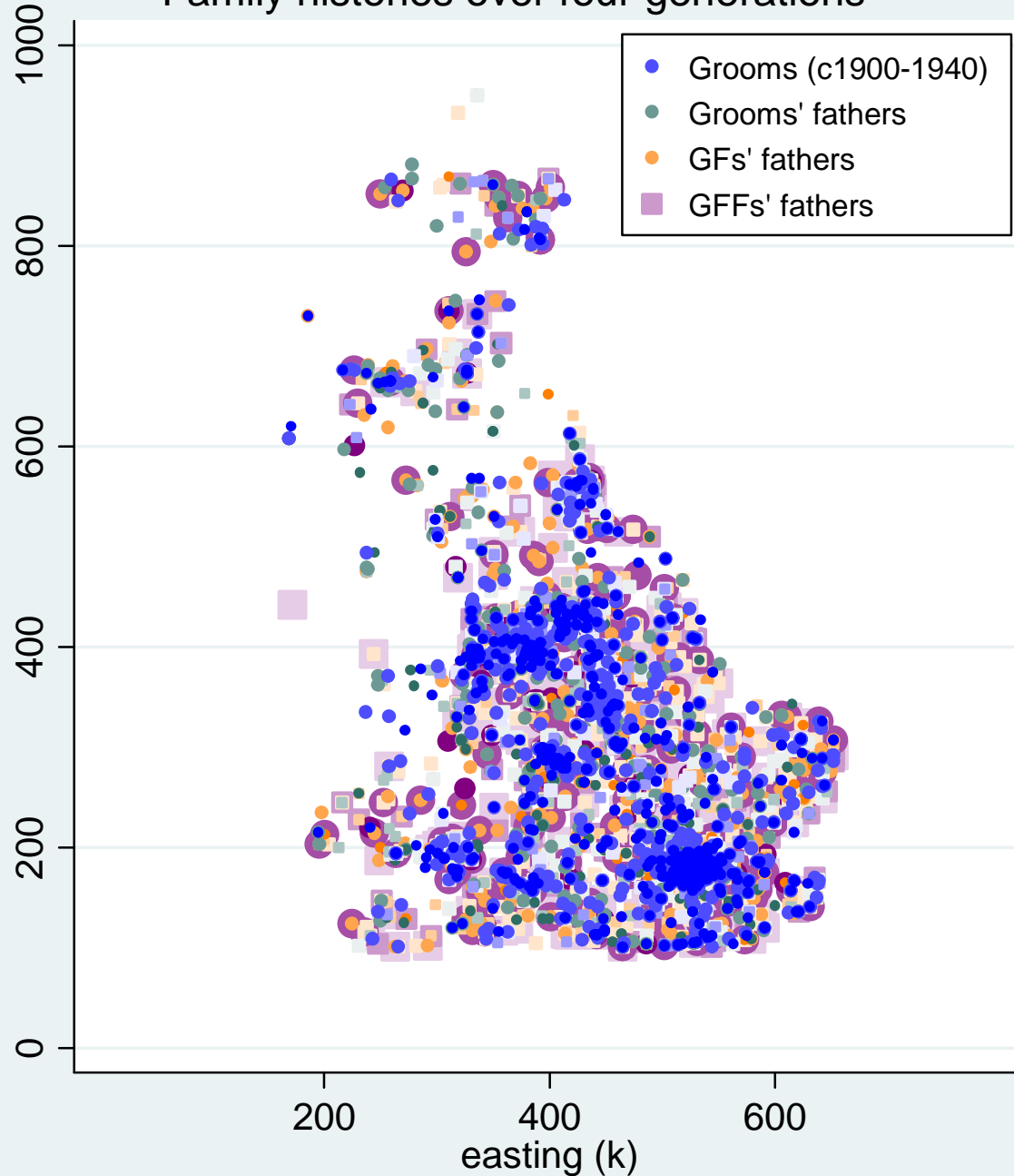
CAMSIS scores by broad occupational groups



Source: NAPP, N=598000 (Intra-household male-female occupational combinations).  
 Panel 1: Dimension scores from correspondence analysis of intra-household occupations  
 Panel 2: Mean scores for males by 'occupational order'.



## Family histories over four generations

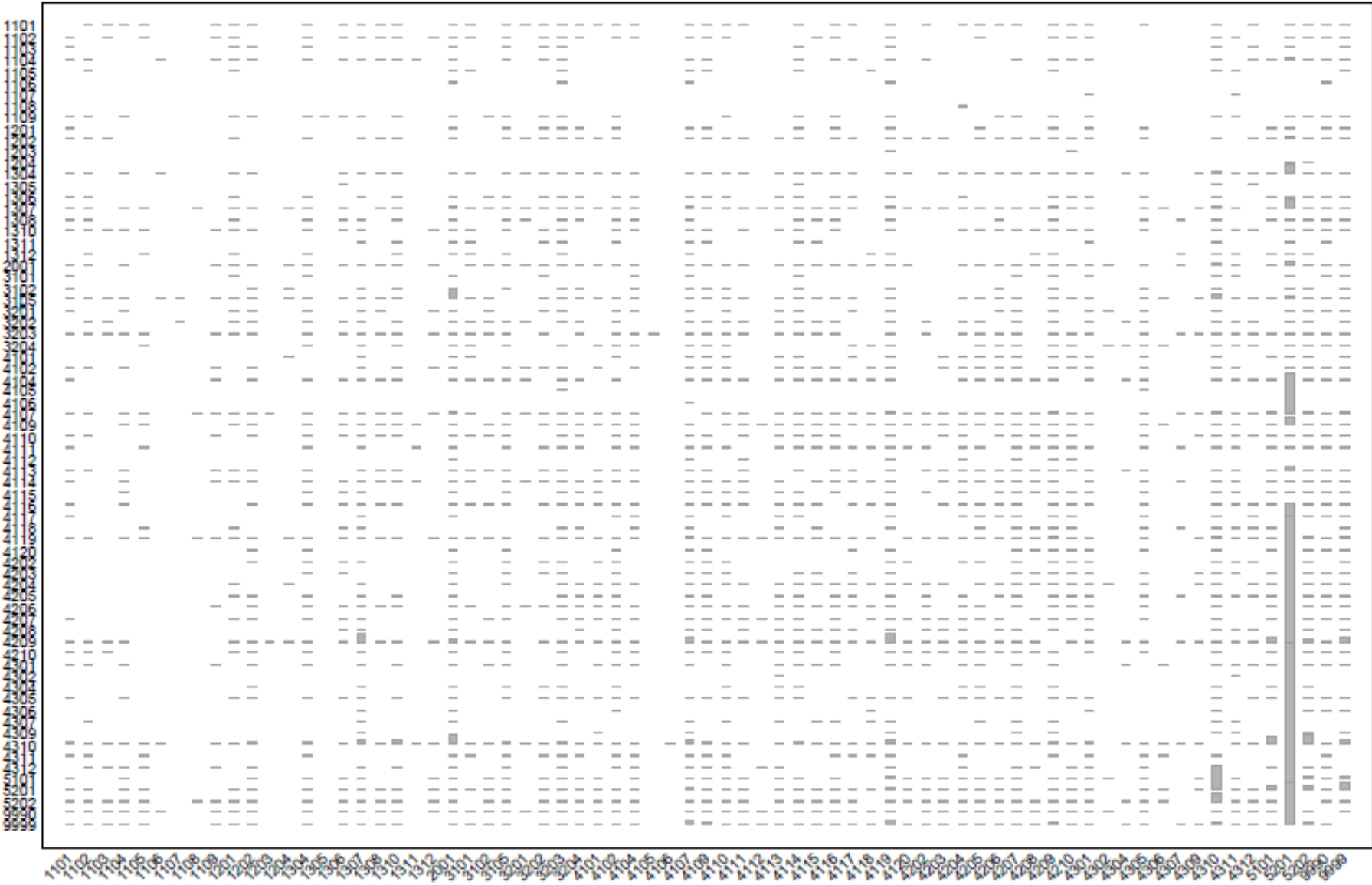


The UK 'Family History Study'  
[Prandy and Bottero 2000]

Source: 2899 marriages c1900-1950, Family History Study (Prandy and Bottero, 2000)  
Shading intensity / symbols reflect HISCAM scale.  
Geographic identifiers for location at time of marriage or child birth

# Microclasses

## Norway, 1865

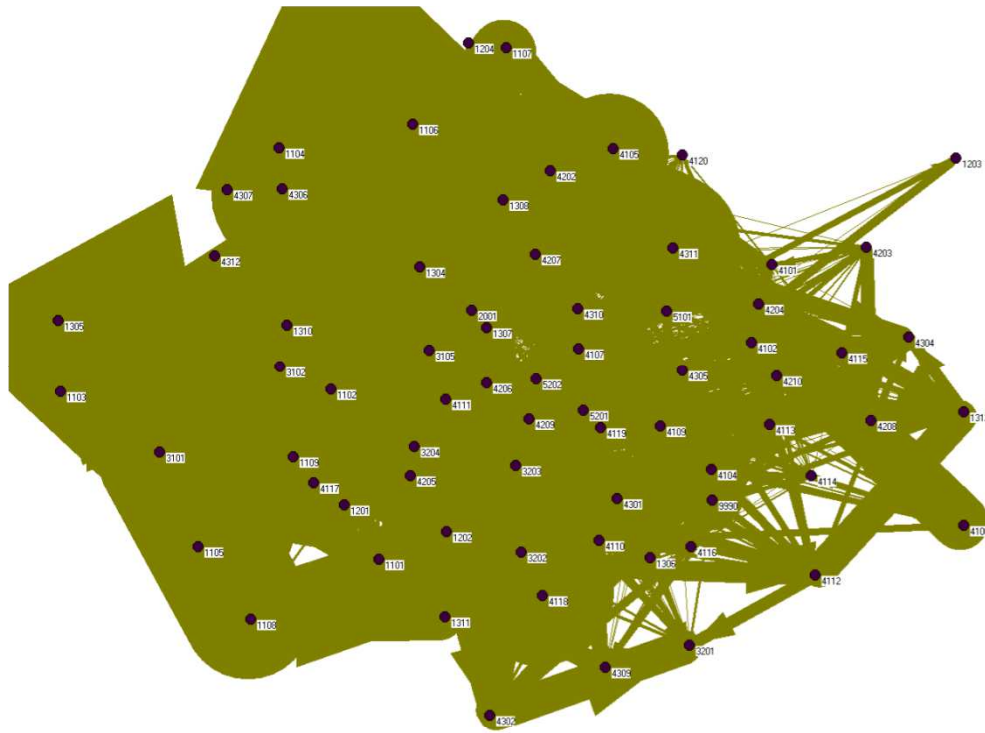


(Excluding diagonals, n\*20)

# Norway 1865

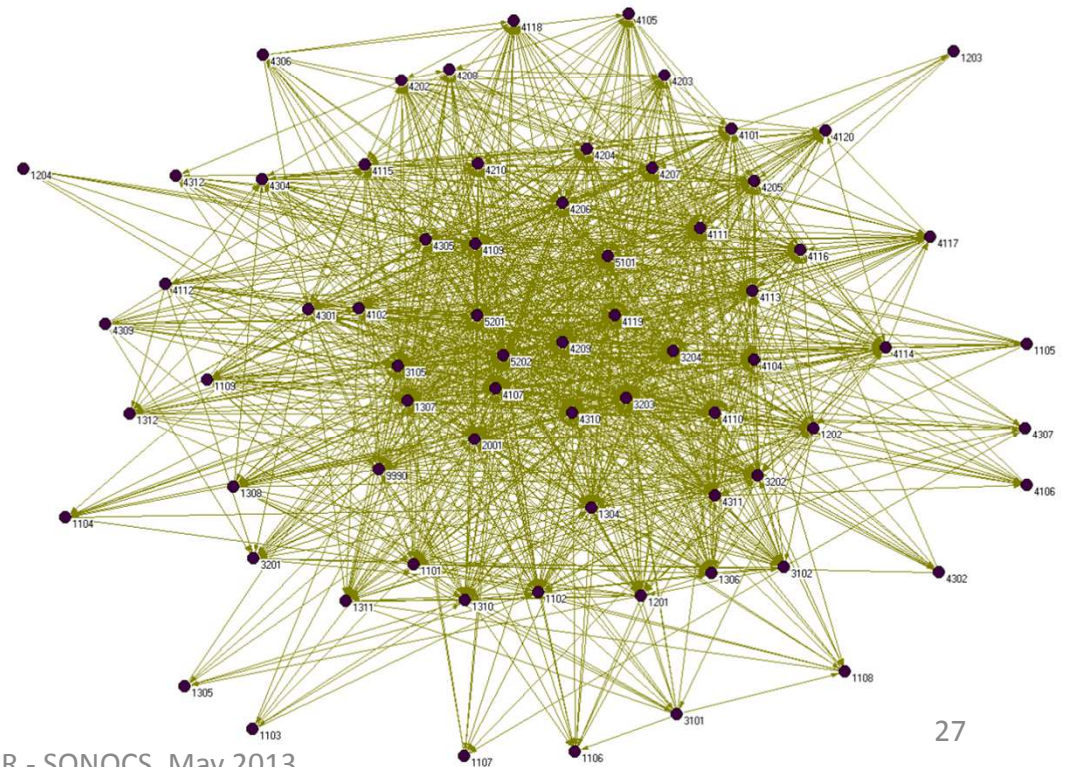
Male-male microclass combinations  
of at least 16 year difference.

Displayed with, and without, lines  
replicating levels of representation.



Networks of raw ties are too  
large to be remotely meaningful.

Sparse ties are created, whilst a  
link with 1 connection has as  
much influence over position as  
a link with 10,000 connections.



# What constitutes a tie between occupations?

- Remove all combinations performing the same occupations

(structural relationships more readily explored by looking at mobility than immobility)

- Over representation: must occur at least **X** times more than expected by chance

(occurs more often than if occupational combinations were allocated randomly)

- Frequency of relationship: must occur in at least **Y,000** combinations

(to exclude cases where over-representation occurs with a small number of cases)

- Apply confidence intervals when identifying over-representation

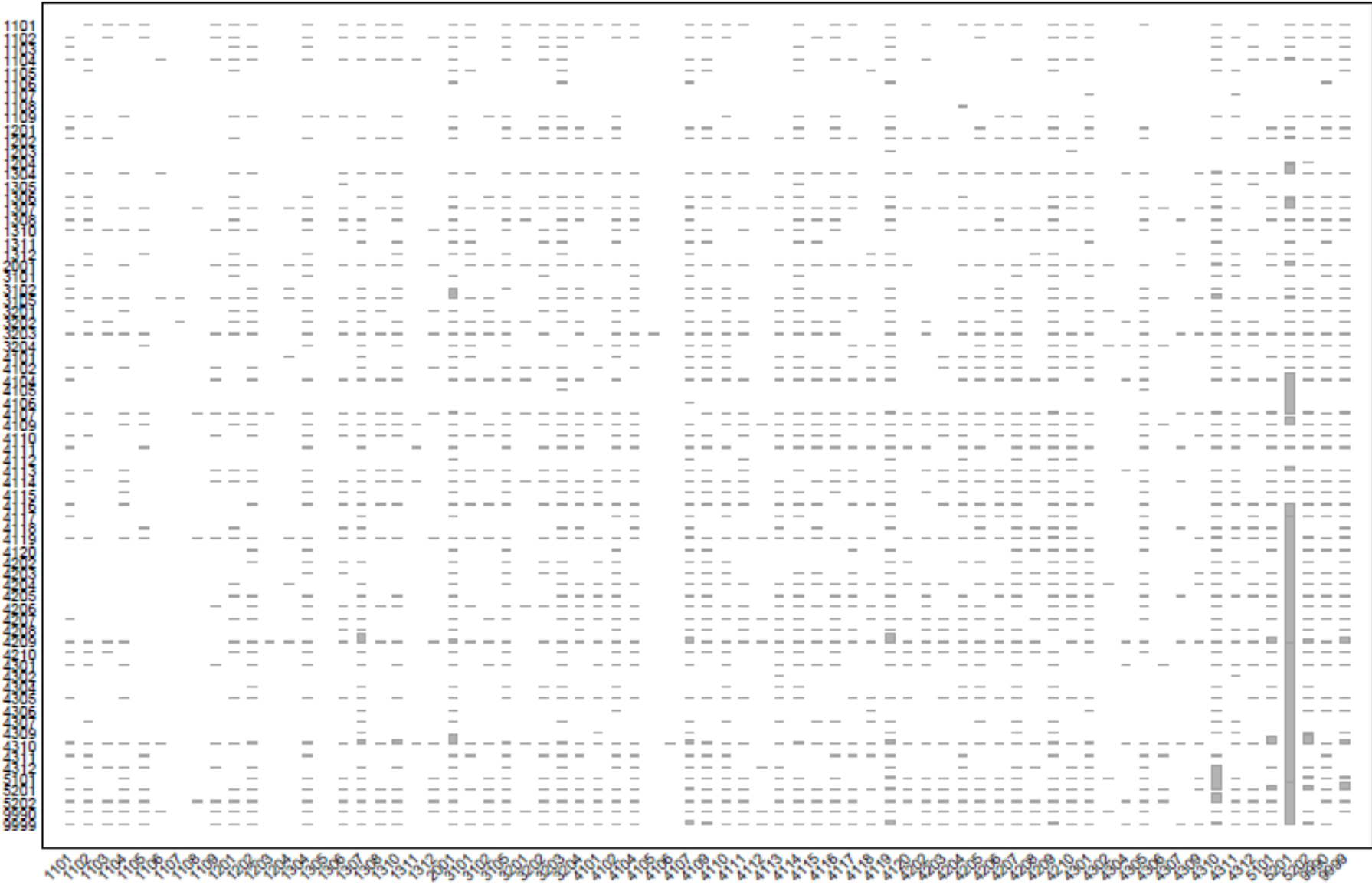
# Typical example of Stata syntax

```
*****Exporting only those linkages which are
** above the expected values
**create frequency dataset
capture drop freq
gen freq = 1
collapse (count) freq, by(hocc wocc)
*****Find total for each category
capture drop tot
egen tot=sum(freq)
*****Find totals for men and women
capture drop nhocc
capture drop nwocc
egen nhocc=sum(freq), by(hocc)
egen nwocc=sum(freq), by(wocc)
****Find percentage for each category for men and women
capture drop phocc
capture drop pwocc
gen phocc=nhocc/tot
gen pwocc=nwocc/tot
*****Calculate expected numbers of women
capture drop ewocc
gen ewocc=pwocc*nhocc
*****create expectation surplus
capture drop value
gen value=freq/ewocc
*****Create standard error predictions
capture drop prop
gen prop = freq/tot
capture drop staner
gen staner = sqrt((prop)*(1 - prop) / tot)
```

```
capture drop pro_obs
gen pro_obs = freq/tot
capture drop pro_exp
gen pro_exp = ewocc/tot
capture drop pro_min
gen pro_min = pro_obs - staner
capture drop pro_max
gen pro_max = pro_obs + staner
capture drop value
gen value = pro_obs / pro_exp
capture drop val_min
gen val_min = pro_min / pro_exp
capture drop val_max
gen val_max = pro_max / pro_exp
*****label variables
label variable tot "total number in sample"
label variable nhocc "total number of males in occupation"
label variable nwocc "total number of females in occupation"
label variable phocc "percentage of men in occupation"
label variable pwocc "percentage of women in occupation"
label variable ewocc "expected number of partnerships"
label variable staner "Standard error for tie"
label variable pro_obs "Observed proportion of all ties"
label variable pro_exp "Expected proportion of all ties"
label variable pro_min "Lower confidence interval of observed proportion"
label variable pro_max "Higher confidence interval of observed proportion"
label variable value "Observed value of representation"
label variable val_min "Value of representation for lower confidence
interval"
label variable val_max "Value of representation for higher confidence
interval"
```

# Microclasses

## Norway, 1865



(Excluding diagonals, n\*20)

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	Label
1.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1101
2.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3203	
3.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1104	
4.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4204	
5.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1105	
6.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5101	
7.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1106	
8.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	9990	
9.	#	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1201	
10.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1307	
11.	#	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1202	
12.	#	.	.	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1304	
13.	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1310	
14.	.	.	#	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1306	
15.	#	#	.	.	.	.	.	.	.	.	#	#	#	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2001	
16.	.	.	#	.	.	.	.	#	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4104	
17.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4111	
18.	.	.	.	.	.	.	.	.	.	#	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4209	
19.	#	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3204	
20.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1102	
21.	#	.	.	.	#	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3102	
22.	#	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3105	
23.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3201	
24.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4114	
25.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4304	
26.	#	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3101	
27.	#	.	.	.	#	.	.	.	#	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3202	
28.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	4305	
29.	.	.	#	.	.	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4101	
30.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4117	
31.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	4206		
32.	#	.	.	.	.	#	.	#	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4102	
33.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4110	
34.	.	.	#	.	.	.	.	#	#	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4107	
35.	#	.	.	.	.	.	.	#	.	.	.	#	.	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	4109	
36.	#	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	#	#	#	.	.	.	.	.	.	.	.	.	.	.	.	4115	
37.	.	.	.	#	.	.	.	.	.	.	.	#	#	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	#	4119		
38.	.	.	.	.	.	.	.	#	.	.	.	#	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	#	4208	
39.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4113	
40.	#	#	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4116	
41.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4118	
42.	.	.	.	.	.	.	.	.	.	.	.	.	#	#	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	#	.	4207		
43.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	4202	
44.	.	.	.	.	.	.	.	.	.	.	.	#	.	.	.	.	.	.	#	#	#	.	.	.	.	.	.	.	.	.	.	.	.	4203	
45.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1312	
46.	.	.	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4205	
47.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5201	
48.	.	.	.	.	.	.	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	.	4210		
49.	#	#	.	.	.	.	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4301	
50.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	.	4309		
51.	.	.	.	.	#	.	#	.	#	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	.	4310		
52.	.	.	.	.	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4311	
53.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4312	
54.	.	.	.	.	.	#	.	.	#	.	.	.	.	#	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	#	.	5202		

Norway 1865

Male linkages with at least 16 year difference

Combinations at least twice as often as expected.

Combinations must occur at least once in every 10,000 pairings.







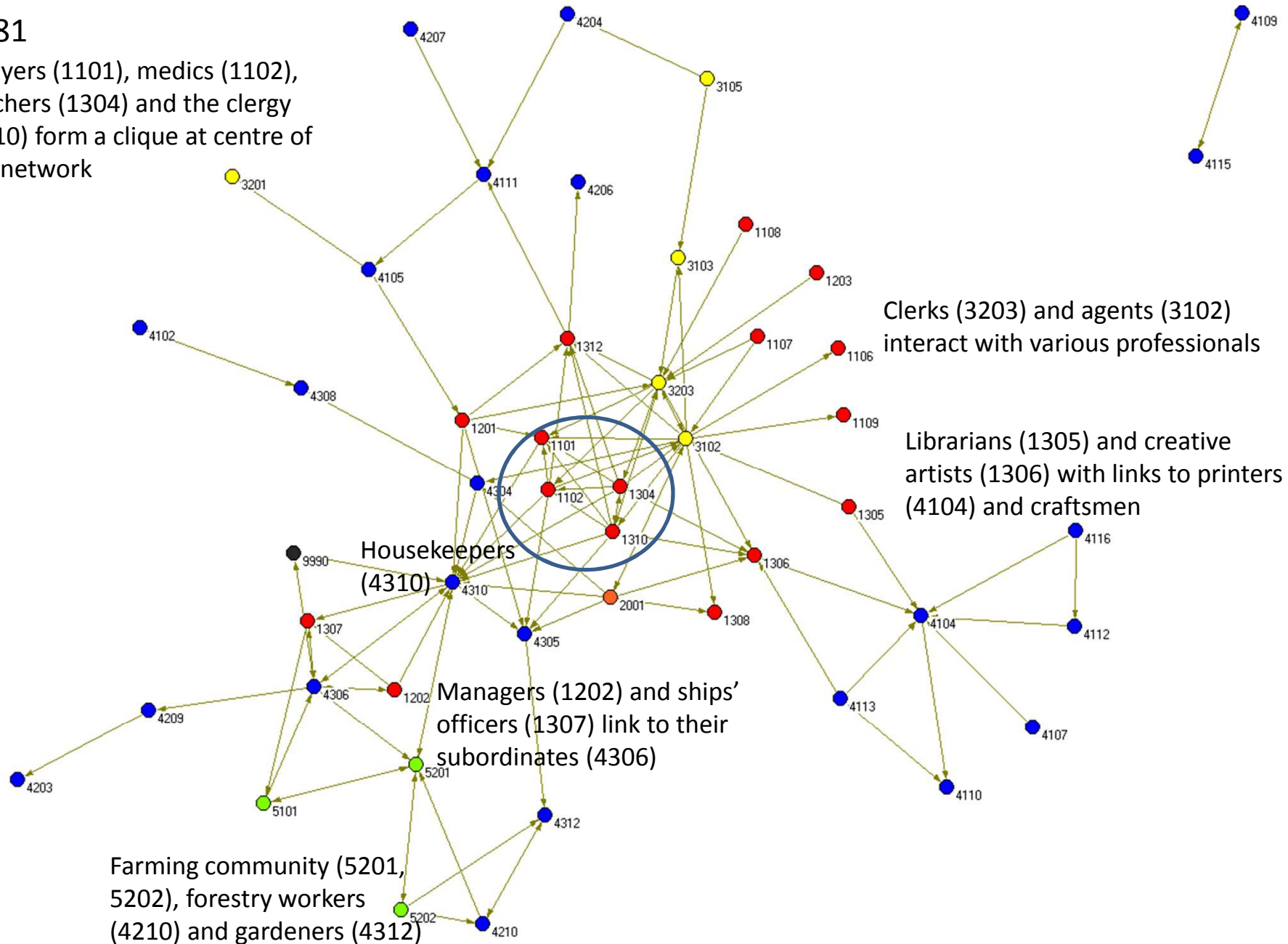
	Canada	Norway	Scotland	USA
Cases	123,749	54,067	261,187	22,349
Links	101	136	111	208
Microclasses (older cohort)	45	50	41	45
Microclasses (younger cohort)	35	38	39	41
Strongest bond (* times expectation)	239	146	19	55
Network: Degree centrality	.10	.14	.12	.18
Network: Closeness centrality	.23	.23	.27	.26
Network: Components	2	1	2	1
Network: Distance	10	12	7	5
Network: average distance	3.8	3.7	3.2	2.6

*Note, for Canada and Scotland closeness centrality refers to largest component only.*

# Scotland

1881

Lawyers (1101), medics (1102), teachers (1304) and the clergy (1310) form a clique at centre of the network



# Inter-generational professional sector or students living together?

wocc	Freq.	Percent	Cum.
1101	175	50.58	50.58
1102	71	20.52	71.10
1304	66	19.08	90.17
1310	34	9.83	100.00
Total	346	100.00	

Older cohort mostly  
teachers (1304) and  
clergy (1310)

Younger cohort mostly  
lawyers (1101) and  
medics (1102)

Stark differences partly, but not wholly, attributable  
to cohort effects of professions

% who are older members	1101	1102	1304	1310
Prof. sharers	11.2	40.8	65.1	81.7
All sharers	17.1	38.0	32.6	73.6

	Lawyers	Medics	Teachers
Medics	36/9		
Teachers	73/7	32/10	
Clergy	73/6	30/3	49/25

Older members in rows, younger in columns

Teachers and the clergy have lots of young lawyers and medics living with them. Lawyers generally younger than medics, but clergy and teachers similar ages.

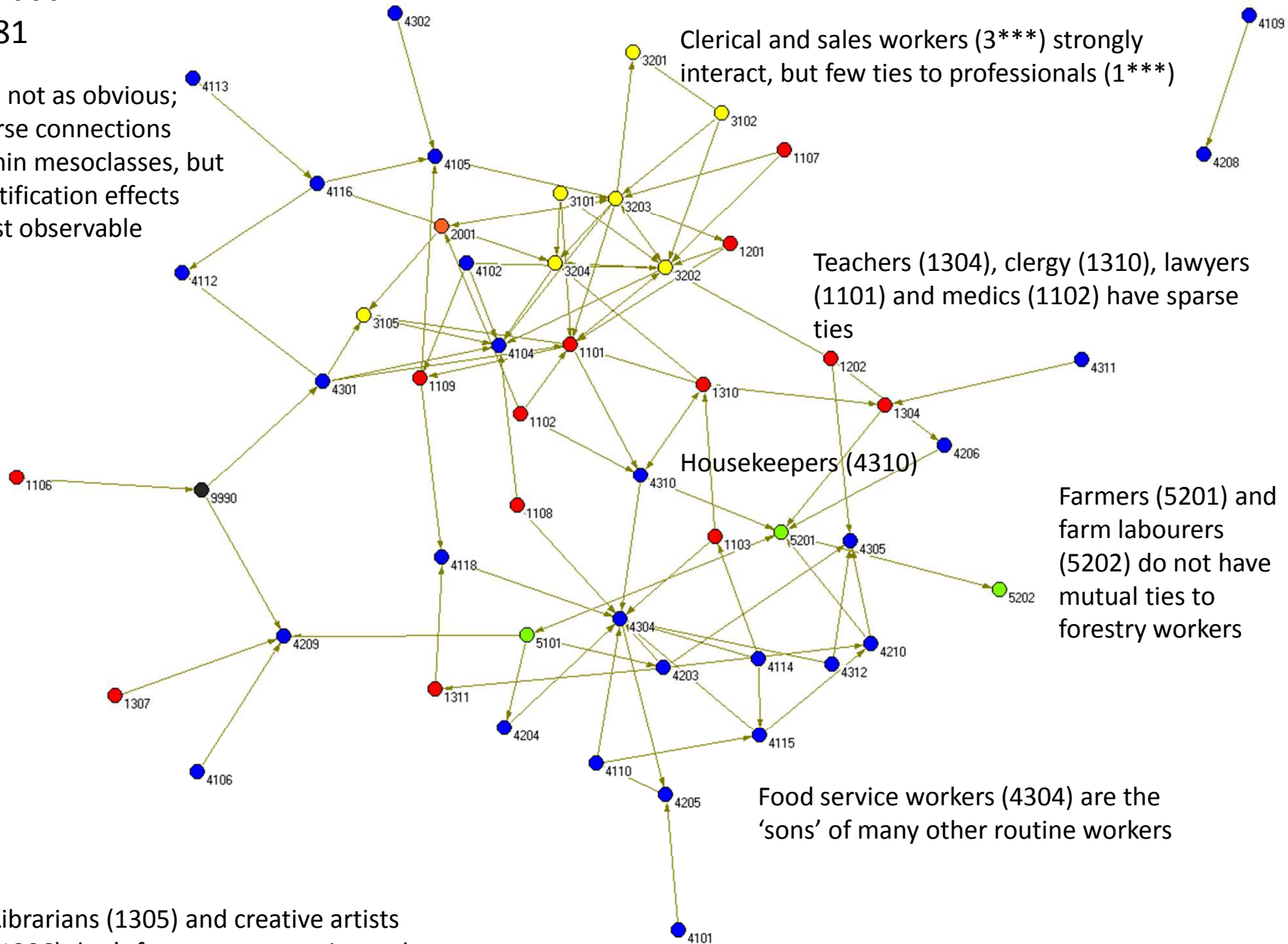
(sum) one	Freq.	Percent	Cum.
1	277	90.23	90.23
2	27	8.79	99.02
3	2	0.65	99.67
9	1	0.33	100.00
Total	307	100.00	

Generally just 1 or 2 professional pairings per household

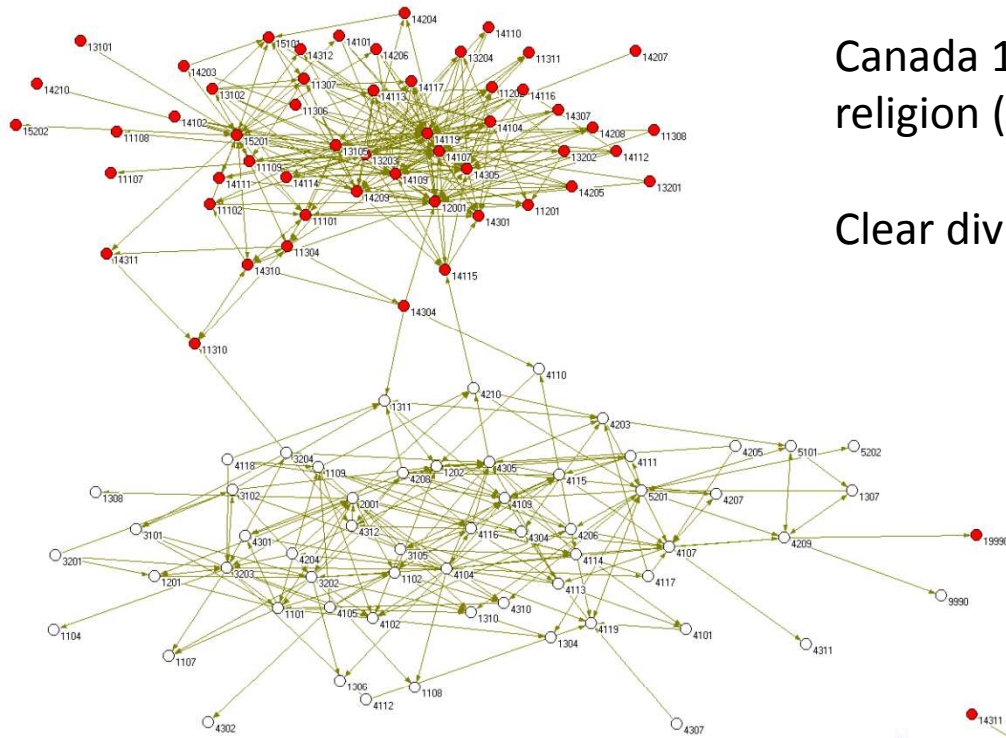
Signs that educated adults have educated children, irrespective of sector?

# Canada 1881

Ties not as obvious;  
sparse connections  
within mesoclasses, but  
stratification effects  
most observable

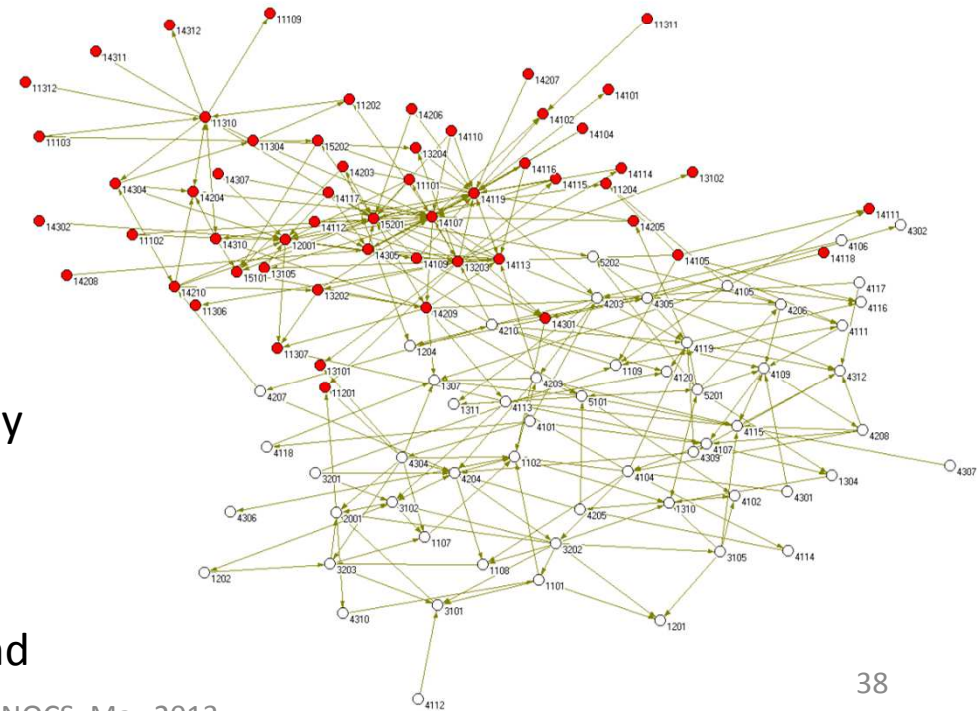


Librarians (1305) and creative artists (1306) don't form any strong ties and aren't represented



Canada 1881 (left) with microclasses split by religion (red=catholic; white=non-catholic).

Clear division on religion grounds in 1881.



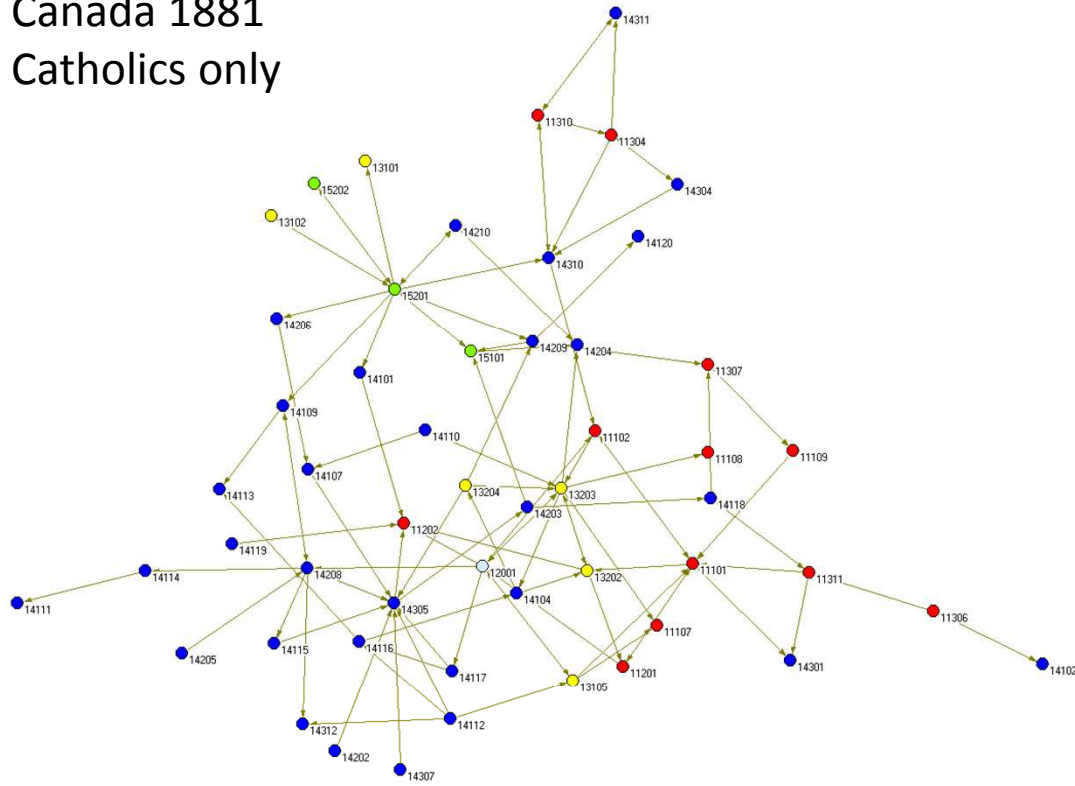
Canada 1891 (right) with microclasses split by religion (red=catholic; white=non-catholic).

Religious divide continues, but much more common for households to cross religions and microclasses.

Canada (by religion)	1881	1891
Cases	92,048	22,084
% Roman Catholic	33.1%	28.6%
% Catholics with Catholic alter	84.1%	60.6%
% non-Catholics with Catholic alter	8.2%	17.4%
Mean HISCAM (All cases) (Standard deviation)	58.0 (10.9)	57.7 (11.4)
Mean difference in HISCAM (all cases) (Standard deviation)	9.2 (11.5)	10.1 (11.6)
% HISCAM difference < 1/2 s.d.		
... (Catholic – Catholic)	52.0%	51.7%
... (non-Catholic to non-Catholic)	51.5%	49.3%
... (Catholic to non-Catholic)	45.5%	44.4%
% HISCAM difference > 2 s.d.		
... (Catholic to Catholic)	11.4%	16.6%
... (non-Catholic to non-Catholic)	12.8%	11.9%
... (Catholic to non-Catholic)	12.4%	11.8%



Canada 1881  
Catholics only

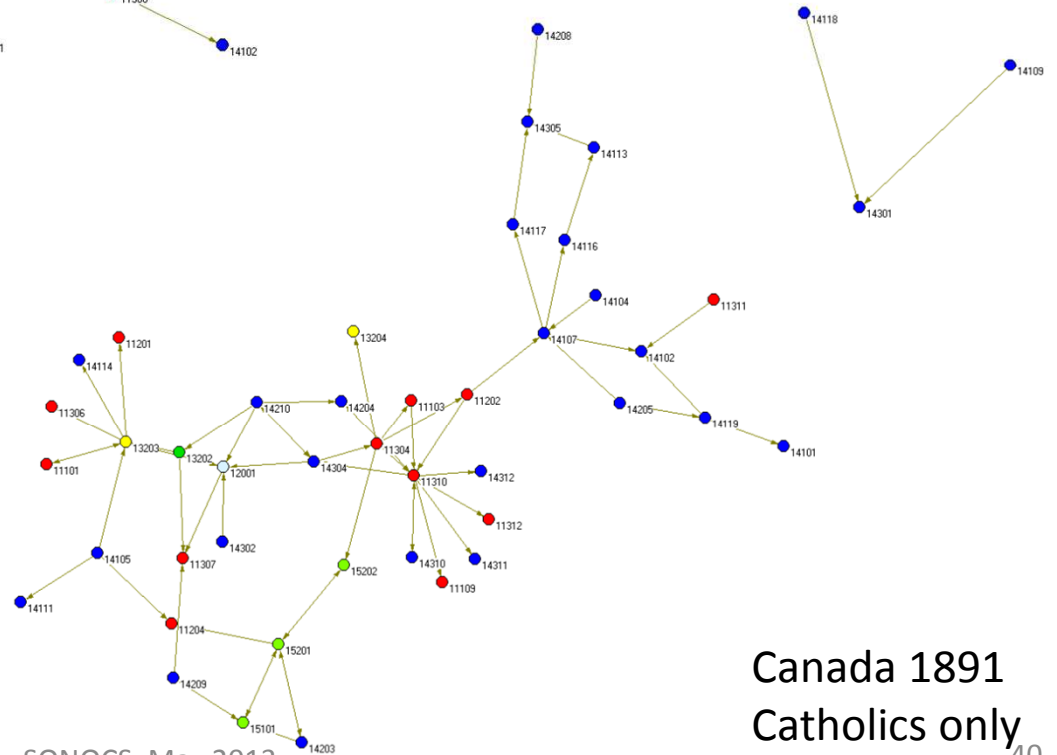


	1881	1891
Degree centrality	.09	.10
Betweenness	.20	.06
Closeness	.16	.22
Ave. distance	4.7	3.0

Story of social change?

Clergy (11310) appear to bring together many occupations in 1891, but structure largely declined since 1881.

Decline of structure shows shorter paths between microclasses, but fewer different paths for accessing links. Links appear to be more on grounds of 'situs'.



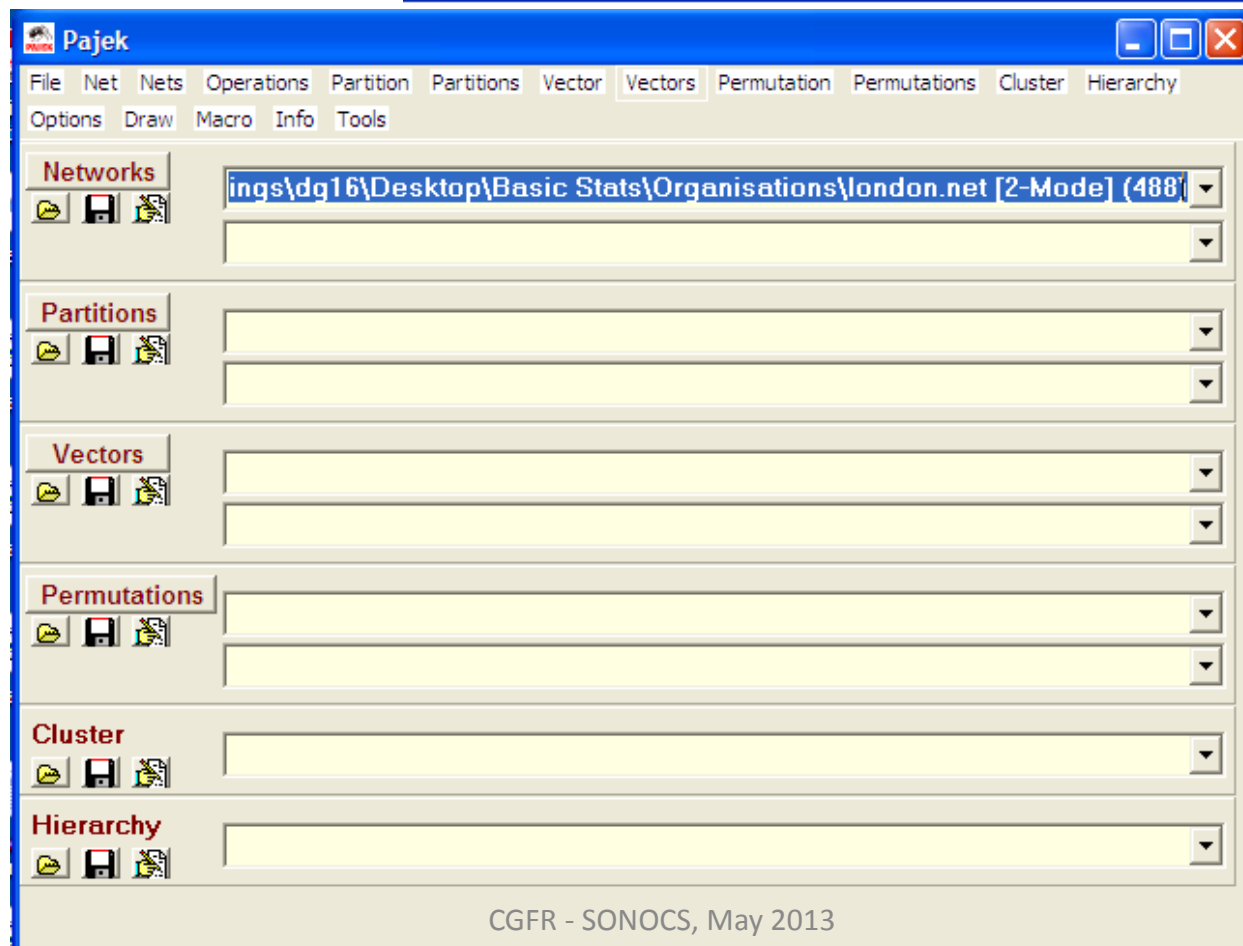
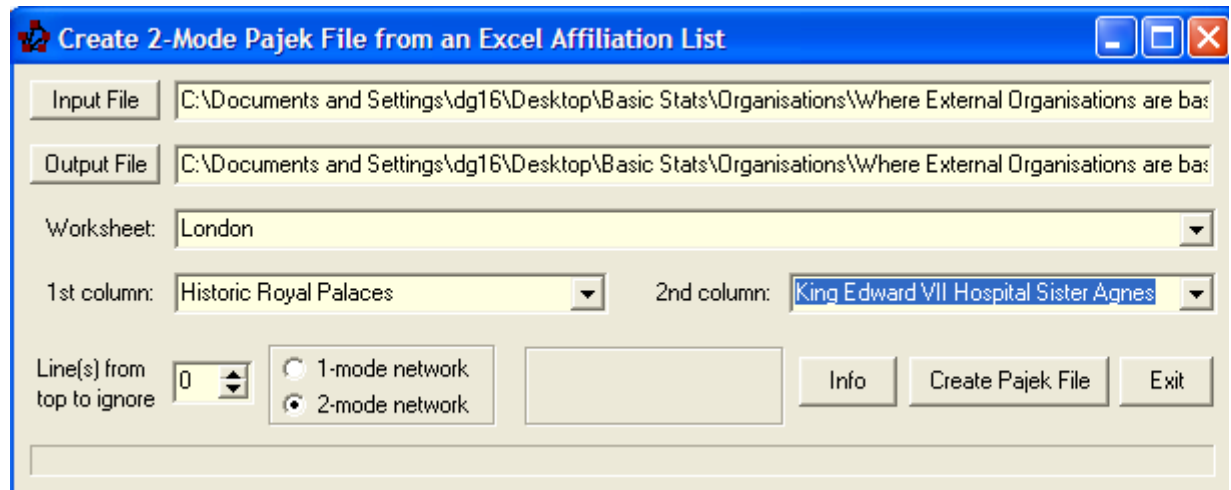
Canada 1891  
Catholics only

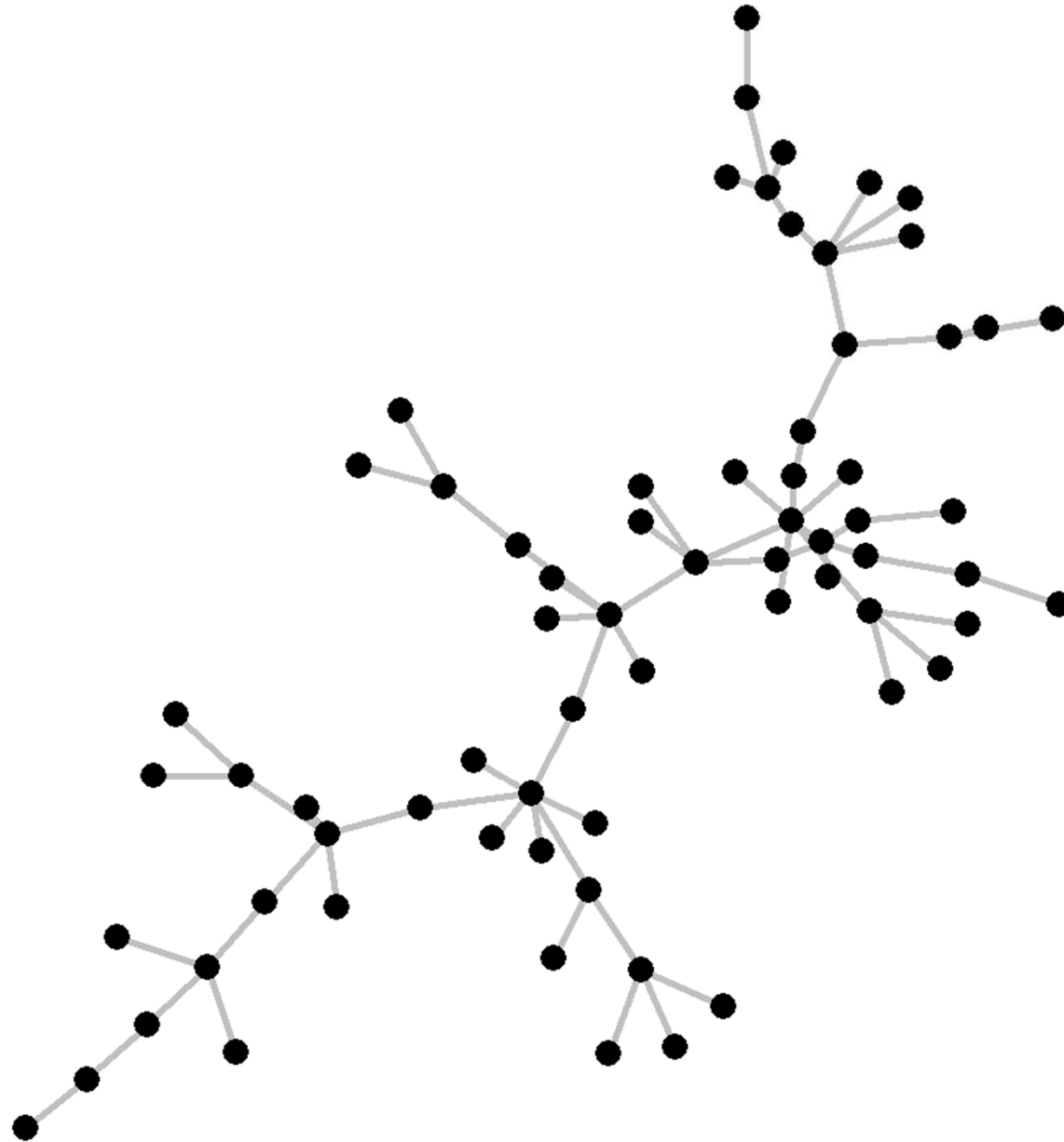


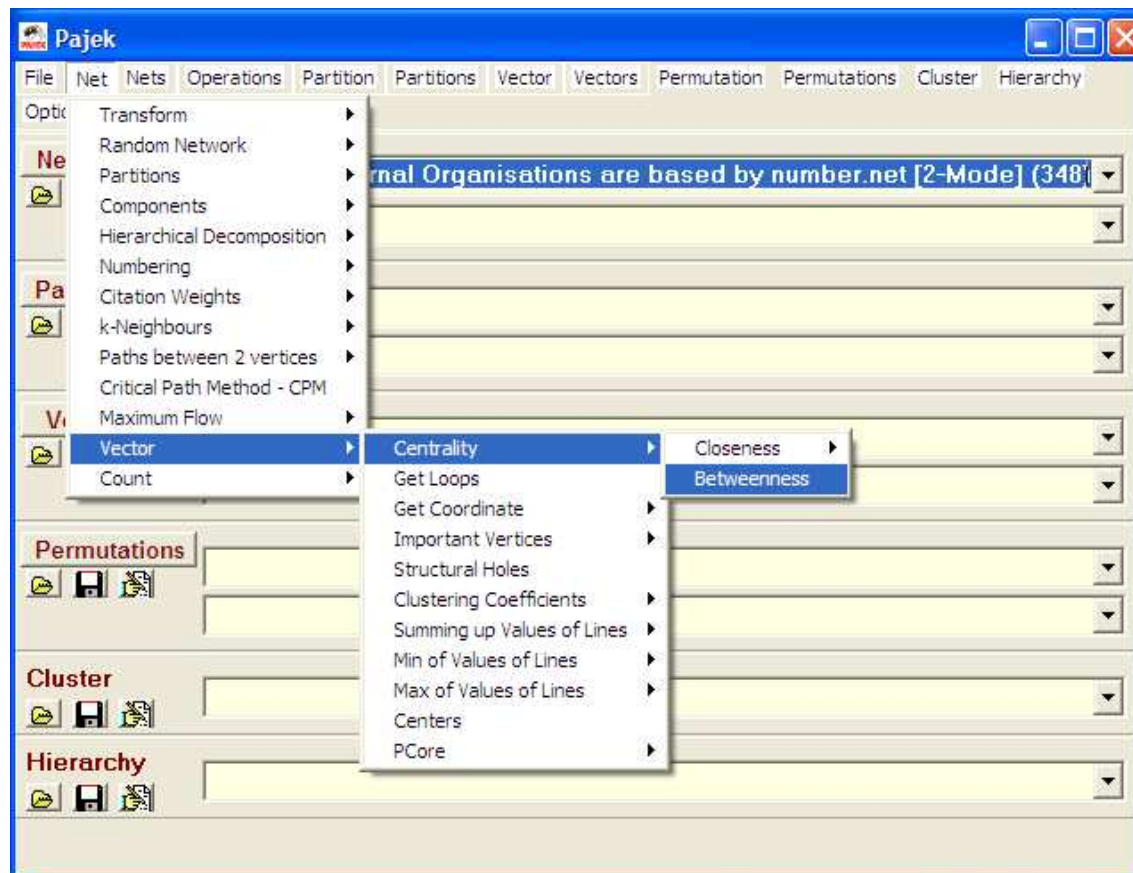
# How do I do it?

- Pajek is a “very simple” software which allows you to create network (from text or spreadsheet packages)
- <http://pajek.imfm.si/doku.php> (type ‘Pajek’ into Google and it will come up)
- This allows you to download the programme for free. You can also download Excel2Pajek.
- The resources section has datasets you can play with if you are particularly interested.

	C	E	I	Q	R	S	T	U	V	W
1	11278	Historic Royal Palaces	King Edward VII Hospital Sister Agnes							
2	10175	English Partnerships	East London University							
3	11379	Economic and Social Research Council	Queen Mary & Westfield College, London							
4	11703	Training and Development Agency for Schools	Petchy Academy							
5	10484	General Teaching Council for England	Richmond Theatre Trust							
6	10484	General Teaching Council for England	St Benedict's - Ealing							
7	11067	Stonebridge HAT	Building Exploratory - Hackney							
8	10124	Cafcass	London Director's Child Care Group							
9	11523	Bank of England	Institute of Employment Studies, London							
10	10578	BBC	New London Orchestra							
11	10578	Pension Regulator	New London Orchestra							
12	10619	Consumer Council for Water	RNH (West End) Ltd							
13	11683	Royal Naval Museum	Chelsea Hospital NHS Trust							
14	10717	Design Council	Architecture Association School of Architecture							
15	10594	Railway Heritage Committee	London Mozart Players							
16	11028	Sector Skills Development Agency	Young Vic Company							
17	11734	Office of the Immigration Services Commissioner	University College, London							
18	11075	Stonebridge HAT	Brent Irish Advisory Service							
19	10943	London Thames Gateway Development Corporation	Stratford Development Partnership Ltd							
20	10943	London Thames Gateway Development Corporation	Stratford Tomorrow							
21	10943	London Thames Gateway Development Corporation	Thames Gateway London Partnership							
22	10943	London Thames Gateway Development Corporation	Newham Primary Care Trust							
23	11709	Training and Development Agency for Schools	Institute of Education, London							
24	10709	Design Council	Royal College of Art							
25	10331	Standards Board for England	Queen Mary & Westfield College, London							
26	10581	BBC	Almeida Theatre - Islington							
27	11268	Higher Education Funding Council for England	Greenwich Hospital							
28	10537	Human Tissue Authority	St George's, London							
29	11530	Bank of England	Finsbury Technology Trust							
30	11530	Bank of England	London Business School							
31	11509	Health Protection Agency	London University							
32	11318	Arts and Humanities Research Council	City University							
33	10174	English Partnerships	Lewisham Hospital NHS Trust							
34	11223	National Biological Standards Board	London University							
35	10254	PITO	Defence Academy							
36	10946	London Thames Gateway Development Corporation	London & Quadrant Housing Trust							
37	10186	English Partnerships	Barking Riverside Ltd							
38	10186	English Partnerships	Basilidon Renaissance Partnership							
39	10186	English Partnerships	London Housing Board							
40	11249	Ofcom	Donmar Warehouse Theatre - London							
41	11273	Higher Education Funding Council for England	North London University							
42	11448	Qualifications and Curriculum Authority	Harrow School							
43	10325	Standards Board for England	Tower Hamlets Synod							
44	10325	Standards Board for England	Queen Mary & Westfield College, London							
45	11290	Historic Royal Palaces	King Edward VII Hospital Sister Agnes							
46	10233	National Forest	Middlesex University							
47	10182	English Partnerships	Whitgift School							







Editing Vector: 1. Betweenness centrality in N1 (348)

Redisplay

Vertex	Val	Label
1	0.0041645	Historic Royal Palaces
2	0.0381151	English Partnerships
3	0.0172746	Economic and Social Research C
4	0.0047565	Training and Development Agenc
5	0.0000167	General Teaching Council for E
6	0.0001666	Stonebridge HAT
7	0.0093930	Cafcass
8	0.1400461	Bank of England
9	0.0542897	BBC
10	0.0000000	Pension Regulator
11	0.0000500	Consumer Council for Water
12	0.0000000	Royal Naval Museum
13	0.0130440	Design Council
14	0.0041645	Railway Heritage Committee
15	0.0000167	Sector Skills Development Ager
16	0.0041645	Office of the Immigration Serv
17	0.0716480	London Thames Gateway Develop
18	0.0041645	Standards Board for England
19	0.0124436	Higher Education Funding Counc
20	0.0000000	Human Tissue Authority
21	0.0279935	Health Protection Agency
22	0.0328152	Arts and Humanities Research C

# Key introductory texts

- Introduction to Pajek software which excellent description of how and why you would use SNA
  - de Nooy, W., Mrvar, A., & Bataglj, V. (2012) *Exploratory Social Network Analysis with Pajek*. Cambridge: Cambridge University Press. 2<sup>nd</sup> edition.
- Introduction to SNA theory
  - Knoke, D., & Yang, S. (2008) *Social Network Analysis*. London: Sage. 2<sup>nd</sup> edition.
  - Scott, J. (2000) *Social Network Analysis*. London: Sage. 2<sup>nd</sup> edition.
- Detailed introduction to various SNA applications
  - Scott, J., and Carrington, P.J. (2011) *The SAGE Handbook of Social Network Analysis*. London: Sage.
- Comprehensive overall of all the underlying statistical theories
  - Wasserman, S., & Faust, K. (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

<http://www.camsis.stir.ac.uk/sonocs/workshops>

- In April 2012 we presented a one-day workshop to the Historical Demography workshop
- Session covered SNA and SID (another statistical method which focuses on similarities rather than connections)
- Slides and copies of workshop materials available on our website
- Data files are available upon request

david.griffiths@stir.ac.uk