Multiple Indicator Models for Social Background

Harry B.G. Ganzeboom
European Survey Research Association
Warsaw, July 1 2009

Common practice

- Common practice:
  - Estimate random error using reliability analysis (Cronbach's alpha)
  - Explore systematic bias (multidimensionality) using factor or component analysis
  - Use component analysis or simple summation to construct an index variable.
- All of this only serves the diagnosis, but hardly ever the correction of errors.

Multiple indicator models - basics

- In measurement of attitudes, it is common practice to use multiple indicators, for the following reasons:
  - Indicators relate to a concept in a part-whole relationship (they cover a part of a broader concept).
  - I.e. indicators of a single concept have a common part, and a unique (but systematic) part, which may cause bias when only one indicator would be used.
  - Indicators are unreliable measurements, i.e. a large part of their variance is random.
- ***All of this also applies to social background variables, although the main point here is unreliability.***

Theses (1)

- The best way to improve the quality of measurement of social background variables is the same as for any variable: use multiple indicators.
- Multiple indicators allow us to estimate and correct random error and systematic error.
- Random error turns out to be a bigger problem than systematic error.
- Even if random error is small (e.g. relative to social attitudes), it makes a big difference to correct for it.

Debates on measurement quality of social background variables

- Unlike social attitudes, social background variables are hardly ever measured with multiple indicator models.
- There seems to be a common assumption that social background indicators are more reliable and more valid than attitude indicators.
- Nevertheless, there are many debates about the measurement of concepts like educational attainment and occupational status. These debates often focus on the "single best indicator" and validity (bias) issues.
- These debates could be resolved by acknowledging that there may be best indicators, but there is no such thing as a perfect indicator.

Theses (2)

- The only way to detect random error is to repeat measurement, i.e. use multiple indicators.
- The only way to detect systematic error (bias) is to repeat the error, i.e. use multiple indicators that vary error sources systematically over – correlated – constructs.
- In other words: MTMM designs are very appropriate to secure fully error-corrected social background effects.
- Having a bad second indicator is much better than no second indicator.

OCCUPATIONS

Occupation – measurement

- Occupations are most often measured using open questions. These questions typically refer to:
  - Industry
  - Job title
  - Description of activities and responsibilities
  - Required qualifications
  - Supervising status
  - Employment (contract) status.
- This information is post-coded into a detailed occupational classification, such as the International Standard Classification of Occupation [ISCO], which contain 100-1000+ categories.

ISCO

- ISCO is a detailed classification for all countries, regularly updated, but with a very stable backbone.
- ISCO is global, stable, and well documented.
- However, how well ISCO is actually coded is unknown, even undocumented.

Occupation - scaling

- The detailed occupational information is then transferred into a less detailed / more manageable measure of occupational status, such as:
  - A limited number of categories, such ISCO first digit, or the well-known EGP "class schema".
  - A metric variable that expresses the general "prestige" or "socio-economic status" of occupations: SIOPS, ISEI.

Murphy (1)

If things can go wrong, they will go wrong.

Murphy (2)

Even if things cannot go wrong, they will go wrong

ISCO coding in ESS

- Knowledge of the the ISCO manual is largely absent.
- Occupation coding is not double checked.
- The occupation strings are not archived.

A second indicator

- We can only assess measurement quality using multiple indicators.
- A practical problem seems to be that it is hard to conceive of another measure or occupation that is acceptable to respondents.
- But in fact, some research has used multiple indicators by combining the detailed occupation question with a precoded, 'crude' indicator.
- A good – be in coincidental -- example is provided by the ISSP 1987.

ISSP1987: self-classification

Here is a list of different types of jobs. Which type did your father have when you were 16 years / [did you have in] the first job you had after you finished your full-time education / [do you have]] in your job now?

1. Professional and technical (for example: doctor, teacher, engineer, artist, accountant)
2. Higher administrator (for example: banker, executive in big business, high government official, union official)
3. Clerical (for example: secretary, clerk, office manager, civil servant, bookkeeper)
4. Sales (for example: sales manager, shop owner, shop assistant, insurance agent, buyer)
5. Service (for example: restaurant owner, police officer, waiter, barber, caretaker)
6. Skilled worker (for example: foreman, motor mechanic, printer, tool and die maker, electrician)
7. Semi-skilled worker (for example: bricklayer, bus driver, tannery worker, carpenter, sheet metal worker, baker)
8. Unskilled worker (for example: labourer, porter, unskilled factory worker)
9. Farm (for example: farmer, farm labourer, tractor driver)

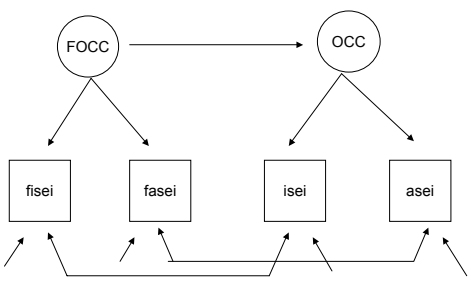
Was your father / were you / are you self-employed, or did he /did you / do you work for someone else?

1. Self-employed, own business or farm
2. Work[ed] for someone else

Reference

- Ganzeboom, Harry B.G. (2005). "On the Cost of Being Crude: A Comparison of Detailed and Coarse Occupational Coding." Pp. 241-258 in: Jürgen H.P. Hoffmeyer-Zlotnik, Methodological Aspects of Cross-National Research, Mannheim: ZUMA-Nachrichten [Special Issue #11], 2005.
- See: http://home.fsw.vu.nl/hbg.ganzeboom

SEM / MTMM model



ISSP87 results

- Measurement relationships:
  - Crude measurement: 0.829
  - Detailed measurement: 0.835
- Education → Earnings
  - Single indicator measurement: 0.13
  - Multiple indicators 0

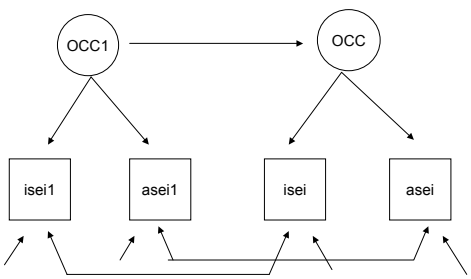
Conclusions ISSP87

- Detailed and crude codes do make some difference to results: random error attenuation decreases by about 9% if you use crude codes.
- However, including both does:
  - Substantively increase explained variance, and
  - Gives more plausible results, in particular no direct effect of education on earnings.
- Attenuation of single indicator measurement relative to multiple indicator model is between 12% and 20%.

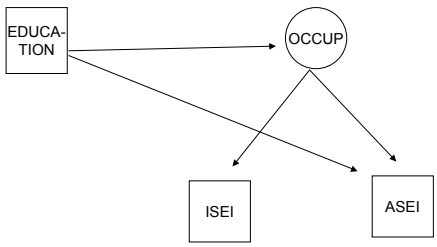
Replication: ISSP-NL data

- In ISSP-NL I have collected double indicator measures for four occupations:
  - Father, mother, first and current
  - The crude question here is the format invented in ISSP87 (nine categories with clear meaning and rank order).
  - Table 1 shows the correlations and N.
- The repetition over four occupations and three persons allows for an MTMM model that assesses to kinds of systematic bias:
  - Echo-effects, I.e. similar response tendencies to similar question formats.
  - Education bias: the tendency to respond to the crude distinction, using educational information.

SEM / MTMM model



SEM: Education bias



ISSP-NL: findings

- With respect to random error, crude occupation measures are almost as good as detailed measures.
- The amount of correlated error ('echo') is limited and does not bias causal coefficients; it is the same for crude and detailed measures!
- The amount of education bias in the crude measure is statistically significant (!) but substantially negligible.

Occupational measurement in ESS

- ESS requires countries to code occupations of respondents and spouses in ISCO. How this is done, remains unknown.
- ESS has decided <u>not</u> to code occupations of fathers and mothers, but in stead:
  - Use a precoded question (of very dubious quality), and
  - Still collect the open information and make it available in verbatim format.
- So by sheer coincidence and lack of foresight ESS has created multiple indicator measurement!

ESS showcard

1. Traditional professionals
2. Modern professionals
3. Clerical and intermediate
4. Senior manager and administrator
5. Technical and craft
6. Semi-routine manual and service
7. Routine manual and service
8. Middle and junior managers

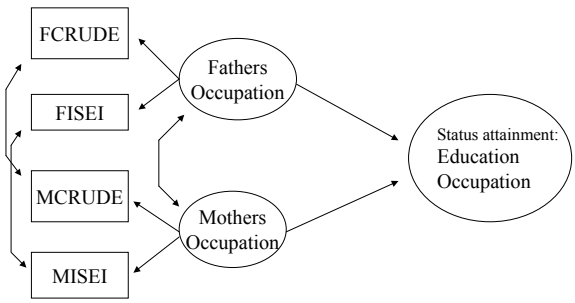
ESS precoded occupations: problems

- The precoded format that ESS has used for parental occupation, has severe problems:
  - The categories are unclear and out of order.
  - The categories miss out or blur the parental occupation that is most prevalent and is know to behave in a extremely peculiar way: farmers!
- However you work with these variables (few people do), it is clear that they underestimate parental occupation effects.

How to operate the precoded question?

- Create optimal order or scale, using criterion variables.
  - This can be done at the cross-national levels, the national level, or even the survey level.
  - Optimization can be chosen to be the same for fathers and mothers, but also be gender specific.
- Quality can be assessed by coding the independently collected open information into ISCO. This has now been accomplished for approximately 30% of the ESS data.
- Because the questions are repeated over fathers and mothers we can estimate both random and systematic (correlated) measurement error.

MTMM-model



EDUCATION

Education

- Education – at a national – level is typically measured precoded (showcard), using a limited number of categories.
- These formats usually abstract away from:
  - Historical changes in the education system (that must be important for many respondents).
  - Educational careers – multiple credentials.
  - International differences, foreign credentials.
- Harmonization is left to the respondents / interviewer.
- This seems to work rather well: education is a dominant variable in predicting almost anything.

Education – comparative measurement

- Two approaches to comparative measurement:
  - Common denominator approach, often using the International Standard Classification of Education [ISCED];
  - Duration: some question on total length of education (alternative: age of leaving education).

ISCED

- ISCED [UNESCO maintained] describes the existing educational system in OECD countries in a certain year (most recently: 1997).
- It organizes the details in seven categories that can be regarded as an ordinal variable.

ISCED

0 – Pre-primary
1 – Primary
2 – Lower secondary
3 – Upper secondary
4 – Post-secondary, but not Tertiary
5 – Tertiary – BA+MA level
6 – Advance Tertiary -- PhD

Problems with ISCED

- It refers to a certain year (1997) and is not historically sensitive.
- It is restricted to a limited number of countries.
- Some important distinctions, at the higher secondary and tertiary levels are not part of the first digit.
- The number of levels distinguished is in practice fairly limited, and may boil down to as few a 3-4. This is unacceptable for within-country use.
- In practice, researchers have a hard time applying ISCED in a correct way (ESS, PISA).

Problems with common denominator harmonization

- Common denominators are hard to find.
- The more data you have to compare, the less detailed your common denominator will be: more data, more problems.
- You will always end up making either very crude distinctions, or compare the incomparable.

Varieties of duration measures

- How many years have you been in education?
  - After age 6?
  - After age 12?
  - After compulsory age?
- At which age did you leave education?
  - For the first time?
  - For the last (most recent) time?
- How many years of education have you completed?
  - This may refer to an institutional duration: it may take you 3 years to complete the first year of university!

Problems with duration

- Many have argued that duration only works in comprehensive systems, generates error and bias in non-comprehensive (tracked, divided) systems.
- Others (human capital theorists) may argue that duration has the advantage of being sensitive to multiple credentials.
- The question itself may impose arithmetic on the respondent – a sure cause of error; in practice it always produces outliers.

Education in ESS

- Education in ESS is measured in <u>two</u> independent ways, that are converted into <u>three</u> measures for data users.
  - EDUYRS Duration
  - EDLVXX Local measure (varies in detail)
  - EDULVL Local measure post-coded in ISCED (common denominator)
- Most users prefer EDULVL as their indicator.
- Education of spouse, father and mother have been measured, but only in ISCED: EDULVLF, EDULVLM, EDULVLP.
- (A similar approach is used in ISSP.)

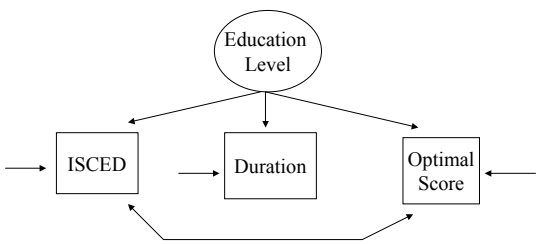
There is no need to use common denominator harmonization

- For many research questions, using a common denominator approach to harmonization is not needed and harmful.
- If you simply accept that education ranks respondents in a single (uni-dimensional) hierarchy (most theories do), an obvious comparative metric is an optimal rankorder, expressed in percentiles.
- This way there is no need to throw away any of the information.

Multiple indicator model

- We should not choose between ISCED and duration, but use them both.
- The three (!) measures can be employed in a multiple indicator model to estimate and correct random error.
- The local measures are rendered comparable by using optimal ordering / scaling.
- ISCED is interpreted as an ordered variable.
- The level of aggregation bias in ISCED can be estimated by adding the duration as a third indicator.

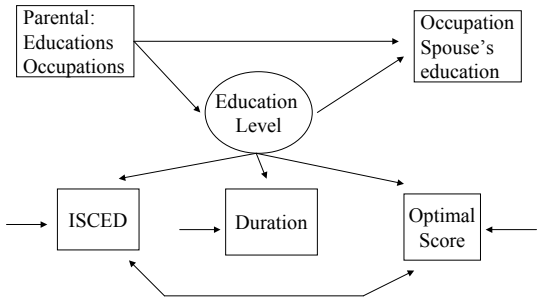
Multiple indicator model



Identification

- The measurement model is not identified by itself.
- But it becomes identified if we add criterion variables:
  - Inputs: (background) variables that produce educational attainment, such as father's and mother's education and occupation.
  - Outputs: (demographic) variables that are produced by education, such as occupation and spouse's education.

Indirect effects model



Example: Germany in ESS

- Germany in ESS is an interesting case to look at:
  - German education is often claimed to be an extreme case of a divided system in which duration measures do not work.
  - In ESS, measurement of German education using EDULVL has gone wrong and leads to very odd results.

Education in ESS-DE

- In 2004 and 2006, the local education variable has 8 categories, that are effectively reduced to 2-3 categories when expressed in ISCED [EDULVL].
- In 2002 I have used a more detailed local education.
- The local information is optimally scaled using parental educations, occupations, respondent's occupation and spouse's education as criterion variables → EDUOPT.

Results ESS-DE

- Measurement relationships:
  - EDUC → EDULVL 0.75
  - EDUC → EDUYRS 0.85
  - EDUC → EDUOPT 0.91
- Structural relationships
  - EDUC → OCC 0.46 0.65
  - MEDUC ←→ FEDUC 0.38 0.72
  - FOCC → OCC 0.20 0.09

Results ESS-DE (1)

- Single indicator models for education all result in downwardly biased effects on and by education.
- The three-indicators model confirms that there is considerable loss of information when using ISCED [EDULVL].
- However, note that the loss in EDUOPT is still 9%.
- Also note that duration [EDDUR] is still a reasonable indicator of the true score, and better than EDULVL.

Results ESS-DE (2)

- If we restrict the model to two indicators of education (always involving EDDUR):
  - The point-estimates remain virtually unchanged.
  - This is also true when we combine EDDUR with EDULVL ("two bad measures").
- However, note the T-values and the loss of statistical power when dropping one indicator.

General conclusions

- The only way to improve measurement of background variables is collect multiple, parallel indicators.
- So: ask more questions, not better (=more complicated?) questions…!
- … and use multiple indicator models to estimate and correct random and – if possible – systematic measurement error.