THE COMPARATIVE MEASUREMENT OF LEVEL OF EDUCATION IN THE ISSP – AN ASSESSMENT AND APPLICATION OF THE ISLED SCALE

Heike Schröder & Harry B.G. Ganzeboom

VU University Amsterdam

This version: June 14 2012 [HG]

In cross-national survey research level of education tends to be measured with either a harmonized qualification measure or with a duration measure. The use of scaling, by contrast, is much less common. In this article we test whether the International Standard Level of Education [ISLED], a scale variable recently developed on data from the European Social Survey [ESS], produces adequate results when applied to the International Social Survey [ISSP]. For this purpose, we apply ISLED scores to the country-specific ISSP variables, which we first convert into ISCED-2011. Conceiving level of education as a latent variable with two measured indicators, we subsequently apply double indicator modeling in a simultaneous equation model. This allows us to assess the quality of the individual indicators as well as to correct random measurement error. We find that ISLED produce adequate results. Its measurement quality is slightly better than that of the ISSP harmonization and surpasses that of the ISSP duration measure by a considerable margin. ISLED measurement quality is, however, topped by double indicator modeling. We conclude that the ISLED can be readily applied to fresh data and holds the promise of becoming a truly standard international measure of level of education.

Introduction

The comparative measurement of level of education in surveys is wrought with problems. These problems concern the primary data collection as well as the need for a common metric. The pertinent source information is by definition country-specific and questions tend to contain a predefined answer format, in which the most representative current and historical educational programmes are listed. The choice of these programmes, however, is where the problem begins. As an exhaustive list of all such programmes is generally not feasible in the limited space of a questionnaire, a choice must be made on which programmes exactly are to be listed. Once this choice is made, the next difficulty concerns the classification of these programmes. Given our interest in the level of education obtained by an individual, such levels need to first be defined and each programme then needs to be allocated to one of these predefined levels.

Precisely which programmes represent the same level, however, is not always self-evident. The decision criteria used tend to be based on the institutional structure of education systems. While some of these criteria are clear and unequivocal, such as the temporal sequencing of programmes or that having passed a given programme is a prerequisite for entering another programme, such decisions are in essence an empirical matter. After all, programmes that are nominally categorized at the same level, may, lead to very different outcomes, such as occupations or incomes in the labor market. This becomes even clearer when historical programmes are to be lumped together with current programmes. Even if current educational programmes nominally resemble historical ones, they may be very different. In other words, no matter how practical they may be, institutional criteria are of limited use for the categorization of educational programmes and some ad-hoc decisions are usually unavoidable.

If it can already be problematic to establish a definitive national classification, the problem exacerbates when national categories are to be compared across countries due to the great structural differences between them. These differences pertain to the basic level structure, to the number and types of programmes discerned per level, to the length of programmes and to access requirements. Virtually the same problems we encounter when classifying national education systems resurface when we try and aggregate the national source classifications into a comprehensive international classification. The level structure may differ between countries and nominally equivalent programmes may lead to entirely different outcomes in different countries. Moreover, systems may differ so much that some programmes simply do not have any equivalent in another country.

There are two commonly used conventional methods to solve the comparability issue: common denominator harmonization and measurement by duration. The idea of harmonization is to look for common elements in the source classifications. Duration measures by contrast, look for a different common ground, namely that it takes a well-defined amount of time to pass through any education systems, irrespective of any structural differences between them.

In this paper we examine the measurement of education level in one of the world's leading academic comparative surveys, the International Social Survey Programme [ISSP], which has implemented both common denominator and duration measurement. We assess the measurement quality of the ISSP common denominator and country specific qualifications and duration measures exploiting the potential of a two-tiered measurement strategy developed by Schröder & Ganzeboom (2012a). This strategy, which was developed on data from the European Social Survey [ESS], consists on the one hand of the scaling of the country-specific [CS] educational categories into the International Standard Level of Education [ISLED] and on the other of double indicator modeling of education conceived. In particular, we test whether the newly developed ESS-bases ISLED scale as proposed by Schroder & Ganzeboom (2012a) can usefully be implemented on a fresh dataset and brings out or even surpasses the explained variance produced by the indigenous common denominator or duration measures.

In order to do so, we develop an optimal score for the CS variables of the ISSP 2009 dataset (that contains the pertinent information on social origins and destinations) and use this optimized scale as a benchmark to judge the other indicators against in a simultaneous equation model. For the time being, the optimal scale that is indigeneous to ISSP 2009, is refered to as OPTED. We then assign ESS-derived ISLED scores to the ISSP CS variables, resulting in a second, alternative scale. Combining these two scales in a double indicator models allows us to assess the measurement quality of ISLED and to compare it to that of the ISSP measures. Double indicator modeling, moreover, yields optimal measurement quality by means of error correction. As measurement quality is known to affect structural regression coefficients (Allison & Hauser, 1991), we finally examine the effects in a status attainment model, contrasting the results produced by the poorest single education indicator with those produced by double indicator modeling.

Conventional approaches to solving the comparability issue

In order to obtain an internationally valid educational classification, the comparability issue is resolved by a common denominator approach and national programmes are allocated to a given level on the basis of the features they have in common. A widely used and highly successful tool is the International Standard Classification of Education [ISCED], developed by UNESCO (1997). This layered classification is comprehensive and provides not only a well-defined classification structure but also extensive documentation on how to map (current) national education programmes onto the harmonized levels. While ISCED is a valuable descriptive tool, it has so far been applied to surveys in a much reduced coarse form, exploiting only the seven main categories, the ISCED main levels. While this reduction in categories is problematic in itself, the problem becomes worse because these categories are highly differentially represented in different countries, in some cases resulting in as few as 3-4 effective categories or when programmes are wrongly classified. Research has shown that this strategy is error-prone and that it may lead to a critically big loss of information (e.g. Schneider, 2009; Kerckhoff & Dylan,1999; Kerckhoff et al., 2002).

Another method of comparing national education systems is to abstract away from their level structure and the different qualifications the levels lead to, and to instead base the comparison on a simple feature that all education systems have in common, namely that it takes a certain well-defined amount of time to pass through a given educational programme. The obvious advantage of this approach is that a measurement based on programme duration has a natural immediately comparable intrinsic metric. Any programme length can be expressed in years and therefore comparability is unproblematic and does not require any mapping or conversion. Instead, a direct question format can be used in questionnaires that is ideally identical for all countries. This method too, however, has been much criticized for inadequately representing some systems types and for producing skewed results (e.g. Hout & DiPrete, 2006; Müller, 2008; Schneider, 2009).

When deriving a common denominator, international projects have increasingly turned to post-harmonization, meaning that the original national source information, which is often much more detailed than the harmonization, remains available in seperate data files as CS variables. This is now the case in the European Social Survey [ESS], the European Value Studies [EVS] and has been the cases for ISSP since its beginning.

It is striking that the CS information is very little, if ever used in comparative analysis. At best, the information is used to check the development and correctness of the common denominator. There appear to be several reasons why users have not accessed this information more often. First, the CS information is often hard to process, as it requires detailed understanding of national educational systems, including their historical developments. Second, the information can be (and should be!) in different languages, but translations are either not available or not meaningful. In fact, they can even be misleading. Abbreviations that are perfectly clear to an insider, across borders and across time episodes very soon lose their meaning. Third, using the CS information requires a cross-national metric, which is difficult to develop.

It is also striking that researchers, while ignoring the CS variables, usually choose *between* the two main methods of comparative measurement, and use either duration or a common denominator, but not both. Researchers seem to be convinced that these methods yield perfectly adequate and unproblematic variables and simply choose the one that is most customary in their field of research. While there is ample evidence that this is simply not true (cf. above) and that educational attainment indicators are just as error prone as any other questionnaire item, we are not aware of any research that has used a (weighted) average of the two indicators.

An integrated empirical approach to solving the comparability issue

In an attempt to readdress the comparative measurement problem, Schröder & Ganzeboom (2012a) have recently proposed a two-step strategy that demonstrably improves measurement quality. In the first step, they have developed a novel education measure labeled the International Standard Level of Education [ISLED], which exploits all the detail contained in the CS education variables. They developed ISLED by scaling all detailed CS educational categories in the ESS (R1-4) to a common metric, that optimizes the role education plays in a basic status attainment model (Blau & Duncan, 1967) for the social reproduction of social background into social destinations.

In a sequel article Schröder & Ganzeboom (2012b) have applied the ISLED methodology to the R5 variables. This time they not only scaled the CS variables but also two new harmonization variables introduced in ESS-R5. One of these harmonized variables, called EDULVLb, is much more detailed than previous harmonizations. Moreover, it is based on ISCED-2011, the latest revised version of ISCED (UNSD, 2011) that has recently been launched, and largely corresponds to its first three digits. This has resulted in ISLED-scores for ISCED-2011, that we list in Appendix A. As any educational category can be converted into ISCED-2011, we believe that this ISLED scores for ISCED-2011 have the potential of becoming a true standard international education variable, with wide applicability in other surveys. In this research, we use the ISSP to put this assumption to a first test.

In the second step, Schröder & Ganzeboom (2012a) have developed and implemented double indicator modeling of the education variable in a simultaneous equation model. Using ESS data, they combined two (independently measured) indicators in a simultaneous equation model. This procedure produces perfect (unattenuated) measurement in the sense that it corrects for random measurement error (provided the assumptions of the SEM measurement models are met). Moreover, a double indicator measurement models enables to actually assess the measurement quality of either indicator used. Here, we apply this strategy to the ISSP data and combine the ISSP duration measure with ESS-bases ISLED, the indigeneous measure OPTED and the ISSP harmonization variable (DEGREE) respectively.

The combination of measuring education level by means of ISLED scaling and of modeling it with double indicators has a number of important advantages. First, the measuring is empirical and avoids ad hoc decisions. Second, any national educational category can actually be ISLED-scaled, resulting in empirically derived national education hierarchies, with ISLED scores of individual programmes being directly comparable across countries. Third, the use of the ISCED-2011 based ISLED scores makes the ISLED easily referable to just any type of education variable found in surveys. Fourth, the method avoids any waste of information and maximally exploits the detail contained both in the cs source variables, which is in the modeling part complemented by the additional information contained in the duration measure.

Data

For our analyses we use the data of wave IV of the Social Inequality module of the International Social Survey Project of 2009. This module deals with subjective perceptions and evaluations of inequality and stratification in 38 countries, 26 of which are European and overlap with ESS countries for which an ISLED measure is available. The particular relevance of this ISSP module compared to others is that Social Inequality IV covers quite a bit of information on the stratification position of the family of origin, in particular father's and mother's occupation when respondent was young, which is not available in other ISSP modules, but is very important to establish the level of education according to the ISLED methodology. The ISSP also has an indicator for the respondent's occupation, which is a standard variable in all ISSP waves, but in this module has been extended by a measure of occupation at entry into the labor market. Altogether, this information allows us to examine the measurement of education in an Origin Education Destination (OED) model, which was also the frameword for generating the ISLED scale. Notice, however, that the stratification information in ISSP 2009 is decidedly poorer than the ESS standard background variables that were used to develop ISLED. In particular, ISSP does not contain measures of education for father, mother and spouse, which were ingredients of the ISLED development.

After eliminating non-European countries as well as those countries which do not fulfill the requirement of two independent measurements of education, we are left with a sample of 21 countries: Belgium, Bulgaria, Switzerland, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Great Britain, Croatia, Hungary, Israel, Netherlands[1], Norway, Portugal, Russia, Sweden, Slovakia and the Ukraine. After excluding people under 25 and over 74 years of age as well as students, we have an effect sample size of 25,999.

ISSP contains two education measures, a duration measure and a common denominator measure. ISSP's common denominator variable is called DEGREE.

[1] The data are the Netherlands are taken from the ISSP data for the modules on Leisure / Religion (2007/2008), as the Social Inequality IV data are not yet available for the Netherlands.

Unlike similar variables in other survey project, DEGREE is not formally defined by a reference to any detailed educational classification, such as ISCED. Instead, ISSP has chosen to harmonize its CS education variables into 6 categories, which give data producers considerable degrees of freedom to code CS categories. These categories are:

- 0 No formal qualification
- 1 Lowest formal qualification
- 2 Above lowest qualification
- 3 Higher secondary completed
- 4 Other qualification above higher secondary
- 5 University degree completed

It is clear that ISSP-researchers had a single hierarchy of education in mind when devising this question format.

While duration is a compulsory question in ISSP, there is no compulsory common question format. Since 2011 the recommended question format has been:

> How many years (full-time equivalents) have you been in formal education? *Include all primary and secondary schooling, university and other post-secondary education, and full-time vocational training, but do not include repeated years. If you are currently in education, count the number of years you have completed so far.*

In practice, however, we have found large variation in the question formats, which have been post-harmonized into the ISSP duration variable EDUCYRS.

An additional problem is that in some ISSP countries the duration question has not been asked independently at all, but is in fact a straight recode of CS qualifications question. We have found this practice for Germany, Austria and Slovenia (???) (as well as Argentina and China outside Europe). This is problematic because in such cases we cannot apply double indicator measurement in an SEM model to obtain unattenuated measurement because that requires independence of measurement (meaning that respondents have the opportunity to and indeed do make errors independently). We therefore have to exclude these countries from our analyses.

Method

We apply both the ISLED methodology and double indicator modeling to the ISSP 2009 data. Concerning ISLED, we do two things. First, we generate toptimal scores for the ISSP CS variables, in much the same way as was previously done for the ESS. This results in OPTED, which we will use as a benchmark for our comparisons. Second, we applied the ESS-based ISLED to the ISSP CS categories, which are for this purpose mapped onto ISCED-2011. We then apply double indicator modeling and combine the ISSP duration measure with OPTED, ISLED and the ISSP harmonization DEGREE respectively. In our analyses, we assess the measurement quality of the ISLED and test whether ISLED, when implemented in data that were not used for its derivation, can compete with or even outdo indigenous optimization, harmonization or duration measures.

Our strategy is to compare ISLED to the best possible scaling of the CS education categories in the ISSP, which we likewise develop with ISLED methodology. The basic procedure is that we standardize four criterion variables (father's occupation, mother's occupation, respondent's first and respondent's current occupation) within each of the 21 countries and then calculate an unweighted average for social origins (father and mother) and an unweighted average for destinations (respondent and spouse). Optimization then involves finding weight that produce a *minimal direct* effect and a *maximal indirect* effect of social origins in destination. Figure 1 shows the basic model for our scaling procedure. The respective national educational categories are interpreted as intervening between a number of input and a number of output variables in the stratification process.

<<< Figure 1>>>

Using these two index variables (restandardized to a common Z-metric) it is easy to find the particular weighted average of the two that happens to maximize the indirect effect, while minimizing the direct effect in the elementary OED model. Despite the criterion variables in ISSP being different from the ESS, the algorithm finds the optimizing weights at about the same as Schroder & Ganzeboomn (2012a, 2012b) for the ESS, 0.60 for origins and 0.40 for destinations.

It may be important to emphasize the difference between the OPTED scale in ISSP 2009 and the ISLED scale developed on ESS. As there are no other educations among the criterion variables in ISSP, the optimal scores can be biased towards occupations. While we think scaling by occupations and scaling by other educations yield very similar results, it is possible that the optimized score produces association that are closer to occupations than to other educations. Then of course, the scores are optimized with respect to the dataset that we use for validation, which will also inflate the associations. Altogether, this means that we use the OPTED scores as a benchmark and do not present it as an alternative to ISLED.

In order to be able to assess the measurement quality of the various education indicators, following classic measurement theory (Bentler, 1980), we model education as a latent variable in a double indicator measurement model, which is measured with two independent indicators, illustrated in Figure 2.

<<< Figure 2 >>>

Using Full Information Maximum Likelihood in LISREL 8.8 (Jöreskog and Sörbom, 1996), we estimate three simultaneous equation models (SEM), in which one of the education indicators is the duration variable, while the other alternates the three qualification measures. This yields measurement coefficients for each indicator. The measurement model is embedded in a larger structural model consisting of three equations, with education, first occupation and current occupations as dependent variables. Figure 3 depicts the full model.

<<< Figure 3 >>>

To bring out the effect of different educational measurement procedure, we compare two different models, contrasting the weakest (duration as single indicator) and the strongest (double-indicators) possible model in order to be able to compare the impact of measurement quality on structural coefficients in the model and the difference in explained variance associated with it. As all coefficients are completely standardized, they are directly comparable with one another.

Results

Table 1 shows the results of the three different measurement models. In each model we combined the duration measure EDUCYRS with one of the qualification variables: in model 1 with the ISSP-generated OPTED, in model 2 with ISLED and in model 3 with the ISSP common denominator variable DEGREE, using a linear scaling. We can assess the measurement quality of the individual indicators by comparing their measurement coefficients (factor loadings). As 1 is the benchmark that would indicates perfect measurement, the difference to 1 signifies the amount of information we lose using the indicator in question in percentage points.

<<< Table 1 >>>

We see that on average across all countries (XX), the ISSP-generated OPTED performs best, closely followed by the ISLED and also the DEGREE variable. The duration measure, by contrast, fares noticeably worse. The measurement coefficients provide an indication of the degree of attenuation each indicator causes:

- OPTED: 0.951 = 4.9% attenuation
- ISLED: 0.941 = 5.9% attenuation
- DEGREE: 0.936 = 6.4% attenuation
- EDUCYRS: 0.867 = 13.3% attenuation

A coefficient of 0.941 for ISLED means that it can be expected that any association with education (in particular when measured by a correlation or regression coefficient) is attenuated by 5.9%. For duration this means that we are losing as much as 13.3% of the variation if we use it as our exclusive indicator. While ISLED causes appreciably less attenuation, we must remember that these estimates can only be made when a second independent measure is available in the data, as imperfect as it may be and that therefore the duration measure is indispensable. The consequences of disattenuated measurement can be seen in the estimated structural coefficients in the model (cf. below).

The overall quality of the education variables is also reflected in the standard deviations of the distribution across countries of their measurement coefficients, which increases with declining measurement quality, meaning that both OPTED and ISLED are more stable in quality across countries than DEGREE or EDUCYRS. The standard deviation for the duration measure is more than double that of the ISLED's.

The results, however, differ per country. In some countries, for example Switzerland, the Netherlands and Norway, we find the same regular pattern as for the cross-country average, with OPTED being better than ISLED and ISLED better than DEGREE. In other countries, for example in Cyprus, the Czech Republic, Israel and Finland, ISLED outperforms OPTED and in the Czech Republic, Croatia, Israel and Sweden the DEGREE variable turns out to work best. Table 1 also shows that the quality of EDUCYRS fluctuates depending on the quality of the second indicator. Our results suggest that the model overestimates its quality when the second indicator appears to be weaker, compensating as it were for the weakness of that second indicator.

Now that we have assessed the quality of the indicators, we illustrate the effect measurement quality has on the structural regression coefficients in an intergenerational status attainment model. Table 2 shows all these effects, to wit the effect of father's and mother's occupation on respondent's education, as well as on their first and current occupation, the effect of education on first and current occupations and, finally, the effect of first on current occupation. Per country, we contrast two sets of effects and the related explained variances in the dependent variables. In the respective top row, education is measured with the duration variable EDUCYRS, the poorest measurement. In the respective bottom row, education is modeled with double indicators, EDUCYRS and ISLED. We observe a clear and ubiquitous pattern that holds in all countries as well as in the cross-country pooled data, namely that all direct effects of inputs on outputs diminish with better measurement, while all indirect effects that run via education increase.

<<< Table 2 >>>

Indirect effects are the parental effects on education and the effect of education on current occupation. We see that all these effects are severely attenuated if the duration measure is used and are much larger with double indicator modeling. Cross-nationally, we observe a large increase by 0.12 points in the effect of education on first occupation, mirroring as it were the likewise large decrease in the direct effect of first occupation on current occupation. The increase in the parental effects on education are with roughly 0.03 points somewhat less marked. The changes in effect size, incidentally, illustrate that the attenuation caused by EDUCYRS is indeed well above 10%. It does, however, fluctuate with the effect type and is much larger for the effect of education on first occupation than for parental occupations on education.

Direct effects are all parental effects on first and current occupation, as well as the effect of first on current occupation. Here we see the reverse picture, namely that the parental effects are virtually halved with double indicator modeling and that in a number of cases the effects become insignificant. In eight countries this is the case for the effect of father's occupation on first occupation and in seven countries for the effect of mother's occupation on first occupation. In another five countries, the effect of father's occupation on respondent's current occupation become insignificant and in two countries, the effect of mother's occupation on respondent's occupation becomes insignificant. The effects of mother's occupation on current occupation even becomes insignificant for the pooled data as well.

In line with the changes in effect size, the explained variance in all the three dependent variables, education, first occupation and current occupation, increases with double measurement of education. In the education variable it increases with 5.4%, in the first occupation with 9% and in the current occupation with 3.5%. Particularly striking is the large increase in explained variance in the first occupation variable. This corresponds to an effect size that on average decreases by almost an entire point, when education is modeled with double indicators. This effect, in other words, can be greatly overestimated when instead EDUCYRS is used as a single indicator. How large the difference is, depends on the individual country. With 2.3 points it is largest in Finland, while with less than 0.2 points it is smallest in Portugal.

Conclusions and Discussion

In this article we have assessed the measurement quality yielded by four different methods of measuring level of education: harmonization, duration, scaling and double indicator measurement, as well as their effects on the structural coefficients in an intergenerational status attainment model. Using ISSP data from 2009, the fourth wave of the Social Inequality module, we compared the ISSP duration and harmonized measures with the International Standard Level of Education [ISLED] and with double indicator measurement.

Our approach involves two consecutive steps. In the first step we derived two alternative sets of ISLED scores. One set we derived by applying the ISLED-methodology to the ISSP country-specific qualification measures. Another set we derived by applying ISLED-scores that were developed by Schröder & Ganzeboom (2012b) on ESS data to the ISSP CS variables. In particular, we mapped all country-specific ISSP-categories onto ISCED-2011 (three-digits) and assigned them the appropriate ISLED scores associated with the respective level. Using indigeneous optimization as a benchmark, we show that ISLED comes remarkably close, with a difference in measurement quality of only one percentage point. This illustrates that ISLED-scores can be readily applied to fresh data and produce adequate results.

It is important to note that ISCED-2011 greatly facilitates the application of ISLED in new data. The alternative, namely the matching of the ISSP CS-measures with the ESS CS-measures and then assigning the appropriate ISLED scores, is very cumbersome, time-consuming and error-prone. In fact what would be needed for that is expert knowledge on the various national education systems as well as some arithmetic in applying weights in cases where ISSP categories do not have a direct equivalent, but are represented by several different ESS categories. Given that ISCED provides country-mappings[2], this method is much easier and requires no extra knowledge or arithmetic.

In the second step we combined each of the qualification scales with the duration measure in three double indicator models. This allowed us to compare the quality of the individual indicators as well as to optimize measurement quality by correcting the measurement error contained in each indicator. Double indicator modeling yields the best measurement quality and unattenuated regression coefficients. The best single indicator is the ISSP-derived OPTED, but closely followed by the ESS-applied ISLED. Remarkably, the ISSP harmonization DEGREE is only marginally inferior, leading us to conclude that this six-category harmonization, works much better than its ISCED-based ESS equivalent (EDULVLa). The latter entails a much greater loss in information. We tentatively suggest that this may be due to the much less strict coding regulations that ISSP imposes, which leaves researchers some more discretion and allows them to group the appropriate national education categories together. Like in the ESS, the duration variable turns out to be the poorest indicator. Moreover, its quality is overestimated to the degree that the qualification indicator is weaker. In any case, it entails an additional loss in measurement quality of about 8%.

The way we measure education level in an intergenerational status attainment model has a clear impact on the structural coefficients in the model. The better the measurement quality, the larger the indirect effects that run via education and by the same token, the smaller the direct effects of inputs on outputs. When education level is optimally measured with two indicators, in a number of countries the effects of parental occupations on respondent's first and current occupations turn out to be non-significant. The effect of father's occupation on respondent's current occupation becomes insignificant even for the pooled data. The effect we find when we measure education with the duration measure instead, must be attributed entirely to bad measurement. The other three parental effects on respondent's occupation remain significant for the pooled data and must therefore be regarded genuine effects.

[2] So far these country-mappings are only available for ISCED-97, the predecessor of ISCED-2011. Due to the amount of detail they contain, they are, however, sufficient for our purposes.

There are some interesting parallels between our results on the European countries in ISSP and the ESS as analyzed by Schröder & Ganzeboom (2012a). First, we find virtually the same measurement coefficients for the optimized OPTED in ISSP as in ESS, which is rather astonishing given that different variables were used[3]. Moreover, the ESS bases ISLED applied to the ISSP CS-variables deviates with only one percentage point. This is clear evidence that the ISLED scale can be readily applied to fresh data. Second, we find ISLED to be only marginally superior to the common denominator variable. Even the optimized scale is merely 1.5% better than DEGREE. The ESS common denominator EDULVLa by comparison, does decidedly worse than ISLED. The difference here is 8% for R1-4 and 6% for R5. Finally, we find that the ISSP duration measure performs marginally better than the one in ESS. Given that the ESS has implemented its comparative question format with much more rigor than ISSP, and requires countries to use the exact same question wordings, this is surprising.

The comparatively high quality of both the ISSP harmonization and of its duration variable are in fact somewhat puzzling. We would conject that it may be precisely be due to the lower level of rigor in ISSP that allows its measurement to be more adaptive to local circumstances. Thus ISSP researchers appear to have taken the liberty to interpret DEGREE categories to suit differences in education level that are specific to their national systems. In ESS data producers have to live by the book (the ISCED manual) or otherwise someone will interfere and correct their harmonization steps.

Something similar may be going on in the measurement by duration. While ESS makes all respondents go through the same arithmetic (a big no-no in questionnaire methodology books), in ISSP researchers can choose a locally appropriate formulation. If true, both results would speak in favor of functional equivalence as a comparative measurement principle. We would caution the ISSP not to follow ESS in rigor in this respect.

[3] For the ESS Schröder & Ganzeboom (2012a, 2012b) used educations of parents and spouse, but did not use first occupation for the derivation of ISLED.

Our experience with these data leads us to another recommendation, however, with respect to coding and archiving. In ESS, as of R5, the CS education variables have to be coded in ISCED 2011, which allows for 26 different categories and almost always exhausts the national classification. In ESS-R5, CS education measures typically contain around 15 categories and the three digit ISCED 2011 is sufficient to code this. If CS data are presented by an internationally documented scheme, this would bring much clarity in the CS variables, which is now obscured by abbreviations, local languages, odd translations, and non-romanized scripts. Implementing ISCED-2011 in its three digit version would bring the measurement of education level much closer to that of occupation, in which the detailed ISCO classification has been successfully used for similar purposes.

References

Allison, P. D. and R. M. Hauser (1991). 'Reducing Bias in Estimates of Linear Models by Remeasuring a Random Subsample.' *Social Methods & Research* 19: 466-492.

Bentler, P. M. (1980). Multivariate Analysis with Latent Variables: Causal Modeling, *Annual Review of Psychology*, 31, 419-56.

Blau, P. and Duncan, O.D. (1967) *The American Occupational Structure*. New York: Wiley.

Ganzeboom, H.B.G., de Graaf, P.M. & Treiman, D.J. (1992) 'A Standard International Socio-Economic Index of Occupational Status', *Social Science Research,* 21: 1-56.

Hout, M. & DiPrete, T.A. (2006) 'What we Have Learned: RC28's Contributions to Knowledge about Social Stratification', *Research in Social Stratification and Mobility,* 24(1): 1-20.

UNESCO (1997) ISCED-97 Manual http://www.uis.unesco.org/Library/Documents/isced97-en.pdf

ISLED (2012) Syntax: http://home.fsw.vu.nl/hbg.ganzeboom/isled.htm.

ISLED (2012) REFERENCE TO BE ADDED.

Jöreskog, K. G. and D. Sörbom, (1996). *LISREL 8 User's Reference Guide*. SSII: Lincolnwood.

Kerckhoff, A. C., Ezell, E. D., & Brown, J. S. (2002). Toward an Improved Measure of Educational Attainment in Social Stratification Research. *Social Science Research, 31*, 99-123.

Kerckhoff, A.C. and Dylan, M. (1999) 'Problems with International Measures of Education', *The Journal of Socio-Economics*, 28: 759-775.

Müller, W. (2008). Foreword. In Schneider, S. L. (Ed), *The International Standard Classification of Education (ISCED-97)*. Mannheimer Zentrum für Europäische Sozialforschung.

Schneider, S.L. (2009) *Confusing credentials: The Cross-nationally Comparable Measurement of Educational attainment.* Dissertation, Nuffield College Oxford.

Schröder & Ganzeboom (2012a) 'Measuring and Modelling Education Levels in European Societies', Manuscript under review at the *European Sociological Review*.

Schröder & Ganzeboom (2012b) ‘Lessons from the ESS education measures: exploring the potential of ISCED-2011 to develop a new international standard indicator of level of education’, unpublished manuscript.

UNSD Classifications Newsletter (2011) http://www.uis.unesco.org/Education/Documents/UNSD_newsletter_27e_ISCED.pdf

Appendix A: ISLED-scores for ISCED-2011

ISCED		ISLED
0	**Less than primary**	**17.3**
10	never attended an educational programme	17.3
20	some early childhood education	17.3
30	some primary education (without level completion)	17.3
100	**Primary**	**19.3**
100	Primary	19.3
200	**Lower secondary**	**31.0**
240	Lower secondary general	30.2
242	partial level completion and without direct access to upper secondary	29.9
243	level completion, without direct access to upper secondary	30.3
244	level completion, with direct access to upper secondary	30.5
250	Lower secondary vocational	31.8
252	partial level completion and without direct access to upper secondary	34.0
253	level completion, without direct access to upper secondary	29.6
254	level completion, with direct access to upper secondary	--
300	**Upper secondary**	**47.8**
340	Upper secondary general	41.4
342	partial level completion and without direct access to tertiary	39.9
343	level completion, without direct access to tertiary	40.4
344	level completion, with direct access to tertiary	44.0
350	Upper secondary vocational	54.2
352	partial level completion and without direct access to tertiary	--
353	level completion, without direct access to tertiary	53.0
354	level completion, with direct access to tertiary	55.4
400	**0 Post-secondary non-tertiary**	**55.4**
440	Post-secondary non-tertiary general	54.2
443	level completion, without direct access to tertiary	56.6
444	level completion, with direct access to tertiary	51.8
450	Post-secondary non-tertiary vocational	56.6
453	level completion, without direct access to tertiary	57.7
454	level completion, with direct access to tertiary	55.4
500	**Tertiary Short-cycle**	**63.5**
540	Tertiary Short-cycle general	69.6
550	Tertiary Short-cycle vocational	57.4
560	Tertiary Short-cycle orientation unspecified	63.5
600	**Bachelor or equivalent**	**74.2**
640	Academic	78.0
650	Professional	70.4
660	orientation unspecified	74.2
700	**Master or equivalent**	**80.9**
740	Academic	83.3
750	Professional	78.4
760	orientation unspecified	80.9
800	**Doctoral or equivalent**	**90.4**
840	Academic	90.4
850	Professional	90.4
860	orientation unspecified	90.4

Source: Schröder & Ganzeboom (2012b, Table 6).

Figure 1: Measuring education levels: an optimal scaling procedure

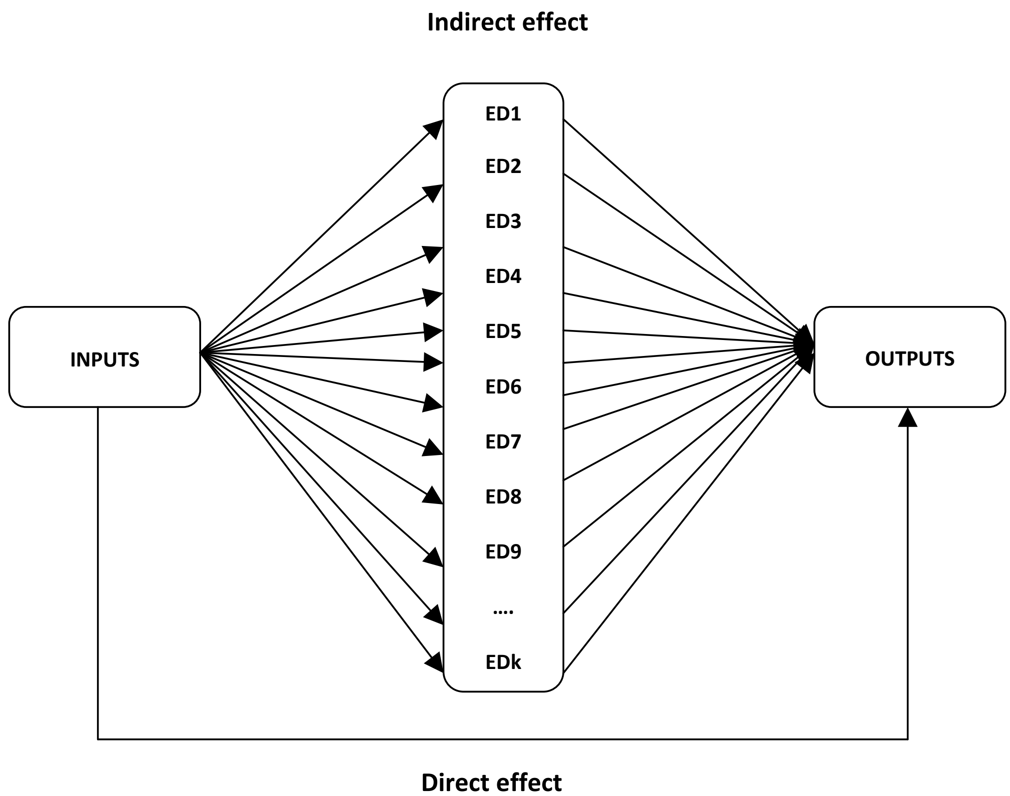


Figure 2: The measurement model

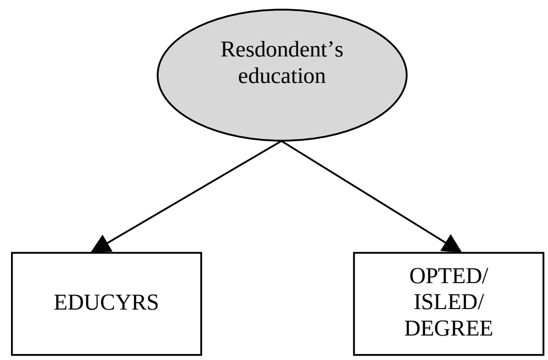
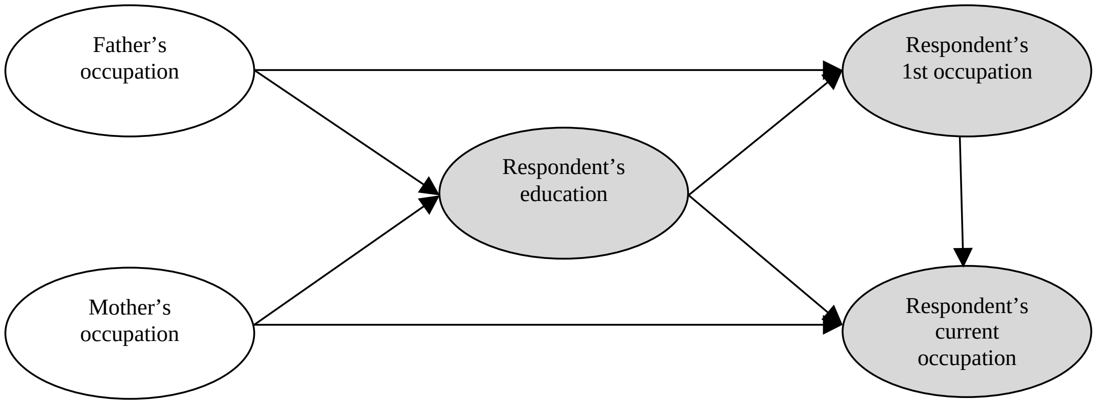


Figure 3: The structural model



Note: dependent variables are marked with grey shading

Table 1: Measurement coefficients (factor loadings) of education measures
N= 25,999

	Model 1		**Model 2**		**Model 3**	
Country	**OPTED**	**EDUCYRS**	**ISLED**	**EDUCYRS**	**DEGREE**	**EDUCYRS**
BE	0.974	0.869	0.977	0.884	0.963	0.924
BG	0.928	0.907	0.924	0.936	0.837	0.981
CH	0.950	0.811	0.941	0.803	0.900	0.785
CY	0.967	0.951	0.982	0.965	0.986	0.980
CZ	0.938	0.814	0.980	0.834	0.984	0.854
DK	0.871	0.622	0.830	0.642	0.841	0.650
EE	0.945	0.874	0.943	0.885	0.928	0.885
ES	1.009	0.774	1.000	0.784	0.998	0.790
FI	0.909	0.661	0.917	0.683	0.886	0.709
FR	0.964	0.875	0.939	0.880	0.938	0.900
GB	0.892	0.790	0.896	0.804	0.872	0.787
HR	0.957	0.910	0.948	0.926	0.960	0.930
HU	0.982	0.896	0.967	0.891	0.982	0.930
IL	0.917	0.894	0.931	0.916	0.929	0.920
NL	0.950	0.796	0.943	0.796	0.939	0.789
NO	0.933	0.721	0.894	0.737	0.821	0.764
PT	0.980	0.950	0.963	0.933	0.975	0.940
RU	0.967	0.921	0.902	0.935	0.946	0.927
SE	0.929	0.807	0.927	0.826	0.936	0.832
SK	0.960	0.838	0.984	0.876	0.983	0.890
UA	0.977	0.878	0.977	0.895	0.956	0.898
XX	**0.951**	**0.848**	**0.941**	**0.857**	**0.936**	**0.867**
SD	0.033	0.088	0.040	0.086	0.053	0.089

Note: XX = cross-country averages
OPED = Optimal scores generated on ISSP-data
ISLED = ESS based ISLED scores applied to country-specific ISSP education variable
DEGREE = harmonized 6-category ISSP variable
EDUCYRS = ISSP duration measure

Table 2: The worst and best models: Structural effects in an intergenerational status attainment model, with EDUCYRS and double indicators as measures of level of education (completely standardized coefficients)

Country	N	Education Indicator	FOCC-EDU	MOCC-EDU	FOCC-OCC1	MOCC-OCC1	EDU-OCC1	FOCC-OCC	MOCC-OCC	EDUC-OCC	OCC1-OCC	R^2 in EDU	R^2 in OCC1	R^2 in OCC
BE	900	EDUCYRS	0.367	0.164	0.102	0.082#	0.502	0.057#	0.031#	0.276	0.450	0.221	0.352	0.475
		Double-In	0.380	0.189	0.049#	0.048#	0.616	0.024#	0.008#	0.444	0.350	0.251	0.438	0.526
BG	731	EDUCYRS	0.272	0.307	0.139	0.026#	0.522	0.020#	0.036#	0.364	0.494	0.271	0.337	0.634
		Double-In	0.289	0.331	0.095	-0.031#	0.652	0.011#	0.009#	0.461	0.442	0.310	0.471	0.654
CH	1014	EDUCYRS	0.368	0.085	0.217	0.053#	0.484	0.052#	0.038#	0.245	0.507	0.173	0.395	0.519
		Double-In	0.464	0.091	0.097	0.036#	0.644	-0.017#	0.028#	0.478	0.323	0.265	0.507	0.586
CY	835	EDUCYRS	0.170	0.408	0.173	0.093#	0.527	0.040#	0.018#	0.348	0.568	0.289	0.471	0.755
		Double-In	0.222	0.399	0.137	0.080#	0.569	0.021#	0.011#	0.392	0.551	0.328	0.492	0.762
CZ	1,012	EDUCYRS	0.181	0.227	0.083	0.072	0.585	0.082	0.042#	0.187	0.614	0.120	0.412	0.624
		Double-In	0.279	0.263	-0.024#	0.009#	0.757	0.032#	0.012#	0.408	0.476	0.210	0.564	0.668
DK	1139	EDUCYRS	0.180	0.115	0.139	0.114	0.350	0.071	-0.012#	0.168	0.522	0.064	0.207	0.397
		Double-In	0.278	0.176	0.019#	0.041#	0.650	0.032#	-0.031#	0.363	0.389	0.152	0.449	0.442
EE	791	EDUCYRS	0.164	0.251	0.040#	0.121	0.455	0.011#	0.043#	0.381	0.315	0.120	0.271	0.388
		Double-In	0.195	0.246	0.007#	0.100	0.551	-0.007#	-0.038#	0.470	0.246	0.134	0.352	0.424
ES	715	EDUCYRS	0.199	0.254	--	--	--	0.241	0.096#	0.345	--	0.167	--	0.299
		Double-In	0.268	0.262	--	--	--	0.147	0.055#	0.546	--	0.228	--	0.428
FI	711	EDUCYRS	0.193	0.163	0.076#	0.190	0.406	0.028#	0.018#	0.165	0.610	0.099	0.282	0.520
		Double-In	0.224	0.329	-0.002#	0.021#	0.708	-0.006#	-0.041#	0.418	0.436	0.241	0.514	0.572
FR	2,303	EDUCYRS	0.293	0.192	0.111	0.081	0.492	0.098	0.041#	0.148	0.532	0.177	0.339	0.473
		Double-In	0.344	0.173	0.055	0.073	0.586	0.069	0.037#	0.262	0.468	0.206	0.412	0.496
GB	611	EDUCYRS	0.168	0.221	0.151	0.006#	0.431	0.107	-0.007#	0.322	0.245	0.107	0.244	0.280
		Double-In	0.257	0.270	0.093#	-0.041#	0.522	0.046#	-0.062#	0.531	0.155	0.194	0.299	0.380
HR	834	EDUCYRS	0.219	0.311	0.059	0.004#	0.567	0.018	0.015	0.238	0.628	0.215	0.352	0.646
		Double-In	0.217	0.337	0.043#	-0.037#	0.643	0.043#	-0.037#	0.334	0.572	0.236	0.416	0.667
HU	885	EDUCYRS	0.197	0.329	0.170	0.118	0.509	0.016#	0.025#	0.184	0.727	0.224	0.448	0.766
		Double-In	0.284	0.343	0.096	0.075#	0.612	-0.005#	0.012#	0.235	0.696	0.315	0.503	0.772
IL	965	EDUCYRS	0.205	0.222	0.116	0.095	0.433	0.077	0.005#	0.341	0.377	0.132	0.277	0.428
		Double-In	0.200	0.334	0.107	0.022#	0.494	0.069	-0.048#	0.412	0.349	0.212	0.305	0.450
NL	2,311	EDUCYRS	0.219	0.164	0.149	0.121	0.362	0.038#	0.054	0.253	0.399	0.114	0.245	0.355
		Double-In	0.261	0.203	0.097	0.078	0.507	0.015#	0.038#	0.367	0.311	0.165	0.344	0.391
NO	1,060	EDUCYRS	0.201	0.179	0.220	0.073#	0.317	0.056#	-0.016#	0.103	0.533	0.094	0.211	0.359
		Double-In	0.296	0.219	0.143	0.010#	0.439	0.044#	-0.027#	0.153	0.493	0.174	0.319	0.370
PT	834	EDUCYRS	0.371	0.233	0.126	0.099	0.554	0.101	-0.009#	0.271	0.503	0.293	0.468	0.586
		Double-In	0.357	0.257	0.122	0.079#	0.582	0.094	-0.025#	0.339	0.435	0.300	0.487	0.604
RU	1,266	EDUCYRS	0.145	0.316	0.158	0.075#	0.486	0.016#	0.038#	0.218	0.543	0.165	0.352	0.506
		Double-In	0.182	0.328	0.130	0.054#	0.537	0.006#	0.027#	0.263	0.518	0.199	0.387	0.516
SE	934	EDUCYRS	0.279	0.166	0.104	0.104	0.316	0.101	0.007#	0.307	0.320	0.144	0.172	0.323
		Double-In	0.329	0.204	0.059#	0.076	0.407	0.042#	-0.029#	0.484	0.250	0.205	0.219	0.409
SK	887	EDUCYRS	0.183	0.277	0.113	0.028	0.546	0.057#	0.052#	0.204	0.574	0.145	0.358	0.562
		Double-In	0.230	0.323	0.053#	0.044#	0.695	0.037#	0.015#	0.350	0.493	0.208	0.488	0.591
UA	1,715	EDUCYRS	0.122	0.331	0.084	0.161	0.482	0.028#	0.042#	0.430	0.353	0.164	0.362	0.524
		Double-In	0.181	0.339	0.033#	0.118	0.598	-0.011#	0.030#	0.571	0.256	0.208	0.450	0.576
XX	**25,999**	EDUCYRS	**0.231**	**0.227**	**0.126**	**0.091**	**0.462**	**0.057**	**0.029**	**0.253**	**0.491**	**0.155**	**0.317**	**0.485**
		Double-In	**0.276**	**0.256**	**0.071**	**0.046**	**0.586**	**0.031**	**0.007#**	**0.376**	**0.410**	**0.209**	**0.407**	**0.520**

Note: XX = cross-country averages
Double indicators: ISLED & EDUYRS
: statistically not significant: t < 2.0.