

A new International Socio-economic Index for the status of ISCO-08 occupations: ISEI-08

Harry BG Ganzeboom

Department of Sociology, VU University Amsterdam

SILC, September 8, 2020

Earlier presentation: Cambridge Stratification Seminar, Edinburg (online), August 26-28, 2020

Take-home points

- SEI-scales are the methodologically preferable and empirally best way to measure the status of occupations.
- The optimal scaling procedure used to develop ISEI gave a conceptual underpinning to the use of SEI scales
- There are three versions of ISEI available: ISEI-68, ISEI-88, ISEI-08, each connected to a different occupational classification: ISCO-68, ISCO-88, ISCO-08.
- ISEI-08 is at present a fake construction, but it works well.
- The paper constructs a new and true version of ISEI-08. However, it does not work as well as the fake version...

HISTORICAL AND BIOGRAPHICAL BACKGROUNDS

Occupational status scales

- Occupational status scales score (detailed) occupations into a unidimensional hierarchical measure (with ordinal or metric properties).
- There exist three main varieties of such scales:
 - Occupational prestige – popular evaluation of the value / attractiveness of occupations.
 - **Socio-economic index (SEI) – weighted sum of (required) education and (typical, expected) earnings.**
 - Social interaction / distance (CAMSIS) – optimizing associations between occupations in social interaction (friendship, marriage).
- (I avoid to include social class measures in this list – because I think the class – status debate is more about *what* is being scaled (class categories rather than occupations), than *how* it is scaled.)

...

- There has been a lot of debate about the relative merits of the three types of scales. The arguments are of two strands:
 - Conceptual / theoretical: what do we want to measure?
 - Methodological / empirical: how do our measures perform as explanandum / explanans (independent, dependent, mediating variables).
- Historically, conceptual issues have been given much weight in this discussion, and gave an edge to prestige and social interaction scales as measures that bring out the **social** meaning of occupations.
- But methodologically and empirically, SEI measures have prevailed, despite repeated complaints about their “circularity”.

Main methodological / empirical advantages of SEI measures

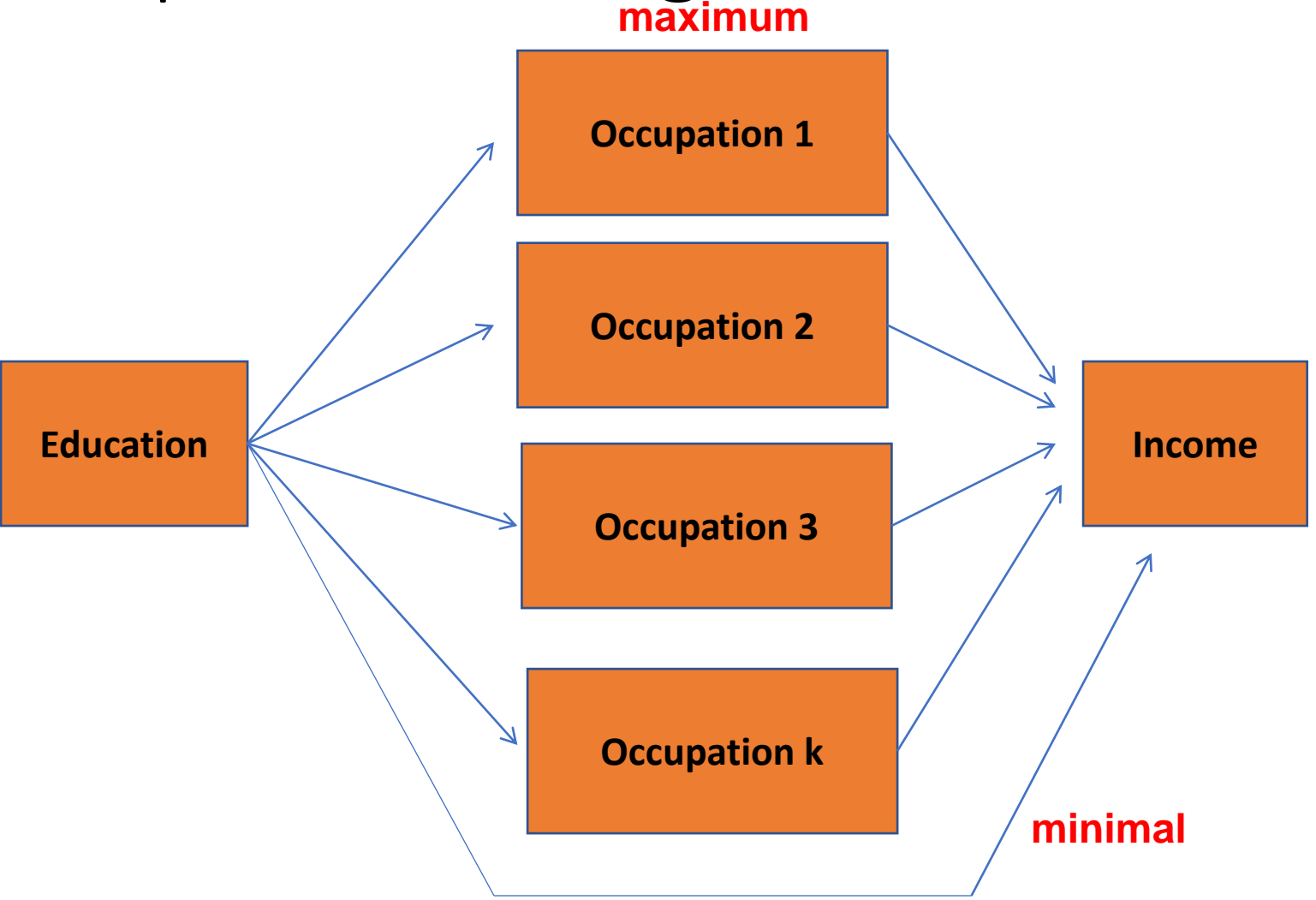
- SEI measures are easy to construct relative to prestige and social interaction scales:
 - SEI can be constructed from existing data that are readily available: education, occupation, income, age, gender.
 - For prestige scales you need a new data collection
 - For social distance scales, you need at least two cross-classified occupations. The diagonal in this table does not provide information on social distance.
- SEI measures tend to correlate more strongly with criteria that matter: education, earnings (but: circularity!!), but also other occupations, offspring's education, etc.

Duncan (1961): blurring prestige en SEI

- Duncan (1961) developed the best known SEI scale using US 1950 census data. (It was not the earliest SEI scale!).
- Duncan developed his scale to generalize prestige scores (from the 1947 NORC prestige survey) to all occupations.
- To do so he established the relationship between the prestige of NORC occupations and the (typical, age-adjusted) education and earnings of their incumbents, using the equation: $\text{Prestige} = .49 * \text{Income} + .47 * \text{Education}$, for 45 occupations – and then derived **expected prestige** for hundreds of occupations.
- This particular procedure has confused many users of SEI, because it suggested that SEI is a ‘proxy’ of the real thing – prestige.
- But in fact, Duncan found out (later firmly re-established by Featherman & Hauser (1976)) that the SEI scores work much better as occupational status indicator than prestige scores, even for the original 45 NORC occupations.
- “Prestige [rather] is an ‘error-prone’ proxy for socio-economic status [of occupations]” (Featherman & Hauser, 1976: p 406).

The best idea I have ever had

The SEI optimal scaling model



Conceptual interpretation

- The optimal scaling model conceives of occupation of what occupation is: the bridge from education to earnings.
- The EDUC → OCC → INCOME chain is a familiar model of the stratification process: how inequality is created.
- Methodologically, it a mediation (or: indirect effect) model.
 - In mediation models, the amount of mediation is crucially dependent upon the measurement quality of the mediation variable.
 - So: conceptualizing occupation as the main mechanism that transfers education into earnings requires an optimally measured (=scaled) mediator.

Development of ISEI

- Ganzeboom, De Graaf & Treiman (1992) used the optimal scaling model to develop the International Socio-economic Index of occupational status [ISEI].
- The indirect effects optimization ('PathAls') problem was solved by De Leeuw (1992) in a separate appendix to the paper.
- Backbone: ISCO-68 (hence: ISEI-68).
- Data: Full-time employed men (N=73.901) from 16 countries, with 271 separately scaled occupations.
- Ganzeboom & Treiman (1996) repeated this for ISCO-88 (hence: ISEI-88), **using the same data**.
- (ISEI-88 became very successful, primarily because it was implemented in PISA as a standard scale to measure social background – and turned out to be a better predictor of student achievement than parental education.)

ISCO AND ISEI

ISCO: 1968 - 1988 - 2008

- ISCO is periodically revised. The 1988 version was a major revision of ISCO-68, building upon different ideas about what constitutes a properly measured occupation and what not.
- ISCO-88 is much better organized and documented than ISCO-68 and became adopted by many cross-national surveys (such as ESS, ISSP and PISA) as well as official (comparative) statistics agencies (ILO, EuroStat).
- Many national surveys and statistical agencies have used ISCO-88 as their standard way to classify occupations.
- ISCO-88 was revised (a minor revision) in 2008, but it took a few years before it was properly documented.

ISCO-88 → ISCO-08: a minor revision?

- ILO announced it as a ‘minor’ revision, but in fact:
 - There exists no one-to-one mapping between the two classifications, even not at the major or sub-major group level.
 - Some parts of the sub-major (two-digit) parts were differently organized, in particular at two locations: Managers (1100, 1200, 1300) and Elementary Occupations (9400-9600).
 - Interestingly, ISCO-08 revived some of the distinctions that had disappeared in the ISCO-68 → ISCO-88 transition (small self-employed).
- ISCO-08 is about 10% more detailed than ISCO-88. When cross-walked, going back to ISCO-88 from ISCO-08 is potentially less harmful than the other way around.

2010

- In 2010, ISCO-08 was about to be implemented in cross-national surveys. PISA/PIAAC implemented it in 2010, ESS in 2012, ISSP in 2014.
- However, at this point there was no instrument available to scale ISCO-08 occupations, and there was no data available to construct an ISEI-08 scale.
- So I faked it...
- Ganzeboom (2010)

A FAKED ISEI-08

Ganzeboom (2010): ISEI-08

- Data: ISSP2002-2007, N=200.000 men and women in 42 countries.
- ISCO-08 codes were created by converting ISCO-88 to the new classification.
 - Logically such a conversion is impossible (or there would be no point in launching a revised classification).
 - However,
 - One-to-many mappings often have a dominant part
 - Some ISEI-08 scores were in fact imputed by referring back to ISEI-68 scores.
- The paper was never formally published, but has received a gratifying number of citations.
- However, the ISEI-08 scores were provisional, and would need to be formally redeveloped at some future date – this date seems to have come.

Ganzeboom (2010): three new elements

- ISEI-08 was constructed on men **and women**, giving in to repeated accusations of ‘intellectual sexism’, ‘male bias’, etc.
 - (The earlier ISEI articles had argued that using (full-time) employed men as a benchmark would be more appropriate to bring out women’s disadvantages in the labour market – by NOT building these into the ISEI scale.)
- As the data used were originally cast in ISCO-88, the **ISEI-08 could be used with both ISCO-88** and ISCO-08 data.
- The paper used a **validation model** to evaluate the relative strength of the old and the new ISEI measures – which used relationships between multiple occupations and between those occupations and other variables to identify a formal measurement model. The new ISEI-08 came out better by a .98/.94 margin.

NEW CONSTRUCTION OF ISEI-08

Data: ISSP 2014-2018 (5 rounds)

- ISSP is a world-wide annual social attitudes survey with a fair amount of representation of European (North-West, East and South), rest-of-the-west, Asian and Latin-American countries. Also: South-Africa, Turkey, Suriname.
- BV includes: detailed (four-digit) occupations, education (qualification and duration), personal and household incomes.
- ISSP adopted ISCO-08 only from 2014 onwards. The most recent (2018) dataset was published (incompletely) last week (August 15 2020).
- Initial sample: 243.452 cases with valid occupations, from 43 countries.
- However, the effective sample is substantially reduced because:
 - Occupations also measure last occupation, for non-employed (women, retired, unemployed).
 - Age restriction: 18-64 year of age.
 - Missing data on earnings (even after missing personal incomes are imputed with household incomes) – but this is not so problematic.
- **Table 1** gives an overview of the selection of the effective sample selection.

Table 1: Effective sample

• Initial (43 countries)	N = 243.452
• With valid occupations	N = 202.122
• After selecting on currently employed	N = 125.505
• After selecting WRKHRS > 20	N = 110.893
• Missing education	N = 384
• Missing income	N = 16.489
• Listwise sample	N = 94.120

Number of occupational groups

		N < 20
• Four-digit groups in the data	404	24
• Three-digit groups in the data	130	??
• Two-digit group	38	0
• One-digit groups	10	0

Issue #1: How to treat earnings

- ISSP would allow us to convert earnings into hourly wages:
 - By dividing earnings by number of hours worked,
 - By adjusting earnings for hours worked.
- (The latter procedure is less sensitive to measurement problems in hours worked, and non-linearities in its effects.)
- The earlier ISEI papers avoided this problem by restricting the data to full-time employed men...
- But is it theoretically obvious to go for hourly wage? Why not total earnings?

Issue #1: How to treat earnings?

- ISSP contains: RINC: personal income, INC: Household income, Work (whether employed), WRKHRS (actual hours worked).
- These variables have various patterns of incompleteness.
- Related problem: some data contain zero incomes. While not impossible, it is implausible that employed workers would have no (monetary) income. I made these missing.
- Local currencies. Made comparable by $\text{LN}(\text{income}/\text{mean}(\text{income}))$
 - Removes local unit of measurement, but maintain local income inequality
 - Makes distribution more symmetric
 - Reduces outliers
- Alternative would be: taking within-country ranks \rightarrow uniform distribution. Lowers R2.
- Possible: make outliers / top-bottom 1% missing.
- We can compare the different alternatives how to treat earnings by evaluating the EDUC \rightarrow INCOME relationship.

Hourly or total wages?

- Earlier SEI scales were developed on full-time employed men and did not adjust for hours worked.
- Women work less hours than men:
 - Raw gap: -5.73
 - After ISCO adjustment: -4.11
- Between ISCO-08 differences in hours worked: adj R2 = 8.7%
- Between ISCO-08 feminization of occupations: adj R2 = 30.9%.
- (Aggregate correlation: $r = -.62$)
- Both hours worked and feminization are 'part of the job' → no adjustment.

Table X: Determinants of Personal Income (z) in ISSP2014-2018

zEDUC	0.345	0.361	0.352	0.338	0.182	0.170	0.161
FEMALE		-0.519	-0.420	-0.388	-0.308	-0.412	-0.326
HRS			0.017				
HRScat				D	D	D	D
ISCO08					D		
zISEI08						0.292	
zISEI08new							0.331
ADJ R2	11.8%	18.3%	23.0%	25.8%	34.6%	31.4%	33.4%
N = 100.019, All employed workers with hours > 8.							

Issue #2: How to include women

- Women are heavily concentrated in a few female-typed occupations.
- Women have about the same level of education as men, but there is a wide gender gap in earnings, both between and within occupations:
 - Raw gender gap in earnings -0.51 (SD)
 - After adjusting for hours worked -0.42
 - + after controlling (detailed) occupations -0.32
- Including women in the scaling procedure invariably ends up with female-dominated occupations receiving lower scale scores...

Issue #3: Minimum size of occupation groups

- The earlier ISEI papers used a cut-off point of $N=20$ to treat an occupation group as a separate unit in the scaling procedure.
- In the present data about xx percent of the cases are in smaller groups.
- The cut-off point of $N=20$ is arbitrary, but can be empirically evaluated.
 - Making the cut-off point too low would increase random variations in the scale scores.
 - Making the cut-off point too high would also increase random variation in the scale scores, because of aggregation error.
- The opposing tendencies can be evaluated in fresh data.

Issue #4: Age adjustment

- Duncan (1961) adjusted for age (cohort) differences in both education and earnings, on the argument that the scores for some occupations are driven by the young age of their incumbents – which he regarded as bias. Think: shelf-fillers in shops.
- I am not so sure whether this is an valid argument (but in Ganzeboom et al. (1992) I thought it was..)
- In fact, age differences between occupations are not very influential: Age differences suppress the EDUC → INC link, but have little influence on the mediated / non-mediated part.

ISCO

- ISCO-88 and ISCO-08 are organized in a four-digit hierarchical code, such as:
 - 1000 Managers (major groups)
 - 1100 Chief executives (sub-major groups)
 - 1110 Senior Officials (minor groups)
 - 1111 Legislators (unit groups, occupation)
- Notice that I added trailing zeroes.
- This hierarchical (nested) system makes it particular convenient to develop an SEI score at a more detailed / crude level.

Education

- ISSP measures level of education in two independent ways:
 - ISCED level 0-6
 - Years of education completed.Correlation: 0.81
- ISCED level is the better predictor of occupational status and earnings. A combined measure is even slightly better.
- The comparative measure is Z-standardized in the scaling procedure.

Earnings

- Earnings are measured as ratio variables.
- Some countries included zero incomes – which seems impossible for households and also for employed workers – I made these incomes missing.
- Transformation: $LNPINC = \ln(PINC/xPINC)$, I which xPINC is the within-survey mean of PINC.
- (This transformation preserves income inequalities between countries, while producing a comparative metric.)
- Like education, the means of LNPINC are Z-standardized in the scaling procedure.

Optimal scaling algorithm

- Optimal scaling is done by searching for the optimal weight in:
 - $OPTI = p * ZEDUC + (1-p) * ZLNPINC$
- The scores are evaluated in the regression equation:
 - $ZLNPINC = B1 * ZEDUC + B2 * OPTI$
- Scales scores are optimal when B1 reaches a minimum, which happens to be the case for $p = 0.35$ ($1-p=0.65$).
- B1 became 0.15, relative to a total effect of 0.35.

...

Table 2: SAT model with optimal scaled occupations

		Beta	T
1	(Constant)		.0
	ZEDUC	.348	114.2
2	(Constant)		.2
	ZEDUC	.154	46.4
	OPTI	.379	113.9
3	(Constant)		51.9
	ZEDUC	.185	56.8
	OPTI	.348	106.6
	FEMALE	-.210	-74.7

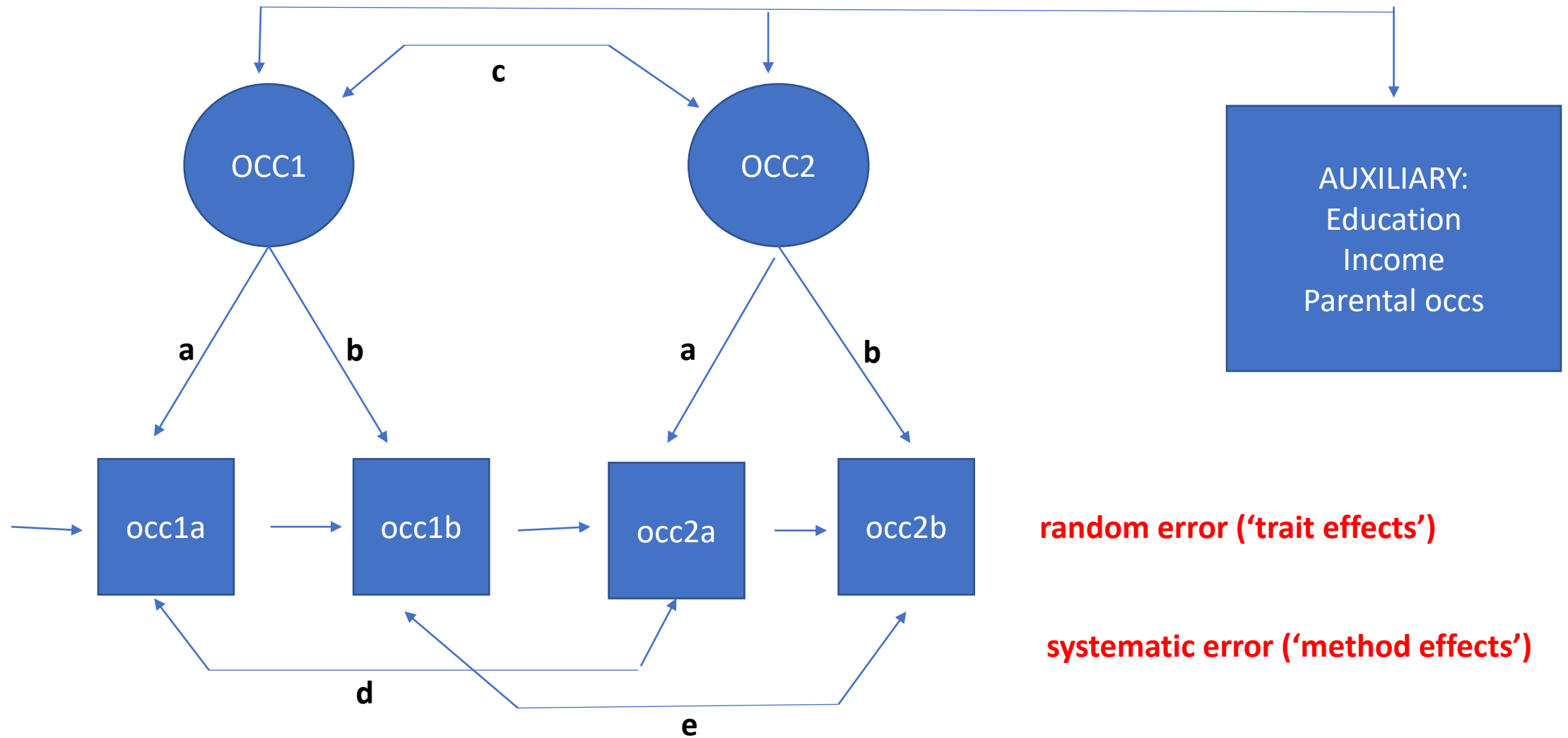
Dependent variable: ZLNPINC

VALIDATION

Validation on ESS R5-R8

- For validation of the new scale, it is best to use fresh data, which contains multiple occupations, originally coded in ISCO-08.
- ESS-R5_R8 is a good source:
 - Cross-national, although not worldwide
 - ISCO-08 codes for respondents and spouses occupations (but only partially for parents).
 - Auxiliary variables: education, household income, father's and mother's occupation (partially) crude classification.
- I use a (confirmatory) factor-analytic MTMM model to identify random and systematic measurement error.

MTMM validation model



random error ('trait effects')

systematic error ('method effects')

Table 3: MTMM correlations (ESS R5-R8)

	osei	nisei	sosei	snisei	educyr	hinc	fisei	misei	female
osei	1.000	.872	.396	.385	.573	.342	.294	.280	.000
nisei	.872	1.000	.390	.367	.556	.359	.277	.262	-.126
sosei	.396	.390	1.000	.887	.405	.334	.272	.263	-.035
snisei	.385	.367	.887	1.000	.402	.334	.262	.251	.101
educyr	.573	.556	.405	.402	1.000	.370	.340	.320	.016
hinc	.342	.359	.334	.334	.370	1.000	.214	.210	-.090
fisei	.294	.277	.272	.262	.340	.214	1.000	.534	-.011
misei	.280	.262	.263	.251	.320	.210	.534	1.000	-.016
female	.000	-.126	-.035	.101	.016	-.090	-.011	-.016	1.000

Results

- Neglegible systematic error (**d** and **e**)
- Coefficient **a** (ISEI-88) 0.947
- Coefficient **b** (ISEI-08 new) **0.924 ← Bad!**
- Coefficient **a** (ISEI-88) 0.939
- Coefficient **b** (ISEI-08 faked) 0.980

Conclusions and discussion

- Much to my surprise, the new ISEI-08 scale is not an improvement over the faked ISEI-08 scale that I created in 2010.
- Possible reasons are:
 - Occupation coding in the ISSP deteriorated over the years
 - ISSP coverage expanded over the years, including countries with very different occupational hierarcies, such as China, India and Thailand.
 - I should not test a world-wide scale only on European data. But what is the alternative.
- What to do?

References

Duncan, Otis Dudley. 1961. "A Socioeconomic Index for All Occupations." In *Occupations and Social Status*, edited by Albert J Reiss, 109–138. New York: Free Press.

Duncan, Otis Dudley. 1961. "Properties And Characteristics Of The Socioeconomic Index." In *Occupations and Social Status*, edited by Albert J Reiss, 138–161. New York: Free Press.

Featherman, David L, and Robert M Hauser. 1976. "Prestige or Socioeconomic Scales in the Study of Occupational Achievement?" *Sociological Methods & Research* 4 (4): 403–22. <https://doi.org/10.1177/004912417600400401>.

Ganzeboom, Harry BG, Paul M De Graaf, and Donald J Treiman. 1992. "A Standard International Socio-Economic Index of Occupational Status." *Social Science Research* 21 (1): 1–56. [https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B).

Ganzeboom, Harry BG, and Donald J Treiman. 1996. "Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations." *Social Science Research* 25 (3): 201–39. <https://doi.org/10.1006/ssre.1996.0010>.

Ganzeboom, Harry BG. 2010. "A new international socio-economic index [ISEI] of Occupational status for the international standard Classification of occupation 2008 [ISCO-08] constructed with data from the ISSP 2002-2007; with an analysis of quality of occupational measurement in ISSP. Paper resented at annual conference of International Social Survey Programme, Lisbon PT, May 1 2010