UNIVERSITÄT
BERN

Faculty of Business, Economics and
Social Sciences

Department of Social Sciences

University of Bern Social Sciences Working Paper No. 5

A new methodological approach for studying intergenerational mobility with an application to Swiss data

Ben Jann and Simon Seiler

January 13, 2014

http://ideas.repec.org/p/bss/wpaper/5.html
http://econpapers.repec.org/paper/bsswpaper/5.htm

University of Bern
Department of Social Sciences
Fabrikstrasse 8
CH-3012 Bern

Tel. +41 (0)31 631 48 11
Fax +41 (0)31 631 48 17
info@sowi.unibe.ch
www.sowi.unibe.ch

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

Ben Jann and Simon Seiler

University of Bern, Institute of Sociology
Fabrikstrasse 8, CH-3012 Bern
ben.jann@soz.unibe.ch
simon.seiler@soz.unibe.ch

January 13, 2014

Abstract

Despite the widespread interest in the topic and a vast international literature, very little is known about the development of intergenerational mobility in Switzerland. Based on a new harmonized database for Switzerland (comprising various surveys such as different waves of the ISSP, EVS, and the ESS), we provide a systematic account of changes in the link between social origin and destination over time (covering birth cohorts from around 1935 to 1980). We analyze effects of parental education and class on own educational achievement and social class for both men and women, using a refined variant of the methodological approach proposed by Jann and Combet (2012). The approach is based on the concept of proportional reduction of error (PRE) and features a number of advantages over more traditional approaches. For example, it provides smooth estimates of changes in social mobility that have a clear interpretation and it can easily incorporate control variables and multiple dimensions of parental characteristics. To evaluate the validity of our approach, we employ the oft-used log-multiplicative layer effect (a.k.a Unidiff) model (Xie 1992, Erikson and Goldthorpe 1992) as a benchmark. Results indicate that our approach performs well and produces qualitatively similar findings as Xie's model. For both men and women, effects of social origin initially decreased, but then, towards the end of the observation period, increased again. This u-shaped pattern, which can be observed with respect to both education and class, appears to be more pronounced for women than for men.

1 Introduction

Equal opportunity is one of the main guiding principles in meritocratic societies. The basic idea is that the social position an individual can achieve should only depend on own effort and merit, not on *ascriptive* characteristics such as social origin or gender. Societies in which equal opportunity is granted are called "open." They are characterized by a high degree of *social mobility*. In the words of Hout (2004: 970): "Mobility is usually understood as 'equality of opportunity'—the outcomes may be unequal, but everyone, regardless of starting point, can have the same opportunity to get a good result."

To evaluate the openness of a society one can therefore analyze, for example, the degree to which the achieved social position of an individual depends on the social status of one's parents, as is done in a vast body of international literature. International research shows that in most countries sizable effects of social origin exist and persist over time, indicating a violation of the principle of equal opportunity. Despite the widespread interest in the topic, however, very little is known about the development of intergenerational mobility in Switzerland. In particular, it is unclear from the existing literature whether social mobility increased in Switzerland—as asserted by modernization theories (e.g. Lipset and Bendix 1959, Kerr 1962, Blau et al. 1967).

Based on a new harmonized database for Switzerland, we provide a systematic account of changes in the link between social origin and destination over time. The database comprises multiple surveys containing information on the social position of the respondents and their parents, such as different waves of the International Social Survey Programme (ISSP), the European Values Study (EVS), or the European Social Survey (ESS), the Swiss Household Panel (SHP), the 2011 wave of the Statistics on Income and Living Conditions (SILC), and a number of further single-wave cross-sectional surveys. In the database we harmonized variables on ed-

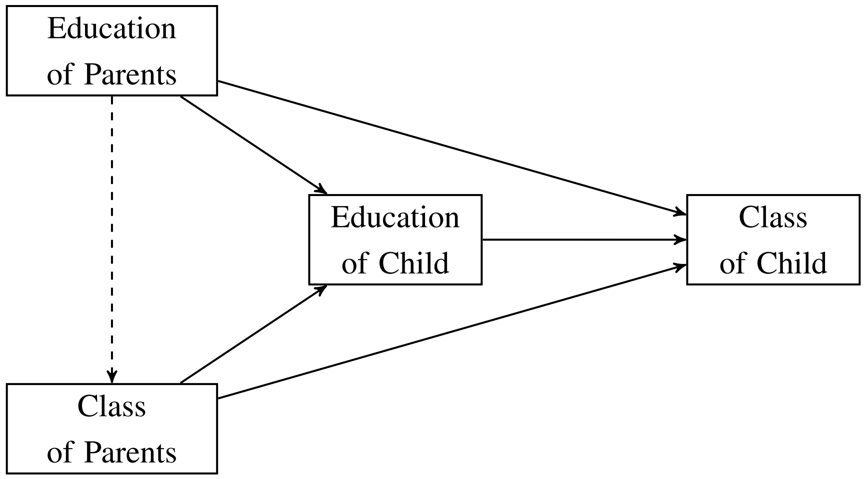

Figure 1: Effects of social origin

ucation and social class of respondents and their parents so that pooled analyses across surveys are possible, covering a wide range of birth cohorts over the 20th century. In particular, as illustrated in Figure 1, we analyze how *educational attainment* and *social class* of respondents depend on education and class of their parents and how the strength of these dependencies changes over time.

As a methodology to analyze the development of social mobility based on categorical variables such as class or educational attainment, Erikson and Goldthorpe (1992: 91-92) 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 (some recent examples are Breen and Jonsson 2005, 2007, Pfeffer 2008, Breen et al. 2009, 2010, Lippényi et al. 2013). The popularity of the model is due to some important advantages over alternative approaches. First, it provides a parsimonious and intuitive way to describe differences in effects of social origin across time (or geography). Second, it allows testing against a null model with time-constant origin effects within the same modeling framework. Third, despite its parsimony, it provides a good fit to empirical data in many applications.

The LMLE model, however, also has a number of limitations. First, it assumes a common baseline pattern of associations that remains constant over time and it will yield misleading results if this assumption is violated. Although it is possible to test the assumption by comparison to a saturated log-linear model, it is unclear how to extend the model to incorporate possible deviations in mobility patterns while, at the same time, preserving ease of interpretation. Second, the model builds upon simple crosstabs, making it difficult to extend the model to more complex settings and, for example, analyze the simultaneous effects of multiple origin variables or incorporate control variables. Third, although the pattern of social origin effects can be interpreted in a meaningful way within a single estimation, comparisons across models are difficult because the absolute level of origin effects remains obscure.

In our study, we therefore contrast the LMLEM with a refined variant of an alternative methodological approach proposed by Jann and Combet (2012). The approach is based on the concept of proportional reduction of error (PRE) and overcomes some of the limitations of the LMLEM. For example, it provides smooth estimates of changes in social mobility that have a clear interpretation and it can easily incorporate control variables and multiple dimensions of parental characteristics. The basic idea behind our approach is that the degree to which information about parents' social status can help predict their children's status reflects the strength of effects of social origin. If the predictions based on parents' characteristics are precise, then origin effects are strong and social mobility is low; if, however, the predictions are imprecise, then social mobility between the generations is high.

Our results indicate that the PRE approach performs well and produces qualitatively similar findings as the LMLE model in situations where similar results can be expected. For both men and women, effects of social origin initially decrease, but then, towards the end of the observation period, increase again. This u-shaped pattern, which can be observed with respect to both education and class, appears to be more pronounced for women than for men. We conclude that equal opportunity in Switzerland increased among the earlier birth cohorts, but then declined again in the second half of the 20th century, particularly among women.[1]

The remainder of the paper is organized as follows. We describe the methodological approaches in the next section and then give a brief account of the data in Section 3. The results section will then provide a comparison between methods and present some extended analyses illustrating the flexibility of our PRE approach. We will conclude the paper with a brief summary of our findings.

2 Methods

2.1 Log-multiplicative layer effect model

Starting point of the log-multiplicative layer effect model is a simple two-way table of origin and destination, called a "mobility table." An example of such a table is given in Table 1. More generally, a mobility table can be conceptualized as shown in Figure 2, where $i = 1, \dots, I$ is the row index and $j = 1, \dots, J$ is the column index. F_{ij} is the observed frequency of the cell defined

[1]Note, however, that social class is measured by individual characteristics in our study. Result may look different if social class is determined based on households characteristics. This seems particularly relevant for women, whose labor market attachment changed significantly over the observation period.

Table 1: Mobility table of respondent's education by parent's education

	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 Section 3. Selection: males, birth cohorts 1969-82

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

Figure 2: Two-dimensional mobility table

by row i and column j. $F_{i.}$ and $F_{.j}$ denote the row and column totals, respectively. A log-linear model expresses the cell frequencies F_{ij} in such a table using a multiplicative function. The saturated model, that is, a model that exactly reproduces the observed frequencies, is given as:[2]

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

The model is called "log-linear" because taking the logarithm leads to a linear expression:

$$\log(F_{ij}) = \log(\tau_{..}) + \log(\tau_{i.}) + \log(\tau_{.j}) + \log(\tau_{ij}) \tag{2}$$

[2] We deviate from standard notation in that we use positioned indices instead of additional row and column superscripts.

In case of independence between rows and columns, all τ_{ij} are equal to one. This is the situation we would expect if origin (parent's education) has no effect on destination (child's education), corresponding to the ideal of a fully mobile society. As soon as, however, some τ_{ij} deviate from one, mobility is constrained. To test whether a society is fully open or not, one can therefore perform a likelihood-ratio test of the saturated model against a restricted model in which all τ_{ij} are constrained to 1.

To find out whether mobility changed over time, we can look at a series of mobility tables across birth cohorts, as depicted in Figure 3. This is, in fact, a three-dimensional mobility table with an additional dimension $k = 1, \ldots, K$ for cohorts. The saturated model for such a three-dimensional table is given as:

$$F_{ijk} = \tau_{...} \cdot \tau_{i..} \cdot \tau_{.j.} \cdot \tau_{..k} \cdot \tau_{i.k} \cdot \tau_{.jk} \cdot \tau_{ij.} \cdot \tau_{ijk}, \quad i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, K \tag{3}$$

In this model full mobility is granted if all $\tau_{ij.}$ and all τ_{ijk} are equal to one. Furthermore, the pattern of dependency between origin and destination remains constant over cohorts if all τ_{ijk} are equal to one. To evaluate how origin effects change over cohorts one could therefore inspect the values of the τ_{ijk} over k, but this would be very tedious and often inconclusive because for each k there are $I \times J$ parameters that would have to be taken into account. To ease interpretation, Xie (1992) and Erikson and Goldthorpe (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 (LMLEM) and is given as:

$$F_{ijk} = \tau_{...} \cdot \tau_{i..} \cdot \tau_{.j.} \cdot \tau_{..k} \cdot \tau_{i.k} \cdot \tau_{.jk} \cdot \exp(\psi_{ij.} \cdot \phi_{..k}), \quad i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, K \tag{4}$$

In this model, the $\psi_{ij.}$ capture the overall pattern of dependencies between origin and destination, and the $\phi_{..k}$ are cohort-specific scaling factors. That is, the higher $\phi_{..k}$, the more pronounced

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

$$\vdots$$

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

$$\vdots$$

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

Figure 3: Three-dimensional mobility table

is the pattern of dependencies and, hence, the stronger is the strength of the associations 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), the constraint we use in our implementation of the model is that the sum over $\phi^2_{..k}$, $k = 1, \ldots, K$ must be equal to 1. From this constraint it is immediately clear, however, that the overall level of the $\phi_{..k}$ in a model is primarily determined by the size of K (the number of cohorts) and does not reflect the strength of the relation between origin and destination in an absolute sense.

2.2 PRE approach

The general idea behind the PRE approach proposed by Jann and Combet (2012) is that strong effects of social origin go hand in hand with high predictive power of the status of the parents for the status of their children. That is, the better the position of children can be predicted based on parents characteristics, the stronger the influence of social origin is and the lower social mobility is. To quantify the predictive power of parents' characteristics for the status of children, one can use the Proportional Reduction of Error (PRE; see e.g. Costner 1965).

Formally, PRE is defined as

$$PRE = \frac{E_0 - E_1}{E_0} = 1 - \frac{E_1}{E_0} \tag{5}$$

where E_0 is the sum of prediction errors under limited information (i.e. excluding information on parents) and E_1 is the sum of prediction errors under full information (i.e. including information on parents). Different error rules can be applied, yielding different PRE measures. Because our dependent variables, education and class, are categorical, however, an entropybased definition (see Theil 1970) appears appropriate. In particular, we use

$$E_m = -\sum_{i=1}^{N} w_i \ln\left(\hat{p}_m(Y = y_i)\right), \quad m = 0, 1 \tag{6}$$

where w_i is the respondent's survey weight and $\hat{p}_m(Y = y_i)$ is the predicted probability of the dependent variable taking on observed value y_i under model m. To estimate these probabilities, we use multinomial logistic regression. That is, the probabilities under restricted information are modeled as

$$p_0(Y = j) = \frac{\exp(\beta_j Z_i)}{\sum_{\ell=1}^{J} \exp(\beta_\ell Z_i)} \tag{7}$$

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

$$p_1(Y = j) = \frac{\exp(\beta_j Z_i + \gamma_j X_j)}{\sum_{\ell=1}^{J} \exp(\beta_\ell Z_i + \gamma_\ell X_j)} \tag{8}$$

where X_i is a vector of parents' characteristics. For each birth cohort, separate models are fit and a separate PRE value is computed. The approach is thus fully flexible and does not assume a single association pattern that is stable across cohorts.[3]

If the number of observations per birth year is small, then multiple birth years can be collapsed into larger cohorts to reduce the variability of estimates, as is common practice for the LMLE model. Such an approach will, however, only provide a crude picture of the changes in social mobility over time. To get a more detailed picture in form of a smoothed curve we propose to compute a PRE value for each birth year including data from surrounding years in

[3] If the dependent variable is continuous, say income or occupational prestige, one could define prediction errors as squared deviations between observed values and predictions from linear regression. This would lead to the R-squared (or the increment in R-squared if the restricted model contains control variables) as the value for the PRE measure. The categorical error measure we use here corresponds to McFadden's pseudo R-squared.

the estimation by means of kernel weights. That is, computations are repeated for each birth year with weights defined as

$$w_i(t^*) = w_i \cdot \frac{1}{h} K\left(\frac{t^* - t_i}{h}\right) \tag{9}$$

where t^* is the target birth year, t_i is observations i's birth year, and $K()$ is a kernel function with bandwidth h. In the applications below we use the Epanechnikov kernel defined as

$$K(z) = \begin{cases} \frac{3}{4}(1 - z^2) & \text{if } |z| < 1 \\ 0 & \text{else} \end{cases} \tag{10}$$

and a bandwidth of $h = 5$. This implies that computations are based on a symmetric data window of plus/minus 4 years where, however, observations from the target birth year receive the highest weight (observations whose birth year is 5 or more years away from the target birth year receive weight 0). To obtain confidence intervals for our estimates, we employ the bootstrap method (Davison and Hinkley 1997).[4]

3 Data

The data required for analysis of social mobility must contain relevant status variables (education, class) for the respondents as well as for their parents. Unfortunately, most Swiss large-scale surveys, such as the official surveys by the Swiss Federal Statistical Office, do not contain information on parents. Nonetheless, we were able to identify a number of surveys that can be used for these types of analyses. The results below are based on a selection of these surveys, as

[4]Stratified by survey; we use normal-approximation confidence intervals based on 500 replications.

Table 2: Surveys included in our study

Survey	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.

listed in Table 2.[5] Figure 4 provides a histogram of the number of available observations from the different surveys by birth years of respondents (age range at time of interview restricted to 30 to 69).[6] Covered are birth cohorts from 1922 to 1982, although pre 1940 and post 1975 the number of observations per birth cohort is low (< 500). The 20-year period starting in 1951 is particularly well-covered using the SILC data.

[5]A few additional surveys are available (especially some older ones) and will be incorporated in a future version of the paper.

[6]All Stata graphs in this paper have been produced by user command "coefplot" (Jann 2013).

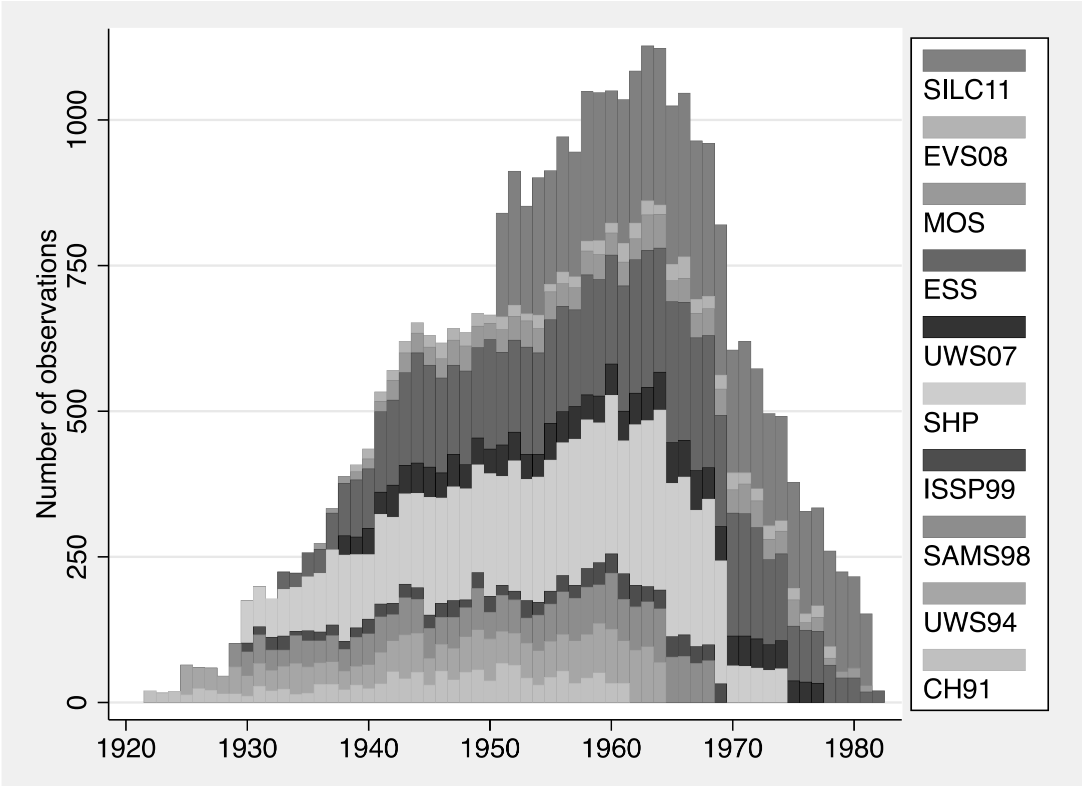

Figure 4: Number of observations by birth year

We harmonized the variables across surveys to build a common database that can be analyzed in terms of birth cohorts. The age range of respondents was restricted to 30 to 69. Our key variables are educational attainment and social class of respondents and their parents. For education we use a five level classification that is common in Swiss statistics and reflects the Swiss educational system with its strong vocational track. A more detailed classification is not possible due to the heterogeneity of the measurement instruments in the different surveys. Table 3 provides an overview of the educational levels. For class we use a slightly simplified EGP scheme based on Erikson et al. (1983: 307). Table 4 provides an overview of the class definitions.

Table 3: 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

Table 4: Social class scheme (EGP)

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

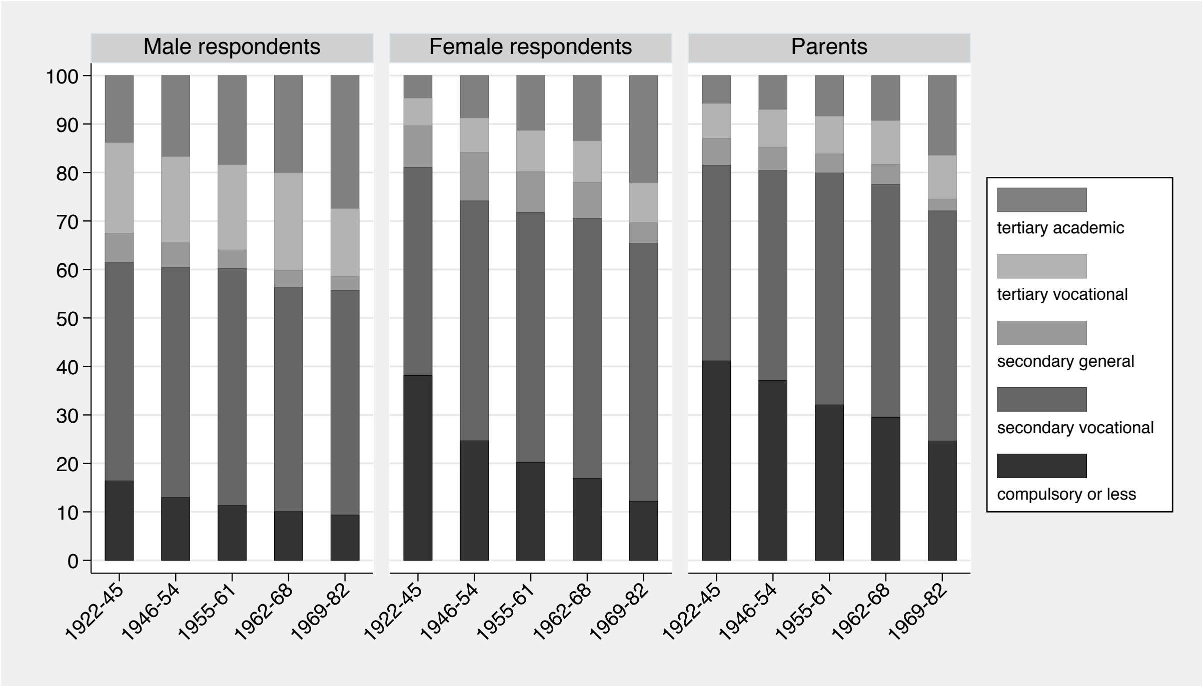

Figure 5: Education of male and female respondents and their parents by birth cohorts

Figures 5 and 6 display the distributions of education and class for male and female respondents and their parents by rough birth cohorts.[7] Survey weights have been employed to compute these statistics as well as all following results. The weights were standardized such that the sum of weights within a survey equals the number of observations used from that survey in a specific analysis.

The educational expansion over time is clearly visible (Figure 5). The proportion of respondents with compulsory education or less decreases over birth cohorts while the proportion with tertiary degrees expands for both men and women. A similar pattern can also be observed for the education of parents. Tertiarization of the labor market with expanding service classes and

[7]Education and class of parents was determined from the better positioned parent if information from both the father and the mother was available.

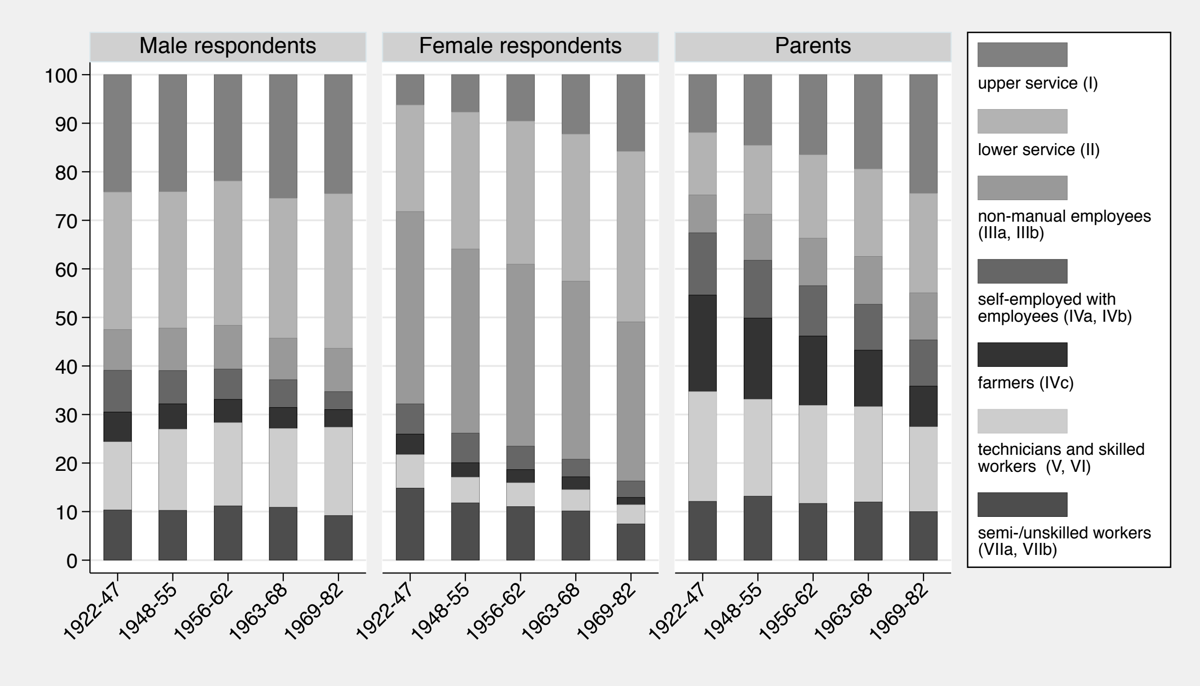

Figure 6: Social class of male and female respondents and their parents by birth cohorts

a decrease in semi- and unskilled workers is evident for female respondents and parents, but much less so for male respondents (Figure 6).

4 Results

4.1 Comparison of LMLEM and PRE

Figure 7 displays the results from the log-multiplicative layer effect model (solid lines, left scale) and the categorical implementation of our PRE approach (dashed lines, right scale) for education and class by gender. For the LMLEM the ϕ^2 parameters are plotted.[8] Spikes indicate point-wise 95% confidence intervals. Dots are placed at mean birth years within cohorts, with

[8]The ϕ^2 parameters always group around 0.2 in the current application because—as discussed above—their sum across the five cohorts is restricted to 1.

a slight horizontal offset to prevent the confidence spikes to be printed upon each other. The upper plots show the effects of parents' education on respondent's education; the lower plots show the effects of parents' EGP class on respondent's EGP class.

Looking at the results for education of males (upper left) we see a close fit between LMLEM and PRE. Only for the last cohort do these curves deviate. Notably, the vertical positioning of the two curves is arbitrary since the two statistics are on separate scales; it is the pattern across cohorts that should be compared, not the overall level. What we can say is that both methods yield a quite similar pattern across the first four cohorts but that for PRE there is a steeper increase in the social origin effect from the fourth to the fifth cohort. As a consequence, the pattern of origin effects exhibits more overall curvature for PRE then for LMLEM.

Also for education of females (upper right), the pattern is more curved for PRE than for LMLEM, the biggest differences in slopes occurring from the first to second and the fourth to fifth cohort. Comparing PRE results for males and females reveals that the origin effects with respect to education are roughly the same for both sexes with PRE values around 10%. Furthermore, the development of origin effects is quite similar between the sexes: a decrease (i.e. an increase in social mobility) among the earlier cohorts and an increase (i.e. a decrease in social mobility) among the latter.

For class, the results from LMLEM and PRE are very similar (lower plots; recall that only the shape of the patterns can be compared, not the overall levels). For both males and females the two methods reveal a u-shaped pattern of origin effects that is slightly more pronounced for females than for males. What is evident from the PRE results, however, is that origin effects on class are considerably stronger for males than for females. For males the PRE values lie around 6.5%, for females around 4%.

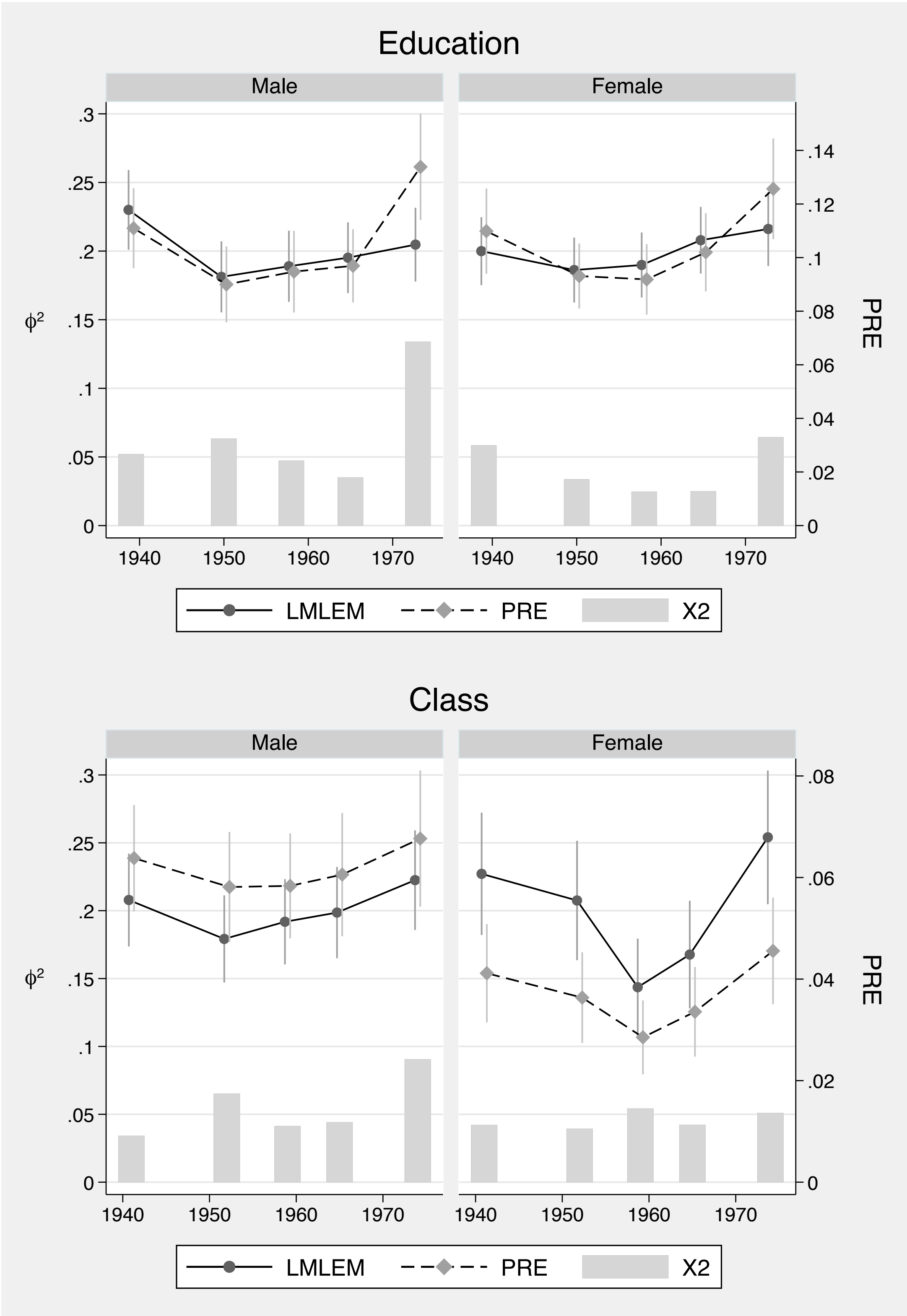

Figure 7: Comparison of LMLEM and PRE results for education and class

As discussed above, the LMLEM assumes a common structure of associations between origin and destination categories that remains stable across cohorts, an assumption that is waived by the PRE approach. Differences between the LMLEM and PRE results may thus be due to a violation of this assumption. To evaluate this point we included, as grey bars, a cohort-specific goodness-of-fit measure for the LMLEM in the plots. The fit measure we use is based on the chi-squared statistic and is defined as follows:

$$\bar{\chi}_k^2 = \frac{1}{N_k} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\left(F_{ijk} - \hat{F}_{ijk}\right)^2}{\hat{F}_{ijk}} \tag{11}$$

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 not included in the plots; what matters are the relative differences of misfit between cohorts.

Comparing the pattern of fit statistics with the deviations between LMLEM and PRE for education (upper plots in Figure 7) reveals striking similarities. For males, it is the last cohort where the LMLEM has a particularly bad fit; for females, it is the first and last cohorts. This is exactly what we would expect from the differences between the LMLEM and PRE results for education discussed above. For class (lower plots), the pattern of fit statistics is less conclusive, which does not come as a surprise as findings from LMLEM and PRE are quite similar.

4.2 Smoothed PRE results

As discussed in the methods section, the PRE results can be refined by smoothing across birth years based on kernel weights instead of collapsing the data into broad birth cohorts. Figure 8 contrasts the categorical PRE estimates from Figure 7 with smoothed PRE curves

Figure 8: Smoothed PRE results

Overall, the smoothed curves follow the same pattern as the categorical results although, naturally, the smoothed curves tend to fluctuate more. Most interesting are probably the results at the boundaries of the observation window where the smoothed curves are more informative than the categorical results. For example, for male's education, it appears that there was an initial increase in origin effects that is not captured by the collapsed PRE.

4.3 Adding control variables

Up to now, we analyzed purely bivariate associations between parents' and respondents' characteristics. Within the framework of the PRE approach, however, it is easy estimate partial associations between origin and destination under control of the effects of additional covariates by including these variables as control variables in the Z_i vector in the multinomial logit models (7) and (8). For example, the data we use stem from different surveys and it might be a good idea to add survey dummies to our models so that social origin effects are identified only based on within-information from the surveys. Furthermore, age at time of interview may have a distorting effect because the age distribution changes across birth cohorts.

Figure 9 compares the bivariate results for education (solid lines) with the partial origin effects on education after controlling for survey dummies and a linear age-at-time-of-interview effect (dashed lines). The upper plots in the graph display the results from the categorical PRE approach, the lower plots contain the smoothed results. Controlling for survey and age slightly reduces origin effects, but the differences are small and the overall patterns remain the same. We therefore conclude that effects of surveys and age do not substantially bias the findings.

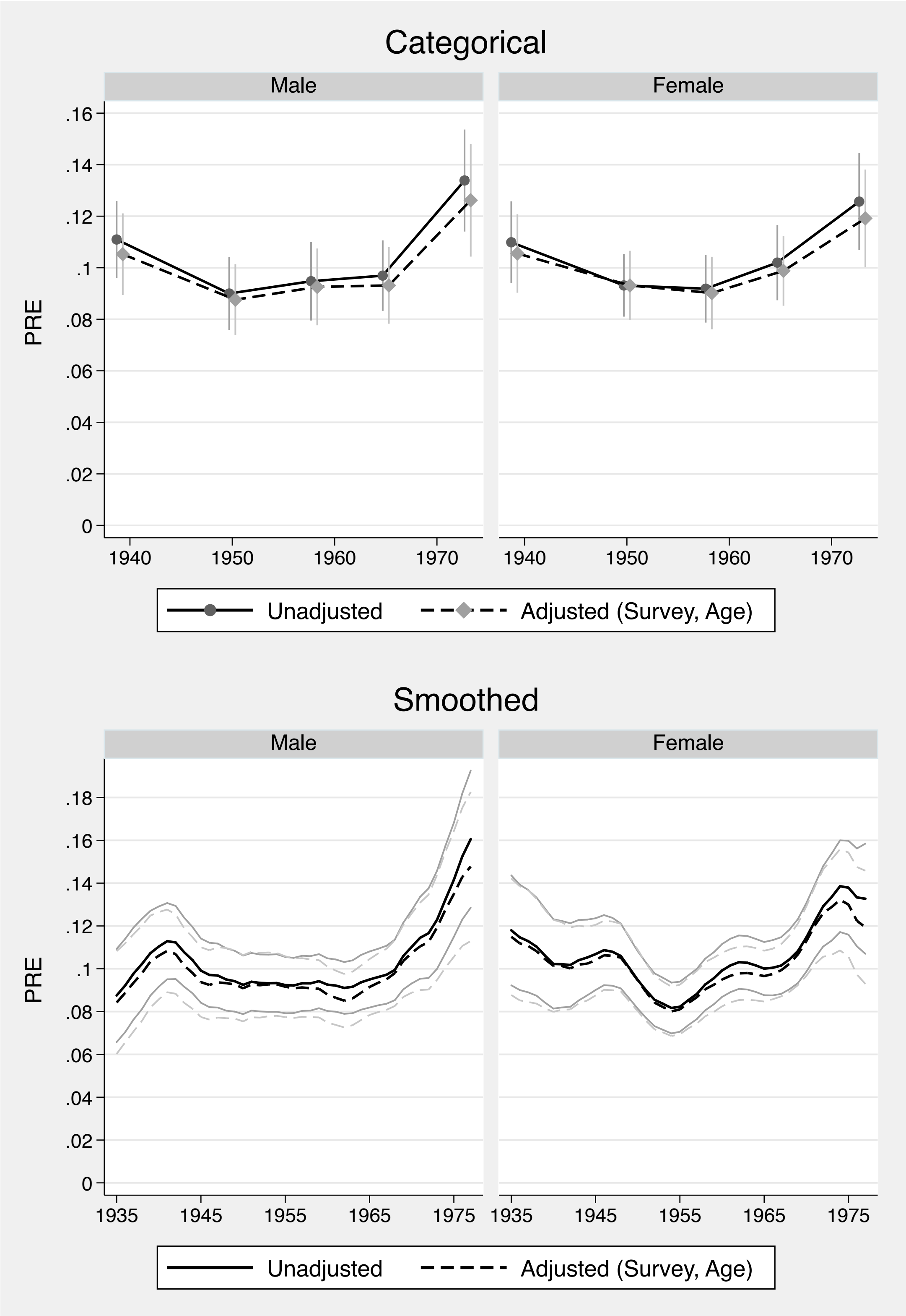

Figure 9: PRE results for education including control variables

4.4 Simultaneous effects of multiple origin variables

To estimate the joint effect of several origin characteristics (e.g. education and class of parents, characteristics of fathers as well as mothers) on a destination outcome, multiple variables can be included in the X_i vector in the multinomial logit model (8). Figure 10 displays the smoothed PRE effects of parents' education (solid lines) and parents' education together with class (dashed lines) on respondents' education (upper plots) and respondents' class (lower plots) (in all models, age-at-time-of-interview is included as a control variable).

Unsurprisingly, we see that parents' class does not add much beyond parents' education in explaining respondents' education, but has a strong independent effect on respondents' class. For respondents' education, the solid-line and dashed-line curves are close together and roughly parallel; for respondents' class, however, there are larger differences in level and shape of the curves. In particular for women, adding parents' class to the model yields a more pronounced u-shaped pattern with decreasing origin effects in the beginning and increasing effects among the more recent cohorts.

4.5 Direct and indirect origin effects on class

Based on the sequential nature of the attainment of educational degrees and the integration into the labor market one would expect that respondent's class is largely a function of respondent's education (see Figure 1). That is, social origin effects on class operate, at least to some degree, as indirect effects through educational achievement. It may thus be interesting to decompose the total effect of social origin on class into an indirect effect through education and a remaining direct effect, as depicted in Figure 11. A positive direct effect indicates that social origin still matters for class position, even if the achieved education has been taken into account.

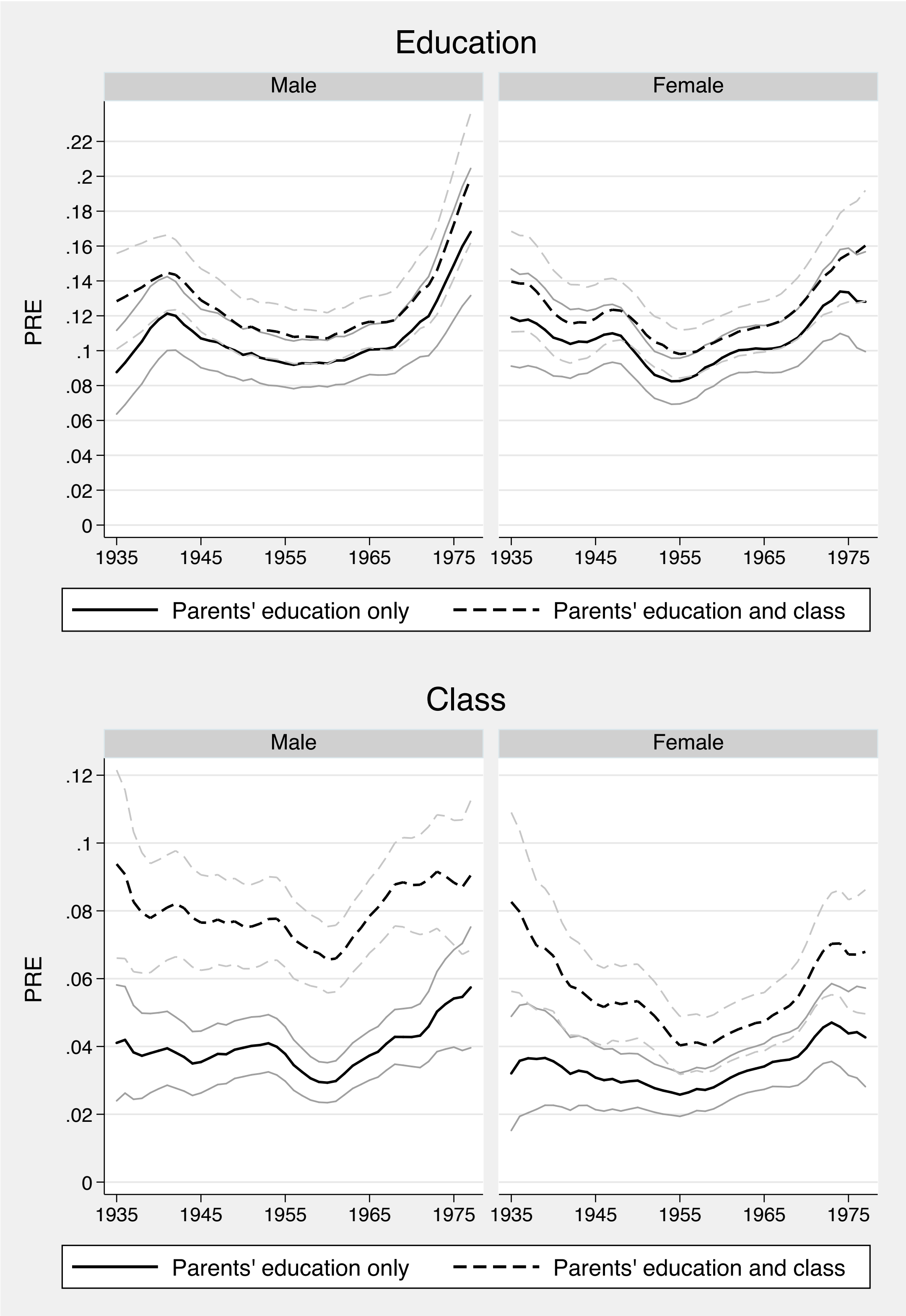

Figure 10: PRE results using multiple origin variables

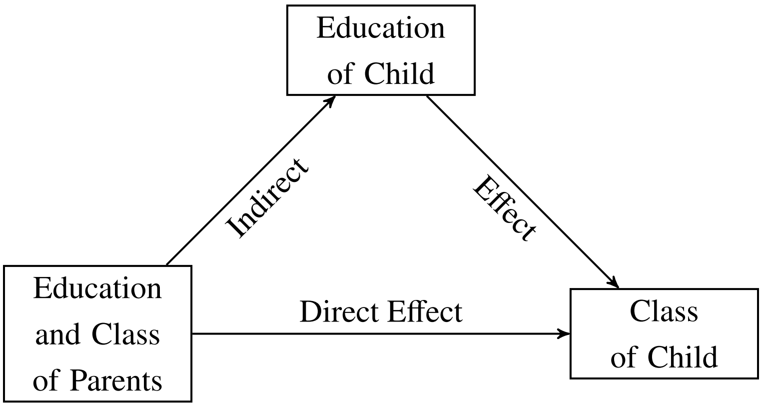

Figure 11: Direct and indirect effects of parents' characteristics on class

Figure 12 displays the decomposition of effects of social origin (parents' education and class) on respondent's class for our Swiss data (again under control of age-at-time-of-interview). The solid lines display the total effects, the dashed lines display the direct effects net of respondent's education. The indirect effects result from the difference between the solid-line and dashed-line curves. As can be seen, a significant positive direct effect exists for all birth cohorts for both men and women. That is, social origin matters beyond educational attainment. Furthermore, the shape of the development of the total effect over cohorts is mostly dominated by the shape of the direct effect. Nonetheless, there is some weak evidence that the indirect effect gained in importance, since the difference between the curves widened somewhat over time for both males and females.

5 Conclusion

In this paper we proposed a new methodological approach for studying social mobility that is based on the statistical concept of Proportional Reduction of Error (PRE). We applied the method to a harmonized data set of Swiss population surveys, covering a wide range of

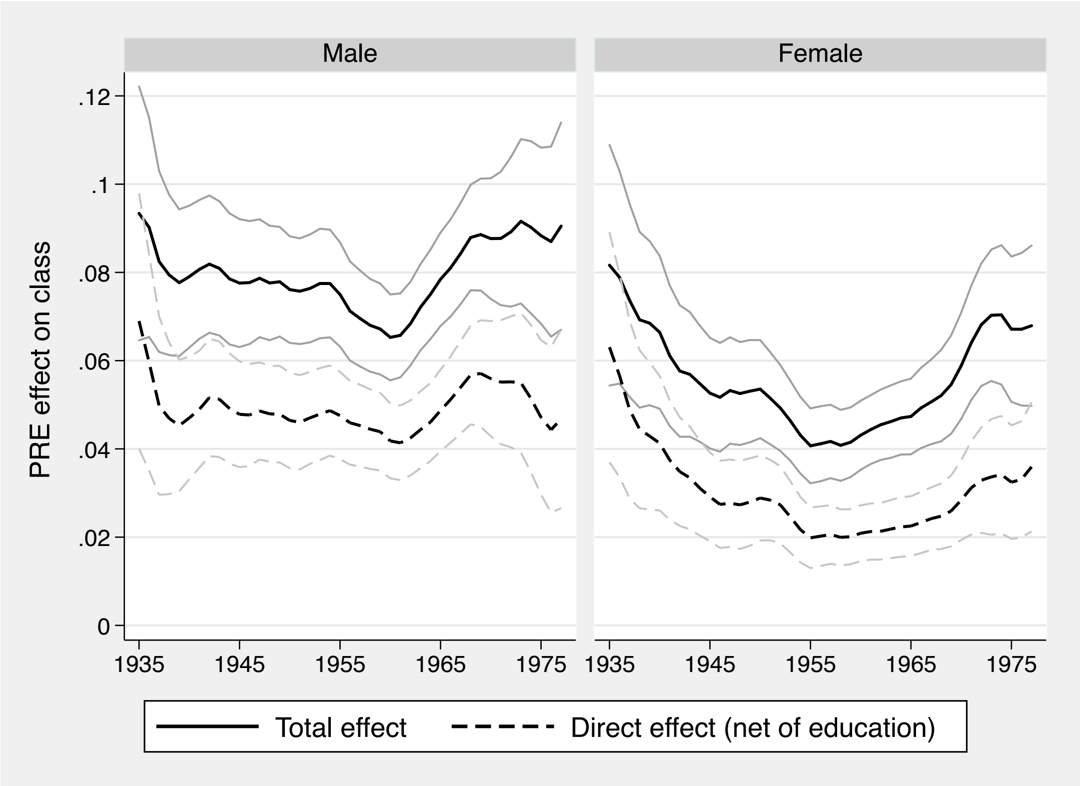

Figure 12: Decomposition of social origin effects on class

birth cohorts across the 20th century, and contrasted the method's results to the classic log-multiplicative layer effect model (LMLEM).

Direct comparison of results from the PRE approach to results from the LMLEM reveals that the two methods yield qualitatively similar findings. However, the PRE approach is more flexible in that is does not assume a stable basic association pattern across birth cohorts, and there is evidence that the observed differences between the PRE and LMLEM results are, at least in part, due to misfit of the LMLEM. Hence, the PRE approach appears to be a viable and flexible method for the analysis of intergenerational mobility. The method closely adapts to the data without restrictive identifying assumptions, can easily be extended to produce smoothed estimates across birth years and incorporate control variables or multiple origin dimensions, and, above all, provides estimates that have a clear substantive interpretation and whose values can be compared across analyses in an absolute sense.

From a substantive viewpoint our analyses reveal that in Switzerland social mobility increased among birth cohorts the mid 1930s to about 1960, but then decreased. This pattern can be observed for both men and women and for both destination outcomes, education and class. The pattern, however, is less pronounced for men's class. Further results are that survey and age-at-time-of-interview effects do not substantially bias results, that after controlling for parents' education, parents' class has only a weak effect on respondents' education but strongly influences respondents' class, and that social origin variables exert considerable direct influence on class net of respondents education, although indirect effects through education seem to have gained in significance over time.

References

Blau, Peter Michael, Otis Dudley Duncan, and Andrea Tyree. 1967. *The American Occupational Structure*. New York: Wiley.

Breen, Richard and Jan O. Jonsson. 2005. "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility." *Annual Review of Sociology* 31:223–243.

Breen, Richard and Jan O. Jonsson. 2007. "Explaining Change in Social Fluidity: Educational Equalization and Educational Expansion in Twentieth–Century Sweden." *American Journal of Sociology* 112:1775–1810.

Breen, Richard, Ruud Luijkx, Walter Müller, and Reinhard Pollak. 2009. "Nonpersistent Inequality in Educational Attainment: Evidence from Eight European Countries." *American Journal of Sociology* 114:1475–1521.

Breen, Richard, Ruud Luijkx, Walter Müller, and Reinhard Pollak. 2010. "Long-term Trends in Educational Inequality in Europe: Class Inequalities and Gender Differences." *European Sociological Review* 26:31–48.

Costner, Herbert L. 1965. "Criteria for Measures of Association." *American Sociological Review* 30:341–353.

Davison, Anthony C. and David V. Hinkley. 1997. *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press.

Erikson, Robert and John Harry Goldthorpe. 1992. *The Constant Clux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.

Erikson, Robert, John H. Goldthorpe, and Lucienne Portocarero. 1983. "Intergenerational Class Mobility and the Convergence Thesis: England, France and Sweden." *The British Journal of Sociology* 34:303–343.

Hout, Michael. 2004. "How Inequality May Affect Intergenerational Mobility." In *Social inequality*, edited by Kathryn M. Neckerman, pp. 969–987. New York: Russell Sage.

Jann, Ben. 2013. "Plotting Regression Coefficients and Other Estimates in Stata." University of Bern Social Sciences Working Papers No. 1 (http://ideas.repec.org/p/bss/wpaper/1.html).

Jann, Ben and Benita Combet. 2012. "Zur Entwicklung der intergenerationalen Mobilität in der Schweiz." *Swiss Journal of Sociology* 38:177–199.

Kerr, Clark. 1962. *Industrialism and Industrial Man: The Problems of Labor and Management in Economic Growth*. London: Heinemann.

Lippényi, Zoltán, Ineke Maas, and Marco H. D. van Leeuwen. 2013. "Intergenerational class mobility in Hungary between 1865 and 1950: Testing models of change in social openness." *Research in Social Stratification and Mobility* 33:40–55.

Lipset, Seymour Martin and Reinhard Bendix. 1959. *Social Mobility in Industrial Society*. London: Heinemann.

Pfeffer, Fabian T. 2008. "Persistent Inequality in Educational Attainment and its Institutional Context." *European Sociological Review* 24:543–565.

Theil, Henri. 1970. "On the Estimation of Relationships Involving Qualitative Variables." *American Journal of Sociology* 76:103–154.

Xie, Yu. 1992. "The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables." *American Sociological Review* 57:380–395.