Trends in Intergenerational Class Mobility and Status Attainment in Argentina

A conditional logistic regression model estimated without and with error-correction

Harry BG Ganzeboom

(with Raul Jorrat)

Buenos Aires, Gino Germani Institute, December 10, 2024

Acknowledgements

- This presentation is in preparation for a paper to be written with Raul Jorrat. However, the current presentation is entirely the responsibility of Harry Ganzeboom.
- Earlier and future presentations:
  - SILC seminar VU University Amsterdam, January 8, 2019
  - RC28 Spring Meeting, Frankfurt DE, March 22, 2019
  - SILC Seminar, VU University Amsterdam, March 12, 2024
  - Gino Germani Institute, Buenos Aires, July 18, 2024
  - Cambridge Stratification Seminar, Stirling UK, August 28, 2024
  - European Consortium for Sociological Research, Barcelona ES, September 13, 2024
  - SILC Seminar, VU University Amsterdam, October 8, 2024

Development of the analysis in earlier presentations

- Frankfurt, SILC, 2019:
  - OLS model for 16 datasets (1961-2016)
  - Gave up on correcting for survey effect using SEM measurement models
  - Survey corrections make a lot of difference for trend estimates
- Dalle et al. (2022): Cohort comparison on current occupation
- Jorrat et al (2024): Period comparison on current occupation
- SILC, BA, Stirling, Barcelona, 2024
  - Only first jobs data (8 from 18 surveys), still 1961-2021 time window
  - EGP11:
  - HG model
  - CMLR model

Issues in this presentation

- ISEI or classes?
- EGP or ISCOcat?
- **Cohort vs Year?**
- **How survey effect corrections change the results**
- **Observed, structural and relative mobility**
- **Include first jobs and current/last jobs?**

Summary & Research Question

- We examine relative mobility (social fluidity) in Argentine, as displayed by 8 surveys, 1961-2021, with first job information. N=18132.
- Historical analysis by cohort comparison.
- Relative mobility is modelled with the Hauser-Goodman [HG] scaled uniform association model, that distinguishes between on-diagonal association (IMM) and off-diagonal association (U): stayers vs movers.
- The HG model is estimated in a Conditional Multinomial Logistic Regression [CMLR] model, which can extend bivariate analysis into multivariate analysis.
- Most importantly, CMLR makes it possible to control survey heterogeneity.
- **RQ: how did relative mobility change between cohorts 1915 – 1993?**

Table xx: Stepwise estimated CMLR models, entry into first job, 16 cohorts, 1916-1992. Eight surveys 1961-2021, N-18.131.

Model specification

D1	Cohort + IMMk + U + U_COHx
D2	D1 + COHORT (marginals)
D3	D2 + IMM_COHx
D4	D3 + FEMALE + FEMALE interactions on U and IMM
D5	D4 + AMBA + AMBA interactions on U and IMM
D6	D5 + survey main effects
D7	D6 + survey interactions U
D8	D7 + survey interactions IMM
D9	D8 = education main effect
D10	D9 + education interaction on U and IMM

Table xx: Stepwise estimated CMLR models, entry into first job, 16 cohorts, 1916-1992. Eight surveys 1961-2021, N=18.131.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
DF	23	173	174	177	180	187	194	201	202	204
U	0.472	0.339	0.356	0.357	0.385	0.381	0.405	0.399	0.294	0.315
U_COHx	-0.258	-0.051	-0.083	-0.098	-0.118	-0.104	-0.206	-0.151	-0.096	-0.135
IMM_COHx			0.215	0.294	0.191	0.201	0.202	-0.180	-0.164	-0.07
Female_ddd				-0.394	-0.394	-0.422	-0.426	-0.426	-0.487	-0.411
U_female				0.038	0.039	0.029	0.015	0.016	-0.019	-0.022
IMM_female				-0.387	-0.393	-0.389	-0.386	-0.399	-0.409	-0.411
AMBA_ddd					0.049	0.059	0.067	0.055	0.038	0.038
U_amba					-0.029	-0.022	0.015	-0.012	-0.038	-0.036
IMM_amba					-0.18	-0.184	-0.184	-0.003	-0.012	-0.015
ZEDUC_ddd									0.639	0.629
U_ZEDUC										0.035
IMM_ZEDUC										-0.108

U: scaled uniform association. IMM: immobility. ddd: destination scaling (Z). COHx: Cohorts scaled between 0 and 1; AMBA: Buenos Aires Metropolitan Area. ZEDUC: years of education (Z). Coefficients in red are not statistically significant.

Conclusions

- There are great benefits of taking **first jobs** as the focus of intergenerational occupational reproduction / mobility studies.
- Most importantly, first jobs allow for **cohort comparison** and **correction of survey effects**
- The **Hauser-Goodman [HG] model** has great benefits over other loglinear models.
- Conditional Multinomial Logistic Regression Model (CMLR) can make traditional Gen3 bivariate mobility designs **multivariate**, and in particular can control survey effects.
- When applied to Argentina 1961-2021 (cohorts 1915-1995) **results are completely as expected from standard modernization theory.**

The BD SAT model

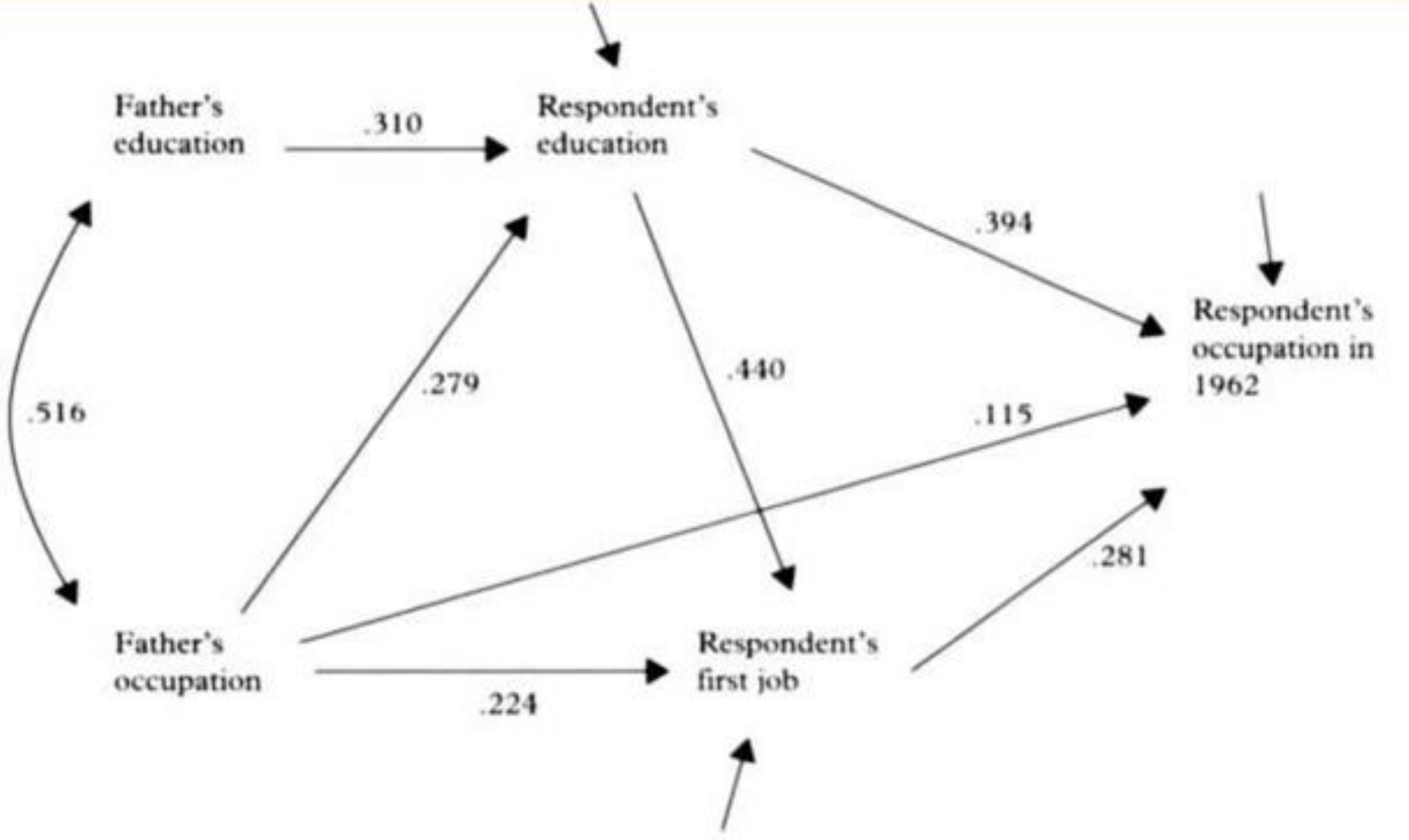
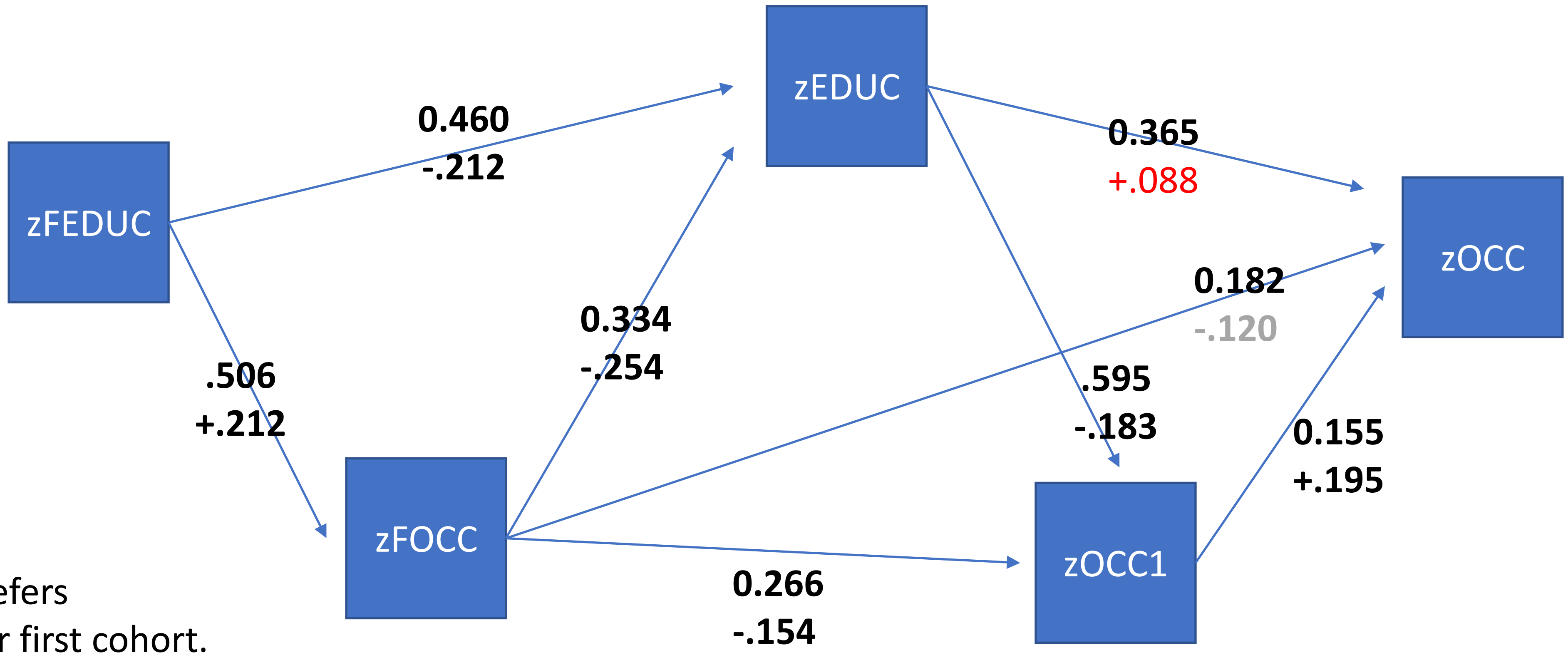


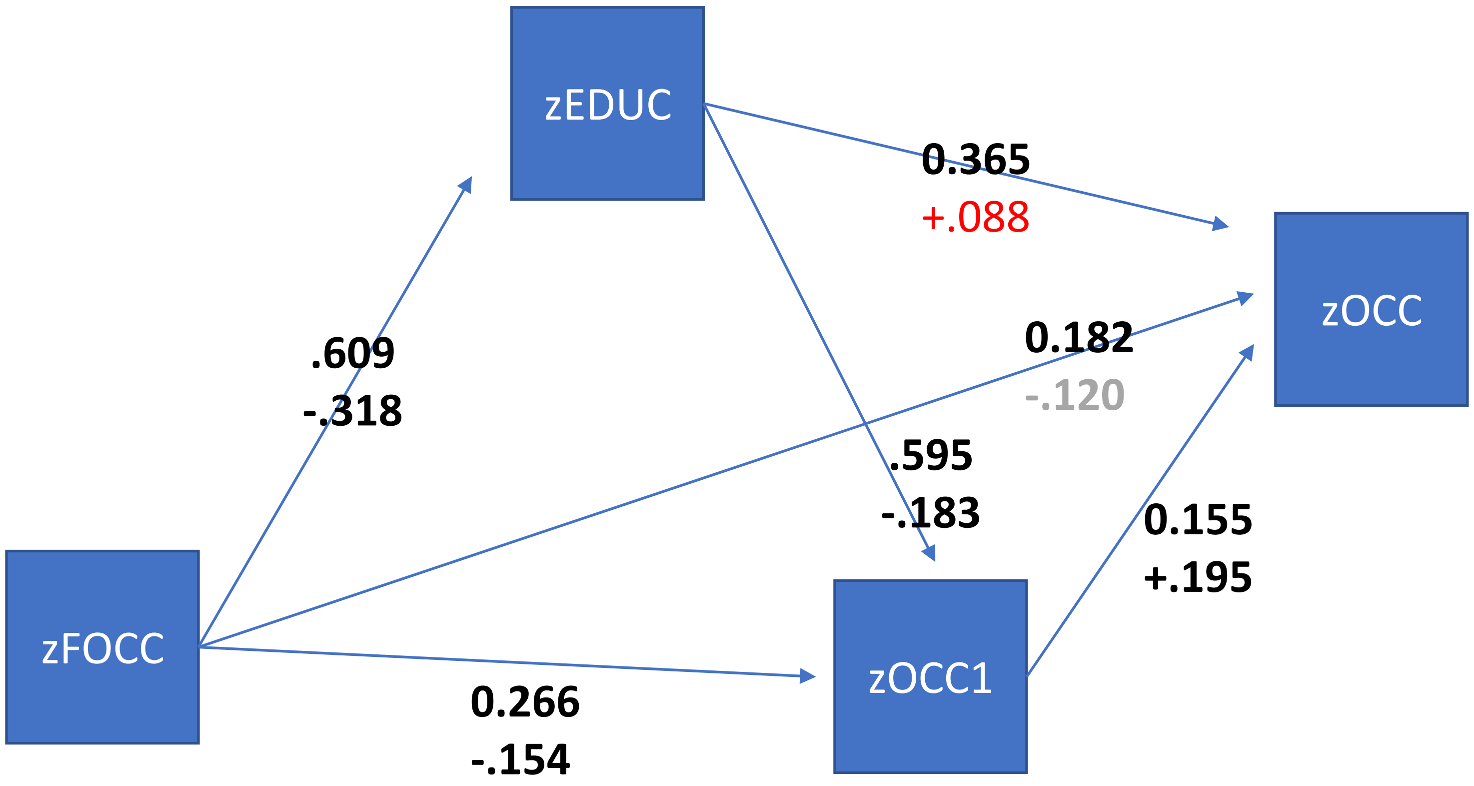
Figure 5.1. Blau and Duncan's (1967) path model of the process of stratification.

SAT model for AR, cohorts 1915-1993



First coefficient refers
To expectation for first cohort.
Second coefficient is change
Between first and final cohort

SAT model for AR, cohorts 1915-1993, simplified



Remarks at SAT analysis

- 8 surveys with first job information, N = 21415 (MLMV).
- Occupations measured in ISEI, education in years.
- 8 cohorts, equal size (=different intervals).
- Models with controls of survey main effects and survey interactions.
- All models controlled for FEMALE and AMBA & interactions.
- Linear OLS models, available data.
- Models with FEDUC refer to 6 surveys.
- Models with OCC refer to mature sample (age 30-55).

SAT and APC

- Interestingly, four of the five variables in BD-SAT model CANNOT change by Age: Father’s Education, Father’s Occupation, Education, First Occupation. Historical changes in these characteristics (and their associations) must be due to Cohort replacement.
- For these four characteristics Cohort and Period effects are separately identified (although collinear).

➔ If a certain cohort behaves differently in another Period, this cannot be because of historical change, but because of data deficiencies (‘survey heterogeneity’).

Period effects: can they exist in social mobility?

- It is hard to think of societal changes that would affect social mobility patterns as period effects:
  - Respondents do not choose their current/last occupation at the point of survey.
- Period effects are confounded with survey heterogeneity:
  - Different samples: sample frame, sampling procedure, fieldwork
  - Different measurement procedure: questionnaire, coding

  While we can try to reduce survey heterogeneity by data harmonization, this can never be perfect.

  ~~➔Combine cohort and period effects in the analysis of current occupation~~

  ➔Do not overinterpret period effects as showing societal change; rather interpret them as controls for controls for survey heterogeneity.

Substantive conclusions on SAT model

- It is NOT true that there are no FEGP effects after labour entry
- However, there are no significant historical (cohort) changes after labour market entry.

FIRST JOBS

First jobs – the eight advantages

- Making first jobs (="occupation at entry into the labour market after completing education") was NOT invented by Blau & Duncan (1967), but they did inspire it much.
- There are eight benefits of using first jobs to study intergenerational mobility.
  1. First job is a major 'determinant' ('pivot') of occupational careers.
  2. Parental occupation matters for the occupational career mostly / exclusively at the beginning of the occupational career.
  3. Education has its strongest effect on occupation career at the beginning. However, education also matters a lot after career beginnings.
  4. Everybody who has ever been employed, has had a first job, including non-participating women (mothers), unemployed, disabled, retired, etc.
  5. Policy makers are most interested in the education to employment nexus.
  6. **First jobs allow for the study of historical trends by cohort comparison.**
  7. **Cohort comparisons provide a wider time window than period comparison**
  8. **In pooled cross-sections, first jobs allow for correction of survey effects.**

First jobs – some disadvantages

- Retrospective data with recall error and selective attrition.
- “First” is a complicated attribute.
- First jobs strongly determine the further occupational career, but there is also a lot of difference between first and subsequent jobs.
- **There is much more data on current / last jobs than on first jobs.**

DATA

Argentina 1961-2021, 8 surveys with first job data

Table 2: Eight AR mobility surveys with first job data

Mean

SURVEY	YEAR	FEMALE	AGE	AMBA	FISEI	ISEI1	AGE1st	EDUC
.0 Germani 1961	1961	10%	47.5	100%	38.7	33.6	14.3	6.8
2.0 CEDOP - UBA, 1995	1995	53%	47.5	100%	34.1	32.6	18.2	10.2
7.0 CEDOP-UBA, 2007, ISSP 2006, 2007	2007	52%	44.4	36%	33.6	31.2	18.9	11.0
9.0 CEDOP-UBA, 2010, ISSP 2009	2010	47%	47.1	35%	30.6	31.3		10.4
11.0 EDSA-UCA, 2010	2010	52%	43.5	59%	34.4	30.7	17.8	11.0
13.0 Chávez Molina, Ipar - IIGG-UBA, 2013	2013	51%	45.4	100%	42.0	33.6	17.7	10.3
15.0 IIGG-UBA, 2016	2015	51%	44.5	100%	38.3	32.5	17.0	13.0
18.0 Covid 2021	2021	53%	43.0	20%	35.1	32.0	18.6	11.4
Total	2006	48%	44.7	56%	35.1	31.8	17.8	10.6

Data

- 8 surveys with first jobs from Argentina, collected between 1961 and 2021 (60 years): Cohorts entered the labour market between 1925 and 1995 (70 years).
- N = 18,370 men and women, organized in 8 / 16 cohorts.
- The surveys vary in quality:
  - Sample coverage (metropolitan (AMBA), non-rural, national).
  - Sampling methods: probability multi-stage, quota-sampling
  - Measurement of occupations and education.
  - Definition of parental: father, father + mother, main provider
  - First jobs definitions
  - More ….
- Survey heterogeneity will likely affect any social mobility estimate and possibly confound them, as more recent surveys probably have higher quality than earlier ones.

OLS results (earlier version)

OLS models with and without survey controls

Model A: No controls

Y = B0 + i.Cohort + X + COHx*X

Model B: Survey main effects

Y = B0 + i.Cohort + i.Survey + X + COHx*X

Model C: Survey main effects + survey associations

Y = B0 + i.Cohort + i.Survey + i.Survey.X + COHx*X

Cohort: discrete; COHx, 1922-1987, scaled as 0..1.0

FISEI → EDUC, trends without and with survey controls (Table 6)

A. Without survey controls:

zEDUC = 0.565*zFISEI – 0.268*zFISEI*COHx
(0.015)

C. With controls of survey main effects and survey interactions:

zEDUC = 0.674*zFISEI – 0.431*zFISEI*COHx
(0.019)

- Main effect of zFISEI refers to expectation in first cohort (1914) averaging over surveys
- Interaction with COHx refers to change until 1992 (after 78 years).
- Controls: Cohort (cat), Female, Female*Cohort, Survey, ZFISEI*Survey

First occupation, without and with survey controls (Table 7-8, 8 surveys, N=17552)

A. Without survey controls:

zISEI1 = 0.450*zFISEI – 0.202*zFISEI*COHx

B. With survey controls:

zISEI1 = 0.572*zFISEI – 0.395*zFISEI*COHx

A. Without survey controls:

zISEI1 = 0.194*zFISEI – 0.048*zFISEI*COHx
+ 0.542*zEDUC – 0.187*zEDUC*COHx

B. With survey controls:

zISEI1 = 0.232*zFISEI – 0.136*zFISEI*COHx
+ 0.547*zEDUC – 0.161*zEDUC*COHx

Current occupation, without and survey controls (Table 5ab, age 35-54: ‘occupational maturity’) (OLD)

Without survey controls:

zISEI1 = 0.386*zFISEI + 0.005*zFISEI* xCOH

With survey controls:

zISEI1 = 0.404*zFISEI – 0.014*zFISEI* xCOH

Without survey controls:

zISEI1 = 0.156*zFISEI + 0.109*zFISEI*xCOH
+ 0.516*zEDUC + 0.093*zEDUC*xCOH

With survey controls:

zISEI1 = 0.100*zFISEI – 0.109*zFISEI*xCOH
+ 0.490*zEDUC + 0.176*zEDUC*xCOH

EGP (CLASS) MOBILITY

Intergenerational occupational mobility

- Classical (Gen2) model: (standardized) regression = pearson correlation:
  - Single (=powerful) parameter to summarize the pattern of association
  - Can easily be expanded to multivariate analysis: the BD SAT model.
  - Parameters can be compared between countries if the samples are comparable.
  - However, estimates are sensititive to non-uniform difference between marginal distributions (=non-uniform structural mobility).
- After the ‘categorical revolution’ (Gen3): odds-ratio models:
  - Can model the pattern of association controlled from full marginal distributions
  - Generate often multiple, if not a multitude of parameters
  - Are intrinsically bivariate. Multivariate analysis requires multidimensional tables, that soon break down on ‘empty cells’ problem.

OBSERVED MOBILITY

	EGP1											
	1 I :Higher Controllers	2 II :Lower Controllers	3 IIIa:Routine Nonmanual	4 IIIb:Lower Sales-Service	5 IVa:Selfempl with empl	6 IVb:Selfempl no empl	7 V :Manual Supervisors	8 VI :Skilled Worker	9 VIIa:Unskilled Worker	10 VIIb:Farm Labor	11 IVc:Selfempl Farmer	Total
1 I :Higher Controllers	198	231	316	267	21	70	14	84	170	6	9	1386
2 II :Lower Controllers	86	316	385	508	26	103	7	172	324	18	11	1956
3 IIIa:Routine Nonmanual	31	105	263	362	8	49	3	141	241	19	3	1225
4 IIIb:Lower Sales-Service	19	67	136	367	4	58	2	118	228	16	5	1020
5 IVa:Selfempl with empl	43	140	154	248	138	65	6	163	176	12	17	1162
6 IVb:Selfempl no empl	47	169	296	592	27	505	6	426	681	50	25	2824
7 V :Manual Supervisors	21	53	46	62	3	16	16	66	115	4	7	409
8 VI :Skilled Worker	25	142	226	599	12	107	5	578	797	46	11	2548
9 VIIa:Unskilled Worker	40	167	233	807	17	164	21	436	1200	113	16	3214
10 VIIb:Farm Labor	11	34	32	180	1	47	1	118	393	251	11	1079
11 IVc:Selfempl Farmer	31	70	98	162	19	103	11	171	317	209	118	1309
otal	552	1494	2185	4154	276	1287	92	2473	4642	744	233	18132

FEP by EGP1 - OUTFLOW

	1 I :Higher Controllers	2 II :Lower Controllers	3 IIIa:Routine Nonmanual	4 IIIb:Lower Sales-Service	5 IVa:Selfempl with empl	6 IVb:Selfempl no empl	7 V :Manual Supervisors	8 VI :Skilled Worker	9 VIIa:Unskilled Worker	10 VIIb:Farm Labor	11 IVc:Selfempl Farmer	Total
1 I :Higher Controllers	35.9%	15.5%	14.5%	6.4%	7.6%	5.4%	15.2%	3.4%	3.7%	.8%	3.9%	7.6%
2 II :Lower Controllers	15.6%	21.2%	17.6%	12.2%	9.4%	8.0%	7.6%	7.0%	7.0%	2.4%	4.7%	10.8%
3 IIIa:Routine Nonmanual	5.6%	7.0%	12.0%	8.7%	2.9%	3.8%	3.3%	5.7%	5.2%	2.6%	1.3%	6.8%
4 IIIb:Lower Sales-Service	3.4%	4.5%	6.2%	8.8%	1.4%	4.5%	2.2%	4.8%	4.9%	2.2%	2.1%	5.6%
5 IVa:Selfempl with empl	7.8%	9.4%	7.0%	6.0%	50.0%	5.1%	6.5%	6.6%	3.8%	1.6%	7.3%	6.4%
6 IVb:Selfempl no empl	8.5%	11.3%	13.5%	14.3%	9.8%	39.2%	6.5%	17.2%	14.7%	6.7%	10.7%	15.6%
7 V :Manual Supervisors	3.8%	3.5%	2.1%	1.5%	1.1%	1.2%	17.4%	2.7%	2.5%	.5%	3.0%	2.3%
8 VI :Skilled Worker	4.5%	9.5%	10.3%	14.4%	4.3%	8.3%	5.4%	23.4%	17.2%	6.2%	4.7%	14.1%
9 VIIa:Unskilled Worker	7.2%	11.2%	10.7%	19.4%	6.2%	12.7%	22.8%	17.6%	25.9%	15.2%	6.9%	17.7%
10 VIIb:Farm Labor	2.0%	2.3%	1.5%	4.3%	.4%	3.7%	1.1%	4.8%	8.5%	33.7%	4.7%	6.0%
11 IVc:Selfempl Farmer	5.6%	4.7%	4.5%	3.9%	6.9%	8.0%	12.0%	6.9%	6.8%	28.1%	50.6%	7.2%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

FEP by EGP1 - OUTFLOW

	1 I :Higher Controllers	2 II :Lower Controllers	3 IIIa:Routine Nonmanual	4 IIIb:Lower Sales-Service	5 IVa:Selfempl with empl	6 IVb:Selfempl no empl	7 V :Manual Supervisors	8 VI :Skilled Worker	9 VIIa:Unskilled Worker	10 VIIb:Farm Labor	11 IVc:Selfempl Farmer	Total
1 I :Higher Controllers	14.3%	16.7%	22.8%	19.3%	1.5%	5.1%	1.0%	6.1%	12.3%	.4%	.6%	100.0%
2 II :Lower Controllers	4.4%	16.2%	19.7%	26.0%	1.3%	5.3%	.4%	8.8%	16.6%	.9%	.6%	100.0%
3 IIIa:Routine Nonmanual	2.5%	8.6%	21.5%	29.6%	.7%	4.0%	.2%	11.5%	19.7%	1.6%	.2%	100.0%
4 IIIb:Lower Sales-Service	1.9%	6.6%	13.3%	36.0%	.4%	5.7%	.2%	11.6%	22.4%	1.6%	.5%	100.0%
5 IVa:Selfempl with empl	3.7%	12.0%	13.3%	21.3%	11.9%	5.6%	.5%	14.0%	15.1%	1.0%	1.5%	100.0%
6 IVb:Selfempl no empl	1.7%	6.0%	10.5%	21.0%	1.0%	17.9%	.2%	15.1%	24.1%	1.8%	.9%	100.0%
7 V :Manual Supervisors	5.1%	13.0%	11.2%	15.2%	.7%	3.9%	3.9%	16.1%	28.1%	1.0%	1.7%	100.0%
8 VI :Skilled Worker	1.0%	5.6%	8.9%	23.5%	.5%	4.2%	.2%	22.7%	31.3%	1.8%	.4%	100.0%
9 VIIa:Unskilled Worker	1.2%	5.2%	7.2%	25.1%	.5%	5.1%	.7%	13.6%	37.3%	3.5%	.5%	100.0%
10 VIIb:Farm Labor	1.0%	3.2%	3.0%	16.7%	.1%	4.4%	.1%	10.9%	36.4%	23.3%	1.0%	100.0%
11 IVc:Selfempl Farmer	2.4%	5.3%	7.5%	12.4%	1.5%	7.9%	.8%	13.1%	24.2%	16.0%	9.0%	100.0%
	3.0%	8.2%	12.1%	22.9%	1.5%	7.1%	.5%	13.6%	25.6%	4.1%	1.3%	100.0%

Cohort by survey

SURVEY by COHORT									
	1915	1942	1954	1963	1970	1977	1984	1993	Total
.0 Germani 1960	1904	168	0	0	0	0	0	0	2072
2.0 CEDOP - UBA, 1995	488	670	361	312	295	85	0	0	2211
7.0 CEDOP-UBA, 2007, ISSP 2006, 2007	41	577	578	569	602	584	311	51	3313
9.0 CEDOP-UBA, 2010, ISSP 2009	40	238	161	163	144	164	136	87	1133
11.0 EDSA-UCA, 2010	162	864	692	803	605	872	991	693	5682
13.0 Chávez Molina, Ipar - IIGG-UBA, 2013	0	31	131	158	130	150	100	0	700
15.0 IIGG-UBA, 2016	0	0	209	193	202	193	179	89	1065
18.0 Covid 2021	5	258	390	576	665	635	798	1912	5239
	2640	2806	2522	2774	2643	2683	2515	2832	21415

Observed mobility AR 1961-2021

IMM UP DOWN * SURVEY

SURVEY		IMM	UP	DOWN
.0 Germani 1960	1799	19%	38%	41%
2.0 CEDOP - UBA, 1995	1787	18%	39%	41%
7.0 CEDOP-UBA, 2007, ISSP 2006, 2007	2688	16%	41%	43%
9.0 CEDOP-UBA, 2010, ISSP 2009	957	16%	42%	41%
11.0 EDSA-UCA, 2010	4850	21%	39%	39%
13.0 Chávez Molina, Ipar - IIGG-UBA, 2013	681	16%	28%	56%
15.0 IIGG-UBA, 2016	859	17%	36%	47%
18.0 Covid 2021	4511	26%	35%	37%
Total	18132	20%	38%	40%

IMM UP DOWN * COHORT

COHORT		IMM	UP	DOWN
1915	2223	19%	38%	41%
1942	2416	20%	41%	37%
1954	2211	19%	40%	40%
1963	2466	21%	39%	39%
1970	2291	20%	38%	41%
1977	2338	21%	36%	42%
1984	2110	22%	35%	42%
1993	2076	22%	35%	42%
Total	18131	20%	38%	40%

Why observed mobility is not so meaningful

- Observed mobility is a result of:
  - Structural mobility = the differences in marginal distributions of origins and destinations
  - Relative mobility (= social fluidity): the association between origins and destinations.
- The art of mobility analysis is in separating structural and relative mobility, not in confusing them.
- Notice: upward and downward mobility can only be asymmetric in structural mobility, not in relative mobility.

Why structural mobility is not so meaningful

- Structural mobility = differences between origins and destinations.
- Trends in structural mobility are produced by changes in both marginals.
- These changes are partly produced by occupational structures, but also by:
  - Fertility patterns: some occupations have more children than others
- A common error in mobility stories is that changes in observed mobility are attributed to changes in the destination marginal, but are in fact stronger connected to changes in the origin marginal.
- If observed mobility is decreasing, it may simply mean that younger cohorts had more advantaged parents than older cohorts: is this a bad thing?
- Structural mobility must be controlled, not researched.

RELATIVE MOBILITY

Odds-ratio

	DESTINATIONS	
ORIGINS	**a**	**b**
	c	**d**

- Odds-ratio:
  - = (a/b) / (c/d)
  - = (a*d) / (b*c)

- Odds-ratio's:
  - are sociologically meaningful
  - are insensitive to changes / differences in marginal distributions

- ➔ interpret loglinear models in terms of odd-ratio's

Mobility models: the Hauser-Goodman model

- There are many varieties of loglinear (and logmultiplicative) models of mobility patterns. I favour the Hauser-Goodman [HG] parametrization, in particular for comparative analyses (such as comparing mobility tables from different periods or cohorts).
- Start with the Uniform Association model: all contiguous odds-ratio's are equal to a single parameter U.
- "Contiguous" assumes order or categories.
- Goodman's RC-II multiplicative model estimates optimal scaling for the row and column (does not assume ordering but produces these): Ui and Uj.
- Scaled Uniform Association model estimates an overall association parameter U.
- Interesting specifications:
  - Constraining row and column scalings to be equal (Ui = Uj) = symmetrical association = there is a single social hierarchy for fathers and offspring.
  - Rescale the scalings to Z-variables: the scaled association parameter U then becomes numerically similar to a pearson correlation.

The addition of Hauser to Goodman’s RC-II model

- Hauser proposed to give the diagonal cells in the mobility table a separate treatment (immobility or inheritance effects).
- Then study the difference in immobility between tables in an over-all mobility effect IMM.
- The Hauser-Goodman model thus gives a different treatment to on-diagonal and off-diagonal association (stayers vs movers): IMM and U.

HG-model: advantages

- Parsimony, power: the model summarizes differences in association between tables in only two parameters .
- Ui = Uj is sociologically interpretable as a social hierarchy (as estimated from the mobility pattern).
- U is the single (standardized, logged) odds-ratio.
- When standardized, U become a familiar number: the pearson correlation.

HG model in practice

- HG model can be estimated in LEM (I have no experience with other softwares).
- In practical application, I estimate the model initially in LEM and transfer it to SPSS GENLOG with fixed scalings.

Hauser-Goodman parameters

Parameters of the Hauser-Goodman model, FEGP * EGP1 pooled table

	I	II	III-a	III-b	IV-a	IV-b	V	VI	VII-a	VII-b	IV-c
Ui=Uj	1.720	1.117	0.947	0.178	0.768	-0.094	0.373	-0.292	-0.521	-2.463	-1.246
U	0.384										
IMMk	.787	.235	.241	.480	2.465	1.234	2.059	.528	.316	-.360	1.972

When Ui = Uj are Z-standardized, the U parameter looks very much like a pearson correlation!

IMM and U by SURVEY, COHORT

IMM U by SURVEY

SURVEY		IMM	se	U	se
.0 Germani 1960	1799	-.382	.086	.309	.024
2.0 CEDOP - UBA, 1995	1787	-.559	.083	.539	.034
7.0 CEDOP-UBA, 2007, ISSP 2006, 2007	2688	-.661	.075	.453	.027
9.0 CEDOP-UBA, 2010, ISSP 2009	957	-.709	.110	.421	.039
11.0 EDSA-UCA, 2010	4850	-.508	.057	.419	.022
13.0 Chávez Molina, Ipar - IIGG-UBA, 2013	681	-.626	.132	.402	.064
15.0 IIGG-UBA, 2016	859	-.621	.114	.419	.047
18.0 Covid 2021	4511	ref		.281	.019

IMM U by COHORT

COHORT		IMM	se	U	se
1915	2223	-.159	.091	.350	.022
1942	2416	-.227	.085	.376	.024
1954	2211	-.249	.087	.427	.028
1963	2466	-.251	.084	.447	.029
1970	2291	-.149	.085	.353	.030
1977	2338	-.104	.084	.358	.030
1984	2110	-.119	.086	.347	.036
1993	2076	ref		.293	.033

Linear trends in HG loglinear parameters U and IMM, by cohort and survey.

- See Table in Excel
- By eight surveys:
    - Diagonal: IMM = IMMk + .665*YR
    - Parameters: IMMk: 0 -0.172 -0.276 -0.325 -0.123 -0.237 -0.234 0.386
    - Off-diagonal: U = .381 - .041*YR
    - Parameters:
- By eight cohorts:
    - Diagonal: IMM = IMMk + .191*YR
    - Parameters
    - Off-diagonal: U= .340 - .007*YR
    - Parameters: Uk: 270 .504 .422 .384 .429 .343 .458 .252

YR: window width: 0 .. 1

No control of survey effects!! We can only integrate the cohort and survey comparison with CMLR.

CMLR: Conditional Multinomial Logistic Regression

How to cure the empty cells problem: CMLR

- It has long been known how to rephrase various loglinear models in a linear regression model, that allows for introduction of control variables (including mediators) and is not sensitive to empty cells: the Conditional Multinomial Logistic Regression (CMLR) model.
- It requires setting up the mobility table data in an individual level data format and can be estimated in Stata with CLOGIT or XTLOGIT.
- (No doubt it can be estimated in other software.)

More about the CMLR model

- Also known as the discrete choice model by McFadden (1974). Nobel prize 2000.
- First introduced into social mobility analysis by Logan, J. A. (1983). A Multivariate Model for Mobility Tables. *American Journal of Sociology*, *89*(2), 324. https://doi.org/10.1086/227868.
- Other applications by DiPrete, Breen.
- More in: Hendricks & Ganzeboom (1998) and Dessens et al. (2003).
- Other applications: choice between political parties, consumer choices.
- Is very similar to a panel model and can be estimated with panel data model.

Setting up a data file for CMLR analysis

- For each unit (respondent) set up a record for each category of the dependent variable (CHOICES).
- Code CHOSEN=1 of respondent is found in a category and CHOSEN=0 otherwise.
- Add covariates for the respondents: survey, father’s occupation, education, gender, etc. Notice that these covariates are constant within cases.
- Add covariates for the dependent variable (the CHOICES), such as the status of an occupation.
- Effects of X-vars on Y are studied as interactions of the X- and Y-covariates.
- I prefer setting this up with the HG parametrization, but this is not necessary. It will also work with other loglinear models.

Part of a CMLR file (first two cases)

SURVEY	ID	CHOICES	CHOSEN	FEGP	EGP1	ooo	ddd	zod	IMM	FEMALE	AMBA	EDUC
0	1	1	0	11	4	1,030	-2,08	-2,14	0	0	1	3
0	1	2	0	11	4	1,030	-1,34	-1,38	0	0	1	3
0	1	3	0	11	4	1,030	-1,14	-1,17	0	0	1	3
0	1	4	1	11	4	1,030	-,204	-,211	0	0	1	3
0	1	5	0	11	4	1,030	-,904	-,931	0	0	1	3
0	1	6	0	11	4	1,030	,129	,133	0	0	1	3
0	1	7	0	11	4	1,030	-,444	-,458	0	0	1	3
0	1	8	0	11	4	1,030	,359	,370	0	0	1	3
0	1	9	0	11	4	1,030	,642	,661	0	0	1	3
0	1	10	0	11	4	1,030	3,043	3,134	0	0	1	3
0	1	11	0	11	4	1,030	1,580	1,627	1	0	1	3
0	2	1	0	4	4	-1,44	-2,08	3,001	0	0	1	9
0	2	2	0	4	4	-1,44	-1,34	1,937	0	0	1	9
0	2	3	0	4	4	-1,44	-1,14	1,643	0	0	1	9
0	2	4	1	4	4	-1,44	-,204	,295	1	0	1	9
0	2	5	0	4	4	-1,44	-,904	1,303	0	0	1	9
0	2	6	0	4	4	-1,44	,129	-,186	0	0	1	9
0	2	7	0	4	4	-1,44	-,444	,640	0	0	1	9
0	2	8	0	4	4	-1,44	,359	-,517	0	0	1	9
0	2	9	0	4	4	-1,44	,642	-,925	0	0	1	9
0	2	10	0	4	4	-1,44	3,043	-4,39	0	0	1	9
0	2	11	0	4	4	-1,44	1,580	-2,28	0	0	1	9

Number of cases read: 22 Number of cases listed: 22

Table xx: Stepwise estimated CMLR models, entry into first job, 16 cohorts, 1916-1992. Eight surveys 1961-2021, N=18.131.

Model specification

D1	Cohort + IMMk + U + U_COHx
D2	D1 + COHORT (marginals)
D3	D2 + IMM_COHx
D4	D3 + FEMALE + FEMALE interactions on U and IMM
D5	D4 + AMBA + AMBA interactions on U and IMM
D6	D5 + survey main effects
D7	**D6 + survey interactions U**
D8	**D7 + survey interactions IMM**
D9	D8 = education main effect
D10	D9 + education interaction on U and IMM

Table xx: Stepwise estimated CMLR models, entry into first job, 16 cohorts, 1916-1992. Eight surveys 1961-2021, N=18.131.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
DF	23	173	174	177	180	187	194	201	202	204
U	0.472	0.339	0.356	0.357	0.385	0.381	0.405	0.399	0.294	0.315
U_COHx	-0.258	-0.051	-0.083	-0.098	-0.118	-0.104	-0.206	-0.151	-0.096	-0.135
IMM_COHx			0.215	0.294	0.191	0.201	0.202	-0.180	-0.164	-0.07
Female_ddd				-0.394	-0.394	-0.422	-0.426	-0.426	-0.487	-0.411
U_female				0.038	0.039	0.029	0.015	0.016	-0.019	-0.022
IMM_female				-0.387	-0.393	-0.389	-0.386	-0.399	-0.409	-0.411
AMBA_ddd					0.049	0.059	0.067	0.055	0.038	0.038
U_amba					-0.029	-0.022	0.015	-0.012	-0.038	-0.036
IMM_amba					-0.18	-0.184	-0.184	-0.003	-0.012	-0.015
ZEDUC_ddd									0.639	0.629
U_ZEDUC										0.035
IMM_ZEDUC										-0.108

U: scaled uniform association. IMM: immobility. ddd: destination scaling (Z). COHx: Cohorts scaled between 0 and 1; AMBA: Buenos Aires Metropolitan Area. ZEDUC: years of education (Z). Coefficients in red are not statistically significant.

Conclusions

- There are great benefits of taking **first jobs** as the focus of intergenerational occupational reproduction / mobility studies.
- Most importantly, first jobs allow for **cohort comparison** and **correction of survey effects.**
- The **Hauser-Goodman [HG] model** has great benefits over other loglinear models.
- Conditional Multinomial Logistic Regression Model (CMLR) can make traditional G3 bivariate mobility designs **multivariate**, and in particular can control survey effects.
- When applied to Argentine 1961-2021 (cohorts 1925-1995) **results are completely as expected from standard modernization theory.**

References

- Dessens, Jos AG, Wim Jansen, Harry BG Ganzeboom, and Peter GM Van der Heijden. 2003. “Patterns and Trends in Occupational Attainment of First Jobs in the Netherlands, 1930-1995: Ordinary Least Squares Regression versus Conditional Multinomial Logistic Regression.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 166 (1): 63–84. https://doi.org/10.1111/1467-985X.00259.
- Hendrickx, John, and Harry BG Ganzeboom. 1998. “Occupational Status Attainment in the Netherlands, 1920-1990 A Multinomial Logistic Analysis.” *European Sociological Review* 14 (4): 387–403. https://doi.org/10.1093/oxfordjournals.esr.a018246.

Adding current / last jobs

- We need to solve the APC problem by assumption:
  - Restricting to MATURE sample == assuming that there are no aging-effects between age 30 (??) and 50 (??).
  - Forget about women??
- Add:
  - Survey interactions with DIA(od), DIA(id); IMM(od) and IMM(id)
  - Survey interactions with U(od) and U(id)
- Not yet finished