LONG TERM TRENDS IN OCCUPATIONAL STATUS ATTAINMENT IN THE NETHERLANDS IN THE 19TH AND 20TH CENTURY: EVIDENCE FROM MARRIAGE RECORDS AND SIBLING DATA

Paper to be presented at the European Social Science History Conference, Congresgebouw, The Hague, February 27 2002.

Earlier presentation:
- Summer Conference of the ISA Research Committee on Social Stratification, Laval University, Quebec City, August 14-15 1997.
- HSN Workshop on Large Databases and Best Practices, Amsterdam, May 2001.

Synopsis & tables

Harry B.G. Ganzeboom

Department of Sociology
Utrecht University
PO Box 80.140
3508 TC Utrecht
NETHERLANDS

email: H.Ganzeboom@cc.ruu.nl
phone: +31 30 253/2101
fax: +31 30 253/4405
http://www.fsw.ruu.nl/soc/HG

Acknowledgement: The initial version of this paper was prepared while the author was a Fellow a the Netherlands Institute of Advance Study (NIAS) in Wassenaar. The historical data were provided by the Stichting Historische Steekproef Nederlandse Bevolking HSN. Its principal investigator is Kees Mandemakers, to whom I am grateful for providing me with an early copy of the data.

Summary

This paper analyzes trends in social reproduction among men according to 19th and early 20th century marriage records in the provinces of Utrecht and Zeeland, using an elementary sibling model with metric variables. The main research question is whether the estimated parameters have changed during the period of investigation (1840-1920) and how they compare to modern data. I use data on occupation of brothers serving as grooms or witnesses in the marriage records to model family effects and to assess the degree to which the father-son correlation underestimates total family effects on occupational attainment. I make a comparison with modern survey data (collected in 1992, but covering marriages contracted between 1940 and 1992), in which occupations of men at marriage and occupations of brothers and fathers are available. I find that there are significant changes in the family effects over time, which rise from the first to the second cohort (from 0.70 to 0.80, but then taper off to a level that is quite comparable to that found for the modern data (0.60). The degree that to which father's occupation determines social background effects is almost identical across the historical and modern data sets and explains less of half of the family variance.

INTRODUCTION

The association between occupations of fathers and sons has remained one of the primary tools of analysis for social analysts who want to look at trends in the openness of societies. While modern multivariate research has brought in a number of additional dimensions, in particular education and life style, to the analysis of intergenerational mobility and reproduction pattern, this has not diminishes the interest in occupational status as a core variable. Many analysts today would still subscribe to the statement of Blau & Duncan (1967), that "occupational status summarizes much of what matters in the system of inequality". While this fully applies to modern societies, it is even more true for historically observed societies. Occupational position has been the most prominent indicator of social inequality in the past, even more than it is today.

One of the primary themes in the study of social inequality has been the study of social mobility, i.e. to assess to what degree occupational positions are maintained from one generation to the next, and if not, to what degree the occupational destinations of one generation still depend upon the occupational positions of the previous generation. A high degrees of social immobility is taken as an indicator of closure of a society. In received sociological theories, the image prevails that social closure was much higher in the (recent) past than it is today. Pre-industrial societies are characterized by a high degree of social reproduction, that is testimony to a pattern of social selection by ascribed characteristics. The evolution of the industrial society has not fundamentally changed this pattern. Of course, industrialization has led to enormous horizontal mobility, mainly by drawing the agrarian population into urban areas and converting farm labor to industrial labor. However, with increased social inequality that characterized early industrialization, this hardly makes for a trend towards more vertical mobility. It is only with the evolution of a larger service segments in the economy that achievement patterns begin to prevail. It is the demand for qualified personnel, such as skilled manual labor, clerical, managing and administrative workers in technologically advanced industries and the rise of a large tertiary sector, including massive government bureaucracies, that would bring about less social closure. The trend toward selection by achievement is crucially dependent upon educational expansion, as it started off in the early twentieth century.

In this paper I analyze the long term trend in occupational status reproduction in the Netherlands, covering a period between the mid 19th century and today. The basic question is whether – in line

with the functionalist ascription/achievement hypothesis – the degree of status reproduction in Dutch society has moved towards levels of less closure, and if so, when such changes have primarily occurred. Or is it the case that status reproduction has remained at the same level or even risen with the coming of a more service oriented economy, like reproduction theorists have maintained? I do so using data not only on intergenerational (father-son) association, but also by taking intragenerational (brother-brother) association into account. As I will argue, studying sibling association gives a fuller account of reproduction patterns and amends our understanding of the intergenerational patterns that are traditionally studied.

DATA AND VARIABLES

The historical data I use are taken from the Historical Sample of the Netherlands (HSN). The HSN aims at bringing together demographic data on the population of the Netherlands in the 19th and early 20th century, using data from official population registration records. The HSN sample itself is drawn from birth certificates, as they reside in local administrations since about 1820. The strategy of the HSN collection is first to search for death certificates and marriage certificates for everybody in the sample, by following the initially sampled person in the population register. In a further stage data will be added from other official records, in particular housing and moving records, tax records and civil service records. Using all these records together it is possible to reconstruct someone's life history. In particular it is possible to record social background (occupation of the father), occupational histories, housing, fertility, marriage, life expectancy, etc.

As the population registration of this period, while principally intact, is organized on a local basis and not indexed, the work involved in setting up the primary data base, is tedious, voluminous and expensive. This is the reason why the work of the HSN proceeds step by step, where the Dutch provinces are taken into the process one by one. While currently work is proceeding on most other provinces, we have already access to data from the two first provinces in which the HSN started its work, Utrecht and Zeeland. Fortunately, these two together can be conceived of a fair representation of the country as a whole (of course, it remains to be seen whether this is true when all the data have been collected), as I will do in this paper.

Among the certificates processed for the HSN database, by far the most informative one for stratification research is the marriage record. It usually records occupation (at marriage) of bride and

groom and their respective parents, if they happened to be present at the ceremony (which was often the case). For either bride or groom, the information on parental occupations can be compared and/or supplemented with the same information at the birth certificate. However, this is only of limited value, since only one half of the marriage couple belongs to the original HSN sample.

There is another interesting feature of marriage records: the presence of witnesses. In the Netherlands, marriages are sealed in the presence of formal witnesses, and it was and is common practice to invite members of the respective families, other than the parents, to perform this ceremonial duty. Very often brothers are selected for the task, and this is the reason why the marriage records often contain additional information on social mobility than the parents-offspring information: as the occupations of all witnesses are recorded, we cannot only expand the father-son information provided by the groom by similar information by his brother(s), if present, but also by comparing the occupation(s) of the brother(s) with that of the groom, and using the sibling correlation to look at status reproduction.

Occupational status

The main piece of evidence used in this paper concerns the **occupations** as supplied by the marriage record. The HSN database records this information verbatim. I used the 1984 classification of Statistics Netherlands [the Netherlands Central Bureau of Statistics], which is closely related to the International Standard Classification of Occupations 1968 (ISCO68) of the International Labor Organization to code this information, since this leaves me with a number of options to recode the occupations into a meaningful classification or scale. Establishing a classification or scale for stratification data is an issue hotly debated among historical researchers, but since I want to look at social mobility using a multivariate design with continuous measures, my choices are in fact limited. The particular measure of occupational status I have chosen here is the 6-layer prestige measure developed by Van Tulder (1962) for the 1950's stratification data in a 1954 prestige survey. This measure not only has the advantages that it is established using a representative national opinion survey in a somewhat distant past and is directly connected to the classification scheme used to code the data, it is also convenient that it contains relatively few categories, and combines metric properties with a categorical view. In analyses not reported here, I have compared the effect of this choice with alternative options, in particular more detailed continuous scales like Treiman's (1977) SIOPS or Ganzeboom & Treiman's (19992, 1996) ISEI scale, and they give the

same account of the data.

Other variables

The other variables I use are relatively unproblematic. For all persons recorded at the certificate, year of birth or age is given. Since occupational achievement is somewhat age related, and the ages of witnesses and the groom are moderately correlated, I introduce **AGE** in the model as an individual control variable. Finally, I use year of marriage as the historical dimension in this analysis. The first marriage in the data was concluded in 1835 and the last one in 1942, although most of the data are concentrated in the 1840-1920 period. In order to maximize statistical power, I decided to organize the data into five cohorts of about equal size.

Sample considerations

The initial HSN sample consists of 2058 marriage records, about equally distributed among the provinces of Utrecht and Zeeland. Not all records contain information pertinent to social mobility analysis: there are some scattered missing values in the occupational information of the groom, but there only about 56% of the records contains useable information on the occupation of the father of the groom.

I extracted all possible male sibling pairs from the marriage records. Most of these are a pair of a groom and his brother (74%), but other pairs could be formed occasionally between pairs of groom's brother, as well as between bride's brothers. Altogether 1259 sibling pairs were extracted and it is fair to say that marriage records are (somewhat) richer with respect to siblings' occupational similarity than with respect to intergenerational occupational similarity. In order to avoid complexity, I took the liberty to force parallelism between the two siblings by allotting all brothers to the sibling 1 and the sibling 2 position on a random basis: by doing so, I can simplify the structural model by much using identity constraints between the two parts of the model (see further below).

Modern comparison data

In order to put the historical patterns observed from the HSN data in a modern perspective, I took a comparison sample from the 1992/93 Netherlands Family Survey (Ultee & Ganzeboom 1993;

Ganzeboom et al. 1995), that records complete occupational histories for 690 ever married men, as well as information on current occupation for a complete sibling roster for half of these men, as well as the brothers of the female respondents in the survey, as well as information on last or previous occupation of the father of the selected men. This information is identically coded with respect to occupation (a detailed code using CBS84, which is converted to the six Van Tulder prestige groups). I extracted the data using a strategy that stays as close as possible to design of the historical HSN data: I recorded respondent's occupation at his marriage, and correlated this with the occupations of all his brothers who were 25 years or older at the point of survey, and father's occupation when respondent was 14 years of age, or father's last or previous occupation, if this was not available.

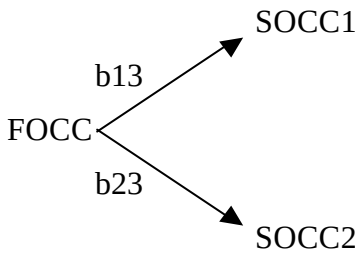
While this design is similar to that of the historical data, it is not identical. First, the occupations of father and in particular of the siblings are not the ones they would have at the marriage ceremony of the respondents. Second, in the modern data the information on fathers is far more complete than in the historical data. Third, in the modern data, all the information is recorded via the "groom", and it may be that he overstates the similarity with his brothers and/or father. Fourth, the modern data has aimed at measuring all brothers, while the historical data only have brothers that have been elected to become formal witness at the ceremony. Fifth and finally, the comparison data are national and not restricted to the two provinces. It is clear that these differences in design prevent a rigorous comparison between the historical data and the contemporary data. I nevertheless believe that it is useful to combine the two sources as it will give us some perspective on the lever of status reproduction in the historical data.

SIBLING METHODOLOGY

The primary question on a society's openness or closure is most often on the similarity between socio-economic characteristics of parents and their children. In particular and for various good reasons, research has concentrated on the similarity of occupation of fathers and their sons: occupational reproduction and occupational mobility are most frequently taken as the primary indicator of a society's openness at a given point in time. If we accept this point as valid, it is easy to see why a sibling model is a logical and important improvement. Assume that occupational status

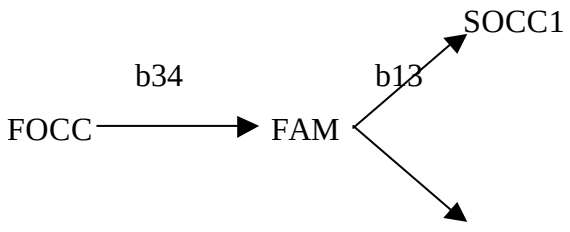
reproduction in a society is adequately summarized by parameter, such as a standardized regression coefficient between father's occupation FOCC and son's occupation SOCC:

If there are no special restrictions on the sampling of the son, we would observe the same number when not one son, but his brother was observed. Now assume that we would collect data on at least two brothers, we would observe the following:



This model implies that the expected correlation between the occupations of the two sons is b11*b21 = b11**2 (if b11=b21). However, in this type of data we also observe the true correlation between the occupations of the two siblings; this correlation is typically higher than the expected correlation constructed from the intergenerational reproduction coefficients. The difference between the expected and observed sibling correlation can be taken as a measure of the degree in which the intergenerational correlation misrepresents the degree to which the family of origin influences a man's occupation. The difference between the two coefficients may be caused by several factors, in particular that father's occupation is not the only factor that is responsible for intergenerational reproduction.

Seen from this perspective, the sibling correlation is a better measure of a society's openness than the intergenerational correlation. Moreover, the following sibling model, that includes father's occupation, would give us additional insights in the patterns of family influence.

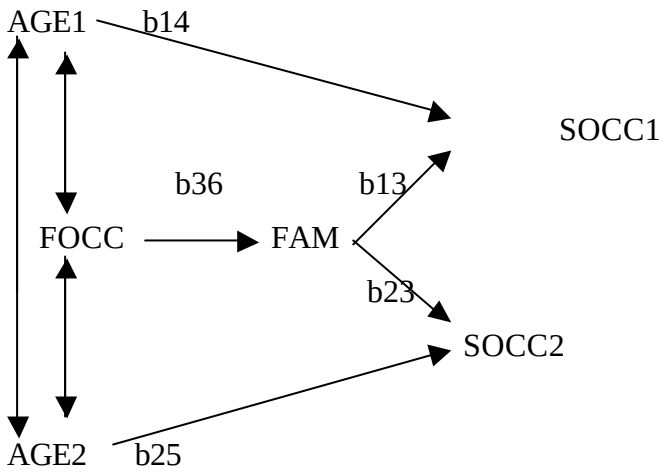


b23

SOCC2

This model separates the two coefficients, and we get two pieces of information: first, the effect of family (FAM --> SOCC) estimates the total family influence, while we have a second coefficient (FOCC --> FAM) that measure the degree to which father's occupation is indeed responsible for the family's effect.

We can detail the model a bit more by introducing other variables in the model. These can be introduced in model in the same position as father's occupation, and thus be a further specification of the family effect. This is appropriate for a variable like region or urban location, that constitutes a background that is shared by the brothers. Alternatively, we can introduce variables as individual effects, that confound the family effect. This is appropriate for the age effects that we observe in our data: older men tend to have higher prestige occupations, and the ages of siblings at marriage certificates tend to be moderately correlated. This has some effect on the similarity of occupations of sibling, that we would not regard as a true family effect. The following model corrects this:



The model can be estimated using a set of linear structural equations (LISREL; Jöreskog 1984, 1992). By introducing the assumption that the two siblings are completely parallel, we can reduce the number of coefficients to be estimated to four, while we have ten degrees of freedom: introducing these constraints brings out the statistical power of the sibling design.

ANALYSIS

Appendix A shows the correlations between the five variables of interest for the five cohorts, as well as for the total sample, as well as for the modern data. In Table 4 the fit statistics of a sequence of models are listed. Model (1) starts off with complete sibling symmetry, but does not use any other cross-cohort constraints. It fits the data quite well, which comes as no surprise, since full sibling symmetry was built into the data by design. The next models introduce cross-cohort constraints in order to test whether certain components in the model can be regarded as historical constants or not. Model (2a) tests the assumption of primary importance, whether the total family influence can be regarded a constant across cohorts. This is not the case by some margin: the chi-squared statistics rises by 11.9 on four degrees of freedom, which is statistically significant and shows that the total family effect is variable across the five historical cohorts. The same is true for the five cohorts in the modern data. Models 2b-2e restricts the coefficients in the model on other ways. Model 2b is not statistically different from Model 1 and shows that the degree two which father's occupational status represents family background can be regarded as constant across the cohorts. Further model specification leads to the conclusion that while age affects can be regarded as constant across cohorts in the historical, they cannot in the modern data. Model 3a is our preferred model. It states that the role of the father in determining status reproduction is constant across cohorts, whereas the

total family effects fluctuate.

Table 5 displays the standardize coefficients that are implied by the models. The first colums displays the constant effect of the father on the family background. It is 0.69 on the historical data and 0.66 in the modern data, which is remarkably similar. It informs is that about 46% of the total family effect is in fact explained by father's occupation. As this effect can be regarded as constant across time, there is no historical bias uncovered by the model. If we would go for the information on father's occupation alone, we could be underestimating the total family effect by more than half, but the bias would not be changing over time.

The second column in Table 5 shows the total family effect itself. For the historical data, we see that it increases when we move from the first cohort to the middle one, and then falls again. This would correspond to a trend towards more social fluidity after1890, and indeed confirm the point that the demise of social closure is primarily brought about by educational expansion. The family background coefficients are more or less stable in the modern data, more or less at the level that the last historical cohort ended. The is a puzzling and disconcerting drop of the family effect in the middle modern cohort. There is a clear unexpected non-linearity here that we cannot readily explain.

Table 1: Overview of the Historical Data

	Zeeland	Utrecht	Total
Number of marriage records:	1154	904	2058
Range of marriage years		1834-1942	
Grooms with valid occupation	1141	881	2022
Brothers of groom with valid occupation	503	298	801
Brothers of bride with valid occupation	477	303	780
Male sibling pairs with groom	503	425	928
Male sibling pairs among groom's brothers	104	54	158
Male sibling pairs among bride's brothers	93	80	173
Total sibling pairs	700	559	1259
Fathers with valid occupation	57.0%	53.6%	55.5%

Note: all counts refer persons with valid occupation score.

Table 2: Overview of the Modern (1992/93) Data

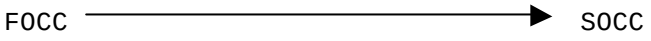
Initial N of ever married respondents	726
Range of year of marriage	1940-1992
"Grooms" with >0 brother	231
"Brides" with >1 brothers	152
Brothers of male respondents	518
Brothers of female respondents	527
Male sibling pairs with male respondent	501
Male sibling pairs among brothers of male respondents	610
Male sibling pairs among brothers of female respondents	570
Total number of sibling pairs	1681

Percentage of sibling pairs with information on father	99.3%

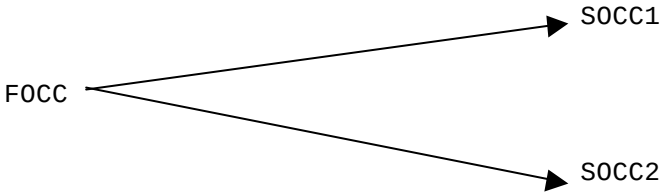
Note: all counts refer persons with valid occupation score.

Figure 3: Sibling Methodology

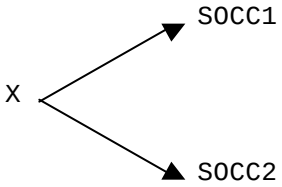
a. Intergenerational reproduction:



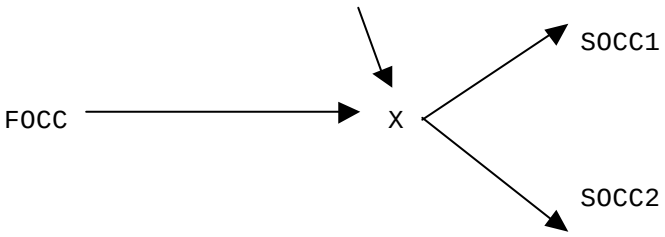
b. Intergenerational reproduction with two sons:



c. Intergenerational reproduction with two sons and no father:



d. Intergenerational reproduction with two sons and a father:



e. Intergenerational reproduction with two sons and a father and individual X-variables

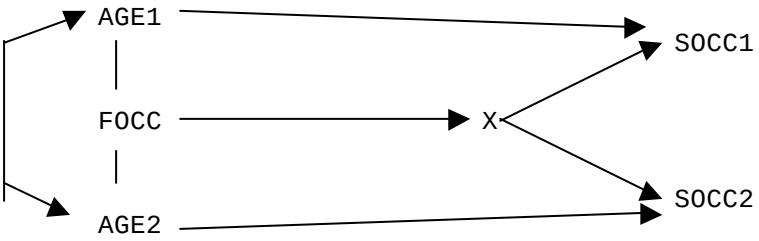


Table 4: Fit statistics sibling-models, HSN data 1835-1940 and Modern data 1940-1992; equality constrains across cohorts

	Historical		Modern	
1. Sibling symmetry	45	30.9	45	35.8
2a. (1) + equal family effects (B13)	49	42.8	49	51.4
2b. (1) + equal father's effect (B36)	49	32.1	49	43.3
2c. (1) + equal age effects (B14, B25)	49	39.6	49	52.2
2d. (2c) + equal age effects (PS46, PS56)	53	50.7	53	58.3
2e. (2c) + equal age correlations (PS45)	57	55.4	57	108.7
3a. (2d/e) + equal father's effects (BE36)	**61**	**56.1**	**57**	**66.1**
3b. (3a) + equal family effects (BE36)	65	71.1	61	79.1
3c. (3b) + equal family effects (PS11, PS33)	69	112.9	65	85.6

Table 5: Parameters of Sibling Model in Preferred Model

	Father's Effect (B36)	Family Effect (BE13) =BE23)	Age Effect (BE14) =BE25)	Age Corr. (PS45)
<u>Historical data</u>				
1835-1875	.687	.715	.077	.472
1876-1898	.687	.832	.077	.472
1898-1913	.687	.800	.077	.472
1914-1926	.687	.604	.077	.472
1927-1942	.687	.686	.077	.472
<u>Modern data</u>				
1940-1965	.661	.587	.099	-.051
1966-1971	.661	.566	.099	.031
1972-1976	.661	.383	.099	.087
1977-1982	.661	.574	.099	.224
1982-1992	.661	.608	.099	.423

Appendix: Correlations observed (A) HSN data 1835-1940: (B) Modern data (Netherlands Family Survey 1992/93)

	Historical data 1835-1942					Modern data 1940-1992				
	SOCC1	SOCC2	AGE1	AGE2	FOCC	SOCC1	SOCC2	AGE1	AGE2	FOCC
SOCC1	1.000	.521	.120	.016	.487	1.000	.369	.012	-.083	.314
SOCC2	.521	1.000	.144	.230	.500	.369	1.000	-.010	-.072	.364
AGE1	.120	.144	1.000	.473	.068	.012	-.010	1.000	-.063	-.007
AGE2	.016	.230	.473	1.000	.116	-.083	-.072	-.063	1.000	-.029
FOCC	.487	.500	.068	.116	1.000	.314	.364	-.007	-.029	1.000
SOCC1	1.000	.710	.173	.118	.580	1.000	.299	.153	.007	.400
SOCC2	.710	1.000	.070	.077	.635	.299	1.000	.095	.116	.448
AGE1	.173	.070	1.000	.397	.209	.153	.095	1.000	.034	.071
AGE2	.118	.077	.397	1.000	.201	.007	.116	.034	1.000	.105
FOCC	.580	.635	.209	.201	1.000	.400	.448	.071	.105	1.000
SOCC1	1.000	.666	.232	.212	.551	1.000	.184	.250	.065	.164
SOCC2	.666	1.000	.183	.205	.558	.184	1.000	.039	.085	.332
AGE1	.232	.183	1.000	.518	.166	.250	.039	1.000	.090	.102
AGE2	.212	.205	.518	1.000	.093	.065	.085	.090	1.000	.081
FOCC	.551	.558	.166	.093	1.000	.164	.332	.102	.081	1.000
SOCC1	1.000	.402	.244	.167	.397	1.000	.323	.202	.000	.392
SOCC2	.402	1.000	.217	.195	.441	.323	1.000	-.035	.138	.374
AGE1	.244	.217	1.000	.484	.283	.202	-.035	1.000	.224	.032
AGE2	.167	.195	.484	1.000	.273	.000	.138	.224	1.000	.055
FOCC	.397	.441	.283	.273	1.000	.392	.374	.032	.055	1.000
SOCC1	1.000	.469	.052	.007	.492	1.000	.384	.155	.202	.430
SOCC2	.469	1.000	-.009	.021	.454	.384	1.000	.083	.133	.411
AGE1	.052	-.009	1.000	.488	.125	.155	.083	1.000	.422	.004
AGE2	.007	.021	.488	1.000	.007	.202	.133	.422	1.000	.042
FOCC	.492	.454	.125	.007	1.000	.430	.411	.004	.042	1.000