A New Methodological Approach for Studying Intergenerational Mobility with an Application to Swiss Data

Ben Jann and Simon Seiler

Institute of Sociology, University of Bern
ben.jann@soz.unibe.ch, simon.seiler@soz.unibe.ch

Spring Meeting of the Research Committee on Social Stratification and Mobility (RC28) of the International Sociological Association (ISA) on "Social Inequality, Cohesion and Solidarity"
Tilburg, May 28–30, 2015

Overview

- Introduction
- Methods
  - Log-multiplicative layer effect model
  - The PRE approach
- Data
- Results
  - Comparison of LMLEM and PRE
  - Smoothed PRE
  - PRE with multiple origin (and control) variables
  - Direct and indirect origin effects
- Conclusions

Introduction

- As a methodology to analyze the development of social mobility based on categorical variables such as class or educational attainment, Erikson and Goldthorpe (1992) and Xie (1992) independently proposed a variant of the log-linear model known as the *uniform difference model* (Unidiff) or the *log-multiplicative layer effect model* (LMLEM).
- The model has been the standard tool in social mobility research since. The model, however, also has some limitations. We therefore propose an alternative approach.
- Our approach is based on the concept of proportional reduction of error (PRE). It quantifies the degree to which information about parents helps predicting the status of the children.
- The approach, we believe, is more flexible than the LMLEM and provides results that are easier to interpret.

Log-Multiplicative Layer Effect Model

- Starting point of the LMLEM is a simple two-way table of origin and destination, called a "mobility table," such as the following:

	Respondent's education					Total
Parent's education	compulsory or less	secondary vocational	secondary general	tertiary vocational	tertiary academic	
compulsory or less	170	299	12	58	63	602
secondary vocational	37	708	27	134	260	1167
secondary general	5	19	3	20	16	62
tertiary vocational	7	51	15	104	52	229
tertiary academic	14	75	12	33	293	426
Total	232	1152	70	348	683	2485

Source: see data section. Selection: males, birth cohorts 1969-82

Log-Multiplicative Layer Effect Model

- Such a two-dimensional mobility table can be formalized as follows, where F_{ij} are observed cell counts and $F_{i.}$ and $F_{.j}$ are row and column totals.

	1	...	j	...	J	Total
1	F_{11}	...	F_{1j}	...	F_{1J}	$F_{1.}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i1}	...	F_{ij}	...	F_{iJ}	$F_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I1}	...	F_{Ij}	...	F_{IJ}	$F_{I.}$
Total	$F_{.1}$	...	$F_{.j}$	...	$F_{.J}$	$F_{..}$

- In a log-linear model, the cell counts are expressed as a multiplicative function:

$$F_{ij} = \tau_{..} \cdot \tau_{i.} \cdot \tau_{.j} \cdot \tau_{ij}, \quad i = 1, \dots, I, \; j = 1, \dots, J$$

Log-Multiplicative Layer Effect Model

- Now think of a table with an additional dimension (e.g. time points or birth cohorts).
- The saturated log-linear model for such a three-dimensional table is

$$F_{ijk} = \tau_{...} \cdot \tau_{i..} \cdot \tau_{.j.} \cdot \tau_{..k} \cdot \tau_{i.k} \cdot \tau_{.jk} \cdot \tau_{ij.} \cdot \tau_{ijk}$$

with $i = 1, \ldots, I$, $j = 1, \ldots, J$, and $k = 1, \ldots, K$

	1	...	j	...	J	Total
1	F_{111}	...	F_{1j1}	...	F_{1J1}	$F_{1.1}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i11}	...	F_{ij1}	...	F_{iJ1}	$F_{i.1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I11}	...	F_{Ij1}	...	F_{IJ1}	$F_{I.1}$
Total	$F_{.11}$	...	$F_{.j1}$	...	$F_{.J1}$	$F_{..1}$

$\vdots$

	1	...	j	...	J	Total
1	F_{11k}	...	F_{1jk}	...	F_{1Jk}	$F_{1.k}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i1k}	...	F_{ijk}	...	F_{iJk}	$F_{i.k}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I1k}	...	F_{Ijk}	...	F_{IJk}	$F_{I.k}$
Total	$F_{.1k}$	...	$F_{.jk}$	...	$F_{.Jk}$	$F_{..k}$

$\vdots$

	1	...	j	...	J	Total
1	F_{11K}	...	F_{1jK}	...	F_{1JK}	$F_{1.K}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	F_{i1K}	...	F_{ijK}	...	F_{iJK}	$F_{i.K}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	F_{I1K}	...	F_{IjK}	...	F_{IJK}	$F_{I.K}$
Total	$F_{.1K}$	...	$F_{.jK}$	...	$F_{.JK}$	$F_{..K}$

Log-Multiplicative Layer Effect Model

- To ease interpretation, Xie (1992) proposed a simplified model in which $\tau_{ij.} \cdot \tau_{ijk}$ is replaced by $\exp(\psi_{ij.} \cdot \phi_{..k})$. This is called the log-multiplicative layer effect model:

$$F_{ijk} = \tau_{...} \cdot \tau_{i..} \cdot \tau_{.j.} \cdot \tau_{..k} \cdot \tau_{i.k} \cdot \tau_{.jk} \cdot \exp(\psi_{ij.} \cdot \phi_{..k})$$
$$i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, K$$

- The $\psi_{ij.}$ parameters capture the overall pattern of dependencies between origin and destination.
- The $\phi_{..k}$ are cohort-specific scaling factors. That is, the higher $\phi_{..k}$, the more pronounced is the pattern of dependencies in cohort k and, hence, the stronger is the association between origin and destination, assuming that there is a stable basic pattern of associations across cohorts.
- To identify the model, constraints have to be placed on $\phi_{..k}$. Following Xie (1992), we use constraint $\sum_k \phi^2_{..k} = 1$.

The PRE Approach

- The LMLEM provides a parsimonious and intuitive way to describe changes in social mobility across time and also allows testing against a null model with time-constant origin effects. However, the model has a number of limitations.
  - First, it assumes a common baseline pattern of associations that remains constant over time. This assumption may be violated so that results are biased.
  - Second, it is difficult to extend the model to include control variables.
  - Third, there is no clear interpretation of the absolute values of $\phi_{..k}$. In fact, the overall level of the $\phi_{..k}$ parameters is meaningless, because the sum over $\phi^2_{..k}$ is restricted to 1. This implies that $\phi_{..k}$ cannot be compared across models.
- We therefore propose an alternative approach that is based on (categorical) regression models and the PRE principle.

The PRE Approach

- General ideas:
  - The stronger the effect of the status of the parents on the status of their children, the lower is intergenerational mobility.
  - The „strength" of an effect is easy to conceptualize for single regression coefficients. Things get more complicated, however, if we have to determine the strength of an effect that comprises multiple parameters.
  - Instead of thinking in terms of model parameters, however, we can ask how "useful" the information on parents is to predict the status of their children.
  - The better the position of children can be predicted based on parents characteristics, the stronger is the influence of social origin and the lower is social mobility.
  - To quantify the predictive power of parents' characteristics we can resort to the statistical concept of the Proportional Reduction of Error (PRE).

The PRE Approach

- Formally:

$$PRE = \frac{E_0 - E_1}{E_0} = 1 - \frac{E_1}{E_0}$$

  where E_0 is the sum of prediction errors under limited information and E_1 is the sum of prediction errors under full information.

- Different error rules can be applied, yielding different PRE measures. Because our dependent variables are categorical, an entropy-based definition (see Theil 1970) appears appropriate:

$$E_m = -\sum_{i=1}^{N} w_i \ln\left(\widehat{p}_m(Y = y_i)\right), \quad m = 0, 1$$

  where w_i is the respondent's survey weight and $\widehat{p}_m(Y = y_i)$ is the predicted probability of the dependent variable taking on observed value y_i under model m.

The PRE Approach

- To estimate $\widehat{p}_m(Y = y_i)$ we use multinomial logit models.
- That is, the probabilities under restricted information are modeled as

$$p_0(Y = y_i) = \frac{\exp(\beta_{y_i} Z_i)}{\sum_{\ell=1}^{J} \exp(\beta_\ell Z_i)}$$

  where Z_i is a vector of control variables (possibly just a constant) and β_ℓ is an outcome-specific coefficient vector.
- Likewise, the probabilities under full information are modeled as

$$p_1(Y = y_i) = \frac{\exp(\beta_{y_i} Z_i + \gamma_{y_i} X_i)}{\sum_{\ell=1}^{J} \exp(\beta_\ell Z_i + \gamma_\ell X_i)}$$

  where X_i is a vector of parents' characteristics.

The PRE Approach

- Categorical variant:
  - Estimate separate models for different birth cohorts and compute a separate PRE value for each cohort.
  - Gives only a crude picture because single birth years usually have to be collapsed into broader cohorts.
- Smoothed variant:
  - Compute a PRE value for each birth year, but include data from surrounding years using kernel weights to stabilize estimation.
  - We use weights

$$w_i(t^*) = w_i \cdot \frac{1}{h} K\left(\frac{t^* - t_i}{h}\right)$$

    where t^* is the target birth year, t_i is observations i's birth year, and $K()$ is the Epanechnikov kernel. We set bandwidth h to 5, so that the data window covers a maximum of ± 4 years around target birth year.
- For both variants, confidence intervals are obtained by bootstrap methods.

Data

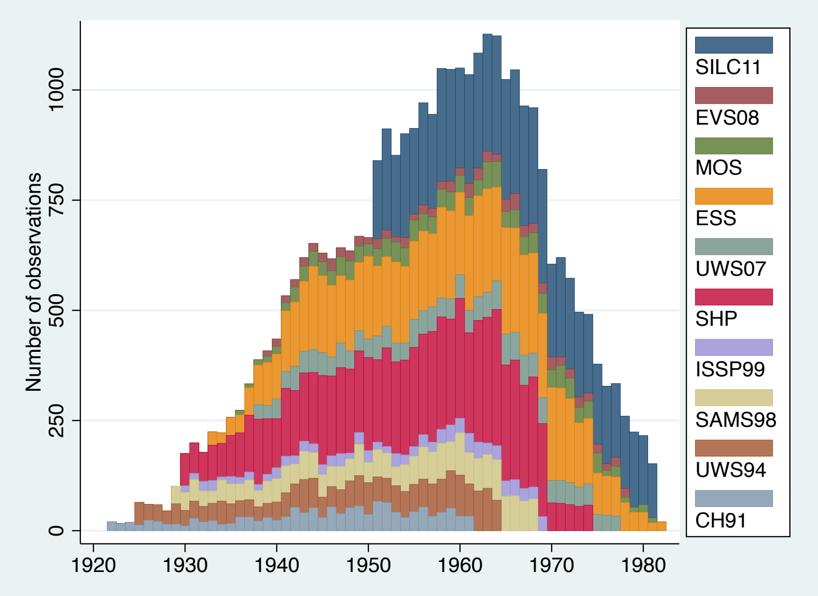
- Required are data that contain the relevant status variables for the respondents as well as information about education and occupation of parents.
- Most Swiss large-scale surveys, such as the official surveys by the Federal Statistical Office, do not contain information on parents.
- Nonetheless, we were able to identify a number of Swiss surveys that can be used for these types of analyses. The results below are based on a selection of these surveys. More surveys are available (especially some older ones) and will be incorporated in future.
- We harmonized the variables in the different surveys to build a common database that can be analyzed in terms of birth cohorts. The age range of respondents we restricted to 30 through 69.

Data: Included Surveys

	Year/Wave	N^a	Label
Les Suisses et leur société	1991	1331	CH91
Swiss Environmental Survey	1994	2233	UWS94
	2007	1973	UWS07
Swiss Labor Market Survey 1998	1998	2340	SAMS98
ISSP "Social inequality"	1999	972	ISSP99
Swiss Household Panel	1999	5365	SHP99
	2004	2420	SHP04
European Social Survey	2002	1450	ESS02
	2004	1457	ESS04
	2006	1267	ESS06
	2008	1187	ESS08
	2010	985	ESS10
	2012	945	ESS12
MOSAiCH (ISSP)	2005	741	MOS05
	2011	819	MOS11
European Values Study 2008	2008	830	EVS08
Statistics on Income and Living Conditions	2011	6753	SILC11
Total		33068	

[a] Number of observations available for our analyses.

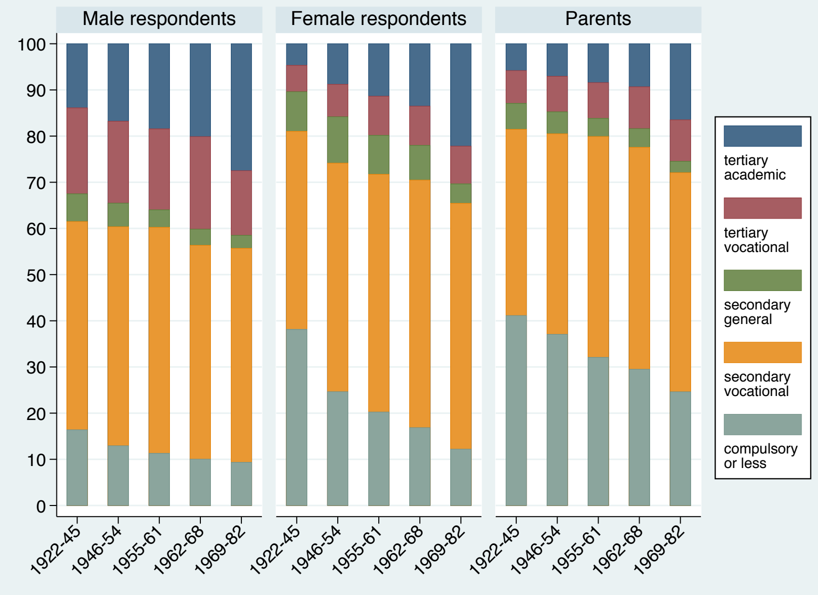
Data: Number of Observations by Birthyear



Data: Classification of Education

Educational level	Included educational degrees
Compulsory or less	No formal education; compulsory education; one year vocational training
Secondary vocational	Vocational training and education; general education without baccalaureate
Secondary general	General education with baccalaureate; vocational baccalaureate; college of education (without university of education)
Tertiary vocational	Professional education and training; advanced federal professional and training diploma; professional education college; university of applied sciences; university of education
Tertiary academic	University; Federal Institute of Technology

Data: Education by Birth Cohort

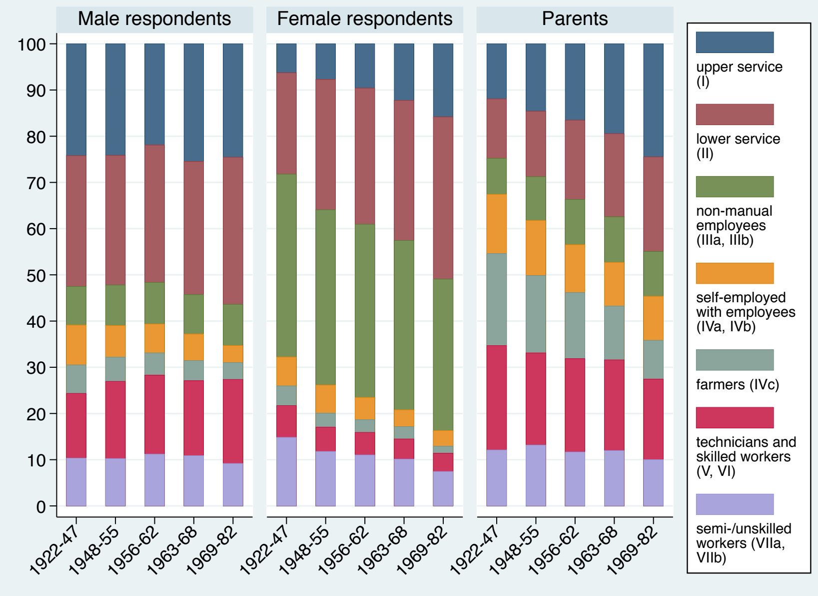


Data: Social Class Scheme

EGP Class		Description
I	Upper service	Higher-grade professionals, administrators and officials; managers in large industrial establishments; large proprietors
II	Lower service	Lower-grade professionals, administrators and officials; higher-grade technicians; managers in small business and industrial establishments; supervisors of non-manual employees
III	Non-manual employees	Routine non-manual employees in administration and commerce; sales personnel; other rank-and-file service workers
IVa,b	Self-employed	Small proprietors, artisans, etc., with employees (IVa); without employees (IVb)
IVc, VIIb	Farmers	Farmers and smallholders, self-employed fishermen (IVc); Agricultural workers (VIIb)
V, VI	Technicians and skilled workers	Lower-grade technicians; supervisors of manual workers; skilled manual workers
VIIa,b	Semi-/unskilled workers	Semi- and unskilled manual workers

Based on Erikson, Goldthorpe and Portocarero (1983: 307).

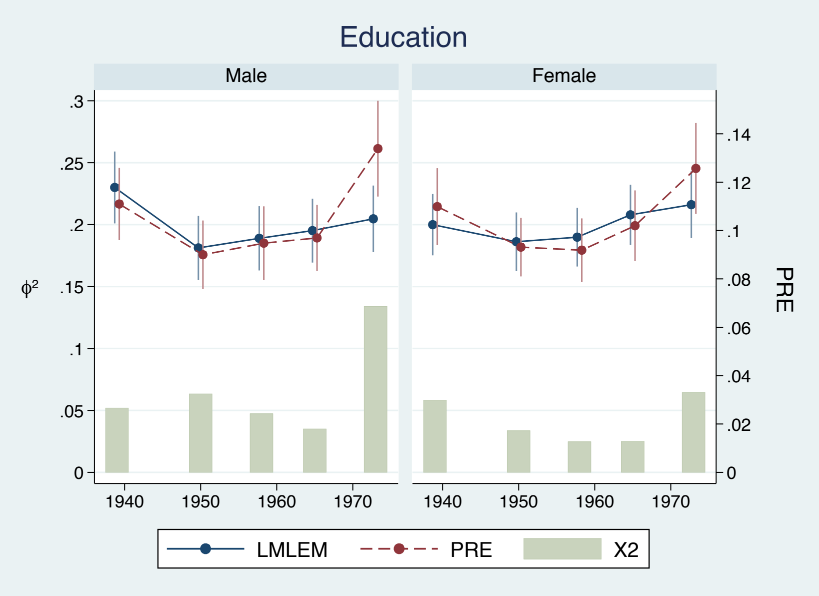
Data: Social Class by Birth Cohort



Results

- Comparison of LMLEM and PRE
- Smoothed PRE
- PRE with multiple origin variables
- Direct and indirect origin effects

Results: Comparison of LMLEM and PRE



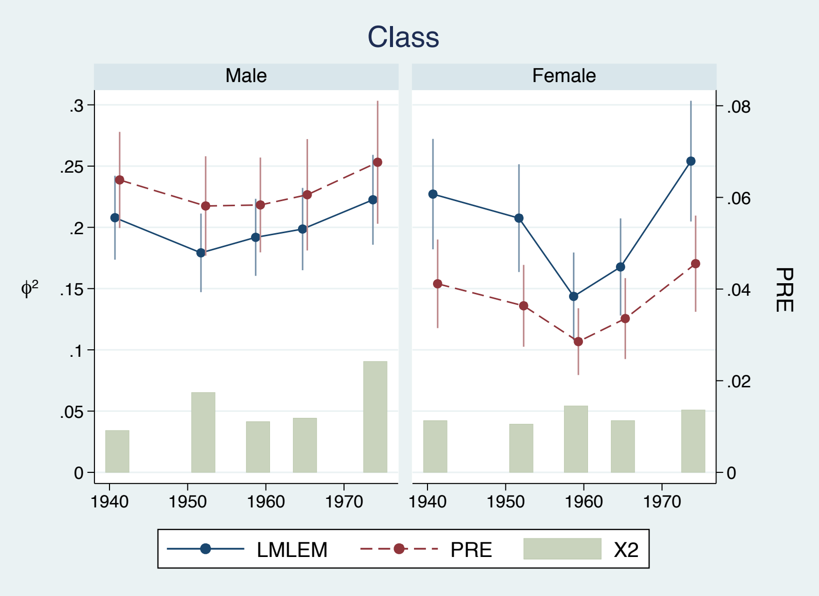
Results: Comparison of LMLEM and PRE

- As discussed above, the LMLEM assumes a common structure of associations between origin and destination categories that remains stable across cohorts.
- Differences between the LMLEM and PRE results may be due to a violation of this assumption.
- We thus included, as grey bars, a cohort-specific goodness-of-fit measure for the LMLEM:

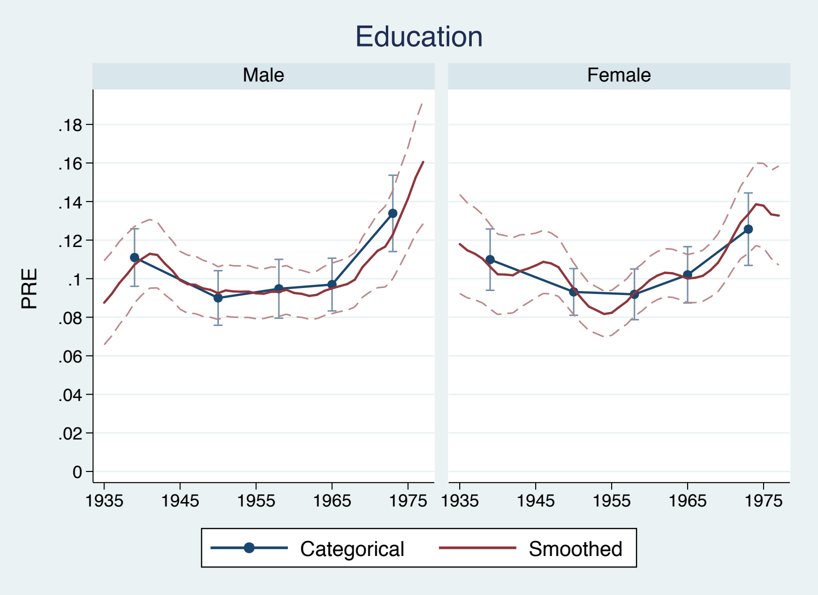
$$\bar{\chi}_k^2 = \frac{1}{N_k} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\left(F_{ijk} - \hat{F}_{ijk}\right)^2}{\hat{F}_{ijk}}$$

with F_{ijk} as the observed cell frequencies, $\hat{F}_{ijk}$ as the cell frequencies predicted by the model and N_k as the number of observations in cohort k. High values of $\bar{\chi}_k^2$ indicate bad fit (the scale of $\bar{\chi}_k^2$ is not relevant here and is omitted).

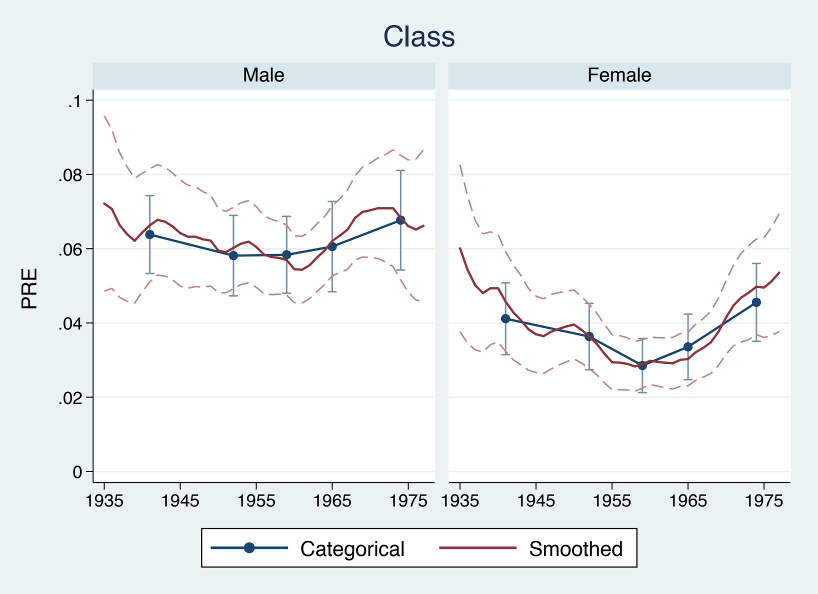
Results: Comparison of LMLEM and PRE



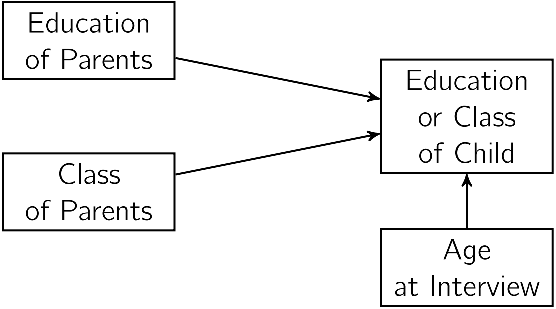
Results: Smoothed PRE



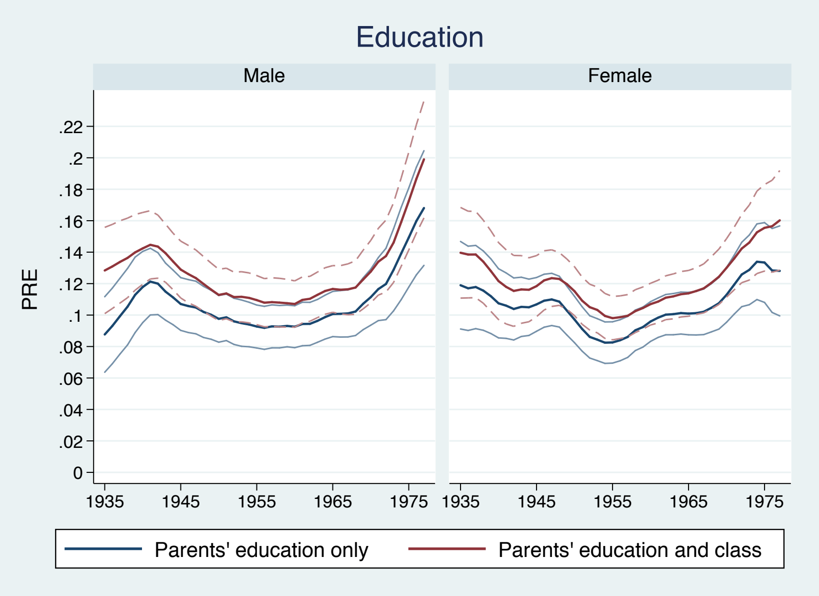
Results: Smoothed PRE



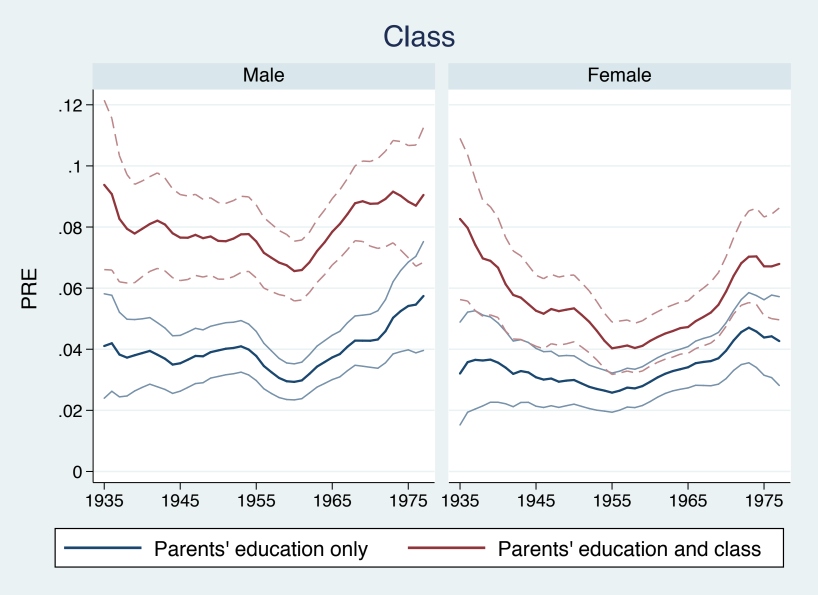
Results: PRE with multiple origin (and control) variables



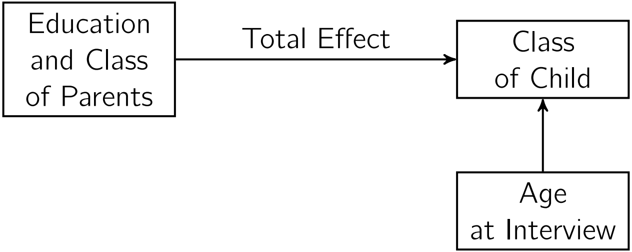
Results: PRE with multiple origin (and control) variables



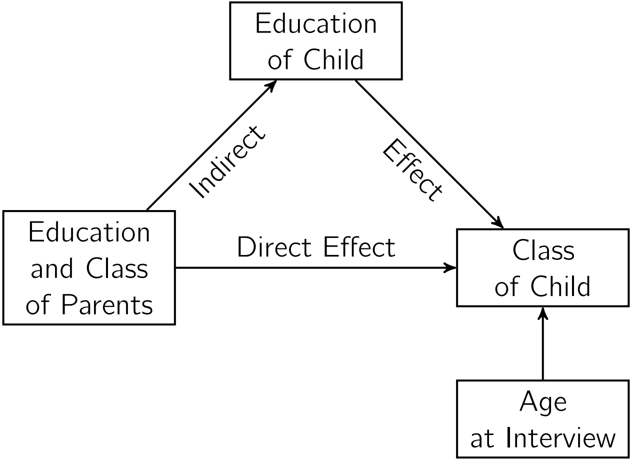
Results: PRE with multiple origin (and control) variables



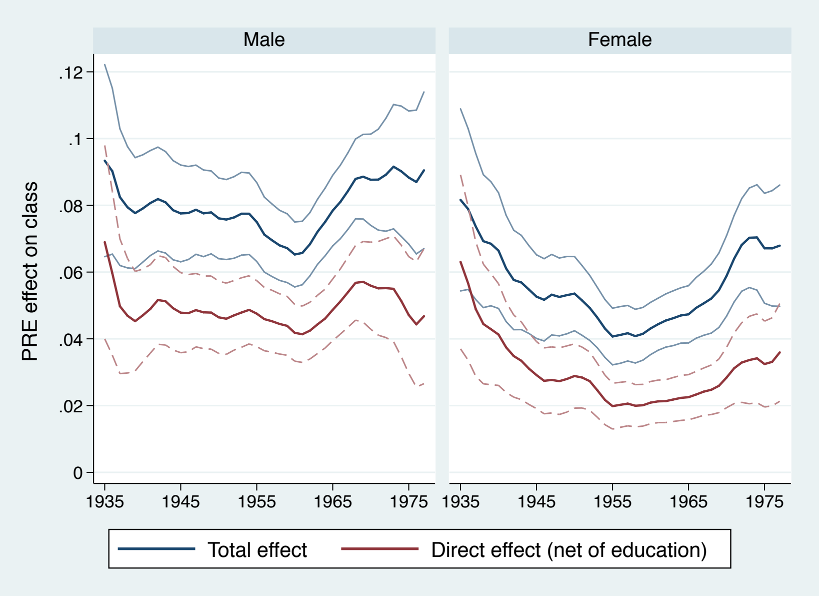
Results: Direct and indirect origin effects



Results: Direct and indirect origin effects



Results: Direct and indirect origin effects



Conclusions

- The PRE approach seems to be a viable and flexible model to analyze social mobility.
  - It produces results that are comparable to the classic LMLEM. Remaining deviations between PRE and LMLEM seem to be mainly driven by misfit of the LMLEM.
  - It can easily include multiple origin variables and control variables.
  - It has a clear interpretation (proportional reduction of prediction errors): How much does the knowledge of parents' characteristics improve the prediction of the child's status?

Conclusions

- Substantive conclusions
  - Our results indicate that social mobility increased among birth cohorts in the mid 1930s to about 1960, but then started to decrease again.
  - In general, this pattern can be observed for both men and women and both education and class. The pattern, however, is least pronounced for men's class.
  - For respondent's education the PRE approach leads to more pronounced results than LMLEM. This indicates that the structure of association changed over time for education.
  - Net of parents education, parents' class still has an effect on both respondent's education and class. As expected, the effect on class is stronger.
  - Parents characteristics have a direct effect on respondent's class, net of respondent's education.

References

- Erikson, Robert, John H. Goldhorpe. 1992. The constant flux. A study of class mobility in industrial societies. Oxford: Clarendon.
- Erikson, Robert, John H. Goldthorpe, Lucienne Portocarero (1983). Intergenerational Class Mobility and the Convergence Thesis: England, France and Sweden. The British Journal of Sociology 34(3): 303–343.
- Theil, Henri. 1970. On the estimation of relationships involving qualitative variables. American Journal of Sociology 76(1): 103–154.
- Xie, Yu (1992). The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables. American Sociological Review 57(3): 380–395.