

How Can We Analyse Change
in Statistical Associations Characterised
by Very Strong Inertia?

Recent Advances
in Log-Multiplicative Modelling

Louis-André Vallet (CNRS)

Quantitative Sociology Laboratory,
CREST, UMR 2773 CNRS & INSEE,
Paris

My contribution will be both statistical and sociological and will address the problem of analysis of trend data. But, at the end, I will clarify that the same idea or framework has also been applied in event history analysis.

More precisely, I would like to *review some methodological advances that occurred in the area of statistical modelling of contingency tables* over the last fifteen years.

These advancements concern the way we can analyse *how the association between two qualitative variables differs among the categories of a third variable* (e.g. among groups, among countries or over time).

And I would also like to demonstrate that *this methodological progress has had important consequences in the field of social stratification studies*.

More precisely, taking advantage of these powerful instruments, *social stratification researchers are now able to discern historical trends* in statistical associations that are characterised by considerable temporal inertia and that were generally regarded as highly stable ten years ago.

AN EXAMPLE

Let me suppose that, using national representative surveys, we are able to observe *the two-way contingency table that cross-classifies*:

- *father's occupational category*
(variable 'origin' O with NO categories)
- *with highest diploma obtained*
(variable 'education' E with NE categories)

for subsets of individuals (i.e. sons or daughters) who belong to NC five-year birth cohorts

(layer variable 'cohort' C with NC categories)

The 'origin' and 'education' variables are defined as qualitative and nominal variables, i.e. there is no total or simple ranking of all occupational classes and all educational diplomas.

In the real application I will present about France, *the NO x NE x NC three-way contingency table is a 8 x 7 x 13 table* with a total sample size of more than 240 000 individuals coming from seven nationally representative surveys.

SUBSTANTIVE QUESTIONS AND STANDARD ANALYSIS

From the end of the 1970s or the early 1980s, social stratification researchers have typically addressed the following questions about such data.

- *Does it exist any statistical association between father's occupational category and son's or daughter's highest diploma in this society?*
- *If it exists, has this statistical association remained constant over birth cohorts in this society?*
- *Or has it changed in the sense of weakening or strengthening over birth cohorts?*

To answer these questions, social stratification researchers have typically used *standard log-linear analysis with so-called hierarchical models*.

A FEW NOTATIONS

Let c be one of the birth cohorts (out of NC), o and o' two different social origins (out of NO), e and e' two different educational destinations (out of NE).

With these notations, m_{oec} is the estimated count in cell (o,e,c) of the three-dimensional table.

And od_c is the odds ratio which measures, in birth cohort c , the intrinsic strength of the statistical association between these social origins and these educational destinations:

$$od_c = \frac{m_{oec}/m_{oe'c}}{m_{o'ec}/m_{o'e'c}}$$

or alternatively:

$$\text{Log}(od_c) = \text{Log}(m_{oec}) + \text{Log}(m_{o'e'c}) - \text{Log}(m_{oe'c}) - \text{Log}(m_{o'ec})$$

Social stratification researchers have routinely considered *three hierarchical log-linear models*.

Null association model or Conditional independence model

$$\text{Log}(m_{oec}) = \lambda + \lambda_o^O + \lambda_e^E + \lambda_c^C + \lambda_{oc}^{OC} + \lambda_{ec}^{EC}$$

Estimated with $NC(NO-1)(NE-1)$ degrees of freedom, this model implies:

$$\text{Log}(od_c) = 0 \quad \text{or} \quad od_c = 1$$

This model assumes that social origin and educational destination are independent in each birth cohort.

It therefore expresses the hypothesis of complete equality of educational opportunity.

And it provides us with a reference for assessing the extent to which more realistic models fit the data more closely.

Constant association model or No three-way interaction model

$$\text{Log}(m_{oec}) = \lambda + \lambda_o^O + \lambda_e^E + \lambda_c^C + \lambda_{oc}^{OC} + \lambda_{ec}^{EC} + \lambda_{oe}^{OE}$$

Estimated with (NO-1)(NE-1)(NC-1) degrees of freedom, this model implies:

$$\text{Log}(od_c) = \lambda_{oe}^{OE} + \lambda_{o'e'}^{OE} - \lambda_{oe'}^{OE} - \lambda_{o'e}^{OE}$$

This model assumes that all the odds ratios that measure the association between social origin and educational destination are constant over birth cohorts.

It therefore expresses the hypothesis of constant inequality of educational opportunity.

Saturated model

$$\text{Log}(m_{oec}) = \lambda + \lambda_o^O + \lambda_e^E + \lambda_c^C + \lambda_{oc}^{OC} + \lambda_{ec}^{EC} + \lambda_{oe}^{OE} + \lambda_{oec}^{OEC}$$

Estimated with no degree of freedom, the saturated model perfectly reproduces the observed data and implies:

$$\text{Log}(od_c) = (\lambda_{oe}^{OE} + \lambda_{o'e'}^{OE} - \lambda_{oe'}^{OE} - \lambda_{o'e}^{OE}) + (\lambda_{oec}^{OEC} + \lambda_{o'e'c}^{OEC} - \lambda_{oe'c}^{OEC} - \lambda_{o'ec}^{OEC})$$

i.e. each odds ratio that measures the association between two social origins and two educational destinations varies freely from one birth cohort to another.

ON THE FRENCH DATA...

Model	G^2	df	DI	rG^2	Bic
<i>All men and women (N=240 367)</i>					
1. Null association model	50 388.8	546	15.9	-	43 623.9
2. Constant association model	4 107.8	504	4.1	91.8	-2 136.7
5. Saturated model	0.0	0	0	100.0	0.0

The null association model is extremely far from the data:

- the likelihood ratio chi-square exceeds 50 000;
- according to the index of dissimilarity, the model misclassifies nearly 16% of the total sample.

Conversely, the constant association model appears to have strong potential for describing French society over 13 five-year birth cohorts:

- this model misclassifies only 4.1% of the total sample;
- it eliminates nearly 92% of the total distance that separates the data from the null association model;
- according to the negative value of the Bayesian Information Criterion, *the constant association model has to be preferred to the saturated model.*

WHAT IS THE PROBLEM?

The constant association model appears relatively close to the data. *This suggests that the association between social origin and educational destination exhibits rather strong inertia over birth cohorts...*

...And sociologists were therefore inclined to conclude in favour of remarkably persistent inequality of educational opportunity (see, e.g., Garnier & Raffalovich, 1984; Shavit & Blossfeld (eds), 1993).

BUT...

The difference between the constant association model and the saturated model is statistically significant at any conventional level.

However, a shortcoming is that *this test disseminates its statistical power over $(NO-1)(NE-1)(NC-1)$ degrees of freedom (504 here!).*

Finally, it is simply unfeasible to examine the 504 independent parameters for the three-way interaction. In other words, we cannot really discover how the association between social origin and educational destination has changed over birth cohorts!

SO, WHAT ABOUT THE LIMITATION
OF STANDARD LOG-LINEAR ANALYSIS?

Except if we decide to focus on one or a few specific odds ratios, *standard log-linear analysis is unable to parsimoniously model how the association between two variables* (here, ‘social origin’ & ‘educational destination’) *varies over the categories of a third variable* (here, ‘birth cohort’).

BUT...

Following a suggestion given by Leo A. Goodman in *International Statistical Review* (1986)...

A SOLUTION HAS BEEN PROPOSED
IN THE CONTEXT OF LOG-MULTIPLICATIVE MODELLING...

THE LOG-MULTIPLICATIVE LAYER EFFECT MODEL,
also called THE UNIDIFF (for uniform difference) MODEL

This model was proposed in 1992, simultaneously by Yu Xie (in *American Sociological Review*) and by Robert Erikson & John Goldthorpe (in their book *The Constant Flux*).

$$\text{Log}(m_{oec}) = \lambda + \lambda_o^O + \lambda_e^E + \lambda_c^C + \lambda_{oc}^{OC} + \lambda_{ec}^{EC} + \beta_c \psi_{oe}$$

Estimated with (NO.NE-NO-NE)(NC-1) degrees of freedom (*i.e. only (NC-1) degrees of freedom less than the constant association model!*), this model implies:

$$\text{Log}(od_c) = \beta_c (\psi_{oe} + \psi_{o'e'} - \psi_{oe'} - \psi_{o'e})$$

i.e. it decomposes each logged odds ratio as the product of a common pattern (the expression with the ψ_{oe} parameters) and a cohort-specific parameter (β_c).

So, assuming a stable structure in the association between social origin and educational destination, this model is able to detect differences over cohorts in strength of association, i.e. in the general level of inequality of educational opportunity.

More precisely, assuming that β_1 is set at 1, estimating β_c as less than 1 (respectively more than 1) for a subsequent cohort will correspond to all estimated logged odds ratios moving towards 0 (respectively away from 0), i.e. will correspond to the association becoming weaker (respectively stronger) than in the first birth cohort.

This model therefore assumes that all odds ratios are moving in the same direction from one cohort to another and it expresses this variation with only one parameter.

As a consequence, it is very powerful to detect a dominant trend in the data, but it may be also rather crude to accurately describe the change that has occurred.

ON THE FRENCH DATA...

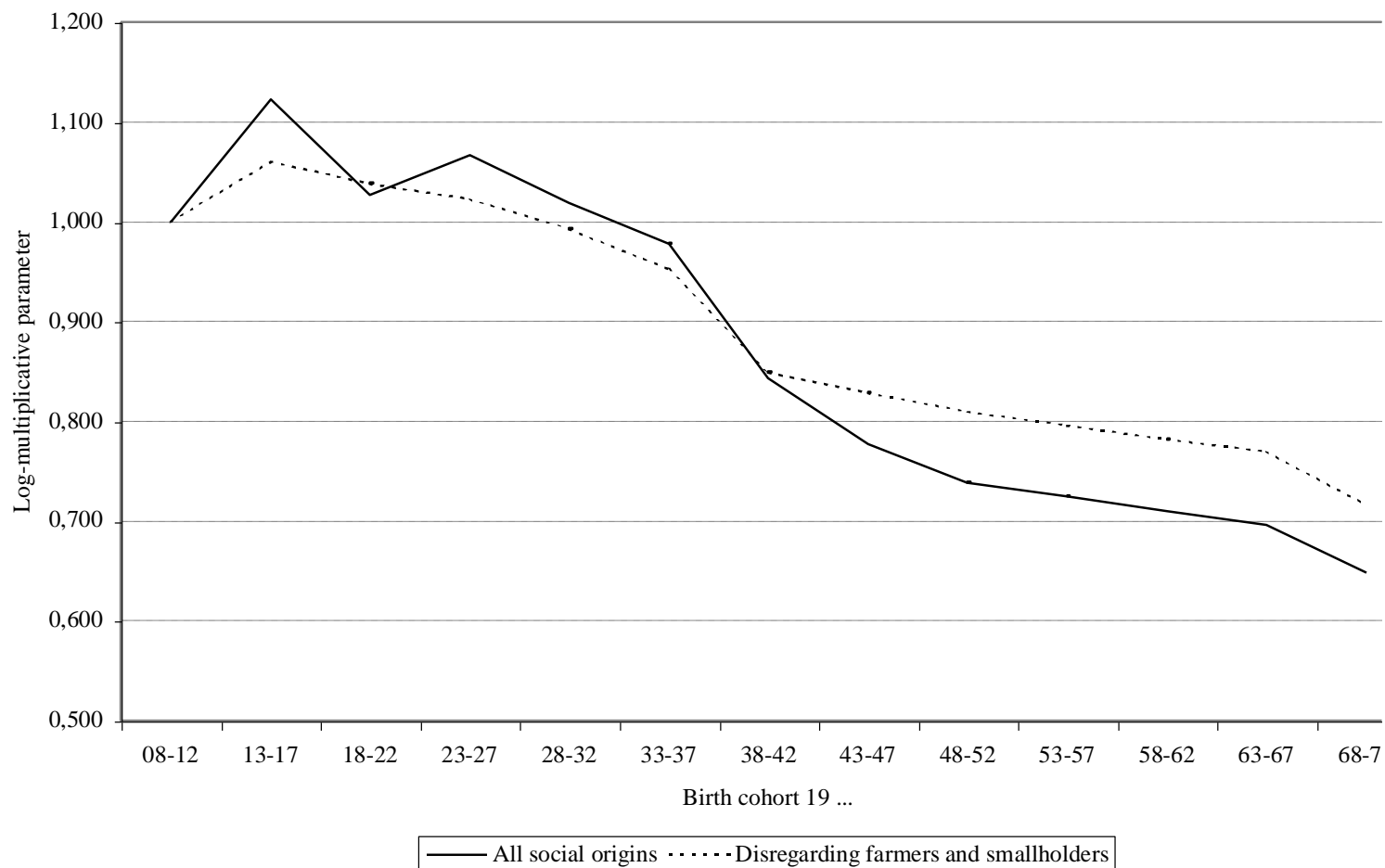
Model	G^2	df	DI	rG^2	Bic
<i>All men and women (N=240 367)</i>					
1. Null association model	50 388.8	546	15.9	-	43 623.9
2. Constant association model	4 107.8	504	4.1	91.8	-2 136.7
3. Multiplicative uniform cohort effect	3 442.1	492	3.9	93.2	-2 653.7
5. Saturated model	0.0	0	0	100.0	0.0

The test of the difference between the ‘multiplicative uniform cohort effect’ association model (or Unidiff model) and the constant association model is highly significant: the difference in the likelihood-ratio chi-square statistic is 665.7 for 12 degrees of freedom.

And, according to the more negative value of the Bayesian Information Criterion, the new model has also to be preferred to the constant association model.

We can therefore analyse the temporal dynamics revealed by the β_c parameters...

CHANGE IN THE GENERAL STRENGTH OF ASSOCIATION BETWEEN SOCIAL ORIGIN AND EDUCATIONAL DESTINATION – ‘MULTIPLICATIVE UNIFORM COHORT EFFECT’ ASSOCIATION MODEL ESTIMATED OVER THIRTEEN FIVE-YEAR BIRTH COHORTS (N=240,367)



We can see that, in French society, the general strength of the association between social origin and educational destination has declined by more than 35% (in the logged odds ratios) over sixty years.

While it has been nearly monotonic, change in the origin – education association was especially sharp between the 1933-1937 and 1948-1952 birth cohorts, then it largely levelled off in the three subsequent cohorts, but took off again in the very last one (1968-1972).

Finally, the interrupted line corresponds to a second analysis that ignores children of farmers and smallholders (using the technique of structural zeroes).

It demonstrates that change in the educational trajectories of children of farmers and smallholders played a major part in accentuating the general trend, but was not the only factor for creating it.

It is however possible to be more specific with a second and more accurate log-multiplicative model...

THE LOG-MULTIPLICATIVE REGRESSION-TYPE MODEL

This model was proposed as a more general one by Leo A. Goodman & Michael Hout in *Sociological Methodology* 1998 & 2001.

$$\text{Log}(m_{oec}) = \lambda + \lambda_o^O + \lambda_e^E + \lambda_c^C + \lambda_{oc}^{OC} + \lambda_{ec}^{EC} + \lambda_{oe}^{OE} + \gamma_c \varphi_{oe}$$

Estimated with (NO.NE-NO-NE)(NC-2) degrees of freedom, this model implies:

$$\text{Log}(od_c) = (\lambda_{oe}^{OE} + \lambda_{o'e'}^{OE} - \lambda_{oe'}^{OE} - \lambda_{o'e}^{OE}) + \gamma_c (\varphi_{oe} + \varphi_{o'e'} - \varphi_{oe'} - \varphi_{o'e})$$

The λ_{oe}^{OE} parameters establish *the baseline pattern of association* between social origin and educational destination.

The φ_{oe} parameters represent *the part of the association that varies over birth cohorts*.

Finally, the magnitude of the γ_c parameter determines *the strength of the adjustment of the association for cohort c*.

As a consequence, this model is able to detect differences over cohorts in both pattern and strength of association.

More precisely, examination of the φ_{oe} parameters will highlight the combined social origins and educational destinations for which change over cohorts has been the most pronounced.

And examination of the γ_c parameters will demonstrate which birth cohorts have been most affected by change in the association between social origin and education.

The following identifying constraints are especially useful:

$$\sum_o \lambda_{oe}^{OE} = \sum_e \lambda_{oe}^{OE} = \sum_o \varphi_{oe} = \sum_e \varphi_{oe} = 0 \quad , \quad \gamma_1 = 1 \quad \text{and} \quad \gamma_{NC} = 0$$

so that *the λ_{oe}^{OE} parameters will represent the pattern of association in the last (youngest) birth cohort and the φ_{oe} parameters will represent the pattern of deviation characteristic of the first (oldest) birth cohort as compared to the last one.*

ON THE FRENCH DATA...

Model	G^2	df	DI	rG^2	Bic
<i>All men and women (N=240 367)</i>					
1. Null association model	50 388.8	546	15.9	-	43 623.9
2. Constant association model	4 107.8	504	4.1	91.8	-2 136.7
3. Multiplicative uniform cohort effect	3 442.1	492	3.9	93.2	-2 653.7
4. Regression-type cohort effect	799.2	451	1.7	98.4	-4 788.6
5. Saturated model	0.0	0	0	100.0	0.0

The test of the difference between the ‘regression-type cohort effect’ association model and the Unidiff model is highly significant: the difference in the likelihood-ratio chi-square statistic is 2 642.9 for 41 degrees of freedom.

According to the more negative value of the Bayesian Information Criterion, the new model has also to be preferred to the Unidiff model.

Finally, the new model is impressively close to the data (or the saturated model): even if the sample size exceeds 240 000, the likelihood-ratio chi-square is less than 800 for 451 degrees of freedom and the index of dissimilarity is only 1.7%.

BASELINE PATTERN OF ASSOCIATION (λ_{oe}^{OE} PARAMETERS)
IN THE 1968-1972 BIRTH COHORT

λ_{oe}^{OE} parameters	No diploma	Primary education certificate	Lower secondary diploma	Lower vocational diploma	Upper secondary diploma	Lower tertiary degree	Upper tertiary degree
Farmers and smallholders	-0.190	0.492	-0.021	0.529	0.033	-0.132	-0.711
Artisans and shopkeepers	-0.036	0.042	0.032	0.125	-0.036	-0.054	-0.073
Higher-grade professionals and managers	-0.479	-1.064	-0.089	-0.640	0.284	0.650	1.339
Teachers and assimilated occupations	-0.583	-1.446	-0.201	-0.589	0.421	0.797	1.602
Lower-grade professionals and technicians	-0.271	-0.319	0.033	-0.203	0.123	0.294	0.342
Routine non manual workers	0.215	0.327	0.035	0.083	-0.072	-0.212	-0.375
Foremen and skilled manual workers	0.508	0.821	0.031	0.268	-0.281	-0.501	-0.847
Agricultural and unskilled manual workers	0.836	1.146	0.181	0.428	-0.472	-0.842	-1.277

In the 1968-1972 cohort, the highest diplomas (upper and lower tertiary degrees) are strongly and positively associated with the most advantaged classes of origin (teachers and assimilated occupations, then higher-grade professionals and managers).

Conversely, they are strongly and negatively associated with origins in the most disadvantaged classes (agricultural and unskilled manual workers, then foremen and skilled manual workers, and farmers and smallholders for upper tertiary degree only).

An opposite pattern characterizes the association between these social groups and the 'no diploma' and 'primary education certificate' destinations.

Finally, getting a lower vocational education diploma is positively associated with origins in the class of farmers and smallholders and the class of agricultural and unskilled manual workers, and negatively associated with the two most advantaged classes.

PATTERN OF DEVIATION (φ_{oe} PARAMETERS) IN THE
1908-1912 BIRTH COHORT (AS COMPARED TO THE 1968-1972 ONE)

φ_{oe} parameters	No diploma	Primary education certificate	Lower secondary diploma	Lower vocational diploma	Upper secondary diploma	Lower tertiary degree	Upper tertiary degree
Farmers and smallholders	2.325	0.883	-0.153	-0.828	-0.915	-1.006	-0.307
Artisans and shopkeepers	0.001	0.170	0.142	-0.032	-0.017	-0.235	-0.029
Higher-grade professionals and managers	-0.709	0.135	0.032	0.067	0.297	0.205	-0.027
Teachers and assimilated occupations	-1.333	0.227	0.544	-0.490	0.302	0.579	0.170
Lower-grade professionals and technicians	-0.847	-0.203	0.189	0.315	0.279	0.095	0.172
Routine non manual workers	-0.531	-0.399	0.107	0.282	0.008	0.279	0.256
Foremen and skilled manual workers	0.320	-0.400	-0.207	0.518	0.058	-0.021	-0.268
Agricultural and unskilled manual workers	0.774	-0.414	-0.653	0.169	-0.013	0.104	0.033

Taking the absolute value of φ_{oe} parameters and computing their average by row and column highlights those classes of origin whose relative position has been transformed most and those educational destinations with the most important contributions to historical change in the pattern of association.

It is especially visible that, in the past as compared to the 1968-1972 cohort, the 'no diploma' category was more strongly positively associated with disadvantaged classes (especially farmers and smallholders) and more strongly negatively associated with the two most advantaged classes and also the class of lower-grade professionals and technicians and the class of routine non manual workers.

Examining the profile of φ_{oe} parameters by row reveals that, as regards their educational destinations, sons and daughters of farmers and smallholders considerably improved their relative position over sixty years while the situation of children of artisans and shopkeepers remained much more stable.

In the early decades of the twentieth century, men and women originating in four classes were characterized by more favorable (relative) educational trajectories than they are in the most recent cohorts: these are the classes of teachers and assimilated occupations, higher-grade professionals and managers, lower-grade professionals and technicians, and routine non manual workers.

Finally, historical change in the educational destinations of the offspring of skilled and unskilled fractions of the working class appears more complex.

Children of foremen and skilled manual workers have improved their relative position more than children of agricultural and unskilled manual workers.

In the first birth cohorts as compared to the last ones, the former were more positively associated with the 'no diploma' and 'lower vocational diploma' categories and more negatively associated with the 'upper tertiary degree' category, but association did not change much between the offspring of the unskilled fraction of the working class and the highest educational destinations (upper secondary diploma, lower and upper tertiary degrees).

STRENGTH OF DEVIATION (γ_c PARAMETERS) OVER BIRTH COHORTS

Birth cohort	1908-1912	1913-1917	1918-1922	1923-1927	1928-1932	1933-1937	1938-1942
γ_c parameters	1	0.891	0.758	0.867	0.772	0.673	0.519
Birth cohort	1943-1947	1948-1952	1953-1957	1958-1962	1963-1967	1968-1972	
γ_c parameters	0.357	0.241	0.200	0.175	0.101	0	

Apart from an exception between the 1918-1922 and 1923-1927 cohorts the strength of the adjustment of the association declines monotonically over the period.

Again, change is especially marked between the 1933-1937 and 1948-1952 birth cohorts.

ON THE UNIDIFF MODEL AGAIN...

In the book *Social Mobility in Europe* edited by Richard Breen (2004, Oxford University Press), the chapter about France demonstrates a declining association between origin class (i.e. father's class, O) and destination class (i.e. respondent's class, D) for men aged 25 to 64.

The log-multiplicative or Unidiff parameter (β_t), fixed at 1 in 1970, was estimated at:

0.953 in 1977

0.887 in 1985

and 0.873 in 1993.

So, *intergenerational social fluidity has increased by about 13% (in the logged odds ratios) over a quarter of a century.*

And, *to explain such a change in social fluidity, it is again possible to use more sophisticated versions of the log-multiplicative layer effect model...*

INTRODUCING UNIDIFF TERMS IN A MULTI-LAYER 'TOPOLOGICAL' MODEL OF THE MOBILITY TABLE

'Topological' or 'Levels' log-linear models of the mobility table (Hauser, 1978) assume that *the structure and strength of the association between origin class and destination class can be captured in a limited number of interaction parameters (or 'density levels')*.

Erikson & Goldthorpe (1992) have proposed such a theoretical representation of social fluidity in a modern society. *In their 'core' model, 'density levels' result from the combination of eight more elementary parameters.*

These parameters capture different effects:

- *Hierarchy: HI1 (short-range mobility) & HI2 (long-range mobility);*
- *Inheritance: IN1 (all classes), IN2 (self-employed classes only) & IN3 (farmers only);*
- *Sector: SE (differentiating between the primary sector and the other sectors);*
- *Affinity: AF1 (negative) & AF2 (positive).*

THEORETICAL REPRESENTATION OF CORE SOCIAL FLUIDITY IN THE 7 x 7 MOBILITY TABLE

D O	I+II	III	IVa+b	IVc	V+VI	VIIa	VIIb
I+II	IN1+IN2	HI1+AF2	HI1+AF2	HI1+SE	HI1	HI1+HI2	HI1+HI2+SE+AF1
III	HI1+AF2	IN1	-	SE	-	HI1	HI1+SE
IVa+b	HI1+AF2	-	IN1+IN2	SE+AF2	-	HI1	HI1+SE
IVc	HI1+HI2+SE	HI1+SE	HI1+SE+AF2	HI1+IN1+IN2+IN3	HI1+SE	SE+AF2	-
V+VI	HI1	-	-	SE	IN1	HI1+AF2	HI1+SE
VIIa	HI1+HI2	HI1	HI1	HI1+SE	HI1+AF2	IN1	SE
VIIb	HI1+HI2+SE+AF1	HI1+SE	HI1+SE	HI1	HI1+SE	SE+AF2	IN1

Hierarchical divisions in the social structure are showed horizontally and sectorial divisions are showed vertically.

The table shows how the elementary parameters combine to build the 'density levels' that are characteristic of the 7 x 7 mobility table.

On the basis of this theoretical representation, it is possible to identify factors responsible for change in social fluidity as follows...

Model (seven-class schema)	G^2	df	DI	rG^2	Bic			
A. Constant social fluidity	300.1	108	2.4	98.6	-881.4			
Core Models								
B. Temporally invariant parameters	568.6	136	3.3	97.4	-919.1			
	<i>H11</i>	<i>H12</i>	<i>IN1</i>	<i>IN2</i>	<i>IN3</i>	<i>SE</i>	<i>AF1</i>	<i>AF2</i>
<i>France 1970, 1977, 1985, 1993</i>	-0.243	-0.570	0.399	0.824	1.140	-0.803	-0.817	0.451
<i>France 1970 (The Constant Flux, p. 147)</i>	-0.24	-0.47	0.41	0.92	1.00	-0.89	-0.75	0.47
C. Temporally changing parameters	451.0	112	2.9	97.9	-774.2			
	<i>H11</i>	<i>H12</i>	<i>IN1</i>	<i>IN2</i>	<i>IN3</i>	<i>SE</i>	<i>AF1</i>	<i>AF2</i>
<i>1970</i>	-0.318	-0.703	0.350	0.975	0.935	-0.931	-0.572	0.459
<i>1977</i>	-0.289	-0.574	0.393	0.801	1.424	-0.665	-1.156	0.470
<i>1985</i>	-0.171	-0.538	0.425	0.742	1.091	-0.777	-0.767	0.426
<i>1993</i>	-0.083	-0.441	0.481	0.809	1.264	-0.687	-0.665	0.407
D. Only H11 changes over time	515.5	133	3.1	97.6	-939.5			
E. Only H12 changes over time	544.7	133	3.3	97.5	-910.2			
F. Only IN1 changes over time	550.0	133	3.3	97.5	-904.9			
G. Only IN2 changes over time	532.4	133	3.2	97.5	-922.6			
H. Only IN3 changes over time	566.4	133	3.3	97.4	-888.5			
I. Only SE changes over time	555.2	133	3.3	97.4	-899.8			
J. Only AF1 changes over time	561.5	133	3.3	97.4	-893.4			
K. Only AF2 changes over time	567.8	133	3.4	97.4	-887.1			
L. H11 and H12 change over time	499.5	133	3.1	97.7	-955.4			
<i>UNIDIFF for H11 and H12</i>	<i>1.000 (1970)</i>	<i>0.830 (1977)</i>	<i>0.619 (1985)</i>	<i>0.509 (1993)</i>				
M. Model L with a linear trend	500.7	135	3.1	97.7	-976.1			
<i>Linear trend per year for H11 and H12</i>	<i>-0.0230</i>							
N. Model M + SE changes over time	489.7	132	3.0	97.7	-954.3			
<i>Linear trend per year for H11 and H12</i>	<i>-0.0227</i>							
<i>UNIDIFF for SE</i>	<i>1.000 (1970)</i>	<i>0.825 (1977)</i>	<i>0.839 (1985)</i>	<i>0.846 (1993)</i>				
O. Model N with equality constraint	489.8	134	3.0	97.7	-976.1			
<i>Linear trend per year for H11 and H12</i>	<i>-0.0226</i>							
<i>UNIDIFF (constrained) for SE</i>	<i>1.000 (1970)</i>	<i>0.834 (1977, 1985, 1993)</i>						

The temporally invariant version of the core model (model B) indeed compares very favourably with the constant social fluidity model (model A): its BIC statistic is better and it misclassifies only 3.3% of the total sample. The eight parameters estimated from the four surveys are also very close to those obtained by Erikson and Goldthorpe for France on the basis of the 1970 survey.

With 24 supplementary parameters, the temporally changing version of the core model (model C) provides a G^2 statistic which is 117.6 points lower. Within the context of the core model, this represents the whole change in social fluidity which has taken place over 23 years, but a more comprehensible account would be provided if we were able to model this variation – or a large part of it – using only a few parameters, in addition to the eight basic effects.

As a first step in this direction I estimated a series of models which incorporate a Unidiff effect – or a log-multiplicative layer effect – over time for *only one* of the eight core parameters. Using the temporally invariant core model as a benchmark, a major improvement in the G^2 statistic is provided when this effect is applied to HI1 (model D), and slightly less marked improvements when it is applied to IN2 and HI2 (models G and E).

In model L, I have therefore imposed a Unidiff effect on *both* the hierarchy parameters *simultaneously*, which affords the best-fitting model ever viewed in the table. Moreover, the monotonic change which is depicted in the Unidiff parameters can be summarised, without any significant loss of information, as a linear trend (model M).

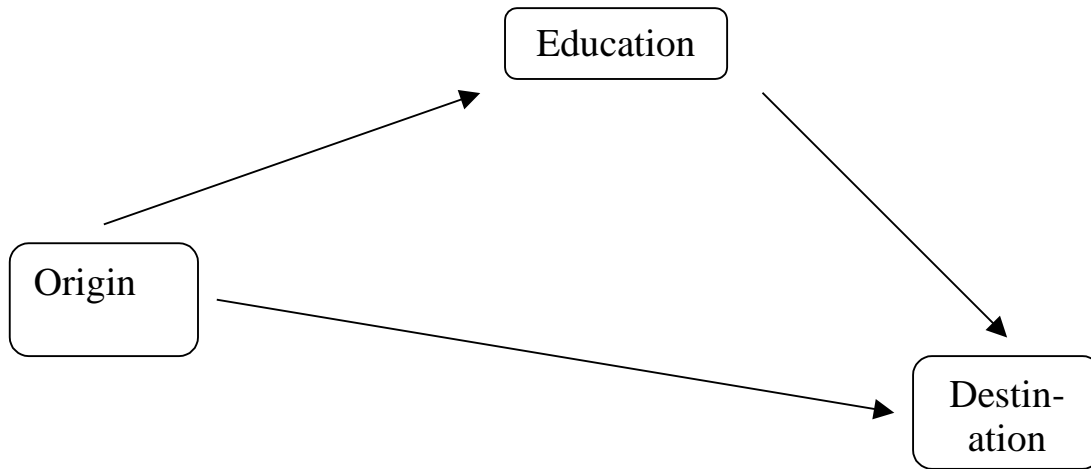
Finally, adding another Unidiff effect to the sectorial parameter significantly lowers the G^2 statistic (model N) and this effect can be represented as a threshold effect which opposes the 1970 survey to the subsequent ones (model O).

Although I investigated a number of supplementary variants of the core model, I was unable to find a more powerful model than model O: with only two parameters, it eliminates 67.1% of the distance between the temporally invariant and temporally changing versions of the model of core social fluidity.

And these parameters provide us with a straightforward understanding of the changing mobility regime in French society: the increase in social fluidity mainly resulted from a progressive weakening in the hierarchical divisions within the class structure which have to be passed through in intergenerational transitions; it also resulted from a reduced distance between the agricultural classes and the other classes.

ANALYSING TRENDS BETWEEN ORIGIN, EDUCATION & DESTINATION

If education is introduced as an intermediate variable, the decline in the association between origin class and destination class could be explained by four different effects:



- a weakening association between Origin and Education;
- a weakening partial association between Education and Destination;
- a weakening partial association between Origin and Destination;
- a compositional effect (the partial association between Origin and Destination is weaker in more educated groups whose relative importance in the labour force has grown from 1970 to 1993).

And the test of these four hypotheses is as follows...

Model	G^2	df	DI	rG^2	Bic
<i>Analysis of the Origin-Education-Time table</i>					
A. Conditional independence {OT}{ET}	12 720.2	192	16.5	-	10 619.8
B. Constant association {OT}{ET}{OE}	418.4	144	2.8	96.7	-1 156.9
C. UNIDIFF	335.9	141	2.5	97.4	-1 206.5
<i>UNIDIFF parameters for {OE} 1.000(1970) 0.886(1977) 0.808(1985) 0.784(1993)</i>					
<i>Analysis of the Origin-Education-Destination-Time table</i>					
D. Conditional independence {OET}{DT}	41 547.0	1 482	32.6	-	25 334.8
E. Constant {OD} association {OET}{DT}{OD}	20 126.4	1 446	20.4	51.6	4 307.9
F. Constant {ED} association {OET}{DT}{ED}	15 920.4	1 434	17.1	61.7	233.3
G. Constant associations {OET}{DT}{OD}{ED}	2 449.1	1 398	5.8	94.1	-12 844.3
H. Only {OD} changes over time	2 428.1	1 395	5.7	94.2	-12 832.5
<i>UNIDIFF parameters for {OD} 1.000 0.956 0.894 0.890</i>					
I. Only {ED} changes over time	2 342.4	1 395	5.5	94.4	-12 918.2
<i>UNIDIFF parameters for {ED} 1.000 0.910 0.855 0.740</i>					
J. Both {OD} and {ED} change over time	2 325.5	1 392	5.4	94.4	-12 902.2
<i>UNIDIFF parameters for {OD} 1.000 0.961 0.902 0.909</i>					
<i>UNIDIFF parameters for {ED} 1.000 0.912 0.859 0.743</i>					
K. Model I + {OD} changes over education	2 238.7	1 387	5.4	94.6	-12 934.3
<i>UNIDIFF (time) parameters for {ED} 1.000(1970) 0.906(1977) 0.851(1985) 0.740(1993)</i>					
<i>UNIDIFF (education) parameters for {OD} 1a 1b 1c 2a 2b 2cgen 2cvoc 3a 3b</i>					
<i>1.000 1.089 1.014 0.988 0.858 0.499 0.814 0.725 0.633</i>					

First, I have analysed the dynamics of the association between origin class and education. The Unidiff model is preferable to the constant association model and its parameters clearly reveal a decline of 21.6% (in the logged odds ratios) in the strength of the Origin – Education association between the population surveyed in 1970 and that surveyed in 1993 (model C).

The negative trend progressively decelerates, which is fully consistent with what I presented previously (most of the change took place among cohorts born between the mid-thirties and the mid-fifties). *The first conclusion is that a decline in inequality of educational opportunity has occurred.*

Second, I have analysed the three-dimensional Origin – Education – Destination tables from a temporal perspective.

The rG^2 statistic clearly suggests that the partial Education – Destination association is stronger than the partial Origin – Destination association (comparing models E and F).

Then, starting with the model that incorporates these two associations (model G), I introduce a Unidiff effect over time on one, the other, or both of them (models H, I and J). The G^2 and Bic statistics clearly favour model I which reveals a decline of 26.0% (in the logged odds ratios) in the strength of the Education – Destination association. *The second conclusion is that a decline in the relative occupational advantage afforded by education has occurred, but that the direct effect of origin on destination has changed little in France over the 1970-93 period.*

Finally, an even better fit is achieved by supplementing model I with a Unidiff effect on education for the Origin – Destination association (model K). The estimated parameters clearly reveal that the direct effect of origin on destination is generally weaker among people with more qualifications – especially from the intermediate general qualification (2b) to the highest tertiary qualifications (3a and 3b). *The third conclusion is that a compositional effect has played a role, progressively increasing the size and influence of the educational categories for which the direct effect of origin on destination is reduced.*

LOG-MULTIPLICATIVE MODELLING

CAN BE EASILY IMPLEMENTED

USING THE

LEM Software

Developed by **Jeroen K. Vermunt (University of Tilburg)**

<http://www.uvt.nl/faculteiten/fsw/organisatie/departementen/mto/software2.html>

- *The log-multiplicative framework has also been successfully applied to event history analysis.*
- **First**, as emphasized in Jeroen Vermunt's book (*Log-Linear Models for Event Histories*, 1997, chapter 4), many EHA models are log-linear hazard models. For instance, this is the case for the piecewise constant exponential model and the Cox semi-parametric model.

When time is discrete and all covariates are categorical, event history data can be summarized in a tabular form.

And EHA models then correspond to a particular subset of log-linear models for contingency tables: *the log-rate models* (where a rate is the ratio between the expected number of events and the exposure, i.e. the number of persons at risk).

- **Second**, in his paper « Log-Multiplicative Models for Discrete-Time, Discrete-Covariate Event-History Data » (*Sociological Methodology* 1994),

Yu Xie has used the log-multiplicative framework to parsimoniously model time-covariate interactions in log-rate models (*see the abstract of the paper on page 301*).