TECHNICAL APPENDIX

To the grant proposal 'Social Networks and Occupational Structure'

An illustrative comparison of Social Distance and Social Network approaches to the analysis of social interactions between occupations

David Griffiths and Paul Lambert (University of Stirling), 23 July 2009

This note outlines findings from our comparison of a social distance analysis and a social networks analysis applied to data on social interactions between the incumbents of occupational positions. We use data on within-household male-female occupational combinations[i] as recorded by two census datasets: the IPUMS International 3% census sample of the United States in 2000 (Minnesota Population Center, 2009); and the NAPP release of the 100% sample from the Scottish population of 1881 (North Atlantic Population Project and Minnesota Population Center, 2008). We choose these extracts because they provide conveniently large samples of husband-wife (or other male-female) occupational combinations recorded at a fine level of occupational detail[ii], and because they illustrate the range of contemporary and historical data sources that we seek to exploit in our proposed application (see 'Case for Support').

Social Distance Analysis

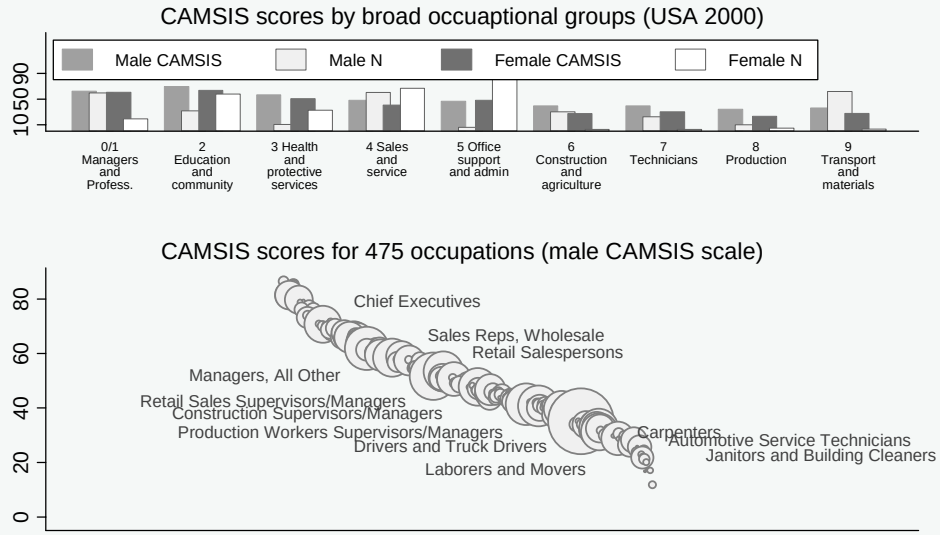
The application of social distance analysis techniques ('SID', for Social Interaction Distance) to occupational data is well documented. Numerous academic publications detail alternative procedures and results (e.g. Chan and Goldthorpe, 2004; Prandy and Lambert, 2003; Luijkx, 1994; Prandy, 1990; Stewart et al. 1980; Lauman and Guttman, 1966). Step-by-step instructions for conducting a SID analysis are provided on the CAMSIS project website (www.camsis.stir.ac.uk/).

SID analyses generally proceed by seeking to identify one or more dimensional structures which provide a good account of the empirical patterns of social distance between occupations. The usual procedure involves taking representative data on the prevalence of pairs of occupations with a social connection between them, and running a technique such as correspondence analysis, association modelling, or multidimensional scaling, in order to identify structured patterns to the relative frequency of those social connections. The important conceptualisation is that a central dimension of the structure of social interaction distance usually emerges which is interpreted as a representation of the structure of social stratification and inequality.

SID analyses are often conducted using a fine level of occupational detail. They result in the construction of datasets of scores for occupational titles, where those scores represent the relative stratification position typically experienced by incumbents of the occupation (as revealed by their typical social interaction patterns). Figure 1 illustrates the distribution of CAMSIS scale scores for male jobs in the two example datasets. The figure tries to show how SID analysis locates occupational positions within dimensions of difference.

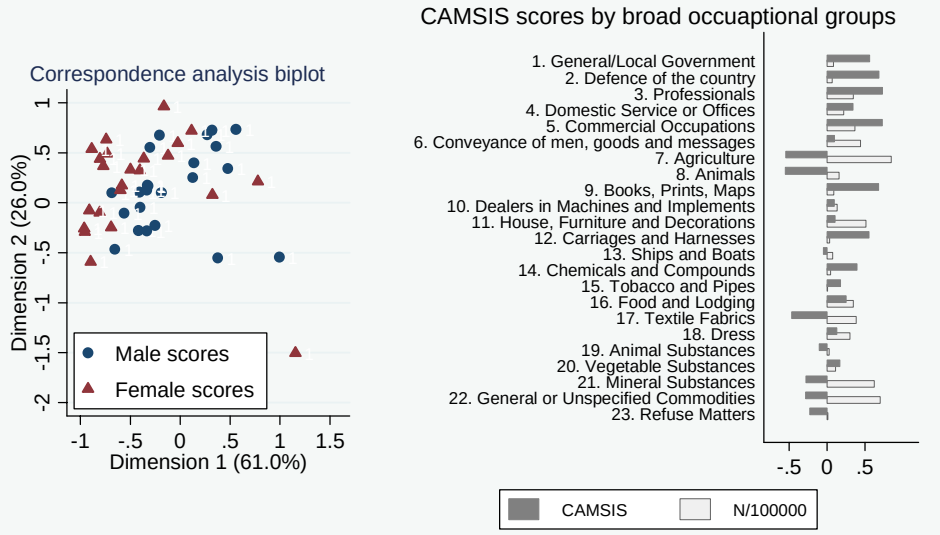
Figure 1: Illustration of SID scales

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.

Scotland, 1881



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'.

The contribution of social distance analysis is therefore to provide a probabilistic account of the overall structure of social stratification as it can be measured using occupations. This can help us to understand the structure of social stratification itself. More instrumentally, scales derived from a SID analysis prove particularly useful, parsimonious measures of occupational structure (e.g. Lambert et al., 2008), and are frequently employed as measures of stratification in further multivariate analyses.

Social Network Analysis

Social Network Analysis (SNA) techniques have not previously been employed to investigate detailed occupational structures. This note will show how our preliminary analysis suggests SNA can add value to the investigation of occupational inequalities.

The principle contribution of a Social Network Analysis is in identifying patterns of connections which occur disproportionately often. In the analysis of social interaction data on occupations, it is possible to calculate statistics which illustrate whether or not particular male-female (or other) occupational combinations occur more often than would be expected if social interactions were random. In our analyses below, we show how an SNA approach may be able to identify key occupational connections in both societies.

Beginning with the data for Scotland, Figure 2 shows the network of ties between occupations with interactions occurring at least double their expected representation. The lines represent links from the husband's to the wife's occupation. A 'centre-periphery structure' exists, but our own analysis of the specific occupations shows no strong stratification gradation in that structure. The high numbers of occupations linked to 'domestic indoor servants' demonstrates this well, with connections to such diverse trades as waterworks servicemen, net makers, dealers in works of art, old clothes dealers, mining engineers, Royal Marines and parliamentarians. Large ranges of links can occur, reducing the overt relation between network structures and social inequality.

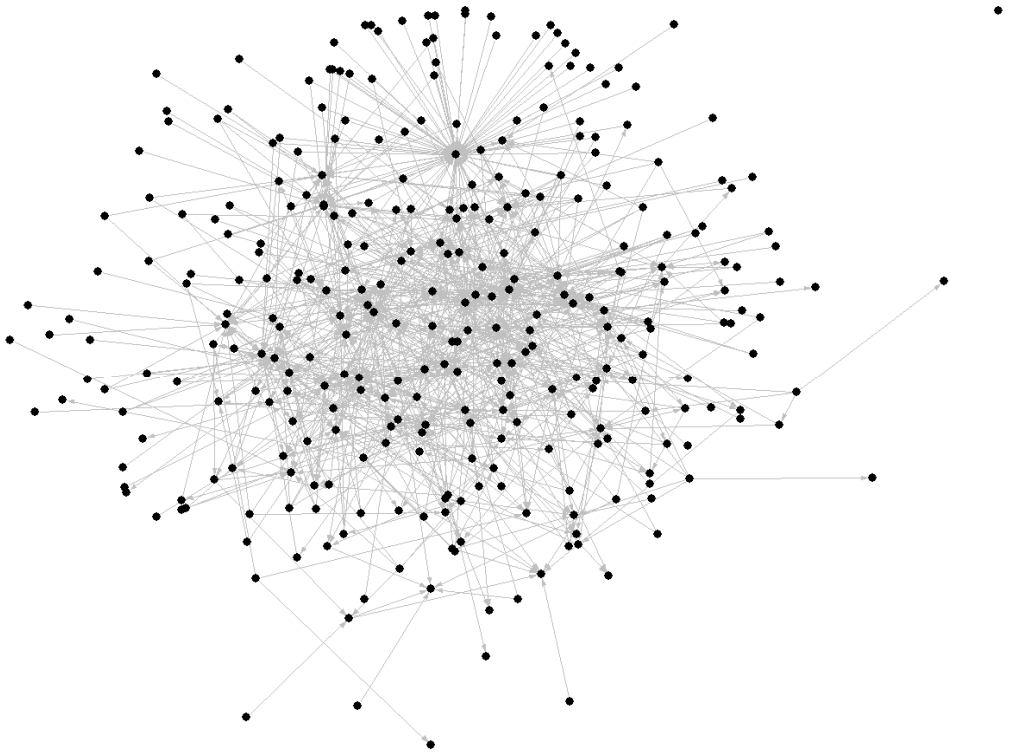


Figure 2: Scotland 1881 pairings occurring at least twice as often as expected

The effects of stratification can be observed through the strongest relationships. Figure 3 shows the central core of the network when over-representation is increased to four times the expectation. The effects of stratification can be identified, with the higher strata professions to the right and the manual work to the left. Whilst the network is connected, it appears the farming trade connects the two groupings. Removing agricultural workers from this network would lead to two unconnected structures. Breaking the data down in this manner enables trends in Networks within this data to be more clearly identified.

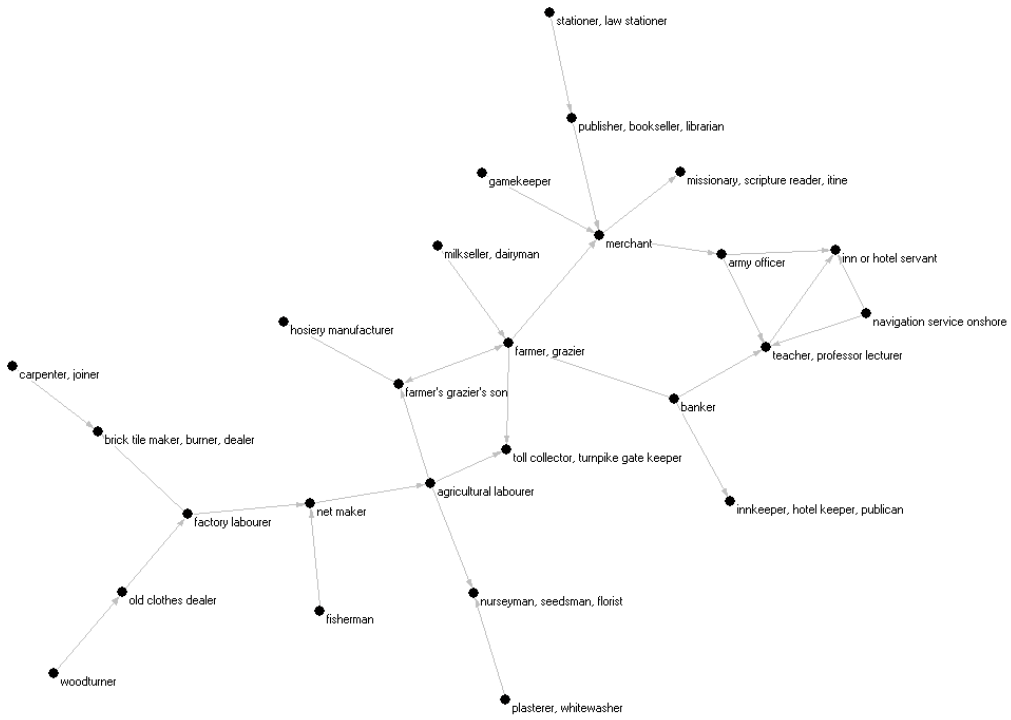


Figure 3: Scotland 1881 core of pairings at least four times the expectation

The 2000 US data shows a clearer stratification effect. Figure 4 show the network of occupational pairings double their expectation, based on size of workforce. The nodes are shaded by the quartiles within the CAMSIS scale, with the darker nodes being the least privileged. The concentration of the lower strata occupations at the top of the network, and the most privileged positions towards the bottom is apparent. The depth of the ties within those two clusters suggests not only is stratification occurring, but also there are large levels of bonding within those groupings. This bonding appears stronger than was identified in Figure 1

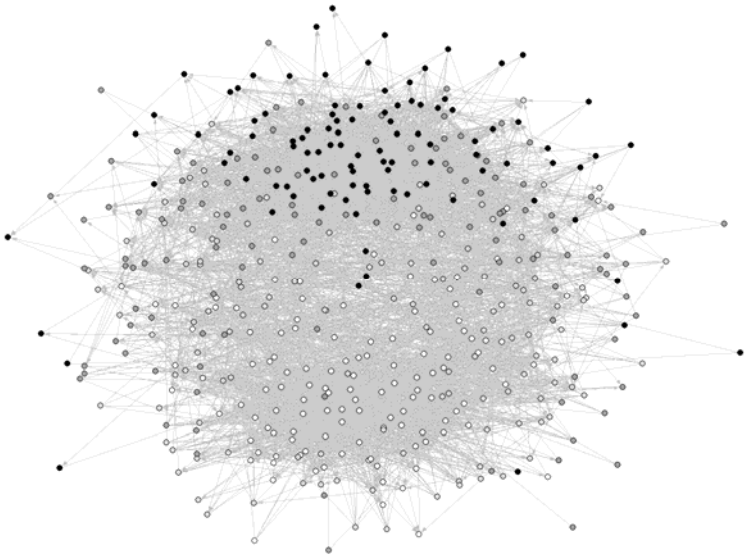


Figure 4: US 2000 ties double their expectation

Figure 5 shows the largest core of the above network. The bi-component structure is apparent, as is the distinction between the most and least prestigious occupations. Whilst these two groupings contain large numbers of connections, they are largely mutually exclusive. This suggests occupational segregation in social networks. Extending the required size of the expectation ratio removes ties and creates a network akin to this core. As well as being the most tightly-bonded set of occupations, these are also those with the strongest over-representation.

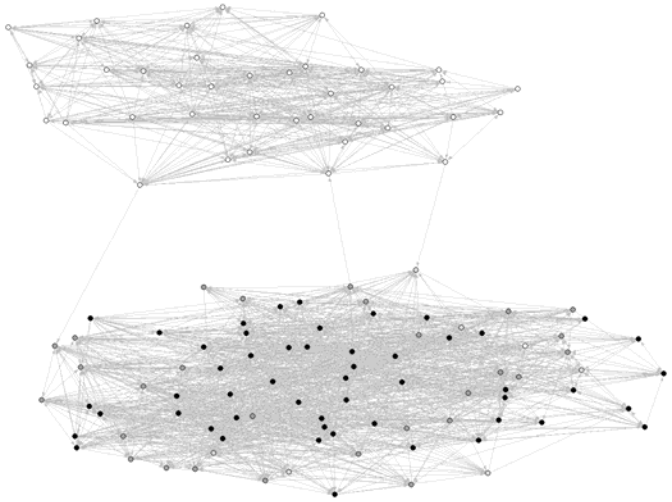


Figure 5: US 2000 core of network of double expected number of relations

Analysing the only three links between these groupings further demonstrates the stratification effect. These are all instances involving occupations with structural connections. One is the link between artists and painting workers. The opportunity for interaction between these occupations is clear. There are 56 other occupations bonded

to artists and 52 to painting workers. Only three of these links are to both, all involved in woodworking or painting. The remaining ties are largely to others within their CAMSIS quartile. A clustering of painters, frame-makers and art preservation staff can be identified. Whilst few barriers exist between those responsible for the creation, production and presentation of art, it appears they move in differing circles outside of their working lives.

Summary

In this note we seek to illustrate how Social Interaction Distance analysis supports empirical evaluations of the overall structure of social inequality in any given society for which suitable data is available, whilst Network analysis supports the detection and exploration of bonds and connections between occupations. The latter approach offers insight into the circumstances of particular occupations which are more tightly and more loosely bonded. For social stratification theory and social policy approaches alike, it is of great value to identify particularly open occupations, or to use network tools to identify clusters of highly bonded occupations (and any evidence of boundaries within the occupational stratification system which may follow).

Whilst Social Interaction Distance analysis reveals similar gradational structures of stratification inequality in both societies, the Social Network Analysis suggests that the USA in 2000 held higher levels of occupational bonding around axes of social inequality than 1881 Scotland. This is an unexpected finding and requires further explanation. There was similarity in the distributions and size of the expectation ratio across both datasets, suggesting structural rather than constructional factors have generated these differences. Further network analyses could be developed – for instance the effects of age and education on occupational networks could be measured and taken into consideration. Nevertheless conclusions based upon comparison of these two datasets would lack generality. Further longitudinal and comparative analyses of occupational networks offer promising tools for studying occupational structures.

[i] This illustrative application uses social interaction data for marriage combinations (for the US dataset, all combinations are married or cohabiting heterosexual couples; for the Scottish 1881 dataset, we use any within-household male-female combination where both adults are aged over 20 and classified to an occupation, meaning that most but not all records are marriage records). Male-female intra-household records have the advantage of being conveniently accessible across a range of societies, but do introduce operational complications associated with patterns of occupational gender segregation (Prandy, 1986, 1990). In social distance analyses, it is widely agreed that numerous alternative measures of social connections between occupations (such as marriage data; inter- and intra-generational occupational links; and friendship patterns) all consistently reveal the same structures of stratification and inequality (Prandy and Lambert, 2003). Whether a similar consistency applies to the results of social network analysis of occupations remains to be established.

[ii] The data for the United States in 2000 covers 1.9 million cohabiting couples and our analysis examines social interactions between 475 different job titles. The data for Scotland in 1881 covers 589109 intra-household pairs and our analysis examines social interactions between 391 different job titles.