HIS-CAM - Presentation and evaluation of an historical occupational stratification scale based upon the analysis of social interaction

Paul Lambert	**Stirling University**
Richard Zijdeman, Ineke Maas	**Utrecht University**
Ken Prandy	**Cardiff University**
Marco van Leeuwen	**International Institute for Social History**

http://historyofwork.iisg.nl/
http://www.camsis.stir.ac.uk/hiscam/

Presentation to: Session D12 on 'Social Mobility' of the **European Social Science History Conference,** Lisbon, 26 Feb – 1 Mar 2008

Social interaction and social stratification

1) **Long observed link between social interaction patterns and social stratification**
   - *[Weber; Bourdieu 1984; Bottero 2003]*
   - *[Blau and Duncan 1967; McPherson et al 1999; Kalmijn 1998]*

2) **Analysis of social interaction patterns can be used to tell us about structures of social stratification**
   - 'Cambridge scale' and CAMSIS: *[Stewart et al 1980; Prandy 1990; Prandy and Lambert 2003]*
   - *[Lauman 1966; Bakker 1993; Chan and Goldthorpe 2004]*

CAMSIS, www.camsis.stir.ac.uk

Lays out a methodology for analysing social interaction 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)*

			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.

Occupations and social structure in history (1800-1938)

- *'Starting from the occupational titles themselves'*
  [HISCO – van Leeuwen, Maas & Miles, 2002, p28]

- *'The historical process …. does of course influence the relative social position of the different groups'* [Thomsen, 2008 – today!]

- Relatively easy to access data on occupations linked through social interactions: inter-generational occupational links from marriage registers etc

HIS-CAM scales can offer

- **Summary measure of occupational positions**
  - Differentiates finer occupational details
    - *Typically 300+ occupational units assigned different scores*
  - Emphasises a hierarchical structure of inequality
  - ***Measures relative advantages 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..?

Data used in HIS-CAM

	1800-1938	**1800-90**	**1891-1938**
	# child-parent data points (% male-male)		
Netherlands*	664311 (47)	412732 (39)	251579 (61)
Germany^{+}	7710 (97)	5499 (99)	2211 (86)
France*	65308 (45)	40931 (47)	24377 (44)
Sweden*	19166 (75)	18079 (74)	1087 (88)
UK*$^{+}$	45517 (78)	28848 (82)	16669 (72)
Canada (Quebec)*	229134 (98)	91680 (99)	137454 (98)
US$^{\#}$	194218 (43)	56310 (20)	137908 (53)
Belgium*	48437 (53)	40320 (54)	8117 (49)

**Marriage/parish registers; $^{+}$Genealogical; $^{\#}$Census*

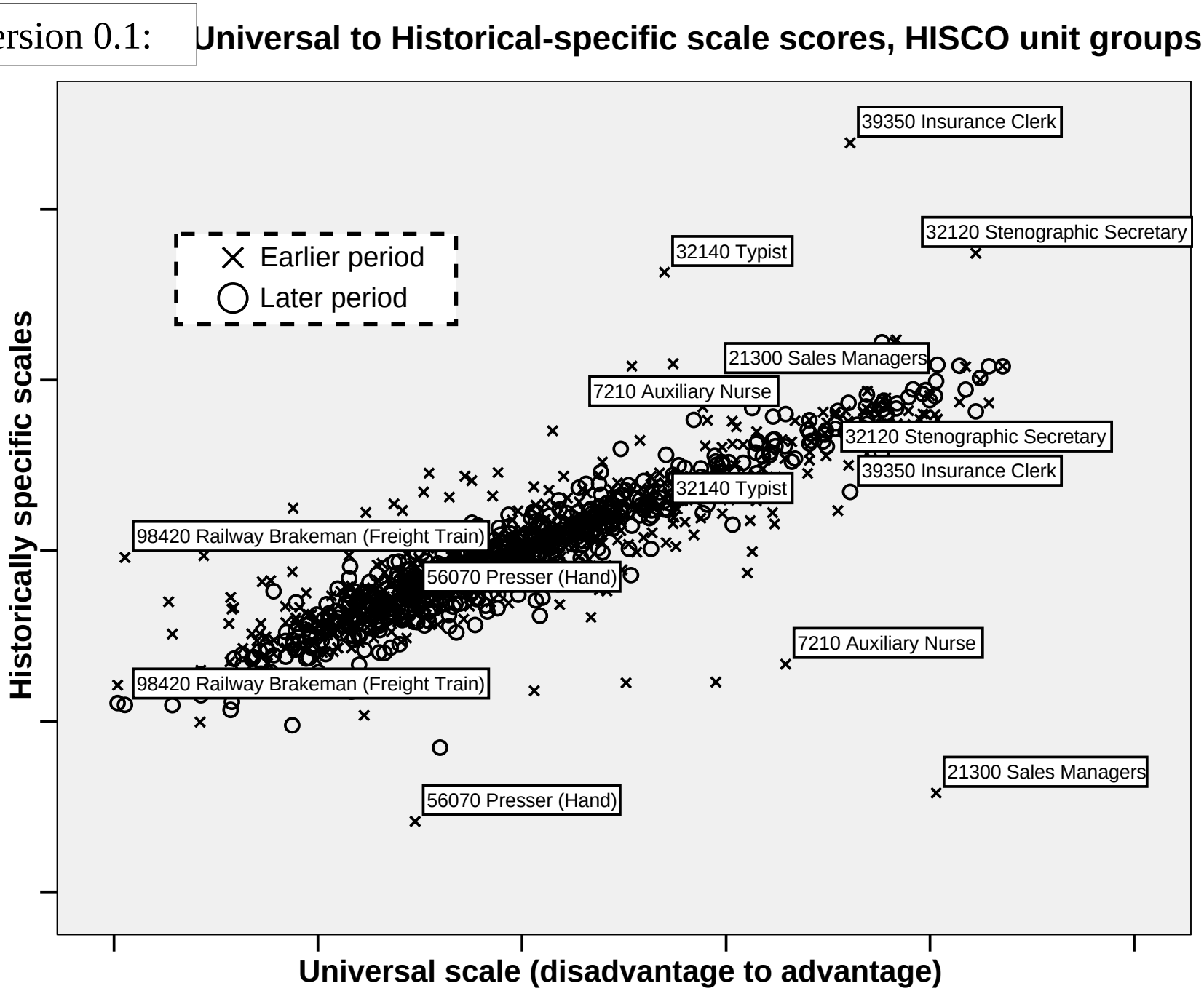
HIS-CAM in summary

- **Version 0.1** (www.camsis.stir.ac.uk/hiscam/, May 2006)
  - Netherlands, Germany, France, Sweden, UK, Canada
  - ***Occupational coding to HISCO standardised across countries***
  - One cross-national scale; 6 national scales (specific), for 1800-1938; scales for 'early' and 'late' periods (c1890)
- **Version 0.2** (February 2008)
  - ***Experimental review producing a great many alterative scales***
  - (combinations of countries*time periods*gender groups)
  - Improved micro-data on 6 core countries (extended coding quality review; increased volume of cases); new data US & Belgium
  - ***Occupational coding at localised levels***
  - ***Automatic scale derivations***
- **Version 1.0** (Spring 2008)
  - **11 scales for public release**
  - ***..See conclusions..***

HIS-CAM scales prove to have very similar properties to contemporary CAMSIS scales

- Clearly reflect an order of stratification advantage / disadvantage in occupations
    - Jobs with educational requirements tend to be highest ranked (Univ. professors)
    - Low skilled labouring jobs tend to be lowest ranked
    - Correlate around 0.7 with prestige scales, class schemes
- Some plausible differences between (some) different specific scales
    - Agricultural jobs show most variation in relative positions between countries
    - Service sector jobs change positions over period

Version 0.1:


Universal to Historical-specific scale scores, HISCO unit groups
Earlier period
Later period
39350 Insurance Clerk
32120 Stenographic Secretary
32140 Typist
21300 Sales Managers
7210 Auxiliary Nurse
32120 Stenographic Secretary
39350 Insurance Clerk
32140 Typist
98420 Railway Brakeman (Freight Train)
56070 Presser (Hand)
7210 Auxiliary Nurse
98420 Railway Brakeman (Freight Train)
21300 Sales Managers
56070 Presser (Hand)
Historically specific scales
Universal scale (disadvantage to advantage)

End point at v0.1

- Combining sparse HISCO unit groups in the same way across countries is problematic

- Universality or specificity
  - (how many scales should there be?)
  - Statistical support for maximum specificity
  - But broad correlations between schemes

Version 0.2: Permutations of occupations

C	**10 national groupings** *(8 countries, plus all countries, plus all countries excl. US)*
L	**5 levels of occupational detail** *(major groups, 1-digit, 2-digit, 3-digit, 5 digit)*
S	**4 gender groupings** *(all occupations combined; male occupations only; female occs based on daughter-father; female occs based on daughter-mother)*
T	**5 time periods** *(whole period; pre- and post- 1891; pre- and post national specific point of transition in agriculture/manufacturing balance)*
10*5*4*5 = 1000 different v0.2 HIS-CAM scales **- In practice, 825 scales were calculated (Feb 2008)**	

Approaching maximum specificity

- Combination of small occupational groups on each of the 825 samples is substantially different
  - *(in v0.1, this was ignored by using common coding in a nested model framework)*

➢ Ideally, national experts in occupational coding and statistical modelling would review coding and categorisations and optimise statistical models

[=>relatively few contemporary CAMSIS scales…]

SPSS Pivot Table - table1

File Edit View Insert Pivot Format Help

	Netherlands				Sweden			
	Child and parent male		Child female, parent male		Child and parent male		Child female, parent male	
	child	parent	child	parent	child	parent	child	parent
1110 Chemist, General	54	7		2				
1200 Physicists	3	2						
1210 Physicist, General	3							
1340 Meteorological Scientist		1						
1400 Physical Science Technicians, Specialisatin Unknown	35	16		1				
2000 Engineer, Specialisation Unknown	220	43			6	3		1
2120 Building Architect	104	102		9		1		
2130 Town Planner	5	8		1				
2140 Landscape Architect	1	2						
2210 Civil Engineer, General	213	77		4	1			
2220 Building Construction Engineer	40	32		5				
2230 Highway and Street Construction Engineer	3	4						
2240 Railway Construction Engineer	4	1				3		
2255 Hydraulics Engineer	54	62		8				
2290 Other Civil Engineers	48	32		1				
2305 Electrical Engineer, General	1							
2310 Electronics Engineer, General	2	1						
2410 Mechanical Engineer, General	14	4		1	1			
2450 Ship Construction Engineer					2	5		1
2510 Chemical Engineer, General	13	1		1				
2620 Extractive Metallurgist	2							
2710 Mining Engineer, General	4	1						
2930 Agricultural Engineer	3							
3020 Land Surveyor	137	86		20	14	11		

v0.2 strategies

- Automated recoding of sparse occupations
  - (to popular or generic subgroup codes)
- Standard model selection criteria
  - (2 dimensional correspondence analysis and limited additional treatments for over-influential cases)

- *The whole process can be automated using Stata (correspondence analysis)*

  - ***825 automatically derived scales now exist***
  - ***What on earth to do with them..?***

Example 1: A well-identified automatic scale

l1_c3_s2_t1 = 1-digit, France, Male-Male, whole time period		
		CA percent inertia:
0 Professionals	85.4	Dim 1=59.0% ; Dim2=36.5%
1 Professionals	71.0	***Correlation with:***
2 Administrative / managerial	73.1	ISEI = 0.73 *(no farm=0.75)*
3 Clerical and related	59.3	Treiman= 0.62 *(no farm=0.88)*
4 Sales workers	61.8	
5 Service workers	47.6	***Father-son correlation:***
6 Agricultural	46.4	0.27
7 Production, transport, labour	46.5	(0.30 ISEI; 0.17 Treiman)
8 Production, transport, labour	48.4	
9 Production, transport, labour	45.6	

Example 2: A well-identified automatic scale

l5_c1_s1_t1 = 5-digit, Netherland, all, whole time period		
		CA percent inertia:
13130 University teachers	99	Dim 1=11.8% ; Dim2=8.1%
13320 First level teacher	68.9	***Correlation with:***
61240 Livestock farmer	47.4	ISEI = 0.81 *(no farm=0.78)*
62105 Farm worker	39.2	Treiman= 0.70 *(no farm=0.78)*
79420 Garment pattern maker	60.1	
79510 Hand or machine sewer	50.6	***Father-son correlation:***
83920 Gunsmith	55.6	0.45
83930 Locksmith	52.6	(0.31 ISEI; 0.12 Treiman)
84230 Precision instrument maker	62.4	
99920 Day Labourer	43.1	

Practical and empirical problems with automatic scales:

Correlations..	**ISEI**	**Treiman**
c0_s2_t1 (all countries)	0.65	0.68
c1_s2_t1 (Neth)	0.86	0.86
c2_s2_t1 (Germ.)	0.64	0.61
c3_s2_t1 (France)	0.86	0.83
c4_s2_t1 (Sweden)	0.55	0.47
c5_s2_t1 (Britain)	0.79	0.77
c6_s2_t1 (Canada)	0.77	0.81
c7_s2_t1 (USA)	-0.06	0.01
c8_s2_t1 (Belgium)	0.43	0.38
c9_s2_t1 (all, excl. US)	0.74	0.75
S2=male-male only; t1=1800-1938; level: 5-digit HISCO		

Summary on v0.2

- V0.2 helps us to evaluate the scale-construction procedure
- Some v0.2 examples show persuasive evidence of specificity (not previously visible in v0.1)

But some problems..

i. Automation produces more results than can easily be reviewed

ii. Automation produces many poor results (?50%)
   - Model estimates not subject to expert review (detecting and highlighting appropriate dimensions)
   - Automated recoding still misses country differences

iii. ***<u>Users of scales</u>*** *probably don't want 825 alternatives..!*

Version 1.0

- There is statistical and substantive evidence for both specificity and for universality
- There are practical limitations to too much specificity
- *Persuaded by a strategy of 'realistic complexity'*

Version 1.0

– Publication of a small number of specific scales (based on key patterns from v0.1 and v0.2)
  - 1 cross-national
  - 8 national specific
  - 1 cross-national pre-1890; 1 cross-national post 1891

– Options for further analyses (using national expertise)

Appendix: how to use HIS-CAM

- Selected scales at www.camsis.stir.ac.uk/hiscam
- Released versions can also be accessed at www.geode.stir.ac.uk

- Data manipulation tasks:
  - Match files in SPSS, Stata or plain text
  - Recode macros in SPSS, Stata

1) Use as a measure

- Interpretation: the relative position typically held by incumbents of the occupational group within the structure of social stratification *[…for context – year, country etc]*
- *Social mobility analysis: no, it's not circular*

2) Use as information about occupations

- New data on relative positions of occupational units