# Using SID to analyse occupational structure

**Paul Lambert & David Griffiths, University of Stirling**

www.camsis.stir.ac.uk/sonocs

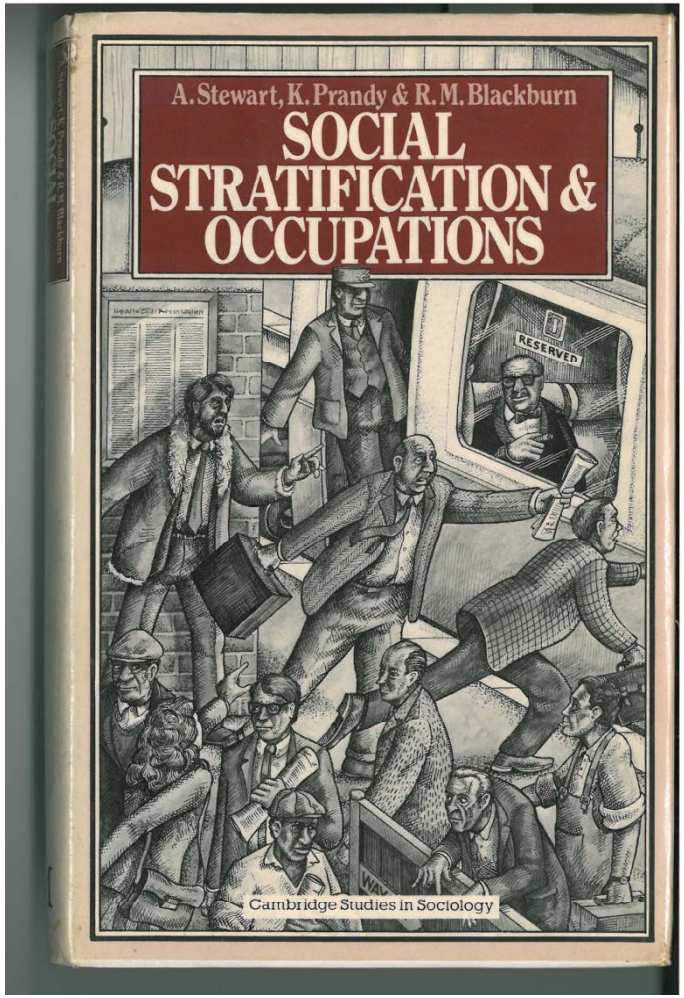
Presented to the workshop of the 'Social Networks and Occupational Structure' project, 12 September 2012, Department of Sociology, University of Cambridge

| | |
|---|---|
| 1) | Some background on CAMSIS |
| 2) | Constructing CAMSIS scales (with examples) |

- Studying social interactions and social connections can help us to understand social structure and social trends and transformations
    - Inequality between the incumbents of occupations
    - Transformations in social inequality through time / over countries
    - Exploiting detailed occupational codes and new (large scale) empirical data

# The Cambridge Scale

*Stewart et al. 1980; Prandy 1990*




- Stewart, A., Prandy, K., & Blackburn, R. M. (1980). *Social Stratification and Occupations. London: MacMillan.*
- Analysis of friendship patterns
- 'White collar' sample within 60 miles of Cambridge, plus a 'general' sample from 4 UK regions (UKDA: 1369).
- Finds an order of social stratification from social interaction distance analysis
- Stewart et al. (1980: 59-68):

| | | | |
|---|---|---|---|
| Employee University teachers | +131 | Employee Guards | -175 |
| Employee Civil Engineers | +102 | Manager Guards | -122 |
| Employer Accountants | +55 | Own account Publicans | -62 |

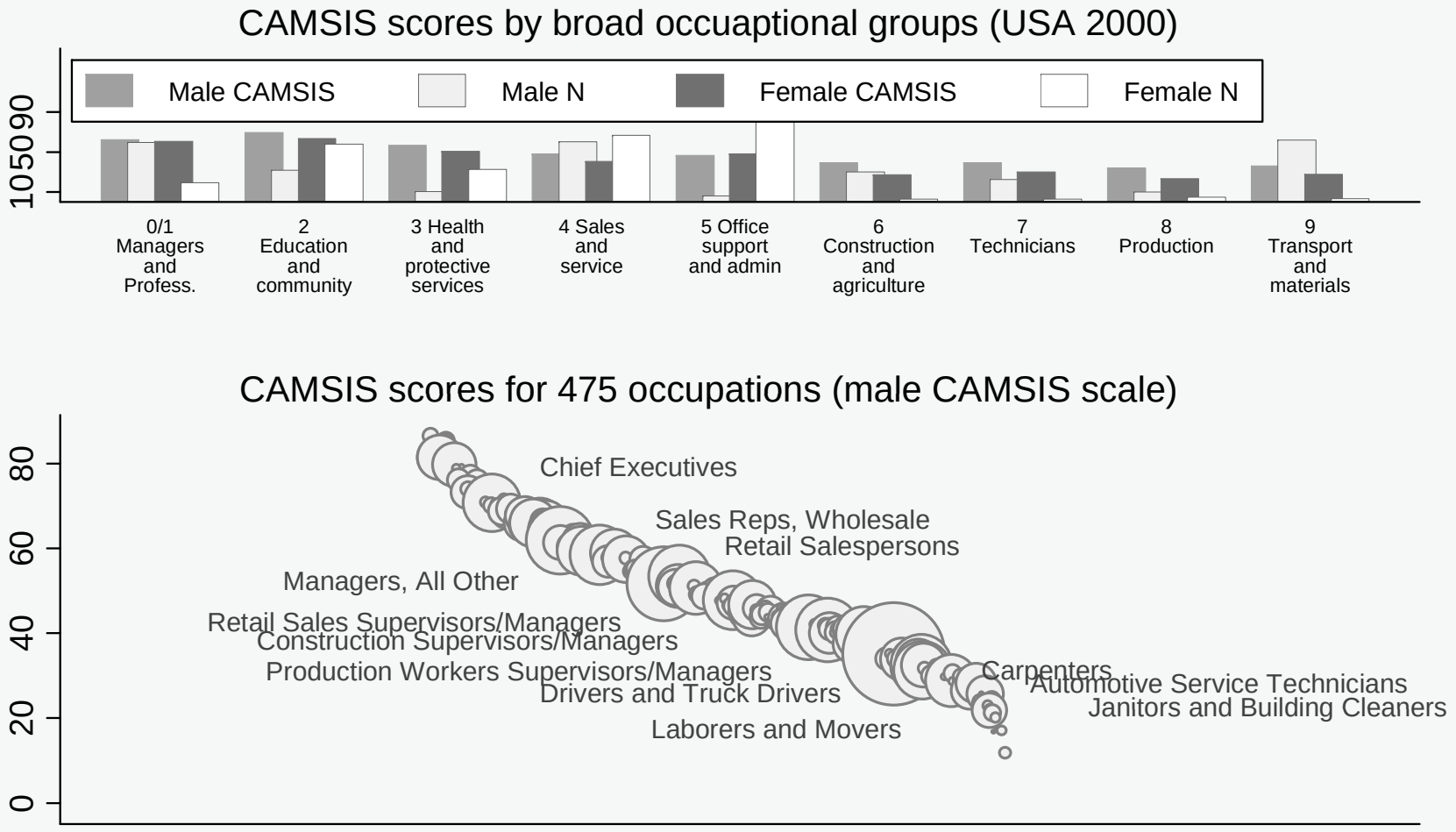
# CAMSIS, www.camsis.stir.ac.uk

*Lays out a methodology for analysing social interactions for the purpose of social stratification research*

- Analyse pairs of occupations linked by a social interaction (marriage; friendship; inter- and intra-generational connections)
- Use correspondence analysis (SPSS; Stata) or RC-II association models (Stata; lEM) on pairs of occupations
- *Tradition of 'specificity': makes an empirical calculation within a 'context' (country; time period)*
- Many other writers are using association models/correspondence analysis for similar structural analytical purposes (e.g. Chan 2010; Bakker 1993; Levine 1990; Laumann and Guttman 1966)

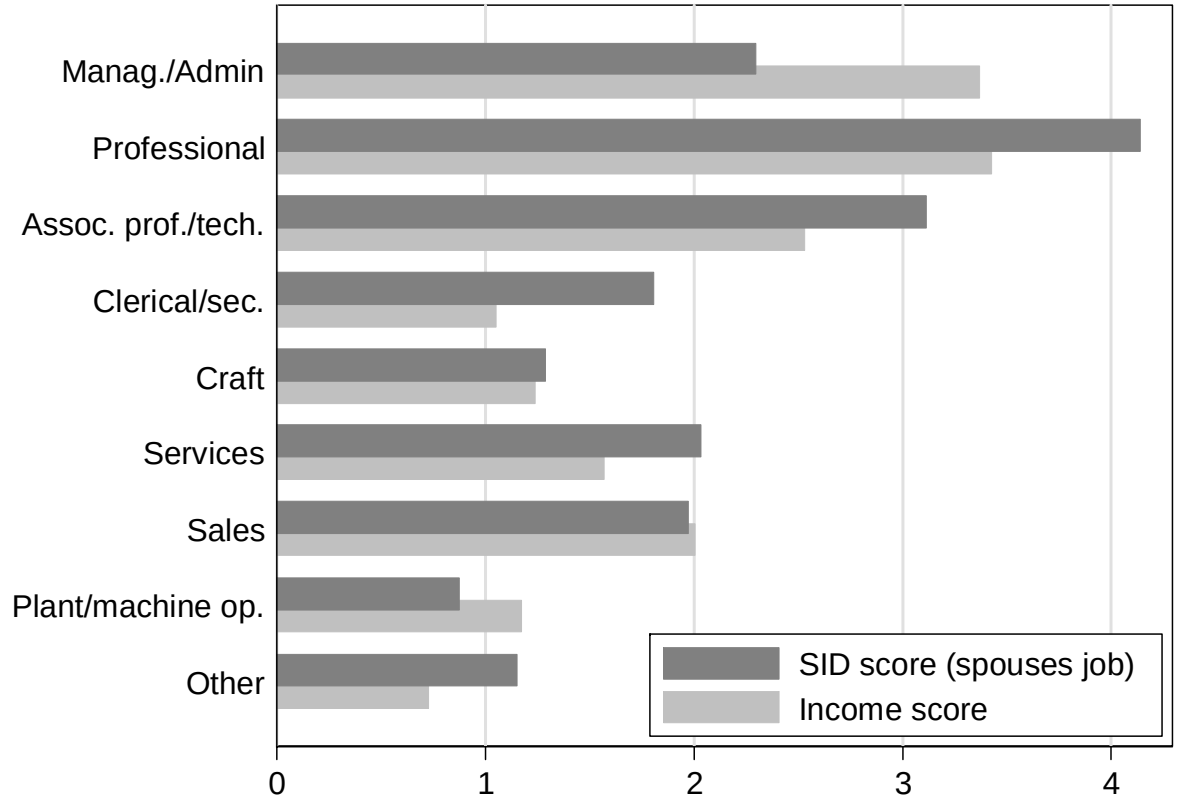
# Figure 1: Illustration of SID scales

## USA, 2000




Source: IPUMS USA, 3% sample, and www.camsis.stir.ac.uk
Panel 1: Occupational groups are first digit of US SOC2000. N is sample N / 3,000,000.
Panel 2: Marker size is proportional to number in occupation. Labels show 15 most common occupations.

# Why characterise structure through interactions?




Source: Analysis of married males in BHPS. Scores mean standardised plus 2.

- Interaction structure is not identical to other empirical summaries
- Interaction structure is theoretically interesting (?the trace of social reproduction)
- Other measures of structure may not exist

# Theorising SID analysis

- Based *only* upon social interaction patterns, CAMSIS scales may *better* represent the true order of social reproduction and stratification
  - Should be more stable over the longer term / across the career (esp. Stewart et al. 1980)
  - Bourdieusian features of capital (Bottero 2009)
  - Superior to 'social class' measures which introduce artificial boundaries and reflect current job content (e.g. Prandy and Blackburn 1997)
  - Debate regarding 'social status' (Chan 2010)

# Sorting out multiple dimensions

- Social reproduction isn't the only force which influences social interactions (e.g. regional proximity; industrial proximity; age and gender; other social institutions such as religion)

- In CAMSIS, which explicitly tries to identify the social reproduction dimension, much work using SID analysis involves differentiating other such forces from the main dimension of interest

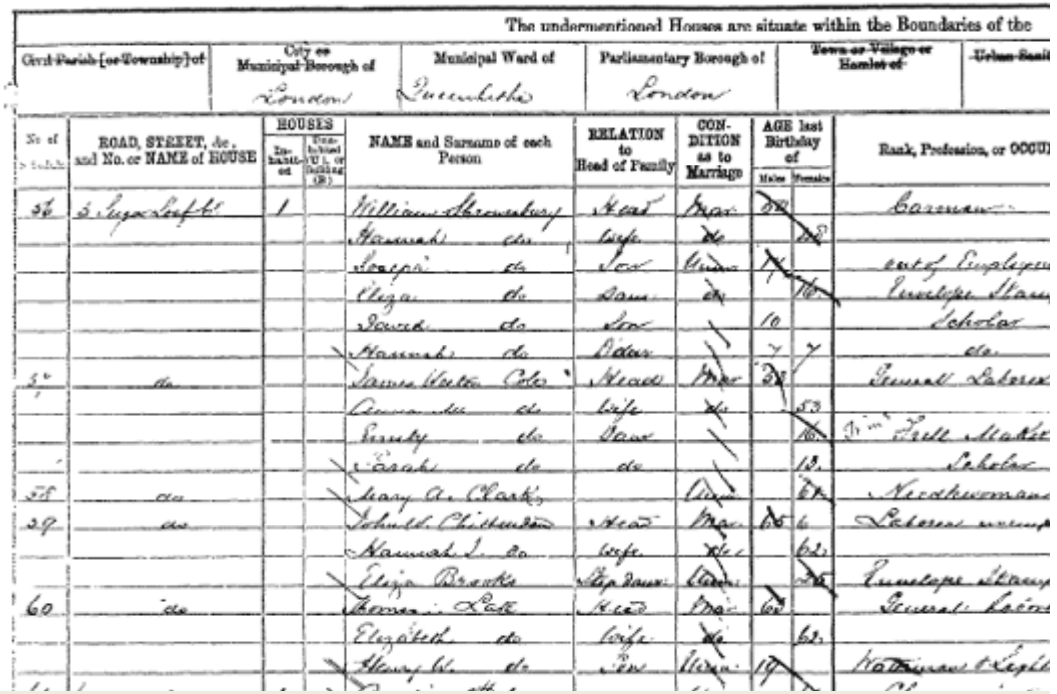
# Constructing CAMSIS scales

Methodological notes at www.camsis.stir.ac.uk

1) Extracting data on pairs of socially connected occupations
2) Association models to detect structure in pairs of connections (and isolate the 'social stratification' dimension)
3) Extraction of scores from model results, typically with re-scaling/standardisation

# 1) Constructing suitable data

Image from:
http://www.uk1881census.com/census_page.php



The undermentioned Houses are situate within the Boundaries of the

| Civil Parish [or Township] of | City or Municipal Borough of | Municipal Ward of | Parliamentary Borough of | Town or Village or Hamlet of | Urban Sanit |
|---|---|---|---|---|---|
| | London | Queenhithe | London | | |

| No. of Schedule | ROAD, STREET, &c., and No. or NAME of HOUSE | HOUSES: Inhabited | HOUSES: Uninhabited or Building | NAME and Surname of each Person | RELATION to Head of Family | CONDITION as to Marriage | AGE last Birthday of: Males | AGE last Birthday of: Females | Rank, Profession, or OCCUP |
|---|---|---|---|---|---|---|---|---|---|
| 56 | 3 Sugar Loaf Ct | 1 | | William Abercrombie | Head | Mar | 50 | | Carman |
| | | | | Hannah do | Wife | Mar | | 52 | |
| | | | | Joseph do | Son | Unm | 14 | | Errand Employed |
| | | | | Eliza do | Dau | do | | 16 | Envelope Stamper |
| | | | | David do | Son | | 10 | | Scholar |
| | | | | Hannah do | Dau | | 7 | 7 | do |
| 57 | do | | | James Watts Clerk | Head | Mar | 32 | | Journal Labourer |
| | | | | Emma do | Wife | Mar | | 33 | |
| | | | | Fanny do | Dau | | | 10 | Dressmaker Apprentice |
| | | | | Sarah do | do | | | 13 | Scholar |
| 58 | do | | | Mary A. Clark | | Unm | | 61 | Needlewoman |
| 59 | do | | | Isabella Chichester | Head | Mar | 28 | 6 | Laborer's wife |
| | | | | Hannah do | Wife | Mar | | 62 | |
| | | | | Eliza Brooks | Step Dau | Unm | | 26 | Envelope Stamper |
| 60 | do | | | Thomas Pole | Head | Mar | 60 | | Journal Laborer |
| | | | | Elizabeth do | Wife | Mar | | 62 | |
| | | | | Henry W. do | Son | Unm | 19 | | Watchmaker's Apprentice |

*Census data: Example from France 1962 (20+yrs) accessed via IPUMS-I*

| serial | pernum | sex | age | fr62a_occ | educf |
|---|---|---|---|---|---|
| 16000 | 1 | Male | 34 | Administrative secretaries | CE |
| 16000 | 2 | Female | 31 | Former workers of the private sector | CE |
| 17000 | 1 | Male | 41 | Skilled workers | CE |
| 17000 | 5 | Female | 40 | Other nonactive people (includes persons age 14 or less) | CE |
| 18000 | 1 | Female | 49 | Smaller merchants | CE |
| 19000 | 1 | Female | 33 | Other nonactive people (includes persons age 14 or less) | None declare |
| 19000 | 2 | Male | 32 | Farmers | CE |
| 19000 | 3 | Male | 24 | Skilled workers | CAP, BE |
| 20000 | 1 | Male | 59 | Office employees | None declare |
| 21000 | 1 | Female | 66 | Other service personnel | None declare |
| 22000 | 1 | Male | 28 | Artists | BAC, BP, BT, BEA, BEC, BE |
| 23000 | 1 | Female | 73 | Skilled workers | None declare |
| 24000 | 1 | Female | 38 | Office employees | None declare |
| 25000 | 1 | Male | 58 | Smaller merchants | None declare |
| 25000 | 2 | Female | 61 | Specialized workers | None declare |
| 26000 | 1 | Female | 37 | Professors, literary and scientific professions | BAC |
| 27000 | 2 | Female | 40 | Other nonactive people (includes persons age 14 or less) | BAC, BP, BT, BEA, BEC, BE |
| 27000 | 3 | Male | 46 | Office employees | BAC |

# For analysis, we usuall convert data into a ‘pairs’ oriented dataset

```
use micro.dta, clear
keep if sex==2
keep serial occ age
rename occ wocc
rename age age_sp
sort serial
sav temp.dta, replace
use micro.dta, clear
keep if sex==1
keep serial occ age
rename occ hocc
sort serial
joinby serial using temp.dta
```

*Stata ‘joinby’ command in this instance matches & keeps all male-female within household pairs*

| serial | hocc | wocc | age | age_sp |
|---|---|---|---|---|
| 1000 | Professors, professional scientists | Skilled industrial artisans | 30 | 32 |
| 2000 | administratve employees | administratve employees | 45 | 45 |
| 8000 | Civil employees, service agents of public function | Service personnel to indivduals | 63 | 61 |
| 9000 | Professional news, arts and shows | Professional news, arts and shows | 40 | 40 |
| 25000 | Businessmen and employees | Skilled driver | 58 | 61 |
| 30000 | Agricultural workers | administratve employees | 68 | 58 |
| 45000 | Service personnel to indivduals | Service personnel to indivduals | 33 | 32 |
| 47000 | Service personnel to indivduals | Service personnel to indivduals | 33 | 29 |
| 52000 | Skilled industrial artisans | Skilled driver | 29 | 36 |
| 53000 | Skilled driver | Skilled driver | 23 | 20 |
| 57000 | administratve employees | Business employees | 40 | 42 |
| 59000 | Skilled driver | Skilled driver | 31 | 27 |
| 63000 | Businessmen and employees | Skilled driver | 30 | 26 |
| 64000 | Skilled industrial artisans | administratve employees | 28 | 30 |
| 65000 | Skilled driver | Service personnel to indivduals | 37 | 37 |
| 68000 | Teachers and other employees | administratve employees | 53 | 60 |
| 77000 | Professional administrative and commercial institutions | Civil employees, service agents of public function | 44 | 40 |
| 78000 | Business employees | Civil employees, service agents of public function | 39 | 36 |
| 85000 | Engineers of technical businesses | Professors, professional scientists | 61 | 54 |
| 95000 | Service personnel to indivduals | Service personnel to indivduals | 53 | 48 |

In turn, we would typically reduce the data into a ‘table format’ record

(Loses other features of microdata but dramatically improves storage/performance)

```
tab hocc
gen freq=1
collapse (sum) freq, by(hocc wocc)
summarize hocc wocc [fw=freq]
tab hocc [fw=freq]
```

| | hocc | wocc | freq |
|---|---|---|---|
| 1. | Agricultural farming, fishermen | Agricultural farming, fishermen | 36853 |
| 2. | Businessmen and employees | Artisans | 10335 |
| 3. | Skilled driver | Skilled driver | 9979 |
| 4. | Skilled industrial artisans | Skilled driver | 5851 |
| 5. | Skilled industrial artisans | administratve employees | 4403 |
| 6. | Skilled driver | Service personnel to indivduals | 3169 |
| 7. | Skilled industrial artisans | Skilled industrial artisans | 3072 |
| 8. | Skilled driver | administratve employees | 2792 |
| 9. | administratve employees | administratve employees | 2540 |
| 10. | Skilled industrial artisans | Service personnel to indivduals | 2400 |
| 11. | Businessmen and employees | Businessmen and employees | 2265 |
| 12. | Skilled driver | Skilled industrial artisans | 2262 |
| 13. | Skilled industrial artisans | Civil employees, service agents of public function | 2172 |
| 14. | Skilled driver | Civil employees, service agents of public function | 2001 |
| 15. | Teachers and other employees | Teachers and other employees | 1866 |
| 16. | Artisans | Artisans | 1678 |
| 17. | Technicians | administratve employees | 1496 |
| 18. | Skilled driver | Businessmen and employees | 1461 |
| 19. | Civil employees, service agents of public function | Civil employees, service agents of public function | 1414 |
| 20. | Skilled industrial artisans | Business employees | 1361 |

# That's all, except…

- Fine-grained or coarse-grained analysis?
  - The more detail the better (but diminishing statistical returns, and costs of practical convenience, to more disaggregation)
  - Occupational units, or occupations cross-classified by employment status?
- Gender specific analysis?
  - Ego-alter analysis without gender, or male-female analysis so that separate male and female scores are available
- Checking for sparse occupational units
  - As a broad guideline, if there are less than 30 cases within an occupation we recommend merging it with another 'similar' category

All of these issues are discussed further at www.camsis.stir.ac.uk

# 2) Fitting the association model

- As soon as a ‘pairs’ dataset is set up, we can fit the association model
- With small numbers of units, this will probably immediately identify an obvious stratification dimension
- With larger numbers of units, there will probably be distortion caused by very sparse units and/or by other dimensional structures

## Husband's Job Units

| Occ Units ↓ → | | | 1 | 2 | .. | 407 |
|---|---|---|---|---|---|---|
| | | *Derived scores ↓ →* | *75.0* | *70.0* | .. | *10.0* |
| **Wife's** | **1** | *72.0* | **30** | **15** | **..** | **0** |
| **Job** | **2** | *72.5* | **13** | **170** | **..** | **1** |
| **Units** | **..** | .. | **..** | **..** | **..** | **..** |
| | **407** | *11.0* | **0** | **2** | **..** | **80** |

- *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.

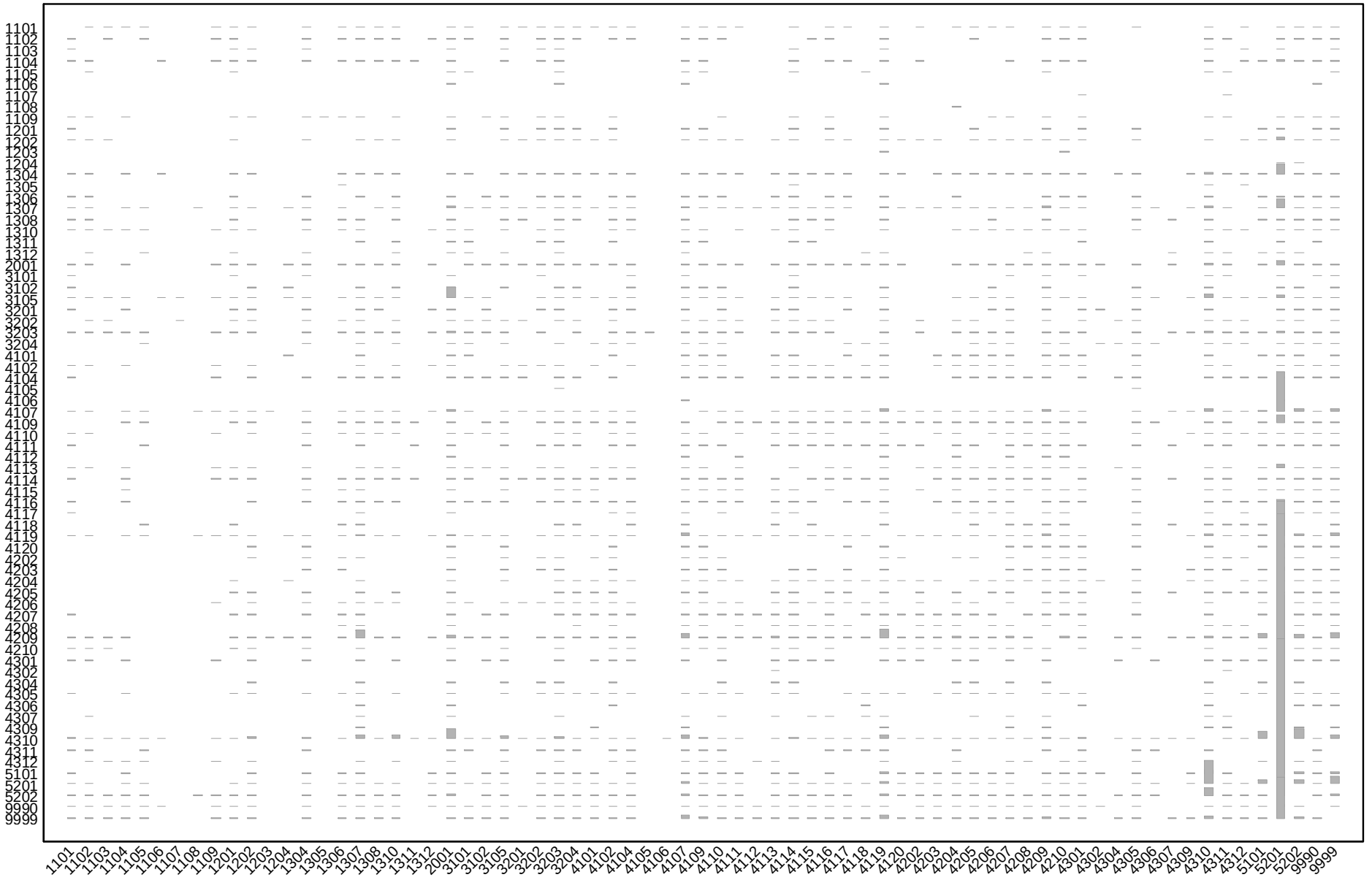
# *In addition you have the option to…*

- Specify ‘diagonals’
- Specify ‘Pseudo-diagonals’
  - Typically exclude or downweight specific occupational combinations which ‘unduly’ influence the SID pattern
  - I.e. if not about stratification and clearly about some other social connection process
- Specify ‘subsidiary dimensions’
  - E.g. explicitly set a gender segregation dimension into the model to separate those effects from the stratification dimension

*By and large, for heterogeneous societies the main dimension story doesn’t change with and without these treatments, but they can add statistical information. For less heterogenous societies, though, the main dimension may well be farming or driven by some similar pseudo-diagonal, and separating out structures is sometimes harder (see more at www.camsis.stir.ac.uk)*

# Unusual cases: Highly skewed historical occupational structures

## Norway, 1865




(Excluding diagonals, n*20)

# Model examples on the website (& some in labs)

› tabdat <- as.data.frame(xtabs(cfreq~hocc+wocc, data=microdata))

› mod <- gnm(Freq ~ hocc+wocc + instances(Mult(1, hocc, wocc),1), family=poisson, data=tabdat, iterMax=10000, ofInterest = "Mult", subset = (Freq!=0))

- ca hocc wocc if (psds==0) [fw=freq], dim(2)

➢ correspondence table=hocc (1 114) by wocc (1 114) /dimensions=2 /measure=chisq /standardize=rcmean /normalization=symmetrical /print=table rpoint cpoint /plot=ndim(1,max) biplot(20).

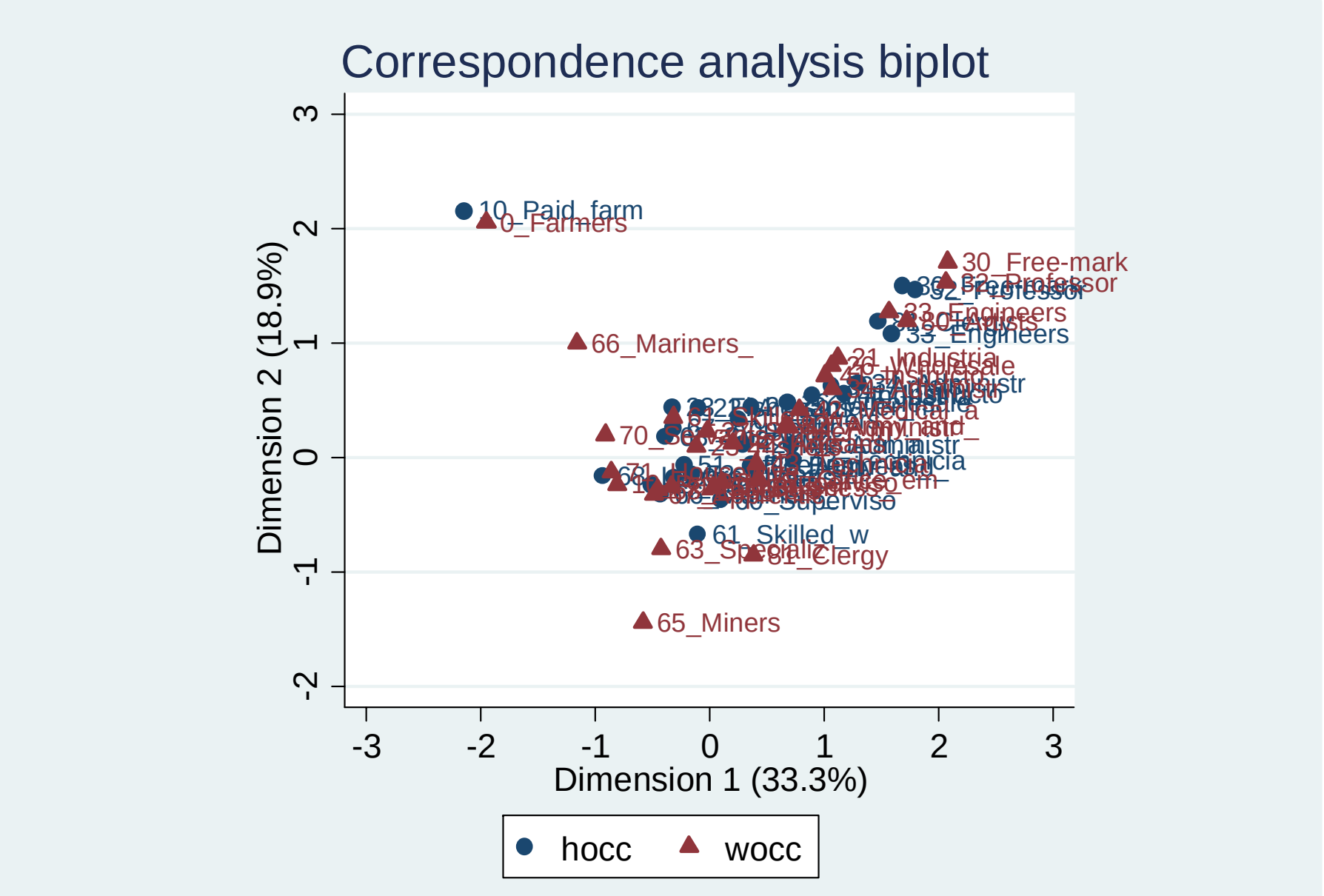
– dim 114 114
– mod {hocc,wocc,ass2(hocc,wocc,5a), ass2(hocc,wocc,5a,3,3),fac(HW,70)}
– rco
– des dmatrix114model1.txt

# *Manual or automated SID scale derivation?*

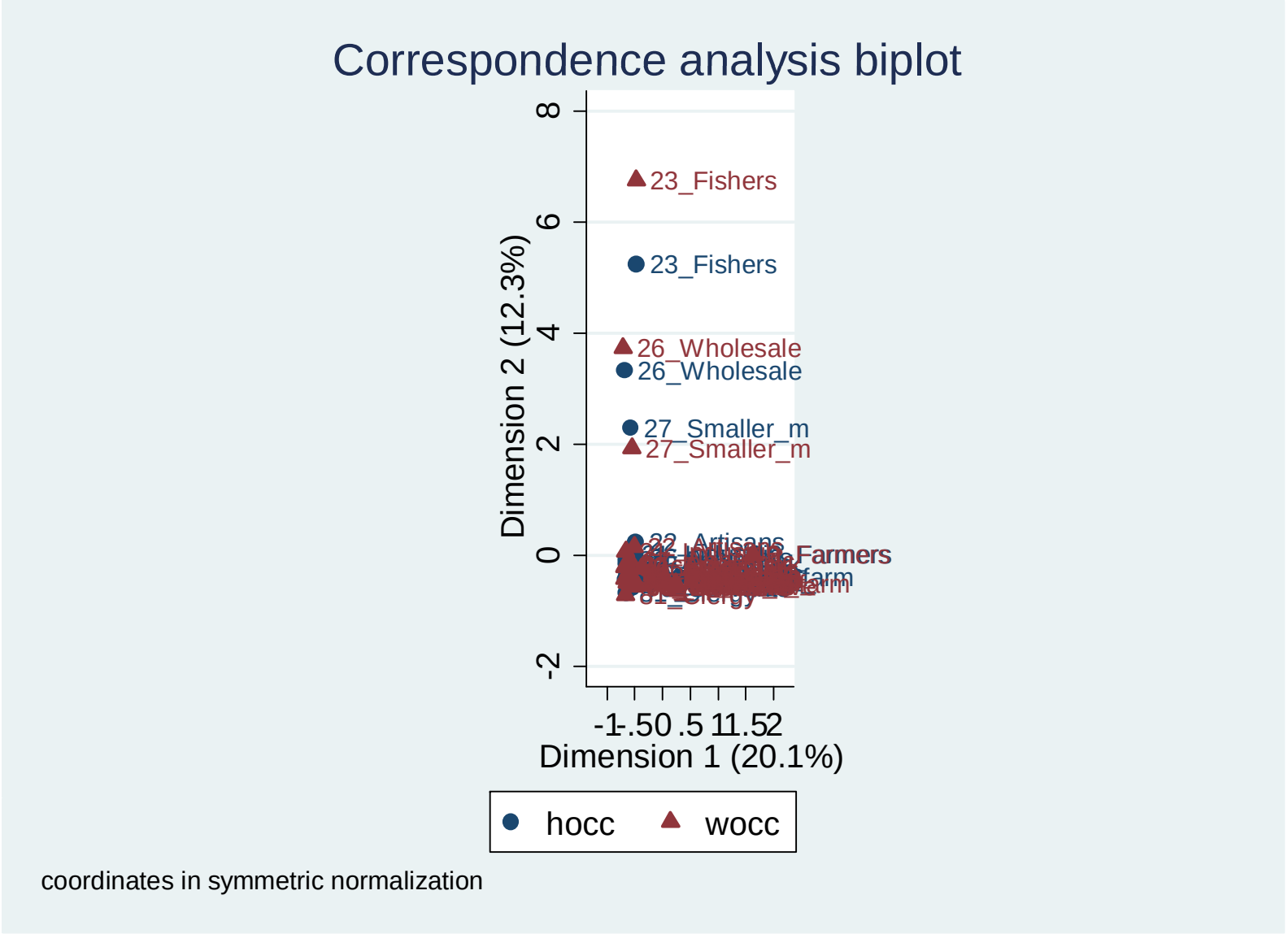
- Sometimes there are many possible scales (different permutations of data, occupational units, and subsidiary dimension coverage)
  - As part of HISCAM, we calculated over 1000 different SID scales (country*time*occupational detail*gender) (Lambert et al. 2012)
  - Each scale takes ~30mins computational time  aside from preparatory time
  - ➢ *Some degree of automation may be desirable, but could be dangerous too*

- **No automation**: case-by-case model development
- **Partial automation**: case-by-case model development sharing model building and subsidiary dimension specification templates between models
- **Substantial automation:** I've written a pre-set routine in Stata which reads in data (hocc & wocc), takes optional psd specification, and runs models and constructs scale scores
  - www.camsis.stir.ac.uk/make_camsis

# (a) You might immediately get an interesting/ convincing structure



# (b) You might get something which doesn't look so smooth



# Solutions to (b)

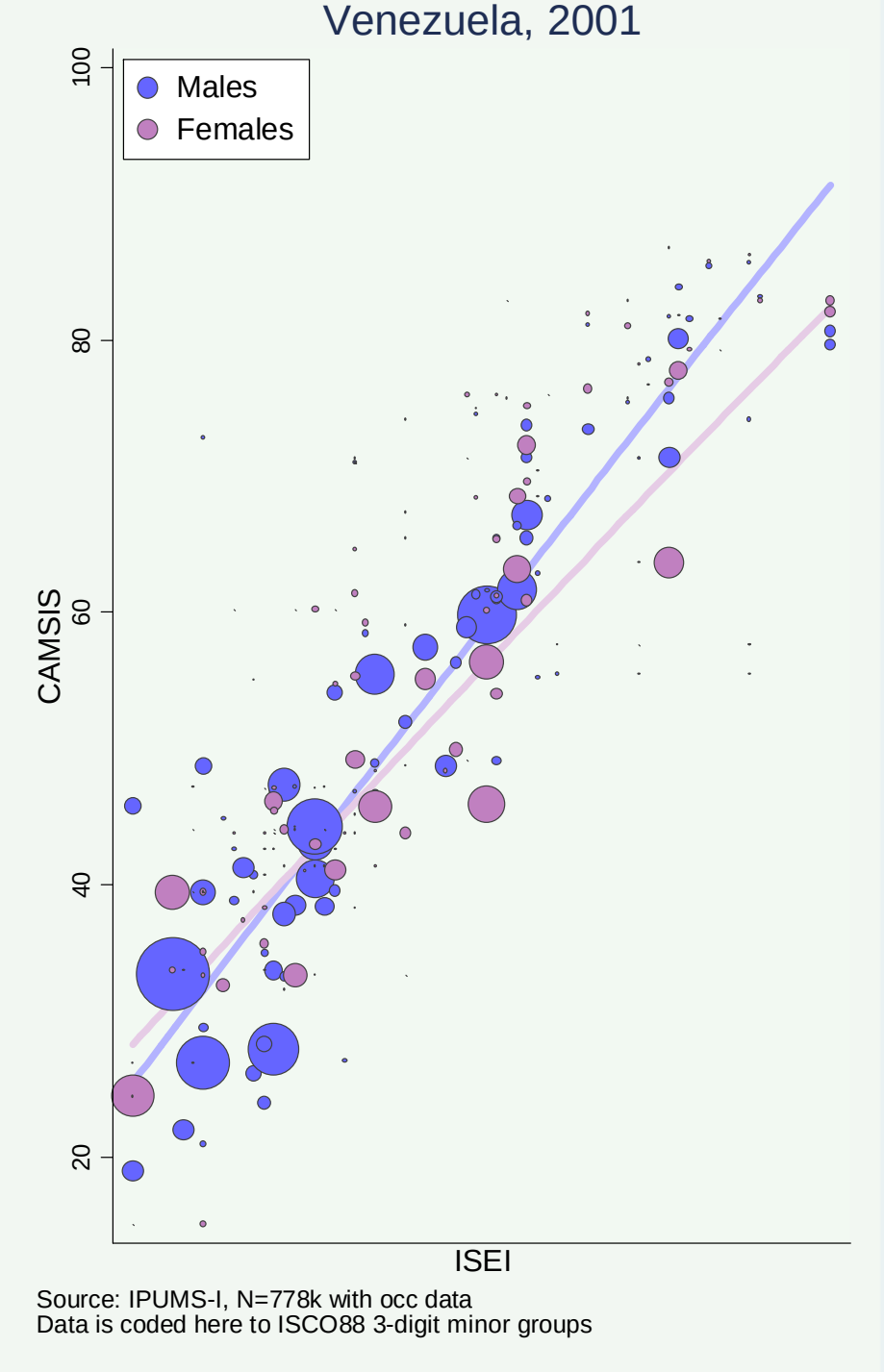
- Check for sparse units (sparse non-diagonal units) and recode if relevant
- Add pre-specified ‘pseudo-diagonals’ to the model to depict common connections between certain jobs which we are confident arise for reasons separate to the stratification dimension
  - PSD’s are either cut out of analysis, or a specific model parameter is fit for them

# 3) Processing/reviewing results

- Retreiving scores tends to be software specific.
  - In CA in Stata, extract the scores with 'predict, rowscore'
  - With lEM, copy and paste the dimension scores
  - With R, export the dimension scores as an object
- *Usually **standardise scores to mean 50, sd 15** for national population, and link them back to all possible occupations (not just the group used in analysis) using subgroup means*

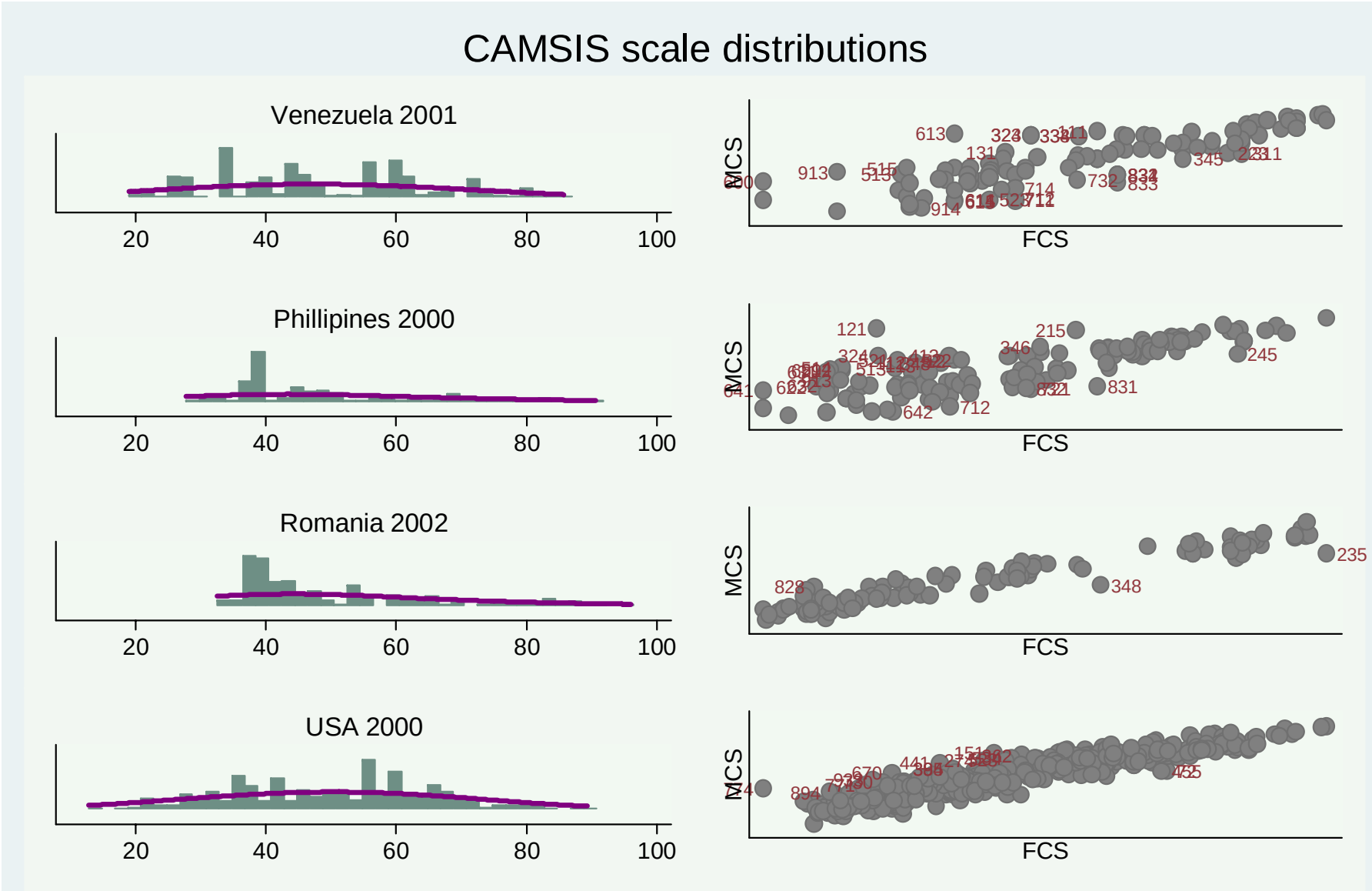
- **Confidence intervals for the scores?**
  - Not traditionally published, but…
  - gnm in R can estimate standard errors for scores which can also be extracted (I've not got this to work on a large scale, though!)
  - otherwise, we recommend calculating approximate standard errors from the mean and sd of the alters distribution on the original data

```
. use http://www.camsis.stir.ac.uk/downloads/gb/gb2010soc2010_details.dta, clear

. de, short

Contains data from http://www.camsis.stir.ac.uk/downloads/gb/gb2010soc2010_details.dta
  obs:         2,214
 vars:            16                          7 Sep 2012 20:09
 size:       154,980 (99.9% of memory free)
Sorted by:  soc2010  es

. list soc2010 gb_csm gb_csf gb_csm_se gb_csf_se if es==5 & ~missing(gb_csm_se) in 400/450

     +-----------------------------------------------------------------------------+
     |                                    soc2010   gb_csm   gb_csf   gb_~m_se   gb_~f_se |
     |-----------------------------------------------------------------------------|
402. |                         2221. Physiotherapists    72.66    73.29   1.282295   .5872164 |
408. |                    2222. Occupational therapists    72.66    73.29   1.282295   .5872164 |
420. |                2229. Therapy professionals n.e.c.    72.66    73.29   1.282295   .5872164 |
426. |                                  2231. Nurses    44.36    52.62   .7649782   .4620625 |
438. |   2311. Higher education teaching professionals    80.02    82.32   .7534429   .7559308 |
     |-----------------------------------------------------------------------------|
444. |    2312. Further education teaching professionals    70.58    74.85   .7893718   .6951411 |
450. |  2314. Secondary education teaching professionals     73.6    69.93   .3925574   .3798914 |
     +-----------------------------------------------------------------------------+

. list soc2010 gb_csm gb_csf gb_csm_se gb_csf_se if es==5 & ~missing(gb_csm_se) in 20/60

     +-----------------------------------------------------------------------------+
     |                                    soc2010   gb_csm   gb_csf   gb_~m_se   gb_~f_se |
     |-----------------------------------------------------------------------------|
 24. | 1122. Production managers and directors in construction    57.74    57.83   .5556703   .8928211 |
 36. |                 1131. Financial managers and directors    66.78     68.3   .5423195   .4818444 |
 42. |                   1132. Marketing and sales directors    67.92    73.79   .5091547   .7752766 |
 48. |                1133. Purchasing managers and directors    56.33    68.16    .850561   1.281459 |
 54. |       1134. Advertising and public relations directors    67.84    90.28     2.7956   1.153109 |
     |-----------------------------------------------------------------------------|
 60. |             1135. Human resource managers and directors    61.68    72.27   .9148157   .5766781 |
     +-----------------------------------------------------------------------------+

. list soc2010 gb_csm gb_csf gb_csm_se gb_csf_se if es==5 & ~missing(gb_csm_se) in 1000/1050

      +-----------------------------------------------------------------------------+
      |                                    soc2010   gb_csm   gb_csf   gb_~m_se   gb_~f_se |
      |-----------------------------------------------------------------------------|
1002. |              3561. Public services associate professionals    57.33    48.24   1.699955   .9856362 |
1008. |    3562. Human resources and industrial relations officers    55.59    60.82   1.027657   .5852029 |
1014. |   3563. Vocational and industrial trainers and instructors    54.73    54.14   .7448601   .7516112 |
1020. | 3564. Careers advisers and vocational guidance specialists    67.14    51.46    1.64791   1.201635 |
1026. |              3565. Inspectors of standards and regulations    50.24    56.54    1.11756   1.726127 |
      |-----------------------------------------------------------------------------|
1032. |                         3567. Health and safety officers    55.16    58.15   1.142406   1.279091 |
1038. |       4112. National government administrative occupations    54.97    46.95   .8143517   .5252942 |
1044. |          4113. Local government administrative occupations    44.81    49.75   1.530865    .597466 |
1050. |           4114. Officers of non-governmental organisations    63.96    56.94   1.608485   .9833232 |
      +-----------------------------------------------------------------------------+

. list soc2010 gb_csm gb_csf gb_csm_se gb_csf_se if es==5 & ~missing(gb_csm_se) in 1990/2030

     +-----------------------------------------------------------------------------+
     |                                    soc2010   gb_csm   gb_csf   gb_~m_se   gb_~f_se |
     |-----------------------------------------------------------------------------|
1992. |                           8221. Crane drivers    29.41     31.7   2.149145          . |
1998. |                      8222. Fork-lift truck drivers    23.35     31.7   .7458186   .9545087 |
2004. |                  8223. Agricultural machinery drivers    26.63     31.7   .9595577          . |
2010. |   8229. Mobile machine drivers and operatives n.e.c.    26.63     31.7   .9595577   .9545087 |
2016. |                                8231. Train drivers    44.86     31.7    1.03836   .9545087 |
      |-----------------------------------------------------------------------------|
2022. |     8232. Marine and waterways transport operatives    44.86     31.7    1.03836   .9545087 |
2028. |                      8233. Air transport operatives    57.18     31.7   1.397962   .9545087 |
      +-----------------------------------------------------------------------------+
```

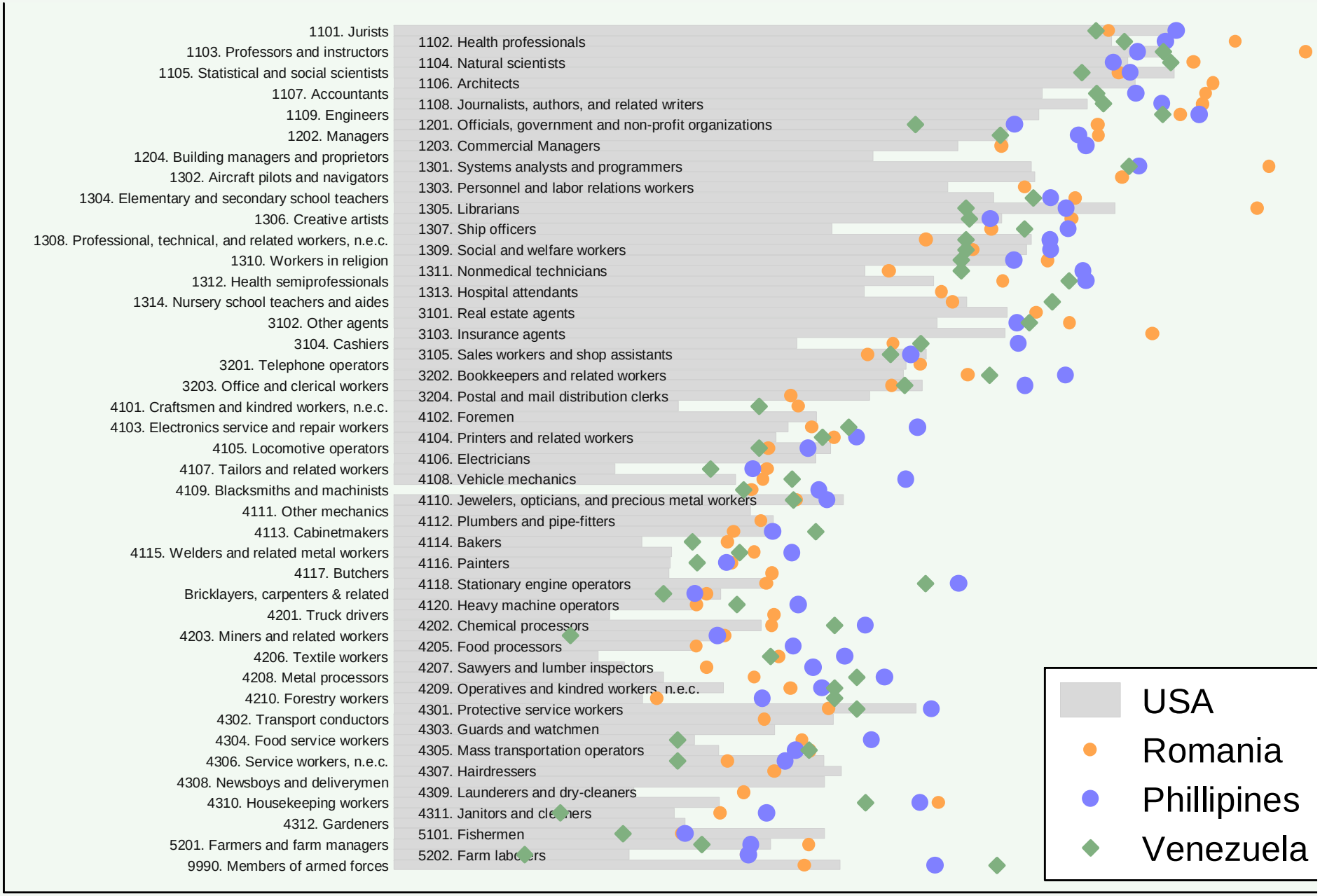



- *Using CAMSIS approaches, [www.camsis.stir.ac.uk](http://www.camsis.stir.ac.uk)*
- First dimension of SID scales is usually 'social stratification'
  - We'd interpret it as the contour of social reproduction
  - Gradational, but 'lumpy' for operational reasons (occ.s)
  - 'Specificity' (many scales!)
  - (Griffiths and Lambert 2012)
- *Dimensions:*
  - *1 main one*
  - *numerous subsidiary patterns*
- *Boundaries:*
  - *None(?)*

# *Dimensions=1; Boundaries= none; or maybe 1 in Ro?*




All microdata from IPUMS-I. CAMSIS scales at www.camsis.stir.ac.uk.
Histograms show distribution of male scale for all adults in work.
Scatterplots show unweighted male-female scores unweighted, ISCO88 3-digit or census SOC for USA




Male CAMSIS scale scores across four countries using 'microclass' units.

# Summary: SID applied to social connections between occupations

- *Connections are central to social organisation of the stratification system [e.g. Bottero 2005]*
- Challenges of data preparation and scale
  - Occupational coding – ISCO; Microclass; etc
  - Identify social connections (within hhld; other)
  - Select/discard some types of connections (e.g. farming)
- Analytical approaches
  - ➢ Model with proxy indicators, random or fixed effects

  *…Focus on the social connection..*

  - ➢ **Association models**
  - ➢ **Network analysis**

# References cited


- **Bakker, B. F. M. (1993). A new measure of social status for men and women: the social distance scale.** ***Netherlands Journal of Social Sciences, 29, 113-129.***
- **Blackburn, R. M., Stewart, A., & Prandy, K. (1980).** ***Social Status in Great Britain, 1974 [computer file]. Colchester, Essex: UK Data Archive [distributor], SN: 1369.***
- **Bottero, W. (2005).** ***Stratification: Social Division and Inequality. London: Routledge.***
- **Bottero, W. (2009). Relationality and Social Interaction.** ***British Journal of Sociology, 60(2), 399-420.***
- **Chan, T. W. (2010). The social status scale: Its construction and properties. In T. W. Chan (Ed.),** ***Social Status and Cultural Consumption (pp. 28-56). Cambridge: Cambridge University Press.***
- **Griffiths, D., & Lambert, P. S. (2012). Dimensions and Boundaries: Comparative Analysis of Occupational Structures Using Social Network and Social Interaction Distance Analysis.** ***Sociological Research Online, 17(2), 5.***
- **Lambert, P. S., Zijdeman, R. L., Maas, I., van Leeuwen, M. H. D., & Prandy, K. (2012). The construction of HISCAM: A stratification scale based on social interactions for historical research.** ***Historical Methods, forthcoming.***
- **Laumann, E. O., & Guttman, L. (1966). The relative associational contiguity of occupations in an urban setting.** ***American Sociological Review, 31, 169-178.***
- **Levine, J. H. (1990). Measuring occupational stratification using loglinear distance models. In R. L. Breiger (Ed.),** ***Social Mobility and Social Structure. Cambridge: Cambridge University Press.***
- **Prandy, K. (1990). The Revised Cambridge Scale of Occupations.** ***Sociology-the Journal of the British Sociological Association, 24(4), 629-655.***
- **Prandy, K., & Blackburn, R. M. (1997). Putting men and women into classes: But is that where they belong? A comment on Evans.** ***Sociology-the Journal of the British Sociological Association, 31(1), 143-152.***
- **Stewart, A., Prandy, K., & Blackburn, R. M. (1980).** ***Social Stratification and Occupations. London: MacMillan.***

*Data from:*

- Minnesota Population Center. (2011). *Integrated Public Use Microdata Series, International: Version 6.1 [Machine readable database]. Minneapolis: University of Minnesota, and https://international.ipums.org/ (accessed 1 July 2011).*
- North Atlantic Population Project and Minnesota Population Center. (2008). *NAPP: Complete Count Microdata. NAPP Version 2.0 [computer files]. Minneapolis, MN: Minnesota Population Center [distributor] [http://www.nappdata.org]*