FORS PAPERS
Working paper series

Entropy Measures of Social Mobility: The Example of the Intergenerational Transmission of Education

Georg P. Mueller

University of Fribourg, Faculty of Economics and Social Sciences

FORS Working Paper 03-2021

How to cite this document:

Mueller, G. P. (2021). Entropy measures of social mobility: The example of the intergenerational transmission of education. FORS Working Paper Series, paper 2021-3. Lausanne: FORS. DOI: 10.24440/FWP-2021-00003

ISSN 1663-523x (online)

FORS, c/o University of Lausanne, Géopolis, 1015 Lausanne, Switzerland, E-mail: paperseries@fors.unil.ch

SUMMARY

Research about intergenerational mobility is traditionally interested in the relation between the class of origin, i.e. the status of the parents, and the class of destination, i.e. the status of their children. In the present article, this research interest is tackled with concepts from information theory like entropy, mutual information, etc. They allow to measure the information gain from knowledge of the status of the parents with regard to the status of their children. Moreover, the proposed methodology allows to address also the reverse question: how much information is in the current status of the children with regard to their social origin? Traditional indicators like chi-squares or phi-coefficients of the related mobility tables cannot distinguish between the two questions, which obviously refer to rather different social phenomena: in the first case the focus is on status inheritance from parents to children, whereas in the second case the research problem is the social exclusion of newcomers from other social strata.

In the second part of this article, the proposed methodology is used for an analysis of the intergenerational transmission of primary, secondary, and tertiary education in sixteen European countries. Based on interview data of the European Values Study it turns out that higher education has a high degree of intergenerational status inheritance but is non-exclusive with regard to newcomers from other strata. The situation of lower education is rather the reverse: it is exclusive with regard to newcomers but displays no intergenerational status conservation. The author argues that this is the result of the secular expansion of medium and higher education in Europe.

Keywords: Intergenerational mobility, inequality, entropy, synentropy (mutual information), education, international comparisons

1. INTRODUCTION AND OVERVIEW

Entropy is a concept from statistical information theory (Mathar, 1996, chap. 3.1), which describes the *uncertainty* about the attribute value of an element randomly taken from a sample with a known statistical distribution: the more heterogeneous the attribute values of the sample, the higher the entropy of the attribute. Thus, for attributes referring to social privileges, entropy is a measure of inequality (Coulter, 1989, chap. 5; Hao & Naiman, 2010, pp. 37 ff.; Mueller, 2004, 2017): if everyone has the same attribute value, entropy and inequality are both equal to zero.

Apart from uncertainty, entropy has also a second meaning, which relates the concept to *information.* The higher the entropy of a privilege distribution, the less we know about the particular life situation of a randomly chosen person. Our statistical information about the privileges of the person may however change by our knowledge of supplementary attributes of the person like age, race, or gender. Information theory coined for this purpose the concept of *conditional entropy,* where the condition is the value of a supplementary attribute like e.g. male gender. By analysing the difference between conditional and unconditional entropy it is possible to determine the *information gain* from knowledge of the supplementary attribute: in many societies male gender is e.g. a source of information about earning a higher salary.

Research about *intergenerational mobility* is traditionally interested in the relation between the class of origin as given by the status of the parents and the class of destination of their children (Breen, 2004a, p. 3). This research interest can obviously be translated into a question about the information gain from knowledge of the status of the parents with regard to the status of their children. However, the methodological approach proposed in this article allows to address also the *reverse* question: how much information is in the current status of the children with regard to their origin? More traditional indicators like chi-squares or phi-coefficients of the related mobility tables (Hout, 1983, chap. 1) cannot distinguish between the two questions, which address indeed different and rather independent social phenomena. In the first case of the information gains from the status of the parents the focus is on the *status inheritance* from parents to children, whereas in the second case of the information gains from the status of the children the focus is on the *social exclusion of newcomers* from other strata. Moreover, the idea of information gains is more *intuitive* than the previously mentioned classical measures.

Information gains from the status of the parents with regard to the status of their offspring point to strata-related inequalities of mobility chances. This article shows that the weighted *mean* of these information gains corresponds to the *synentropy (mutual information)* concept used in abstract information theory (Mathar, 1996, p. 28). Hence, the information theoretic approach proposed in this article allows to directly compare the inequality of mobility chances with the total inequality of the status-related privileges, mentioned in the first paragraph of this section: both concepts have highly commensurable scales based on the *bit,* the standard unit of information theory.

In order to show the applicability of the previously mentioned entropy concepts in mobility research, the author uses the European Values Study (2008) for comparative analyses of the *intergenerational transmission* of educational status in 16 European countries. I.e., the focus of this investigation is on the "relationship between parents' education and their children's educational achievements" (van Doorn et. al., 2011, p. 94). According to the O(rigin) – E(ducation) – D(estination) triangle (Breen & Müller, 2020b, Fig. 1.1; Breen & Luijkx, 2004, pp. 392–394; Hauser & Featherman, 1977, Fig. 1.2), this association has major implications for general class mobility. If the intergenerational conservation of the educational status is strong, as predicted by the inheritance of cultural capital (Powell et al., 2004, pp. 116–117), general class mobility is expected to be weak. If the intergenerational conservation of education is weak, e.g. due to educational policies leading to an expansion of the educational system (OECD, 1996), education becomes a "vehicle" of general mobility (Pollak & Müller, 2020). Thus educational mobility matters also for general (occupational) mobility.

The high international standardization of the data source of this study allows to compare countries with regard to status conservation and the exclusiveness of their primary, secondary, and tertiary education. As a major and not so unexpected empirical result, the decline of primary schooling as the highest educational attainment has important consequences for *structural mobility* (Boudon, 1973, pp. 17 ff.): it allows the children of parents with low education to move up to medium and high levels of education.

2. INFORMATION THEORETIC INDICATORS OF INEQUALITY AND MOBILITY [1]

Shannon and Weaver (1962) defined *entropy* as the average number of decisions in a binary search tree, which are required in order to identify an element of a non-continuous random variable X with a finite number of *n possible values* x_i and the related probabilities prob(X = x_i) = p_i. This intuitive conception of uncertainty leads to the following definition of the *entropy* of a variable X:

$$H(X) = - \Sigma_{i=1,..,n} [prob(X=x_i) * \log_2(prob(X=x_i))] \qquad (1a)$$

(see Theil, 1967, pp. 24). According to Theil (1967, p. 26), H(X) varies between 0 and $\log_2(n)$. Consequently, the *normalised entropy* with a minimum 0 and a maximum 1 is defined [2] as

$$|H|(X) = H(X) / \log_2(n) \qquad (1b)$$

|H|(X) = 1, if X has a rectangular statistical distribution and all possible values of X have the same probability 1/n. To the contrary, |H|(X) converges to *zero,* if the values of X are more and more concentrated in one single category.

Tab. 1 illustrates the varying values of |H|(X) for different stratifications of a variable X with three privilege levels *low, medium,* and *high:* the type 1 stratification is highly unequal with a standardised entropy |H|(X) = 1, as compared to the highly equal type 3a,b,c stratifications, where |H|(X) = 0. Between these extremes is the type 2 inequality with an intermediate value |H|(X) = 0.790. Thus Tab. 1 justifies the use of the normalised entropy |H|(X) as a measure of the *inequality* of the privilege X.

[1] For a better understanding of the formulas of this and the following sections, please refer to the *glossary* in the appendix of the article.

[2] The *standardised* entropy |H|(X) (see formula (1b)) must not be confused with the *absolute* entropy |H(X)|.

Table 1. Five exemplary statistical distributions with different inequalities and entropies H(X) and |H|(X).

Stratification	Type 1	Type 2	Type 3a	Type 3b	Type 3c
X = low	1/3	1/6	0	3/3	0
X = medium	1/3	2/3	3/3	0	0
X = high	1/3	1/6	0	0	3/3
		real-world			
Total	1	1	1	1	1
	Equality of opportunity				equality of outcome
Entropy H(X)	1.585	1.252	0.0	0.0	0.0
Entropy |H|(X)	1.000	0.790	0.0	0.0	0.0

Due to its scientific roots, |H|(X) is also a (reverse) measure of *information* about the privilege X. In Tab. 1 little is known about a person randomly chosen from a type 1 stratification. For the case of a type 2 distribution, a representative person is more likely to belong to category X = medium than to any other category. For the stratifications type 3a, 3b, or 3c of Tab. 1, the identification of the privilege of just *one* exemplary person supplies full information about *all other* individuals.

Table 2a. Information gains $|G|(X|Y=y_i)$ and $|G|(Y|X=x_i)$ with regard to X and Y for an exemplary case with completely absent status inheritance.

Strata	y_i = low	y_i = med	y_i = high	N of obs.	$\|H\|(Y\|X=x_i)$	$\|G\|(Y\|X=x_i)$
x_i = low	100	100	100	300	*1.000*	*0.000*
x_i = med	100	100	100	300	*1.000*	*0.000*
x_i = high	100	100	100	300	*1.000*	*0.000*
N of obs.	300	300	300	900		
$\|H\|(X\|Y=y_i)$	*1.000*	*1.000*	*1.000*			
$\|G\|(X\|Y=y_i)$	*0.000*	*0.000*	*0.000*			

<u>Note</u>: |H|(X) = 1.000, |H|(Y) = 1.000

Table 2b. Information gains $|G|(X|Y{=}y_i)$ and $|G|(Y|X{=}x_i)$ with regard to X and Y for an exemplary case with perfect status inheritance.

Strata	y_i = low	y_i = med	y_i = high	N of obs.	$\|H\|(Y\|X{=}x_i)$	$\|G\|(Y\|X{=}x_i)$
x_i = low	300	0	0	300	*0.000*	*1.000*
x_i = med	0	300	0	300	*0.000*	*1.000*
x_i = high	0	0	300	300	*0.000*	*1.000*
N of obs.	300	300	300	900		
$\|H\|(X\|Y{=}y_i)$	*0.000*	*0.000*	*0.000*			
$\|G\|(X\|Y{=}y_i)$	*1.000*	*1.000*	*1.000*			

Note: $|H|(X)$ = 1.000, $|H|(Y)$ = 1.000

The partitioning of a population into subgroups, based on homogeneous *secondary* characteristics $Y = y_1$, $Y = y_2$, ..., generally modifies the information $|H|(X)$ about the privilege X. Thus, for analysing this effect we have to consider for every $Y = y_i$ the *conditional entropy*

$$H(X \mid Y{=}y_i) = - \Sigma_{j=1,..,n} [prob(X{=}x_j \mid Y{=}y_i) \cdot \log_2(prob(X{=}x_j \mid Y{=}y_i))] \tag{2a}$$

as well as its *normalised* equivalent

$$|H|(X \mid Y{=}y_i) = H(X \mid Y{=}y_i) / \log_2(n) \tag{2b}$$

From knowledge about $Y{=}y_i$ results a (normalised) *information gain*

$$|G|(X \mid Y{=}y_i) = |H|(X) - |H|(X \mid Y{=}y_i) \tag{3}$$

which is based on the comparison with $|H|(X)$. It is generally *positive* — although for particular characteristics $Y{=}y_i$ negative values are also possible. Applied to intergenerational mobility research, where X is the social class of the parents (= origin) and Y of their children (= destination), $|G|(X|Y{=}y_i)$ is the information gain about the parents, as inferred from the class $Y{=}y_i$ of their children. If intergenerational status inheritance is *completely missing,* the information gain $|G|(X|Y{=}y_i)$ about the parents is for all classes y_i equal to the minimum *zero,* as Tab. 2a demonstrates. Tab. 2b shows the opposite case of *perfect* intergenerational status inheritance. Here, the information gain $|G|(X|Y{=}y_i)$ about the children's origin is at the maximum level 1, as intuitively expected. Finally, Tab. 2c displays an exemplary situation of *intermediate* status inheritance, where children with high status (y_i = high) are not only from parents with high (x_i = high) but also with low status (x_i = low). Consequently, the information gain $|G|(X|Y{=}y_i{=}high)$ = 0.335 about the parents of children with high status y_i is at an *intermediate* level between 0 and 1 (cf. Tabs. 2a and 2b).

Table 2c. Information gains $|G|(X|Y=y_i)$ and $|G|(Y|X=x_i)$ with regard to X and Y for an exemplary case with intermediate status inheritance.

Strata	y_i = low	y_i = med	y_i = high	N of obs.	$\|H\|(Y\|X=x_i)$	$\|G\|(Y\|X=x_i)$
x_i = low	200	0	200	400	*0.631*	*0.335*
x_i = med	0	300	0	300	*0.000*	*0.966*
x_i = high	0	0	200	200	*0.000*	*0.966*
N of obs.	200	300	400	900		
$\|H\|(X\|Y=y_i)$	*0.000*	*0.000*	*0.631*			
$\|G\|(X\|Y=y_i)$	*0.966*	*0.966*	*0.335*			

Note: $|H|(X) = 0.966$, $|H|(Y) = 0.966$

In the present example of intergenerational mobility one might also ask about the information gain regarding the status destination Y of the *children*, if the status $X=x_i$ of the parents (= origin) is known. In analogy to equation (3), this kind of (normalised) *information gain* is defined as

$$|G|(Y \mid X=x_i) = |H|(Y) - |H|(Y \mid X=x_i) \tag{4}$$

The more the group x_i deviates from the general mobility pattern of the surrounding society, the higher the information gain $|G|(Y|X=x_i)$. If intergenerational status inheritance is *completely missing,* the information gain $|G|(Y|X=x_i)$ about the situation of the children is for all origins x_i equal to *zero* (= minimum). Moreover, from completely missing status inheritance follows $|G|(Y|X=x_i) = |G|(X|Y=y_i)$ for all classes i (see Tab. 2a). As shown in Tab. 2b, this *mathematical relation* also holds for the situation of *perfect* intergenerational status inheritance. However, here the information gains are for all status groups i equal to the maximum 1. For *intermediate* levels of status inheritance, there is often an *asymmetry* between the information gain about the origin and the destination of the intergenerational mobility of a social class. In the example of Tab. 2c, little is known about the destination of the children of lower class parents ($|G|(Y|X=x_{low}) = 0.335$) but it is nearly certain, that the parents of the lower class children belong also to the lower class, since $|G|(X|Y=y_{low}) = 0.966$. Given the particular structure of the fictitious mobility table Tab. 2c this is quite plausible. Similarly, also the index values $|G|(Y|X=x_{high}) = 0.966$ and $|G|(X|Y=y_{high}) = 0.335$ of the upper class mirror the structure of the mobility table Tab. 2c: according to this table, the upper class tends to reproduce itself but is open for newcomers from the lower class. In sum, the index $|G|(Y|X=x_i)$ is generally *positively correlated* with the *status conservation* of a class i, whereas the index $|G|(X|Y=y_i)$ may be used as a *proxy* for its *exclusiveness* with regard to newcomers from other classes.

The fact, that there are as many indices $|G|(Y|X=x_i)$ as there are analysed social classes calls for a summarizing index. Formal analyses show that the information theoretic concept of (normalised) *synentropy* $|S|(X,Y)$ (= mutual information) corresponds to the mean information gains $|G|(Y|X=x_i)$, weighted by the size of the different classes with $X=x_i$. Thus [3, 4, 5]

$$|S|(X,Y) = \Sigma_{i=1,..,n} [prob(X=x_i) * |G|(Y \mid X=x_i)] \quad (5)$$

Equation (5) suggests that the synentropy |S|(X,Y) is a measure of the *inequality of the mobility chances* of a society: the higher the synentropy |S|(X,Y), the higher average information gain from the parental background with regard to the children's social class. Thus, the synenropy |S|(X,Y) (= |S|(Y,X)) is an *inverse* measure of *social fluidity*. Contrary to the corresponding indices of Breen (2004b), it is not based on odds ratios of intergenerational mobility but on comparisons of entropies. |S|(X,Y) is also related to other forms of inequality, in particular to the total inequality of the children's generation and the mean inequality among the children with the same social origin. In particular, formal reasoning shows for the mentioned inequalities the following mathematical relationship [6]

$$|H|(Y) = |S|(X,Y) + \Sigma_{i=1,..,n} [prob(X=x_i) * |H|(Y \mid X=x_i)] \quad (6)$$

Thus, total inequality |H|(Y) of the generation of the children can be split into two components:

a) The synentropy |S|(X,Y), which can be interpreted as the average *inter-strata inequality of mobility chances*. The higher it is, the more is the total inequality |H|(Y) influenced by structural problems of a society.

b) The mean inequality $\Sigma_{i=1,..,n} [prob(X=x_i) * |H|(Y \mid X=x_i)]$ among the children with the same parental origin $X=x_i$, weighted by the size of the different classes i = 1, 2,...., n of the parents. The higher this component is, the more is the total inequality |H|(Y) influenced by *inter-individual differences* between children with the same strata-specific mobility chances.

Thus, equation (6) corresponds to the more general entropy decomposition of Theil (1972, p. 21), applied to the particular phenomenon of intergenerational mobility and the related social inequalities.

---

[3] By substituting Eq. (4) into Eq. (5) follows: $\mathbf{\Sigma_{i=1,..,n} [prob(X=x_i) * |G|(Y \mid X=x_i)]}$ =
$\Sigma_{i=1,..,n} [prob(X=x_i) * (|H|(Y) - |H|(Y \mid X=x_i))] = |H|(Y) - \Sigma_{i=1,..,n} [prob(X=x_i) * |H|(Y \mid X=x_i)]$ =
= $|H|(Y) - |H|(Y \mid X)$ = **|S|(X,Y)** (see Mathar, 1994, p. 27, definition 3.3, and p. 31, lemma 3.2. b)).

[4] Eq. (5) shows the relation of the information gains $|G|(Y \mid X=x_i)$ to Theil's U = S(X,Y) / H(Y) (Wikipedia, 2021). Thus, $|G|(Y \mid X=x_i)$ is a weighted *component* of Theil's U. The latter is a *summary* statistic and its standardization is different from ours: instead of H(Y) we use the logarithm $\log_2(n)$ (see Eqs. (3), (1b), and (2b)).

[5] Since |S|(X,Y) = |S|(Y,X) (Mathar, 1994, p. 31, lemma 3.2.b) it is also possible to decompose $|S|(X,Y) = \Sigma_{i=1,..,n} [prob(Y=y_i) * |G|(X \mid Y=y_i)]$ by the information gains $|G|(X \mid Y=y_i)$ about the origins.

[6] From the equations (4) and (5) follows: **|S|(X,Y)** = $\Sigma_{i=1,..,n} [prob(X=x_i) * |G|(Y \mid X=x_i)]$ =
= $\Sigma_{i=1,..,n} [prob(X=x_i) * (|H|(Y) - |H|(Y \mid X=x_i))]$ = $\mathbf{1 * |H|(Y) - \Sigma_{i=1,..,n} [prob(X=x_i) * |H|(Y \mid X=x_i)]}$ --> equation (6).

3. AN EMPIRICAL ANALYSIS OF THE CHANGING EUROPEAN EDUCATIONAL STRATIFICATION

3.1 INTRODUCTION

As an empirical application of the previously introduced information theoretic indicators we are going to analyse in this section 3 the educational stratification and the related intergenerational mobility in 16 West-European[7] countries. Educational mobility is insofar of special interest, as it is in modern societies instrumental for economic upward mobility (Hahn, 2016, p. 26; Ashenfelter & Rouse, 2000, pp. 93 – 95). Consequently, as shown in a comparative diachronic eight-nation study, edited by Breen and Müller (2020a), changes in the educational stratification have important consequences for the social fluidity of the class system. Especially crucial in this respect are two secular trends: the equalization of the educational attainments and the expansion of non-primary education (Pollak & Müller, 2020).

Although educational mobility depends on many different factors like the type of schooling (e.g. private vs. public or religious vs. secular, etc.) (Suna et al., 2020), the quality of schools (Kirst, 2007), or public educational policies (OECD, 1996), we are using here a sociological approach, focusing on the intergenerational transmission of education from the parents to their children (Chevalier et al., 2009; Siraj & Mayo, 2014, chap. 6–7). Since the influence of the parental education varies considerably from study to study, Fleury and Gilles (2018) made a *meta-analysis* of 25 related publications that appeared between 2002 and 2014. The reported results depend among others on the *gender* of the analysed parent and his/her child, the parent's *socio-economic status, age-cohort,* and the *world region* of his/her domicile. By a meta-regression analysis Fleury and Gilles were able to control these confounders. This way they could "proof" a statistically significant transmission of the education of the parents onto their children, which was operationalised by the global number of years of schooling. To the contrary, the present study differentiates only between primary, secondary, and tertiary education, which are "naturally" defined by the institutionalised exit points of educational careers. Moreover, it analyses not only the information that can be gained from the parents' education with regard to the educational status of their children but deals also with the reverse question: what can we learn from the children's status about the educational attainments of their parents?

Due to its high standardization and international comparability, we use as the data source for the present analysis the European Values Study 2008, available as dataset ZA-4800 of Gesis (2020). It contains personal interviews collected in 2008 – 2010 with respondents reporting their own education as variable V336 and the education of their fathers as variable V355.[8] By recoding the original data, we were able to identify for each of the analysed countries the *fathers* and *children*

a) with *primary schooling* as the highest educational attainment (codes 0 to 1 of the variables V355 or V336), which is in the article denoted as *low* education;
b) with *secondary schooling* as the highest educational attainment (codes 2 to 4 of the variables V355 or V336), which is in the article denoted as *medium* education;

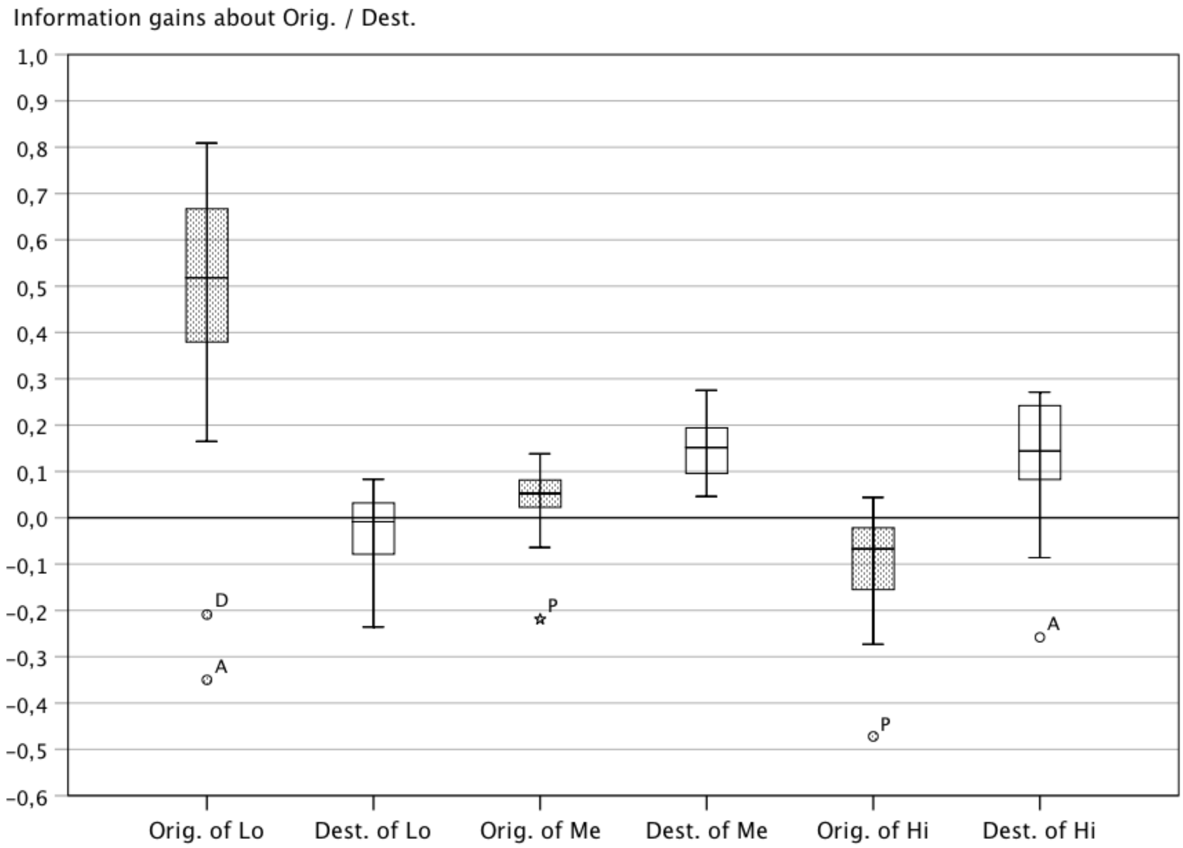
[7] Eastern Europe is excluded due to its very important societal discontinuities at the end of the communist period.

[8] V336 and V355 are both coded on the basis of the International Standard Classification of Education (ISCED).

c) with *tertiary schooling* as the highest educational attainment (codes 5 to 6 of the variables V355 or V336), which is in the article denoted as *high* education.

3.2 INTERGENERATIONAL EDUCATIONAL MOBILITY

In a society with *no* status inheritance from one generation to the next, the information gains about the origins and destinations should for all strata of parents and children be zero (see Tab. 2a). As Fig. 1 (see "Orig. of Low") shows, this is obviously not the case for the relatively high information gain about the *origin* of the children with *low* education. This points to the



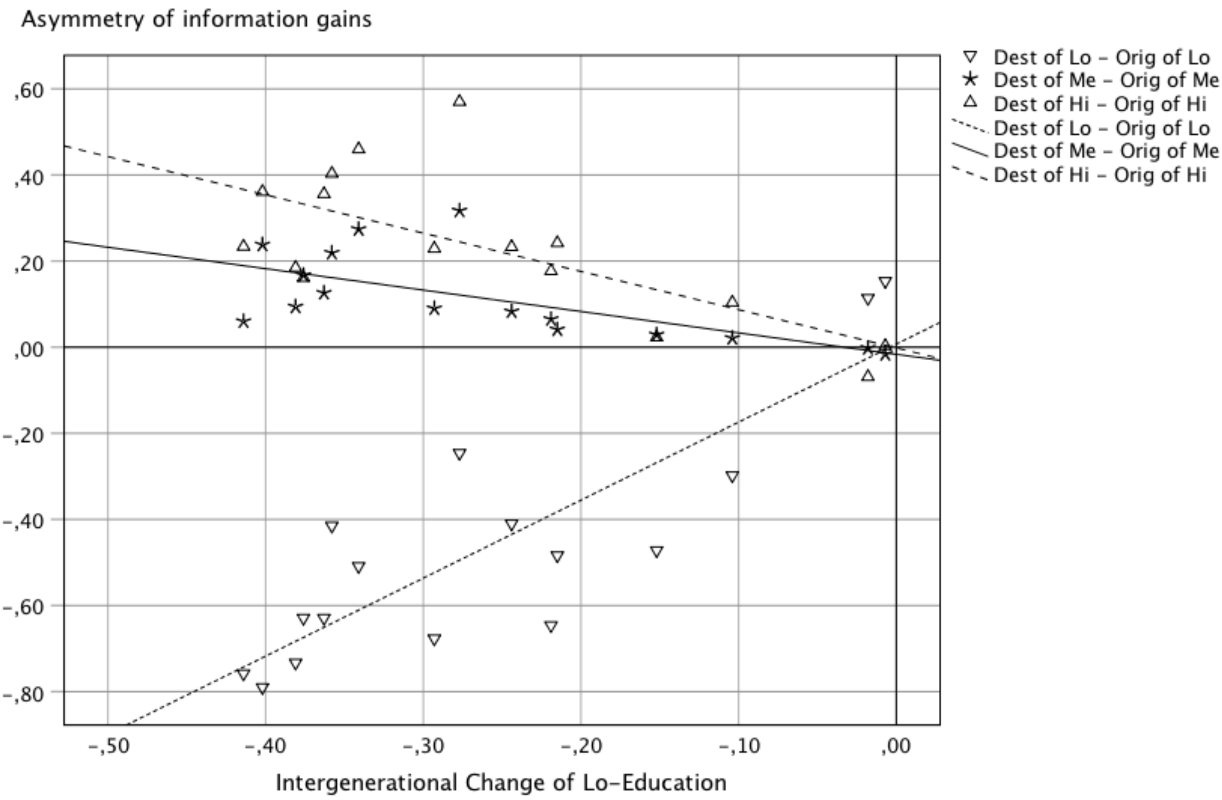
Legend: Outliers: A = Austria, D = Germany, P = Portugal. Orig. of Lo: Parental origin of children with *low* education Y; Orig. of Me: Parental origin of children with *medium* education Y; Orig. of Hi: Parental origin of children with *high* education Y. Dest. of Lo: Children's destination of parents with *low* education X; Dest. of Me: Children's destination of parents with *medium* education X; Dest. of Hi: Children's destination of parents with *high* education X.

Figure 1. Boxplots[9] *of the normalised information gains about the origins (Orig.) and destinations (Dest.) of mobility, by level of education.*

exclusiveness of lower education: in 14 out of 16 countries the most important origin of children with low education are parents, who had also only low education. This closedness protects children of parents with middle and high education against intergenerational downward mobility to the mentioned stratum of low education. According to the box plots of Fig. 1 (Cramer & Howitt, 2004, pp. 17–18), the mentioned exclusiveness is around zero for medium education

[9] Main elements of boxplots: Bottom and ceiling of a box: Percentiles 25 and 75. Horizontal division inside a box: Median value (= percentile 50). Whiskers: Minimal and maximal values, which are not outliers. Small stars and circles: Outliers.

(see "Orig. of Me") and further decreases to negative values for higher education (see "Orig. of Hi"). Thus, higher education seems to be open for children of parents with any educational background. Nonetheless, there is a positive information gain with regard to the *destination* of children of parents with higher education (see Fig. 1, "Dest. of Hi"). This suggests intergenerational *status conservation* of families with higher education: indeed, in 11 out of 15 countries higher education is the most important destination of children of parents with higher education. For parents with medium education, there is a similar positive information gain about the destination of their children (see Fig. 1, "Dest. of Me"), which too points to status conservation. To the contrary, for parents with low education the respective information gain is *zero* (see Fig. 1, "Dest. of Lo"). This lack of status conservation is a consequence of the already mentioned *relative* openness of medium and higher education, which receive a lot of children from families with low educational background.



Legend: See legend of Fig. 1.
Note: One-tailed Pearson correlations of *Intergenerational Change of Lo-Education*:
With (Dest of Lo - Orig of Lo): 0.849 ($p < .001$); with (Dest of Me - Orig of Me): -0.647 ($p < 0.01$);
with (Dest of Hi - Orig of Hi): -0.686 ($p < 0.01$).

Figure 2. The influence of the changing size of lower education on the asymmetry between the information gains about the origins (Orig.) and the destinations (Dest.) of mobility.[10]

In sum, there are interesting asymmetries between the knowledge that can be derived from the education of the parents with regard to their children and the reverse information, which is derived from the education of the children with regard to their parents. For *lower* education there is a lot of information about the origin but not too much about the destination. Consequently, lower education is exclusive but not status conserving. For *medium* education there

[10] Real data and related linear regression lines.

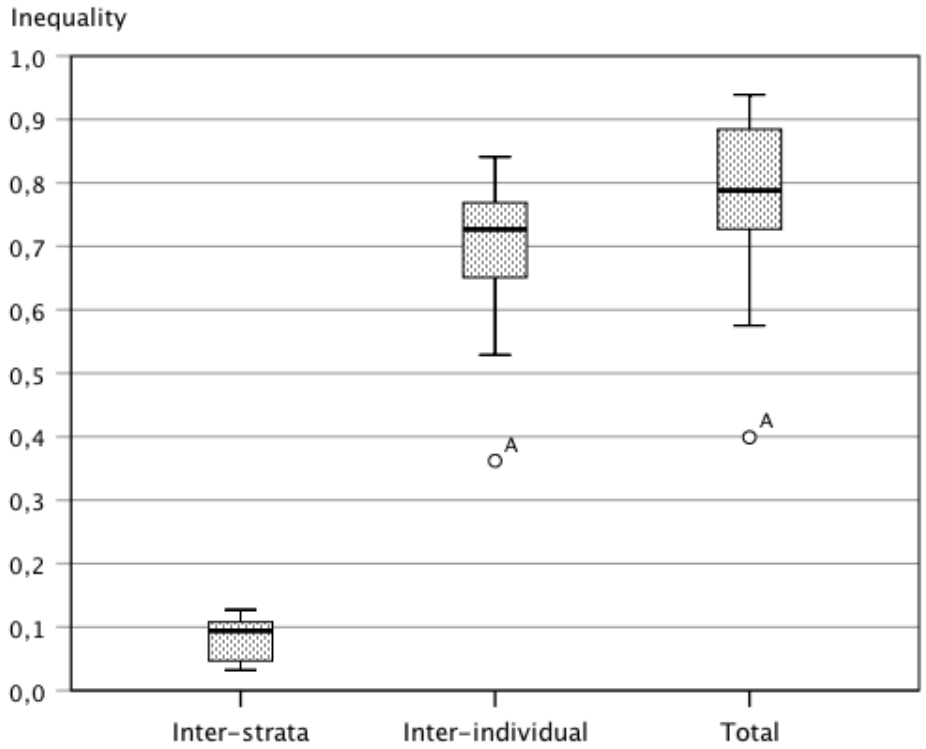
is more information about the destination than about the origin. Thus *medium* education is mainly status conserving. For *higher* education there is much information about the destination but a loss of information about the origin. Consequently, higher education is status conserving, but with regard to newcomers from other strata it is rather inclusive than exclusive. These asymmetries are probably the result of changes of the educational system between the generation of the parents and their children: school leaving with only primary education is in most West European countries no more possible such that medium and higher education expanded (van Zanden et al., 2014, pp. 95 ff.) and consequently induced a (forced) *structural mobility* in the educational hierarchy (Boudon, 1973, pp. 17 ff.). For medium or higher education this facilitated intergenerational status conservation and avoided exclusiveness, whereas for lower education the effects were just the reverse. In order to check this hypothetical explanation, we plotted in Fig. 2 the intergenerational change of lower education against the asymmetries of information gains. The data pattern and the highly significant correlations seem to confirm the hypothesis: the stronger the *decrease* of lower education, the higher the (absolute) asymmetries between the information gains about the origins and the destinations. Remarkably, there are *no* asymmetries with regard to *these* information gains, if there is *no* change of lower education (see regression lines of Fig. 2 for values on the horizontal axis near 0). Thus the observed asymmetries are most probably the result of the mentioned structural mobility.

3.3 THE DIFFERENT FORMS OF EDUCATIONAL INEQUALITY

As we learned earlier from equation (6), the total educational inequality |H|(Y) of the children, which is varying between 0 and 1, can be split into two mutually exclusive components:

a) The average *inter-strata* inequality of mobility *chances*, which is operationalised as the normalised synentropy |S|(X,Y) between the education X of the parents and Y of the children. It varies between 0 and 1 and the more this indicator deviates from 0, the less equal are the mobility chances of the different strata.
b) The mean *inter-individual* inequality between children of parents with the same educational background, which too varies between 0 and 1. It is operationalised as the mean normalised conditional entropy $\Sigma_{i=1,..,n}$ [prob(X=x_i) * |H|(Y | X=x_i)] and mirrors factors like inter-personal differences in achievement orientation, personal "luck", etc. It may however also be influenced by neglected structural factors like discrimination on the basis of gender or ethnic background.

Fig. 3 presents the decomposition of the educational inequality |H|(Y) of the children for the 16 countries, which are analysed in this article. It seems that the total inequality is mainly influenced by inter-individual factors of inequality and to a much lesser extent by the discrimination of educational strata. The minor importance of the second factor is very consistent with Fig. 1: the information gains about the destination of the children from the shrinking lower educational stratum are negative (see Fig. 1, "Dest. of Lo") and those of the middle and higher educational strata only slightly positive (see Fig. 1, "Dest. of Me" and "Dest. of Hi"). Consequently, the average |S|(X,Y) of these gains, which corresponds to the inter-strata inequality of Fig. 3, remains *small* although it still exists, as observed by Chevalier et al. (2009). Planned future analyses on the bases of more recent data from the European Values Study 2017 will show, whether in more recent time inter-strata inequality has further decreased: this would the "fruit" of the educational policy of many European countries, aiming at a democratisation of higher education.



Legend: Total inequality = |H|(Y) of children; Inter-strata inequality = |S|(X,Y);
Inter-individual inequality = |H|(Y) - |S|(X,Y) = $\Sigma_{i=1,..,n}$ [prob(X=x_i) * |H|(Y | X=x_i)].
Outlier: A = Austria.

Figure 3. Boxplots[11] *of the decomposition of the total inequality of education into an inter-strata and an inter-individual component.*

The considerable importance of the other, inter-individual factor is plausible for the achievement-oriented education system of the meritocratic societies of Western Europe (Arrow et al., 2000). However, it deserves further consideration because of the possibility that it includes also *structural* components like parental income, ethnic background, etc. In order to cross-check this possibility, one might consider to modify the parental status X such that it contains not only three levels of education but is additionally combined with the mentioned structural components such that X comprises more than three categories of parents.

[11] For a description of the main elements of boxplots refer to footnote 9.

4. CONCLUSIONS AND DISCUSSION

On the grounds of his previous positive experiences with entropy-based social indicators (Mueller, 2004, 2017), the author proposes in the present article to make use of these indicators in order to analyse intergenerational mobility. As compared to the more traditional approaches, mentioned in Hout (1983) or Breen (2004b), it has two major advantages, which have been demonstrated in the previous section 3:

First, by using entropy concepts we are able to distinguish between the information gain $G|(Y|X{=}x_i)$ from knowledge of the parental status X and the analogous information gain $G|(X|Y{=}y_i)$ from knowledge of the children's status Y. Thus, we differentiate in this article between status conservation $G|(Y|X{=}x_i)$ and exclusiveness $G|(X|Y{=}y_i)$. As Tab. 3 shows, a dichotomization (low vs. high) of these gains leads to a four-fold typology of different status groups with regard to the intergenerational mobility from and to these groups. According to this typology we identified in Fig. 1 lower education as exclusive but non-conserving, whereas medium and high education are rather the reverse, i.e. non-exclusive but status-conserving.

Second, the use of entropy concepts enabled us to compare different forms of mobility and inequality. On the one hand, this was possible because of the common scale of these concepts, which is based on the *bit* as the standard unit of measurement in information theory. On the other hand, many concepts of this article are borrowed from a mathematical discipline, which explored their mutual formal relations (Mathar, 1996; Stone, 2015). In particular we profited from this formal heritage when interpreting the equation (6) as the decomposition of the total inequality into a structural inter-strata inequality and an inter-individual inequality between persons with the same parental background (see Fig. 3). However, this interpretation revealed also some problems of the correspondence between sociological and mathematical concepts, which still have to be solved by further research: is the high inter-individual inequality really the result of differences in fortune and varying personal achievements? Or are there hidden structural factors, which can be identified by a refinement of the proposed information theoretic methods?

Table 3. A four-fold typology of status groups, based on dichotomised information gains $|G|(X \mid Y{=}y_i)$ and $|G|(Y \mid X{=}x_i)$ about the origins and destination of their mobility.

	Gain about destination:	
Gain about origin:	$\|G\|(Y \mid X{=}x_i) = \text{low}$	$\|G\|(Y \mid X{=}x_i) = \text{high}$
$\|G\|(X \mid Y{=}y_i) = \text{low}$	non-conserving, non-exclusive	conserving, non-exclusive
$\|G\|(X \mid Y{=}y_i) = \text{high}$	non-conserving, exclusive	conserving, exclusive

REFERENCES

Arrow, K., et al. (2000). Introduction. In Arrow, K., S. Bowles & S. Durlauf (Eds.). *Meritocracy and Economic Inequality* (ix-xv). Princeton: Princeton University Press.

Ashenfelter, O. & Rouse, C. (2000). Schooling, Intelligence, and Income in America. In Arrow, K., S. Bowles, & S. Durlauf (Eds.). *Meritocracy and Economic Inequality* (chap. 5). Princeton: Princeton University Press.

Boudon, R. (1973). *Mathematical Structures of Social Mobility.* Amsterdam: Elsevier.

Breen, R. (2004a). The Comparative Study of Social Mobility. In R. Breen (Ed.). *Social Mobility in Europe* (chap. 1). Oxford: Oxford University Press.

Breen, R. (2004b). Statistical Methods of Mobility Research. In R. Breen (Ed.). *Social Mobility in Europe* (chap. 2). Oxford: Oxford University Press.

Breen, R & Luijkx, R. (2004). Conclusions. In R. Breen (Ed.). *Social Mobility in Europe* (chap. 15). Oxford: Oxford University Press.

Breen, R. & Müller, W. (Eds.) (2020a). *Education and Intergenerational Social Mobility in Europe and the United States.* Stanford: Stanford University Press.

Breen, R. & Müller, W. (2020b). Introduction. In Breen, R. & W. Müller (Eds.). *Education and Intergenerational Social Mobility in Europe and the United States* (chap. 1). Stanford: Stanford University Press.

Chevalier, A., et al. (2009). A Multi-Country Study of Intergenerational Educational Mobility. In Dolton, P. et al. (Eds.). *Education and Inequality Across Europe* (chap. 12). Cheltenham: *Edward* Elgar.

Coulter, Ph. B. (1989). *Measuring Inequality: A Methodological Handbook.* Boulder: Westview Press.

Cramer, D. & Howitt, D. (2004). *The SAGE Dictionary of Statistics.* London: Sage Publications.

Fleury, N. & Gilles, F. (2018). The intergenerational transmission of education. A meta-regression analysis. *Education Economics 26* (6*),* 557–573.

Gesis (2020). *ZA4800: European Values Study 2008: Integrated Dataset.* (https://search.gesis.org/research_data/ZA4800). Cologne: Gesis.

Hahn, S. (2016). *The Risk of Downward Mobility in Educational Attainment.* Wiesbaden: Springer VS.

Hao, L. & Naiman, D. Q. (2010). *Assessing Inequality.* Thousand Oaks: Sage Publications.

Hauser, R. M. & Featherman, D. L. (1977). *The Process of Stratification.* New York: Academic Press.

Hout, M. (1983). *Mobility Tables.* Beverly Hills. Sage Publications.

Kirst, M. (2007). Secondary and Postsecondary Linkages. In Dickert-Conlin, S. & R. Rubenstein (Eds.). *Economic Inequality and Higher Education* (chap. 3). New York: Russell Sage Foundation.

Mathar, R. (1996). *Informationstheorie (Information Theory)*. Stuttgart: Teubner.

Mueller, G. P. (2004). Entropy Measures for Comparative Analyses of Social Justice and Discrimination. In Grözinger, G. & A. van Aaken (Eds.). *Inequality: New Analytical Approaches* (pp. 15–31). Marburg: Metropolis-Verlag.

Mueller, G. P. (2017). Gender Inequality Under Different National Welfare Regimes: An Empirical Evaluation with Entropy Measures from Information Theory. *The Senshu Social Well-being Review 4,* 39–51.

OECD (1996). *Mapping the Future: Young People and Career Guidance*. Paris: OECD Publications.

Pollak, R. & Müller, W. (2020). Education as an Equalizing Force. In Breen, R. & W. Müller (Eds.). *Education and Intergenerational Social Mobility in Europe and the United States* (chap. 6). Stanford: Stanford University Press.

Powell, B. et al. (2004). Macro Causes, Micro Effects. In Conley, D. & K. Albright (Eds.). *After the Bell – Family Background, Public Policy, and Educational Success* (chap. 4). London: Routledge.

Shannon, C. & Weaver, W. (1962, [1949]). *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Siraj, I. & Mayo, A. (2014). *Social Class and Educational Inequality: The Impact of Parents and Schools.* Cambridge: Cambridge University Press.

Stone, J. V. (2015). *Information Theory.* Sheffield: Sebtel Press.

Suna, H. E., et al. (2020). Socioeconomic Status and School Type as Predictors of Academic Achievement. *Journal of Economy Culture and Society 61,* 41–64.

Theil, H. (1967). *Economics and Information Theory.* Amsterdam: North-Holland Publishing Company.

Theil, H. (1972). *Statistical Decomposition Analysis.* Amsterdam: North-Holland Publishing Company.

van Doorn, M. et al. (2011). Intergenerational Transmission of Education Across European Countries and Cohorts. *European Societies 13,* 93–117.

van Zanden, J. L. et al. (Eds.) (2014). *How was Life? Global Well-Being Since 1820.* Paris: OECD Publishing.

Wikipedia (2021). *Uncertainty Coefficient (https://en.wikipedia.org/wiki/Uncertainty_coefficient)*. Accessed on August, 25, 2021.

GLOSSARY OF MATHEMATICAL TERMS

$\|G\|(X \mid Y=y_i)$:	Normalised information gain about X of parents, if children's $Y=y_i$.
$\|G\|(Y \mid X=x_i)$:	Normalised information gain about Y of children, if parents' $X=x_i$.
$H(X)$:	Entropy of parents' status X.
$H(Y)$:	Entropy of children's status Y.
$\|H\|(X)$:	Normalised entropy of parents' status X.
$\|H\|(Y)$:	Normalised entropy of children's status Y.
$H(X \mid Y=y_i)$:	Conditional entropy of parents' status X, if children's $Y=y_i$.
$H(Y \mid X=x_i)$:	Conditional entropy of children's status Y, if parents' $X=x_i$.
$\|H\|(X \mid Y=y_i)$:	Normalised conditional entropy of X, if $Y=y_i$.
$\|H\|(Y \mid X=x_i)$:	Normalised conditional entropy of Y, if $X=x_i$.
$\log_2(x)$:	Binary logarithm (base 2) of value x.
n:	Number of status categories of X or Y.
prob(E):	Probability of event E.
prob(E|C=c):	Conditional probability of event E, if variable C=c.
$\|S\|(X,Y)$:	Normalised synentropy (mutual information) between statuses X and Y.
X:	Status (education) of parents.
Y:	Status (education) of children.
$\Sigma_{i=1,..,n}(x_i)$:	Sum $x_1 + x_2 + ... + x_n$

DATA APPENDIX

Table 4a. Cross-tabulation of the education X of the parents vs. the education Y of their children: Contingency table of the number of cases.[12]

Country	X	Y = low	Y = med	Y = high
Austria	low	16	45	2
	med	17	1114	103
	high	1	54	34
Belgium	low	143	340	77
	med	12	331	229
	high	0	78	161
Finland	low	48	173	253
	med	3	142	174
	high	1	58	159
France	low	186	368	114
	med	9	196	191
	high	3	56	115
Germany	low	15	44	2
	med	26	1270	278
	high	2	168	160
Greece	low	426	417	104
	med	20	249	106
	high	1	55	45
Ireland	low	135	277	54
	med	0	195	98
	high	0	26	38
Italy	low	192	491	44
	med	5	415	146
	high	0	36	36
Luxembourg	low	201	314	81
	med	41	335	165
	high	6	90	141

[12] Legend and sources: See end of the table.

Table 4a continued.

Country	X	Y = low	Y = med	Y = high
Netherlands	low	110	261	71
	med	23	459	240
	high	1	67	150
Norway	low	15	123	41
	med	3	375	270
	high	0	92	122
Portugal	low	821	334	78
	med	6	73	30
	high	1	23	20
Spain	low	462	425	102
	med	16	182	94
	high	7	35	52
Sweden	low	60	294	128
	med	3	301	153
	high	1	61	89
Switzerland	low	49	124	27
	med	16	603	150
	high	2	66	105
UK	low	55	162	49
	med	8	359	146
	high	0	63	100

Legend: X = Fathers' education, where low = primary, med = secondary, and high = tertiary education; Y = Children's education, where low = primary, med = secondary, and high = tertiary education.
Source: European Values Study 2008, for details refer to section 3.1.

Table 4b. Inequalities of education and educational mobility.

Country	|H|(X)	|H|(Y)	|S|(X,Y)	|H|(Y) - |S|(X,Y)
Austria	0.382	0.399	0.037	0.362
Belgium	0.942	0.859	0.125	0.734
Finland	0.956	0.762	0.032	0.730
France	0.886	0.916	0.112	0.804
Germany	0.533	0.575	0.045	0.530
Greece	0.737	0.925	0.106	0.819
Ireland	0.809	0.855	0.127	0.728
Italy	0.779	0.758	0.110	0.648
Luxembourg	0.940	0.910	0.098	0.812
Netherlands	0.906	0.831	0.099	0.732
Norway	0.840	0.689	0.035	0.654
Portugal	0.376	0.811	0.085	0.726
Spain	0.692	0.939	0.098	0.841
Sweden	0.909	0.763	0.049	0.714
Switzerland	0.780	0.696	0.090	0.606
UK	0.903	0.765	0.086	0.679

Legend: |H|(X): Entropy of parents' educational status X; |H|(Y): Entropy of children's educational status Y; |S|(X,Y): Synentropy between educational statuses X and Y.

Table 4c. Information gains about educational origins and destinations of intergenerational mobility.

	$\lvert G\rvert(X \mid Y = y_i)$			$\lvert G\rvert(Y \mid X = x_i)$		
Country	y_i = low	y_i = med	y_i = high	x_i = low	x_i = med	x_i = high
Austria	-0.350	0.074	-0.189	-0.236	0.072	-0.258
Belgium	0.694	0.073	0.019	0.017	0.163	0.284
Finland	0.669	0.033	-0.025	-0.089	0.093	0.209
France	0.647	0.075	-0.085	0.018	0.201	0.271
Germany	-0.209	0.093	-0.089	-0.056	0.077	-0.086
Greece	0.557	-0.064	-0.206	0.048	0.210	0.254
Ireland	0.809	0.037	-0.121	0.019	0.275	0.240
Italy	0.671	0.028	-0.034	0.042	0.195	0.127
Luxembourg	0.432	0.049	-0.024	0.022	0.132	0.209
Netherlands	0.450	0.095	0.002	-0.034	0.136	0.244
Norway	0.430	0.017	0.044	-0.042	0.046	0.067
Portugal	0.329	-0.219	-0.472	0.083	0.098	0.098
Spain	0.480	-0.026	-0.273	0.065	0.193	0.130
Sweden	0.665	0.056	-0.069	-0.068	0.150	0.116
Switzerland	0.165	0.138	-0.065	-0.133	0.159	0.039
UK	0.556	0.088	-0.019	-0.090	0.153	0.158

Legend: $\lvert G\rvert(X \mid Y=y_i)$: Normalised information gain about parents' educational status X, if children's educational status $Y=y_i$; $\lvert G\rvert(Y \mid X=x_i)$: Normalised information gain about children's educational status Y, if parents' educational status $X=x_i$.