

# Application: Using SNA to analyse occupational structure in the past

Dave Griffiths & Paul Lambert, University of Stirling

With additional inputs from Richard Zijdeman, Ineke Maas, Marco van Leeuwen (University of Utrecht) and Ken Prandy (University of Cardiff)

[www.camsis.stir.ac.uk/sonocs](http://www.camsis.stir.ac.uk/sonocs)

Presented to the Historical Demography Scientific Research Network workshop, 26 April 2012, Utrecht

# Microclasses

Norway, 1865



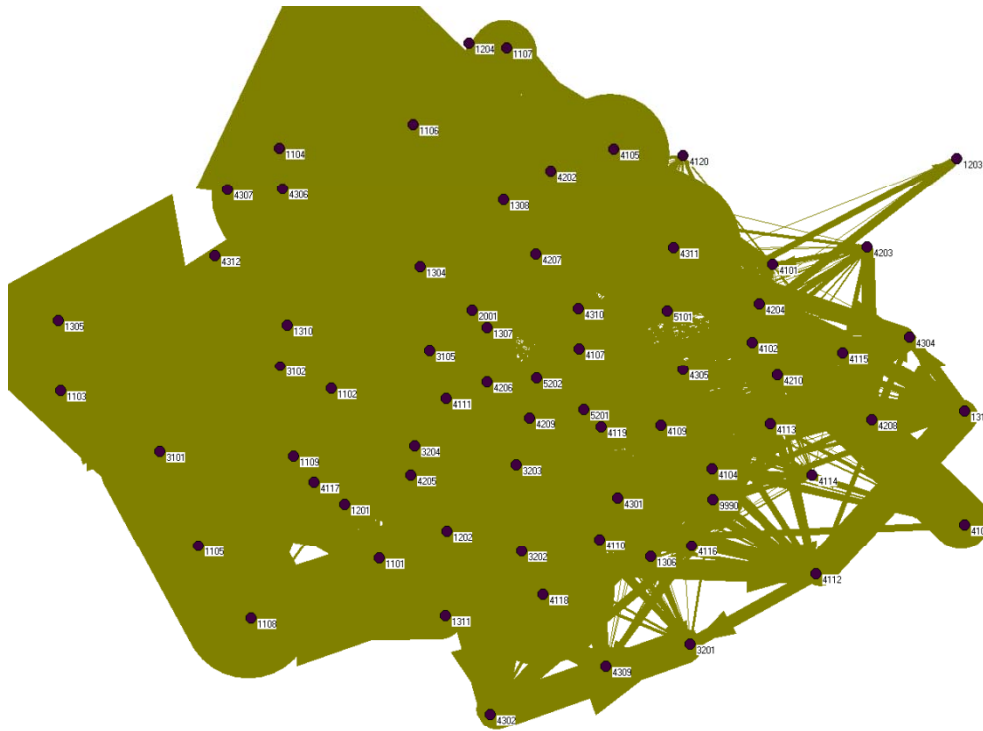
(Excluding diagonals,  $n^*20$ )

SONOCS/WOG, April 2012

# Norway 1865

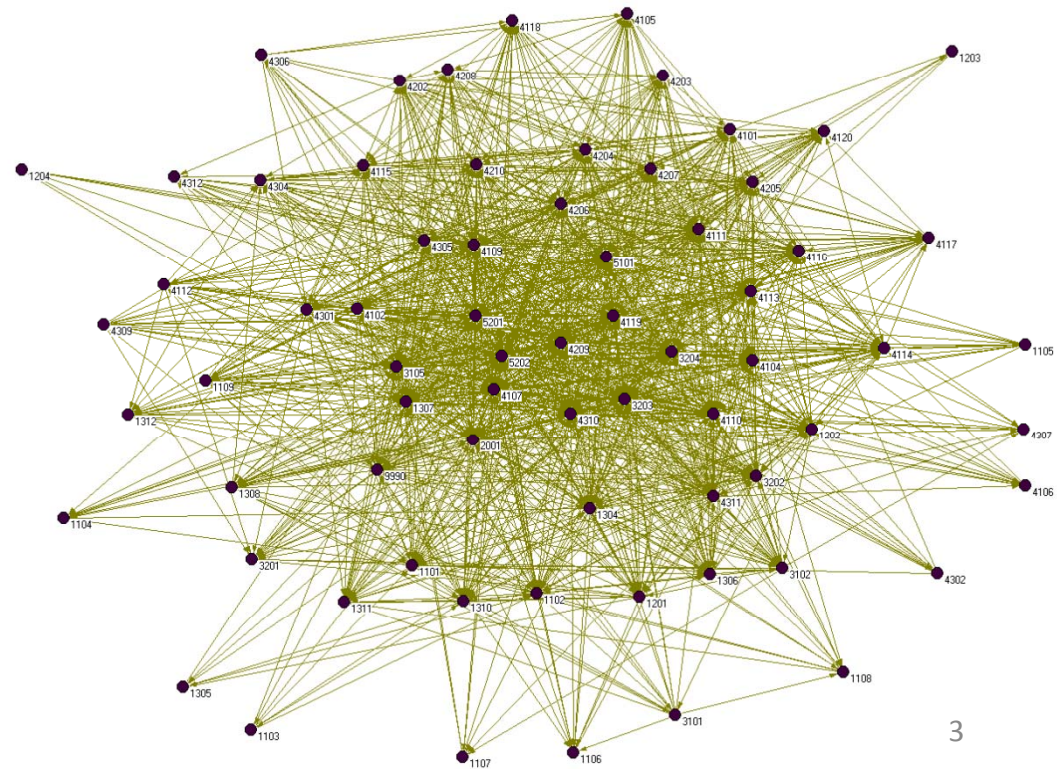
Male-male microclass combinations  
of at least 16 year difference.

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?

- Remove all cases of people performing the **same occupations** (structural relationships more readily explored by looking at mobility than immobility)
- Value of relationship: must occur at least **X** times more than expected by chance (occurs more often than if they was no occupational structure)
- Frequency of relationship: must occur in at least **Y,000 combinations** (exclude cases where over-representation occurs with a small number of cases to avoid artificially finding a tie due to low expected values and also to remove those combinations contributing little to occupational structure)
- Apply confidence intervals to the value of the **relationships** (only include cases where we are confident there is over-representation, rather than cases where it might exist)

# Required 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)
```

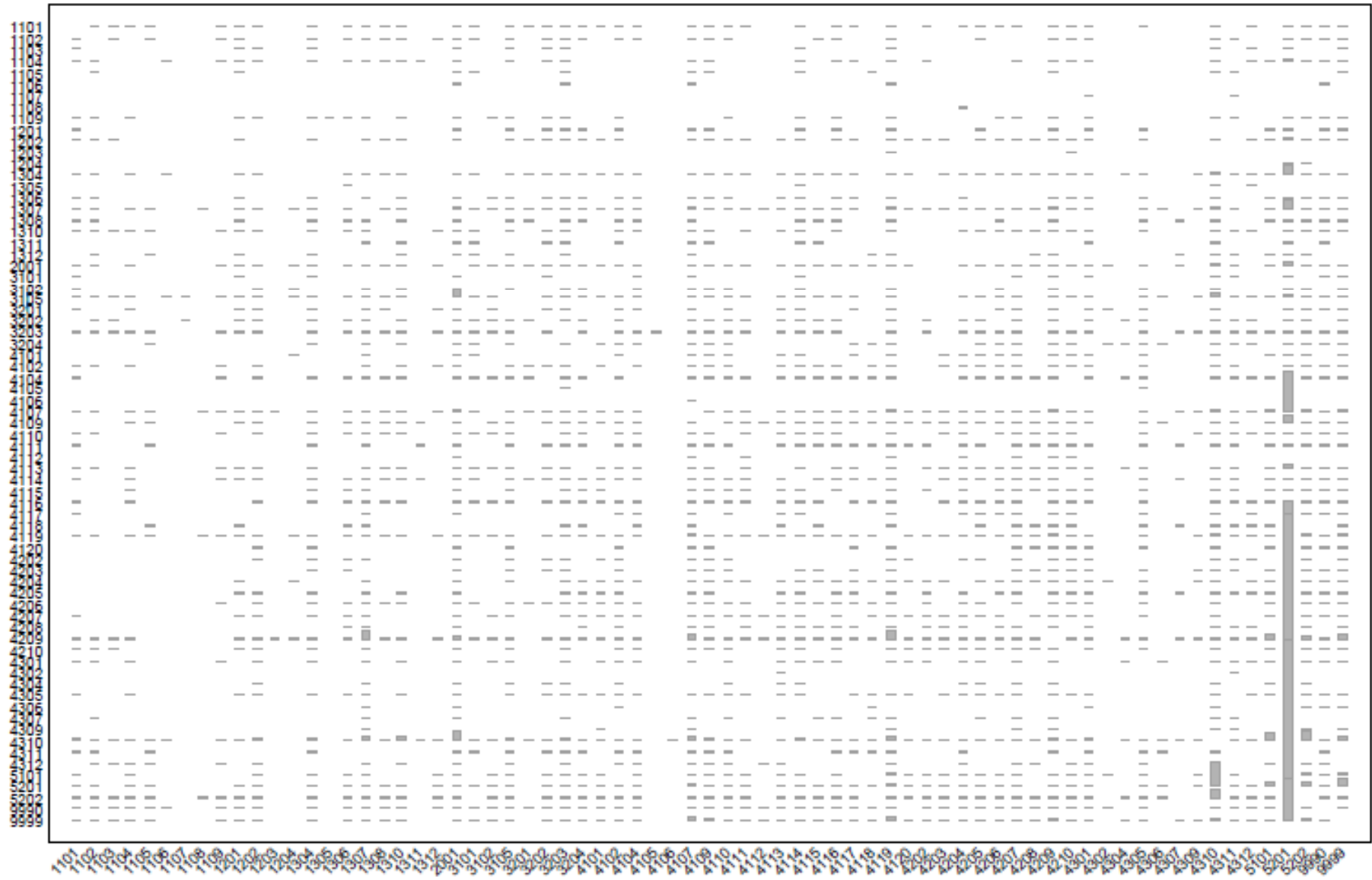
```
capture drop pro_obs
gen pro_obs = freq/tot
capture drop pro_exp
gen pro_exp = ewocc/tot
capture drop pro_min
gen pro_min = pro_obs - staner
capture drop pro_max
gen pro_max = pro_obs + staner
capture drop value
gen value = pro_obs / pro_exp
capture drop val_min
gen val_min = pro_min / pro_exp
capture 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"
```

# Limitations

- Each identified over-represented link produces less opportunity for further linkages (If 30% of ties are to an occupation only 3% of people perform, there are only 70% of links left for remaining 97%; a combination needs to occur 2.77 times more commonly after those structural to pass a 2.0 times more common threshold)
- ‘False’ combinations can therefore mask real combinations in the occupational structure (If housekeepers are commonly linked to through employment rather than natural cohabitation this can influence the potential for identifying other links)
- Setting the criteria for observing a link can define the observed structure (It is plausible a different interpretation will be made if the thresholds were altered and different linkages were analysed)

# Microclasses

Norway, 1865



(Excluding diagonals, n\*20)

SONOCS/WOG, April 2012

	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

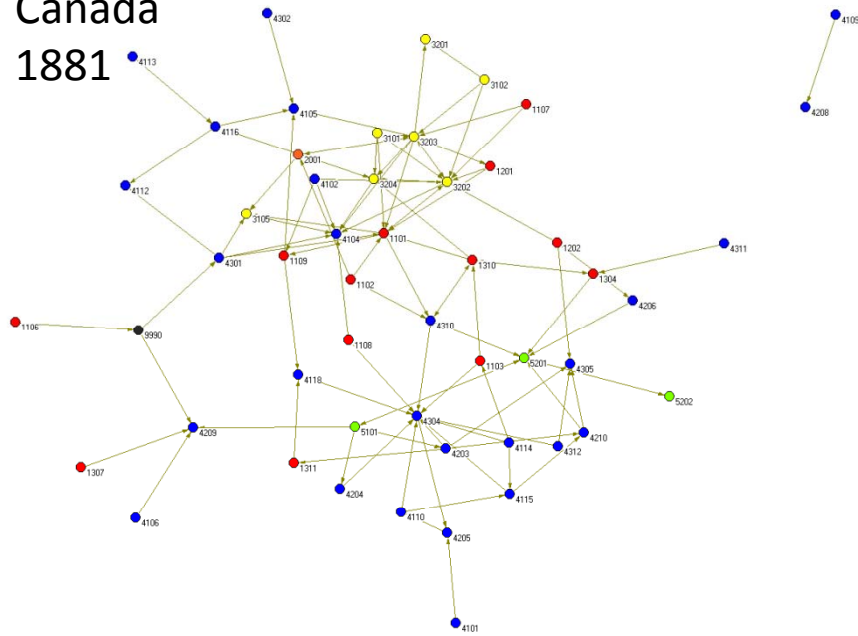
Male linkages with  
at least 16 year  
difference

Combinations at  
least twice as often  
as expected.

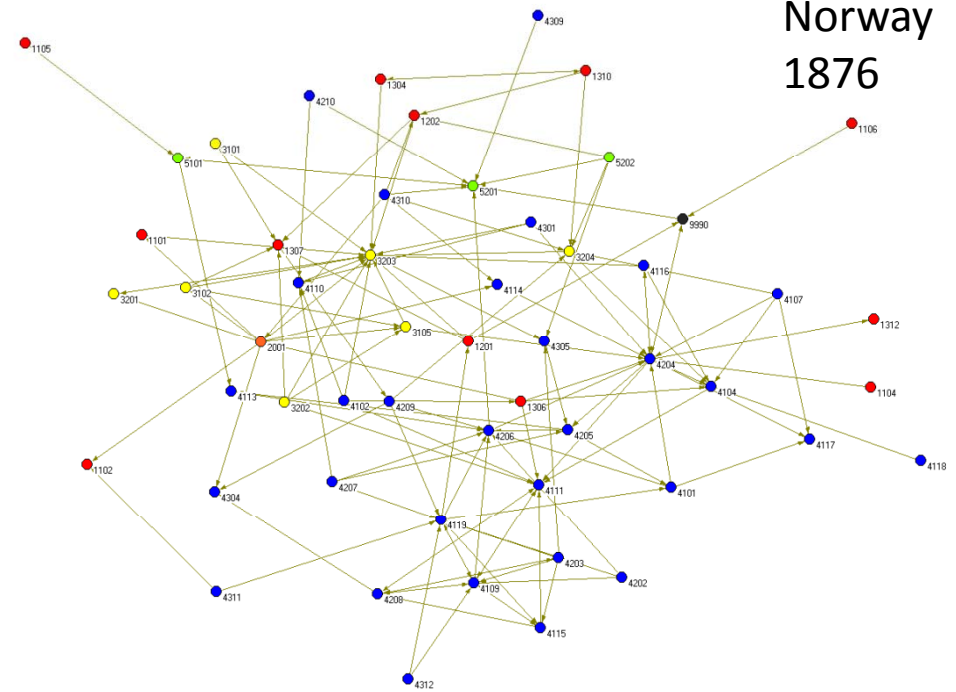
Combinations must  
occur at least once  
in every 10,000  
pairings.



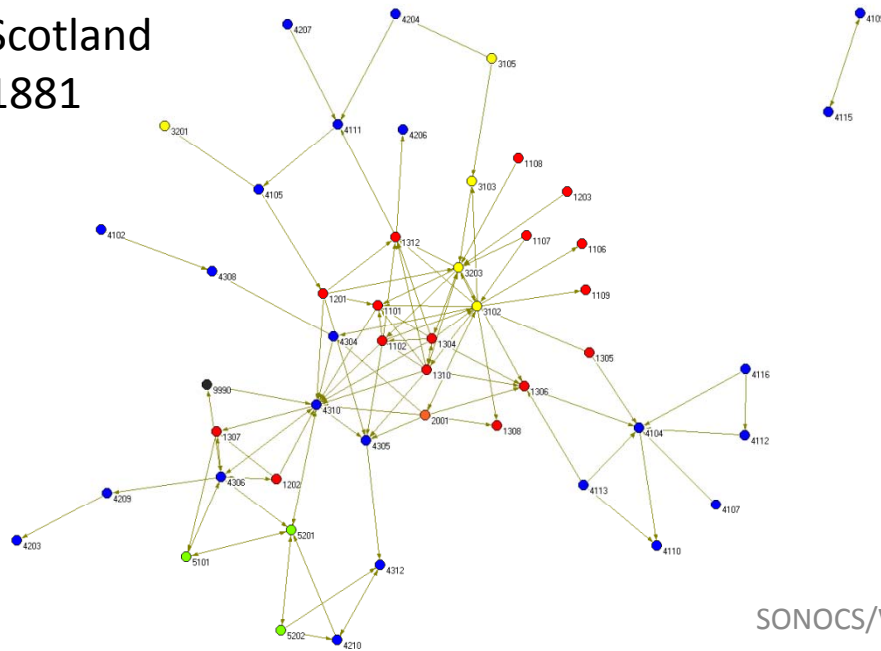
Canada  
1881



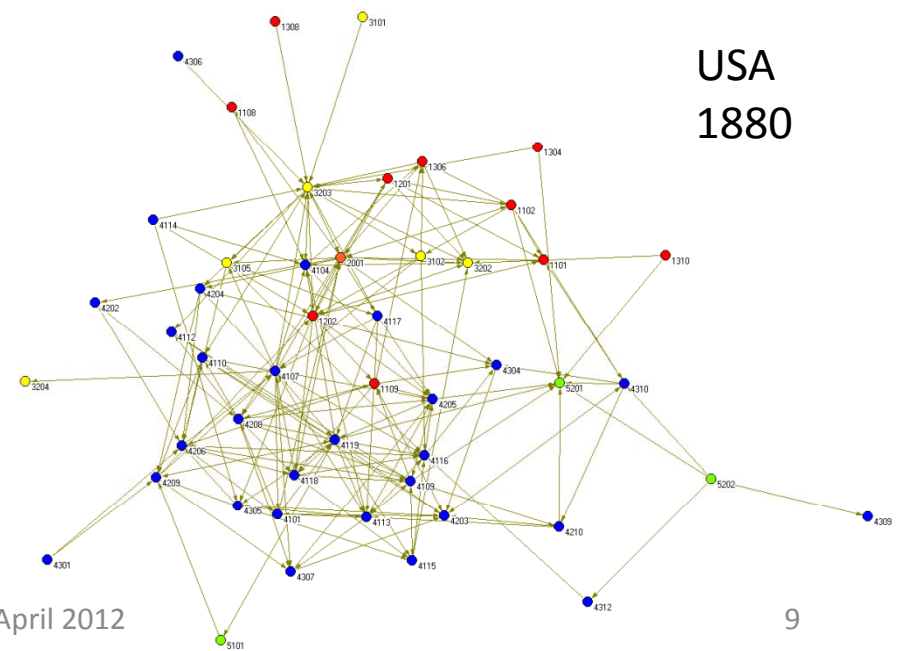
Norway  
1876



Scotland  
1881



USA  
1880

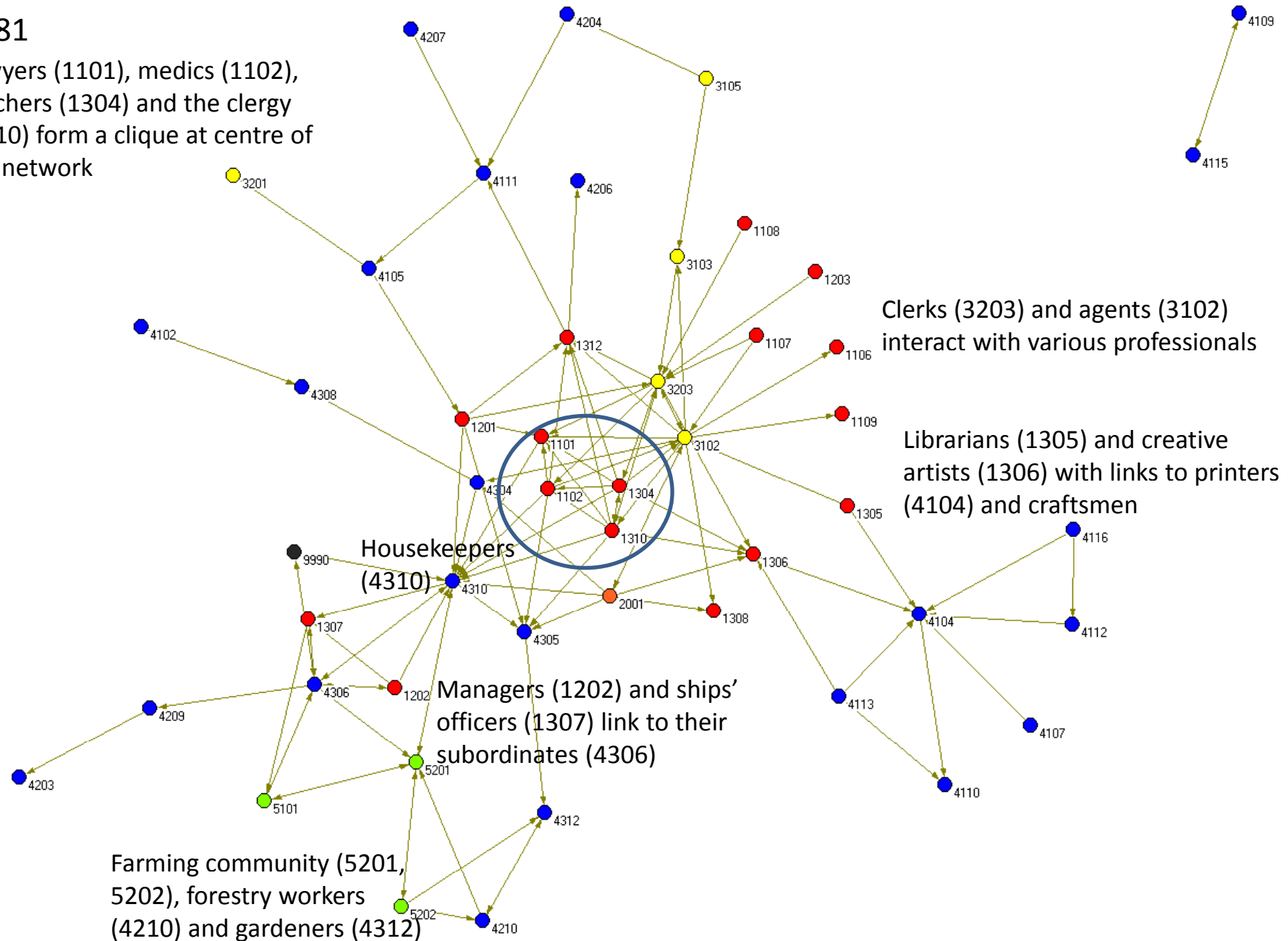


	Canada	Norway	Scotland	USA
Cases	123,749	54,067	261,187	22,349
Links	101	136	111	208
Microclasses (older cohort)	45	50	41	45
Microclasses (younger cohort)	35	38	39	41
Strongest bond (* times expectation)	239	146	19	55
Network: Degree centrality	.10	.14	.12	.18
Network: Closeness centrality	.23	.23	.27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

# 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?

hocc	Freq.	Percent	Cum.
1101	22	6.36	6.36
1102	49	14.16	20.52
1304	123	35.55	56.07
1310	152	43.93	100.00
Total	346	100.00	

Older cohort mostly  
teachers and clergy

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	

Younger cohort mostly  
lawyers and medics

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?

# Housekeepers (4310) in right position?

5 times housekeepers are older members,  
11 times younger members.

	hocc	wocc	freq	ewocc	val_mi n
1.	1101	4310	85	13. 13266	5. 770497
6.	1102	4310	53	16. 89044	2. 706895
14.	1201	4310	216	53. 90065	3. 734819
17.	1202	4310	148	56. 07312	2. 422515
25.	1304	4310	58	22. 95769	2. 194693
37.	1310	4310	119	35. 85548	3. 014707
45.	2001	4310	113	45. 11293	2. 269242
86.	4304	4310	118	52. 06091	2. 057968
91.	4306	4310	319	120. 4251	2. 500728
94.	4310	1307	35	3. 142614	9. 254817
95.	4310	4305	92	12. 94056	6. 368352
96.	4310	4306	279	41. 32797	6. 346929
97.	4310	5201	122	28. 01082	3. 961228
101.	5201	4310	1776	773. 5155	2. 241716
108.	9990	4310	218	37. 32336	5. 44542

Some of these patterns are believable (i.e., to service workers) but seems high levels of housekeepers having a place of employment, but living.

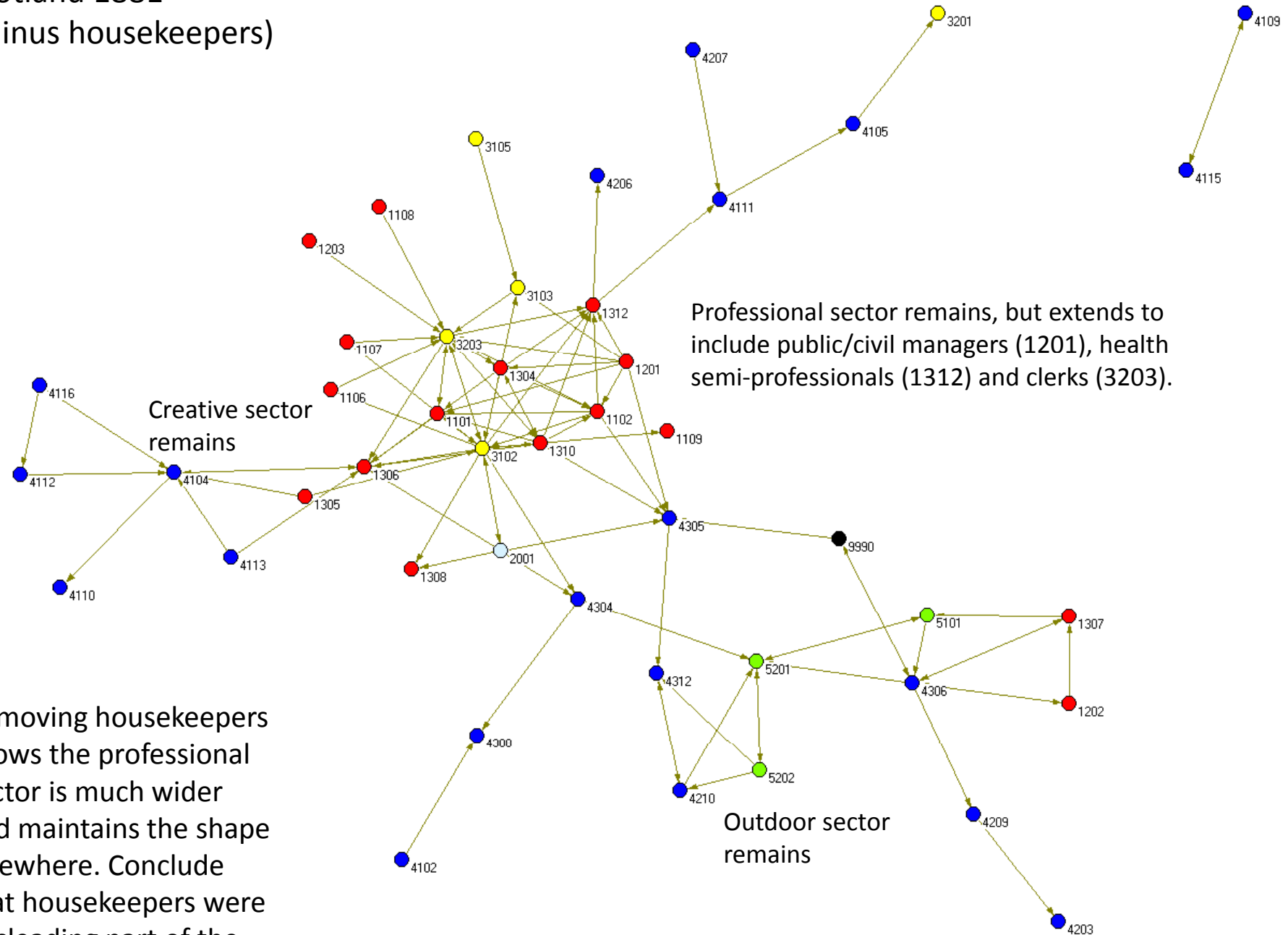
Strongest links are to ship officers (1307), mass transportation operators (4305), other service workers (4306), jurists (1101) and members of the armed forces (9990).

Are these seafarers, drivers and the military who generally work away from home?

Ties also to older health professionals (1102), public/civil/private sector managers (1201/2), teachers (1304), clergy (1310), proprietors (2001) and farmers (5201).

Ties to farmers (5201) and service workers (4306) work from older to younger and younger to older.

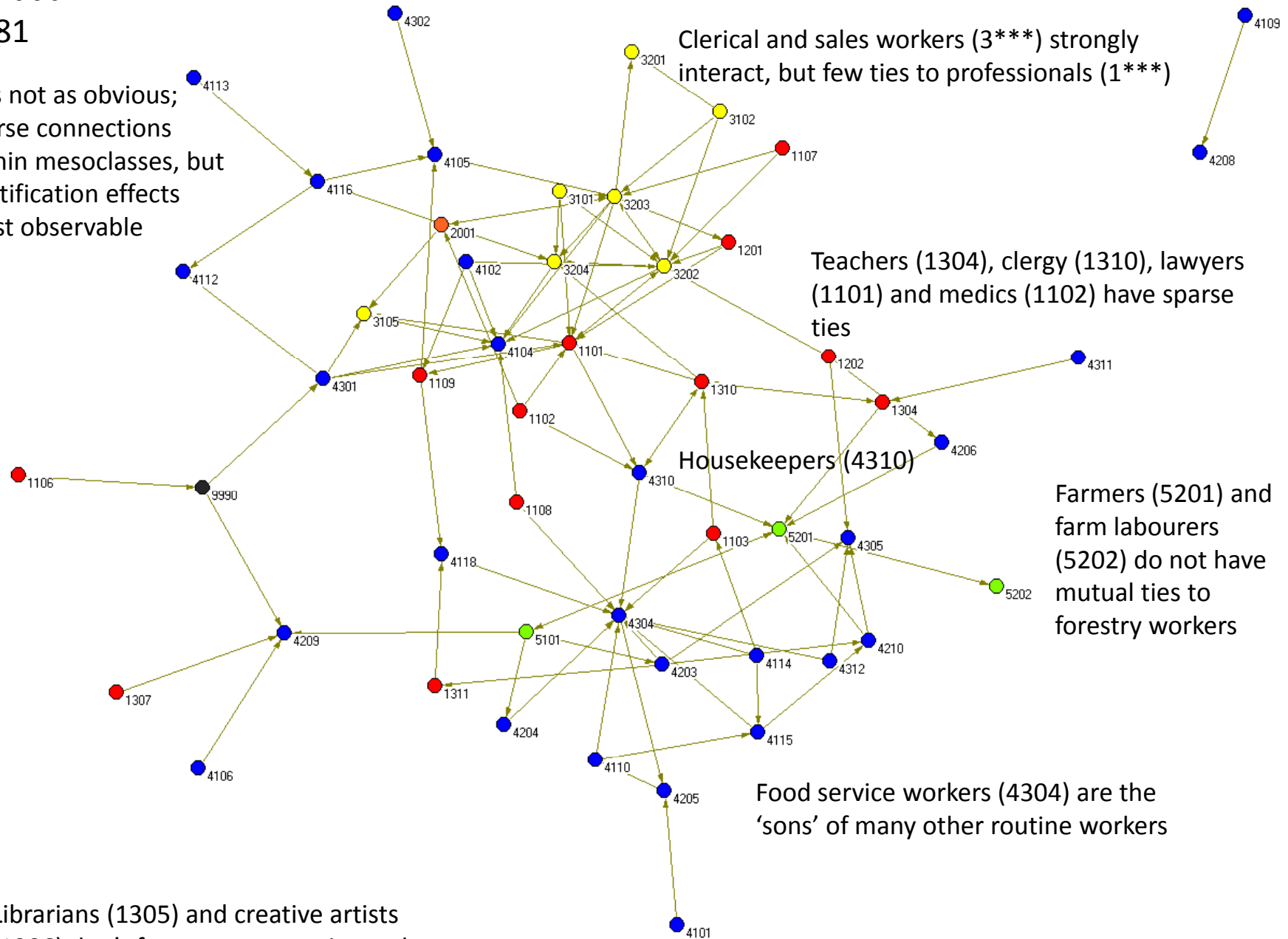
Scotland 1881  
(minus housekeepers)



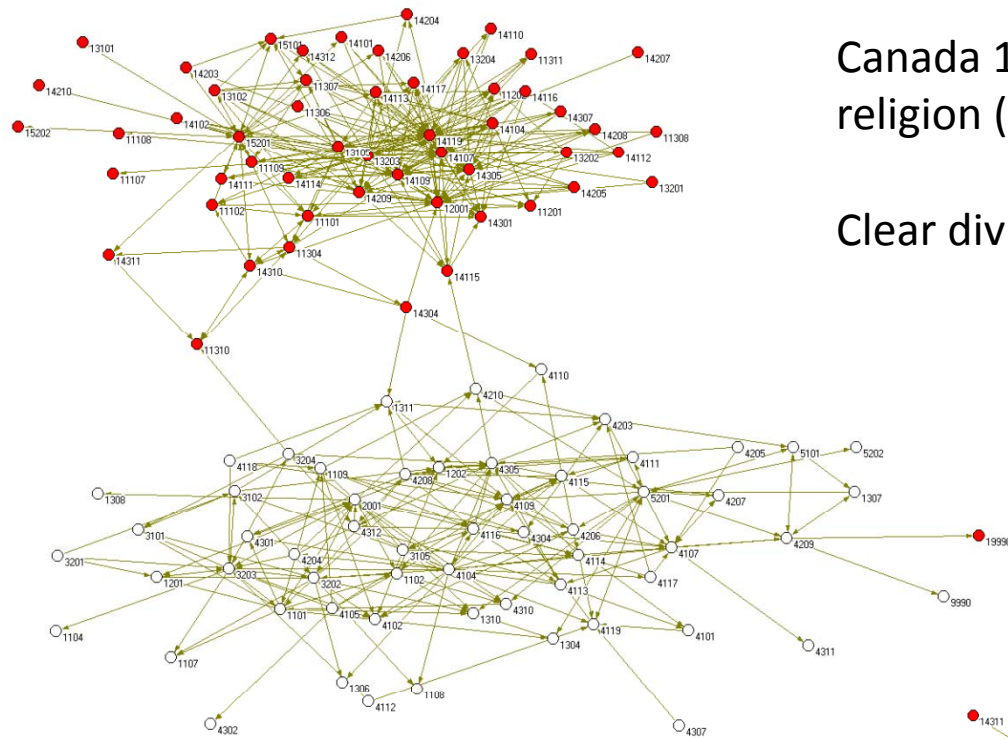
Removing housekeepers shows the professional sector is much wider and maintains the shape elsewhere. Conclude that housekeepers were misleading part of the network

# Canada 1881

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

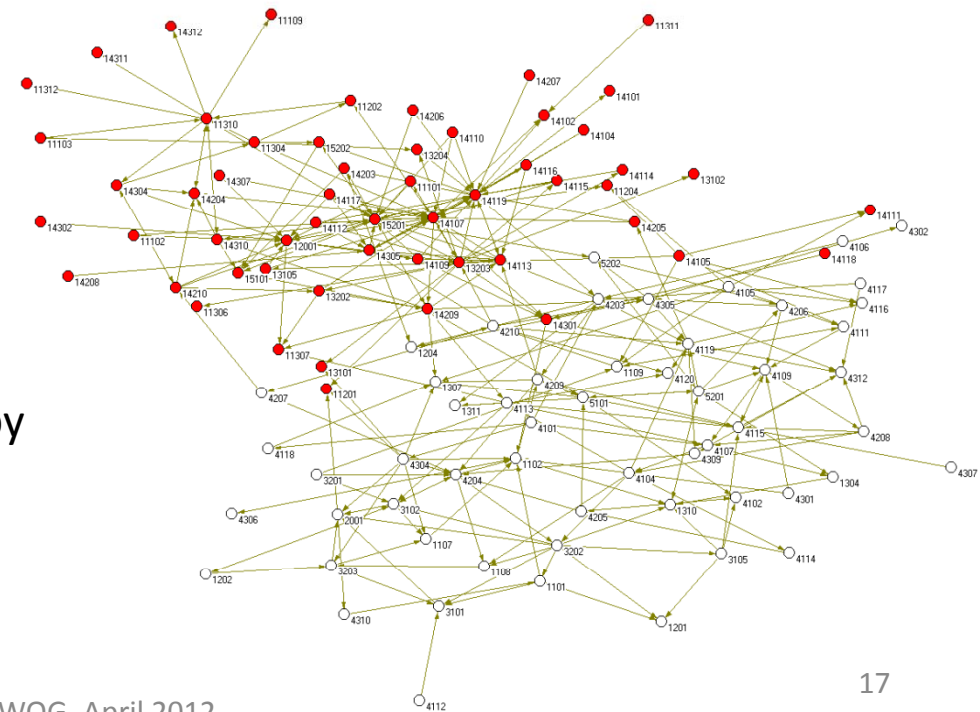






Canada 1881 (left) with microclasses split by religion (red=catholic; white=non-catholic).

Clear division on religion grounds in 1881.

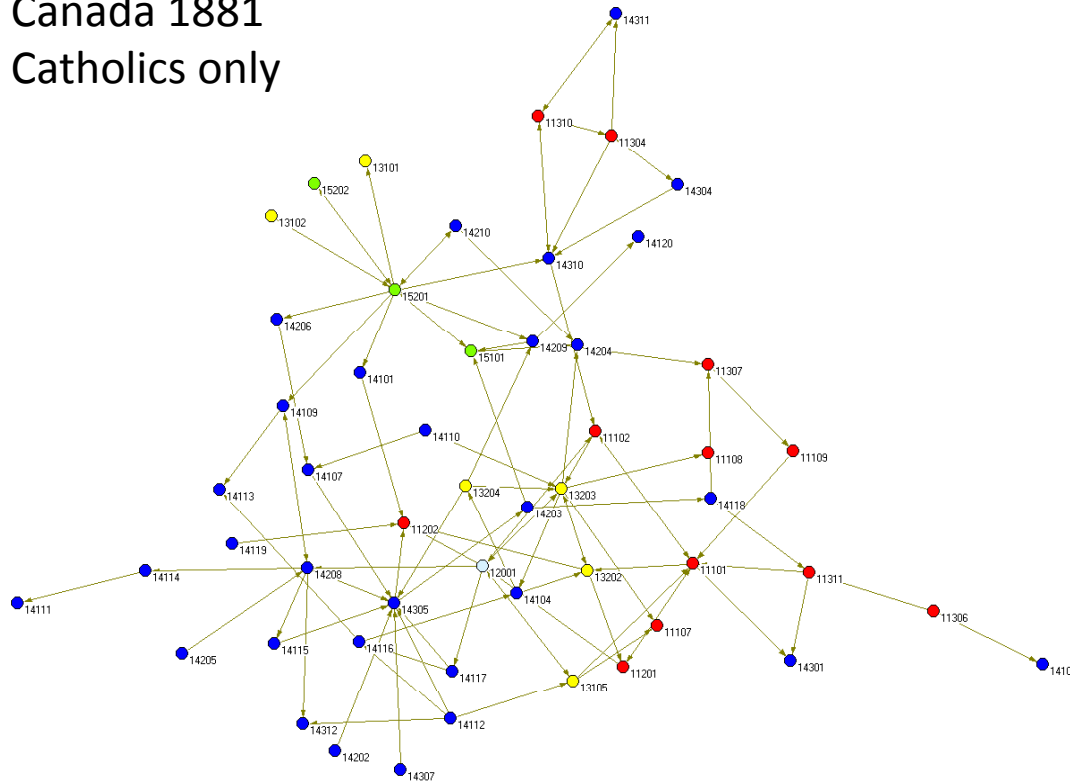


Canada 1891 (right) with microclasses split by religion (red=catholic; white=non-catholic).

Religious divide continues, but much more common for cross-religion and microclass households.

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 < $\frac{1}{2}$ 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%

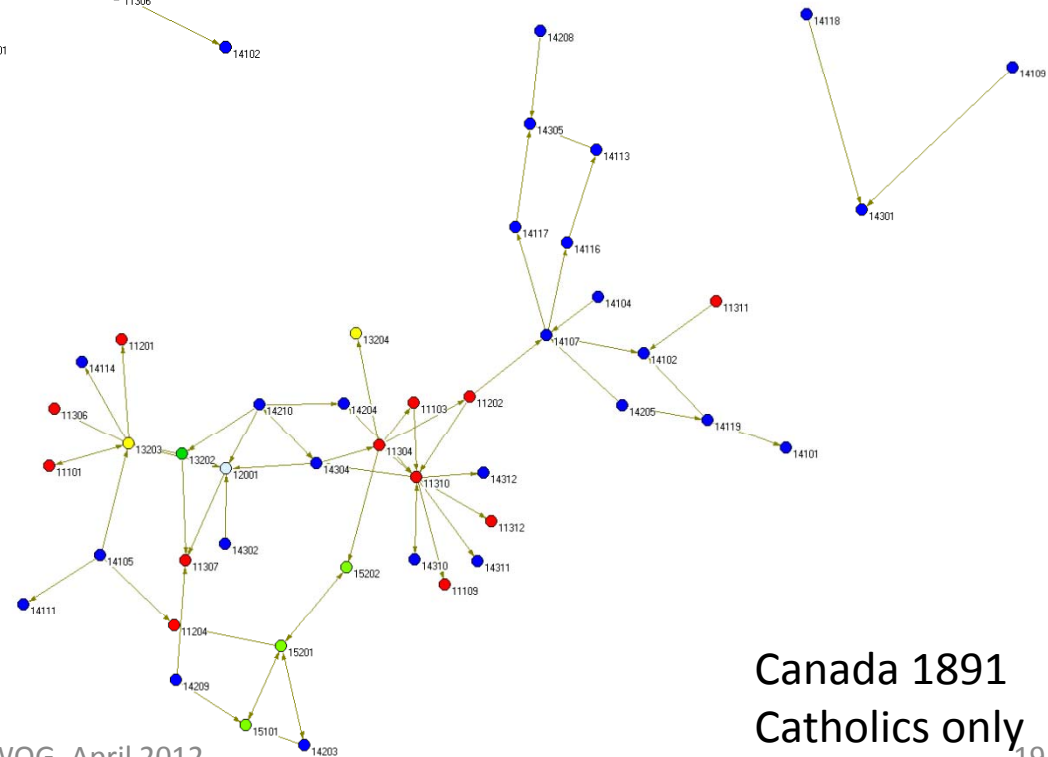
## Canada 1881 Catholics only



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

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

# Summary

- ‘Disproportionately common’ criteria for network ties between occupations is effective way of revealing social structure
- Network depictions can vary considerably
  - Value of sensitivity analysis and checking particular combinations
  - Value of combining graph and statistical summaries
- Occupational and other structures (e.g. religion) can be compared
- Networks don’t provide the answers to how societies were structured, but provide the clues for further exploration.