*European Sociological Review*, 2015, Vol. 31, No. 3, 312–325
doi: 10.1093/esr/jcv005
Advance Access Publication Date: 20 February 2015
Original Article

# Estimating Social and Ethnic Inequality in School Surveys: Biases from Child Misreporting and Parent Nonresponse

**Per Engzell[1,*] and Jan O. Jonsson[1,2]**

[1]Swedish Institute for Social Research (SOFI), Stockholm University, SE-106 91 Stockholm, Sweden and [2]Nuffield College, Oxford University, OX1 1NF Oxford, UK

*Corresponding author. Email: per.engzell@sofi.su.se

Submitted December 2013; revised: December 2014; accepted December 2014

## Abstract

We study the biases that arise in estimates of social inequalities in children's cognitive ability test scores due to (i) children's misreporting of socio-economic origin and (ii) parents' nonresponse. Unlike most previous studies, we are able to draw on linked register data with high reliability and almost no missingness and thereby jointly consider the impact of measurement error and nonresponse. Using data on 14-year-olds ($n = 18{,}716$) from a new survey conducted in England, Germany, the Netherlands, and Sweden (Children of Immigrants Longitudinal Survey in Four European Countries), we find that child reports on parental occupation are well aligned with parents' reports in all countries, but reports on parental education less so. This leads to underestimation of socio-economic disparities when child reports of education are used, but not occupation. Selective nonresponse among parents turns out to be a real problem, resulting in similar underestimation. We also investigate conditional estimates of immigrant–non-immigrant disparities, which are surprisingly little affected by measurement error or nonresponse in socio-economic control variables. We conclude that school-based surveys on teenagers are well advised to include questions on parental occupation, while the costs for carrying out parental questionnaires may outweigh the gains.

## Introduction

A central concern for social stratification research is the lingering impact of socio-economic origin (SES) on educational and social class attainment, as well as on access to other scarce goods, such as income and wealth (Breen and Jonsson, 2005). Fundamental for such studies is the measurement of SES characteristics such as mother's and father's education, occupation, and social class. Yet, few studies reflect on the amount of measurement error (ME) that can occur due to recall problems, misclassifications, or the fact that the respondent does not know the answer. Previous methodological analyses suggest that such variables may lack in quality (Looker, 1989; Breen and Jonsson, 1997), which can affect the conclusions we draw from empirical analyses.

Our aim is to estimate the bias produced by ME and nonresponse in survey information on parental education and occupation, when these SES variables are used as predictors of cognitive ability—a variable that is strongly correlated with educational success and thus with several crucial stratification outcomes. Our take on this issue is one that closely follows the development in the field of social stratification, namely, the increasing reliance on school-based surveys (van de Werfhorst and Mijs, 2010; Le Donné, 2014). In such surveys, children report characteristics of their parents or home environment, which is often complemented by direct information from parents in special interviews. The potential problems with this set-up are twofold: children may not be able to accurately report their parents' characteristics, creating ME in children's data; and nonresponse in parental data is often considerable and systematically related both to SES variables and outcomes.

ME in children's proxy reports of SES has been an active research area at least since the 1970s, and appears to have gained renewed interest as of late. But few if any of these studies have considered the impact of ME and nonresponse jointly. This is unfortunate, as the school survey lends itself to a potential trade-off between these two sources of bias in choosing between child and parent reports. We extend on previous research by examining this trade-off, using data that are unusually suited for the purpose.

Our first contribution is to study misreporting by comparing the information on parental SES from parents with that from children, using a recent comparative data set (Children of Immigrants Longitudinal Survey in Four European Countries, CILS4EU) collected in 2010–2011 in England, Germany, the Netherlands, and Sweden on nationally representative samples of 18,716 pupils aged 14–15 years and their parents (Kalter *et al*., 2013). We use these data to calculate reliabilities and to analyse the difference in ordinary least squares (OLS) regression coefficients of test scores when using parental and child reports, respectively. Our second contribution is to take the validation one step further, and address the issue of selective nonresponse by parents. We are able to conduct this unique study because we have access to register data on parental education for Sweden, so we can compare the estimates of parents' education on children's cognitive ability when the former variable is measured by (i) child reports, (ii) (responding) parents' own reports, and (iii) by high-quality administrative data (with almost no nonresponse). Our third contribution is to move beyond bias in the bivariate case to the consequences of ME when SES variables are used as controls, in which case the amount and direction of the bias in the predictor of interest—in our case, immigrant background—is less obvious.

Our analyses lead us to address the more general question of how to elicit information on SES. Is there a trade-off between bias due to misreporting by children and that due to selective nonresponse by parents? Our results suggest that there is not, necessarily—and for some analyses, ME tends not to be a big problem. We end the article by suggesting how school-based surveys are perfectly able to deliver SES variables of respectable quality at a relatively low cost.

## Previous Research

Previous validation studies of SES reports in school surveys have largely focused on the reliability of pupil reports as compared with those of parents. Looker (1989) reviews the literature up to that date—based mainly on data collected in the 1960s and 1970s—and concludes: children appear to be more accurate in their proxy reports about parents' occupation than about their education; reports from older children are more reliable than from younger; and there are no marked differences by child's gender.

While subsequent studies have expanded on this research (Ensminger *et al*., 2000; Lien, Friestad and Klepp, 2001; West, Sweeting and Speed, 2001; Vereecken and Vandegehuchte, 2003; Kreuter *et al*., 2010; Jerrim and Micklewright, 2014), these conclusions seem to hold. More recent efforts have been centred on indices based on reported home possessions (Currie *et al*., 1997, 2008; Andersen *et al*., 2008) fuelled by concerns about child nonresponse on standard indicators. Results from this approach have so far been mixed (cf. Wardle, Robb and Johnson, 2002; Traynor and Raykov, 2013) and, while promising, we think it has yet to develop into a viable alternative.

Crucial for comparative research is not only how reliable different measures are *on average*, but how that reliability might vary from one country or survey to another. In a recent study on PISA and PIRLS, Jerrim and Micklewright (2014) analyse data for a large number of countries on three measures: parental education, occupation, and number of books in the home. They find that while agreement between pupils and parents on parental education differs markedly between countries, it is uniformly high for occupation, and for number of books it is uniformly low. These results confirm that parental occupation is a more reliable indicator than education, adding the important insight that its performance also appears more consistent across countries.

Few studies have considered the biasing effect of nonresponse in tandem with that of ME as we do here. There has been some research on the results of school and pupil nonresponse in surveys such as PISA (Hanushek and Woessmann, 2011; Micklewright, Schnepf and Skinner, 2012), which has however not extended to parental nonresponse. There are also studies investigating nonresponse in household surveys (Groves, 2006; Groves and Peytcheva, 2008), showing that survey response rates and nonresponse bias appear weakly related, but this needs to be ascertained for parental reports. Existing data are not ideally suited to answer this question because the ME in pupil reports means that they are imperfect as a point of comparison.

## Implications of Misreporting and Nonresponse

Data gathered from school children tend to come with good response rates because pupils who answer the survey in a school setting face few disincentives to participate. Data from parents are more difficult to collect and response rates consequently lower. If we were confident in pupils' ability to report their parents' characteristics, response rates would suggest abandoning parental reports in favour of pupils' (if SES information is what we want to get out of parent questionnaires). However, a number of reasons lead us to expect parents' reports to be more reliable. Parents will have better knowledge about their own education and occupation than their children, they will know the key words to describe this accurately, and as adults will be generally better at the cognitive tasks involved in answering a questionnaire.

We cannot expect parent reports to be wholly free from error, as others have stressed before us (Kreuter *et al.*, 2010; Jerrim and Micklewright, 2014) and studies of adult respondents show (Black, Sanders and Taylor, 2003)—but this error is likely much smaller than for child reports. Thus, we expect that ME will be a greater problem for child reports, while parental reports will be more susceptible to nonresponse.

As is well known, ME in an independent variable (here, SES) typically has the effect of attenuating the regression slope, biasing the coefficient towards zero. For the simplest case involving only two variables, given the classical assumptions (that the ME is mean zero, normally distributed, and uncorrelated with the true values as well as the regression residual), the resulting bias will just be a function of the proportion of variance in the mismeasured variable owing to ME (Bound, Brown and Mathiowetz, 2001).[1] Letting $\sigma_x^2$ denote the variance of the variable when measured without error, and $\sigma_u^2$ the variance of the ME, this function can be written as follows:

$$p\lim_{n\to\infty}\hat{\beta}_{\mathrm{OLS}} = \beta\left[1 - \frac{\sigma_u^2}{\sigma_x^2 + \sigma_u^2}\right]. \quad (1)$$

We illustrate this in Figure 1. Hollow markers represent true and unobserved values of the independent variable, and solid markers the same individuals when observation is subject to a random error. The true regression function is drawn as a dashed line, whereas the regression slope in the presence of error is represented by the solid line—that this is flatter means that the regression coefficient is underestimated, that is, downwardly biased.

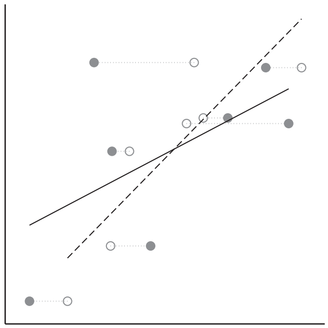

**Figure 1.** Illustration of linear regression with classical ME in x-axis

This illustration depicts a bivariate relationship, but often we want to estimate models that include several explanatory variables. We have chosen, as our example, the common regression of an outcome (here, cognitive ability) on immigrant background, controlling for SES. We assume that children report immigrant background accurately because this characteristic is likely to be more salient to them than parents' education or occupation (cf. Nordahl *et al.*, 2011; Parameshwaran and Engzell, 2014).

Because immigrant background and SES tend to be correlated, intuitively, the error in the latter should reflect on estimates for both. The expected consequence is that the attenuation for the mismeasured variable becomes worse, whereas the estimate for the covariate will be subject to an opposite, upward bias (Bound, Brown and Mathiowetz, 2001), which, in our example, would lead us to overestimate the gradient of immigrant background.

An additional scenario is that a correlation obtains not (only) between the true values of the mismeasured variable and covariates, but that the *measurement error* is systematically related to a covariate. For example, children of immigrants may have greater trouble reporting their parents' education because of language problems or difficulties in 'translating' a foreign qualification to its host-country equivalent. This will lead to a similar overestimation of the immigrant–non-immigrant gap (Black, Sanders and Taylor, 2003).

Because misreporting of parental characteristics is assumed to be more common among children, a standard way to investigate its consequences is to compare

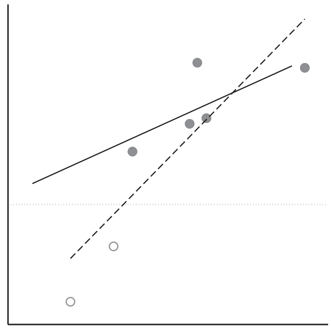

**Figure 2.** Illustration of linear regression with selection on y-axis

regression estimates between separate models for child and parent reports, with the same number of respondents across analyses (Jerrim and Micklewright, 2014). Again, this approach obscures that parent reports may be subject to a different kind of bias due to nonresponse, which in many school surveys is severe. If parent nonresponse is selective on the outcome, the expected consequences are similar to those of ME (Berk and Ray, 1982). This is illustrated in Figure 2, where the correct values are now observed but some observations with low values on the dependent variable are lost, again leading us to underestimate the regression slope.

Potentially, the empirical researcher is faced with two bad choices: to retain all cases, using child reports and risk bias because of misreporting, or to drop cases with no parental response and risk bias because of selective nonresponse (and also risk inefficient estimates because of the smaller sample size). Our aim here is to find out whether one of these options is better than the other.

We do not, in this article, focus on any more sophisticated missing data methods than casewise deletion, i.e. retaining only ‘complete cases’, or observations for which no variable in the analysis is missing. While rudimentary, this method is standard in most statistical software and by far the most commonly used in the field.

## Data and Analytical Plan

We use the first wave of the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU), conducted in the school year 2010–2011 in England, Germany, the Netherlands, and Sweden on pupils aged approximately 14 years (Kalter *et al.*, 2013; www.cils4.eu). The 18,716 pupils are nested within 480 schools (952 classes). Samples are nationally representative of the target cohort but stratified to yield an oversampling of schools with a high proportion of children of immigrants.[2] Participating pupils sat in for a 2-hr paper-and-pencil survey administered during school hours, completing a number of additional tasks including tests of cognitive ability. Parents or other guardians were interviewed by telephone or postal questionnaires.

For the Swedish sample, we also have linked household data from the Swedish Register of Education (*Utbildningsregistret*, see Statistics Sweden, 2011) containing the highest completed education attained by the biological parents of responding children. We expect these data to be of higher quality than survey reports because they are recorded at a much higher level of detail, in many cases directly by the institutions where the education was obtained. The backbone of the register is self-reported data from the population census 1990. Since then, completed educations are reported to Statistics Sweden on an annual basis by (almost all) Swedish educational institutions and added to the register. For foreign-born who might have completed their highest education abroad, special surveys were conducted in 1995 and 1999, and annually since 1999 for newly arrived immigrants (Statistics Sweden, 2011: 24–28).

In the first step of our analyses, we concentrate on the *extent* of nonresponse and misreporting. We tabulate the number of valid reports on parental education and occupation, separately by respondent (child or parent), survey country, and immigrant background, and we show the average test score differences between those with SES information present and missing. To assess reliability, we inspect bivariate correlations (Pearson’s $r$) among child reports, parent reports, and, where available, register data. We complement these estimates with alternative measures of agreement.

In the second step, we go on to investigate the *consequences* of data deficiencies. Initially, we concentrate on estimates for parental education in Sweden, where the availability of register data allows a detailed view of the respective biases owing to ME and nonresponse. We do this by estimating the same basic model but varying the source of SES and the selection criteria for the sample. Third, we extend the analyses to the other survey countries, and the additional indicator of parental occupation, to put the in-depth results for Sweden in a comparative context. Throughout, we display how estimates of the immigrant–non-immigrant gap are affected by data issues in the reported SES.

## Variables

We take pupils’ cognitive test scores as the outcome of interest. As predictors we include parents’ education, parents’ occupation, and pupil’s sex, age (in months), and immigrant background. By using the higher of parents' education and occupation (rather than mother's and father's separately) we avoid problems of collinearity, increase efficiency in estimates, and minimize missing data.[3]

### Cognitive Test Score

The cognitive test used in the CILS4EU is non-verbal and aimed at fluid intelligence, often considered the main component of general intelligence and largely independent of immigrant background (Kvist and Gustafsson, 2008). It focuses on visual puzzles ('patterns'), which do not require verbal ability, therefore minimizing cultural bias (Weiß, 2006). The test comprised 27 items and was clocked at 7 min. Test scores are z-standardized on the national level, zero representing the mean score in each survey country and the unit of measurement being (countrywise) standard deviations.

### Parental Education

The survey measures of parental education contain four categories: less than primary education, primary education, secondary education, and university. For most of our analyses, we assign the approximate number of years based on the higher of mother's and father's education, following the procedure used in PISA 2009 (Appendix A). In initial analyses for Sweden, we also use the categorical variable, and here we collapse the lowest two categories to overcome problems with small numbers.

Administrative and anonymized data on household education, defined as the higher of the biological father's and mother's level of education, were provided by Statistics Sweden and matched to pupils. Information is available for nearly all pupils participating in the survey, with only some missing data owing to information missing from registers (Table 1). Although the register information is given on a more detailed level than survey reports, we recode it to the same four-category schema for comparability.

### Parental Occupation

Parental occupation is based only on survey reports. Both pupils and parents were asked to provide the title and a brief description of the job of each parent. This information was coded according to the 2008 International Standard Classification of Occupations (ISCO-08) and converted into the interval-scale ISEI-08 occupational status (Ganzeboom and Treiman, 1996; Ganzeboom, 2010).

In plotting the results, we multiply coefficients to reflect a move of 40 steps along the scale (which goes from 10 to 90), roughly covering the interquartile range based on our sample. This amounts to contrasting, for example, occupations found around score 30 such as machine operators, service workers, or food vendors with those in the vicinity of 70 such as secondary school teachers, administrators, or engineers.

**Table 1.** Per cent retained reports from pupil and parental respondents, and available register data, as a proportion of pupils taking the cognitive test, separately for children of immigrant and non-immigrant background

| Sample | Parental education | | Parental occupation | |
|---|---|---|---|---|
| | Non-immigrant (per cent) | Immigrant (per cent) | Non-immigrant (per cent) | Immigrant (per cent) |
| England | | | | |
| Pupil report | 96 | 89 | 92 | 84 |
| Parent report | 41 | 22 | 37 | 19 |
| *N* non-immigrant = 2,994, *N* immigrant = 1,055, *N* total = 4,049 | | | | |
| Germany | | | | |
| Pupil report | 99 | 94 | 97 | 95 |
| Parent report | 79 | 69 | 72 | 51 |
| *N* non-immigrant = 3,060, *N* immigrant = 1,549, *N* total = 4,609 | | | | |
| The Netherlands | | | | |
| Pupil report | 99 | 90 | 97 | 92 |
| Parent report | 80 | 43 | 76 | 28 |
| *N* non-immigrant = 3,275, *N* immigrant = 859, *N* total = 4,134 | | | | |
| Sweden | | | | |
| Pupil report | 97 | 87 | 97 | 86 |
| Parent report | 64 | 37 | 63 | 34 |
| Register | 100 | 96 | – | – |
| *N* non-immigrant = 3,151, *N* immigrant = 1,565, *N* total = 4,716 | | | | |

### Sex, Immigrant Background, and Age

Information about pupils' sex, immigrant background, and age is based on the pupil's own report. We define 'immigrant background' as including all pupils whose parents were born abroad, or the only parent about which we know. In some analyses, we further distinguish between 'Western' and 'non-Western' immigrant background, the latter encompassing origins in the Middle East, South America, Africa, or Asia.

## How Much Nonresponse?

We are able to cover several types of nonresponse where selectivity is likely to be of consequence: item

**Table 2.** Correlation between pupil and parent reports of parental education, recoded to number of years, and differences in estimated average number of years of education, separately for children of immigrant and non-immigrant background

| Sample | Correlation | *n* child | *n* parent | *n* complete cases | $\bar{x}$ child | $\bar{x}$ child[a] | $\bar{x}$ parent |
|---|---|---|---|---|---|---|---|
| England | | | | | | | |
| Non-immigrant | 0.55 | 2,862 | 1,227 | 1,188 | 12.78 | 12.85 | 12.99 |
| Immigrant | 0.59 | 934 | 236 | 221 | 13.24 | 13.23 | 13.20 |
| Germany | | | | | | | |
| Non-immigrant | 0.51 | 3,015 | 2,404 | 2,374 | 13.61 | 13.63 | 12.85 |
| Immigrant | 0.37 | 1,460 | 1,065 | 1,013 | 12.75 | 12.81 | 12.85 |
| The Netherlands | | | | | | | |
| Non-immigrant | 0.61 | 3,247 | 2,632 | 2,616 | 12.69 | 12.69 | 12.47 |
| Immigrant | 0.47 | 771 | 366 | 336 | 11.94 | 12.14 | 12.33 |
| Sweden | | | | | | | |
| Non-immigrant | 0.46 | 3,053 | 2,022 | 1,977 | 13.46 | 13.61 | 13.49 |
| Immigrant | 0.42 | 1,360 | 585 | 528 | 12.73 | 12.77 | 12.97 |

| | Correlation reg.—child | correlation reg.—parent | correlation reg.—child[a] | *n* register | $\bar{x}$ register | $\bar{x}$ register[a] |
|---|---|---|---|---|---|---|
| Sweden | | | | | | |
| Non-immigrant | 0.44 | 0.72 | 0.46 | 3,140 | 13.20 | 13.51 |
| Immigrant | 0.36 | 0.53 | 0.42 | 1,495 | 11.67 | 12.31 |

[a]Calculated only over complete cases (i.e. complete parent–child pairs).

nonresponse among pupils, and unit and item nonresponse among parents.[4] After deleting pupils who did not complete the test and/or lacked data on key variables other than SES, we are left with $N = 17{,}508$ (94 per cent). Table 1 shows the number of valid, retained reports of parents' education and occupation as a proportion of this workable sample. Item nonresponse—failure to answer items about SES—is the cause of pupil nonresponse reported in Table 1. For parents, on the other hand, missing data are almost entirely due to unit nonresponse: failure to respond altogether.

The rates are satisfactory at the child level, both for education and for occupation, whereas the numbers for parents are more dire, especially for immigrant-background respondents. Although some of these numbers are discouraging, they are not exceptional compared with similar surveys. That ethnic minorities tend to have lower response rates is also well documented (Feskens *et al.*, 2006; Laganà *et al.*, 2013). Appendix B reports on the test score difference for pupils with missing and present SES information, indicating the selectivity of nonresponse. These differences are mostly in the range of 0.25–0.60 standard deviations, and similar to those that Jerrim and Micklewright (2014: Table 3) report for PISA and PIRLS.

Response rates are somewhat lower for occupation than education, probably for several reasons (cf. Currie *et al.*, 1997: pp. 387–388). Father absence and/or unemployment would seem like potential explanations. They are unlikely to be important for our study, however, where the questionnaire explicitly asked about occupation last held if unemployed, and where our variable incorporates information on both parents. A more plausible explanation is that questions about occupation are open-ended. This will lead to higher nonresponse because of a greater respondent burden, and also to some answers being discarded as uncodeable.

## How Much Misreporting?

What is the extent of divergence between child and parent reports on SES variables? Tables 2 and 3 address this question, using Pearson's $r$ and differences in average years of schooling and occupational status, respectively. To ensure that our conclusions do not depend on what metrics or scale levels we use, we report alternative measures in Appendix C.

The leftmost column in Table 2 displays the correlations between children's and parents' reports of parental education (converted to years of schooling); the three columns to the right instead compare average number of years. At the bottom of the table, we find comparisons for Sweden obtained from registers. The correlations range from 0.46 to 0.61 for the majority population, and 0.37 to 0.59 for those with immigrant background. The two rightmost columns show that children report similar values on average compared with parents, so disagreement appears driven mainly by random noise. (The

**Table 3.** Correlation between pupil and parent reports of parental occupation, recoded to ISEI status, and differences in average occupational status, separately for children of immigrant and non-immigrant background

| Sample | Correlation | *n* child | *n* parent | *n* complete cases | $\bar{x}$ child | $\bar{x}$ child[a] | $\bar{x}$ parent |
|---|---|---|---|---|---|---|---|
| England | | | | | | | |
| Non-immigrant | 0.63 | 2,768 | 1,114 | 1,067 | 54.23 | 55.20 | 55.82 |
| Immigrant | 0.71 | 882 | 197 | 180 | 50.77 | 50.53 | 48.99 |
| Germany | | | | | | | |
| Non-immigrant | 0.75 | 2,967 | 2,195 | 2,152 | 47.66 | 48.11 | 49.14 |
| Immigrant | 0.61 | 1,465 | 785 | 753 | 38.24 | 36.24 | 35.90 |
| The Netherlands | | | | | | | |
| Non-immigrant | 0.69 | 3,183 | 2,480 | 2,435 | 50.94 | 51.10 | 54.68 |
| Immigrant | 0.47 | 789 | 242 | 223 | 44.26 | 39.43 | 39.97 |
| Sweden | | | | | | | |
| Non-immigrant | 0.72 | 3,041 | 1,992 | 1,940 | 54.15 | 56.31 | 59.12 |
| Immigrant | 0.57 | 1,339 | 529 | 484 | 43.29 | 44.98 | 44.64 |

[a]Calculated only over complete cases (i.e. complete parent–child pairs).

only marked difference is a slightly higher average among children compared with parents in the German majority group.)

The comparison with registers in the Swedish case reveals that reliability is higher for parent than pupil reports, confirming a long-standing assumption (provided that register information is somewhat of a 'gold standard'). When we look at the immigrant-background sample, register data are less closely aligned with parent reports. We hesitate to interpret this as signalling lower reliability among foreign-born parents, as register information is less reliable in their case.

Table 3 shows inter-reporter correlations for parental occupation (ISEI-08). These tend to be stronger than the ones obtained for parental education, from 0.63 to 0.75 among the majority, and 0.47 to 0.71 for immigrant-background pupils. The difference in average values between pupils and parents is again minor, especially when we take the differing scales into account: the standard deviation of education is around 1.5–3 years depending on the country, and on the order of 20-scale points for occupation, so the differences are in most cases <0.10 standard deviations. Finally we observe that across countries and measures, inter-reporter correlations tend to be slightly lower for those of immigrant background, so we cannot wholly dispel the worry that measures of SES are less accurate here.

## Bias from Misreporting and Nonresponse—The Swedish Case

We have seen evidence of more misreporting among pupils, but a larger amount of missing data, and hence greater worry about selectivity bias, among parents. We expect both these errors to downwardly bias the gradient of SES on outcomes. For estimates of the immigrant–non-immigrant gap, we expect it to widen when error-ridden child reports of SES are used, but narrow when observations with nonresponding parents are discarded. To test this, we estimate a linear OLS regression of the child's cognitive test scores on SES, immigrant background, and control variables:

$$s_{ij} = \alpha + \Sigma_k \beta_k SES_i + \Sigma_m \gamma_m Imm_i + \delta\ sex_i + \lambda\ age_i + \mu_s + \varepsilon_{ij}, \tag{2}$$

where $s_{ij}$ is the test score of individual $i$ in school class $j$, $SES_i$ is a set of indicators for parental education (secondary education, university; primary education or less being the reference), $Imm_i$ is the measure of immigrant background (Western, non-Western; no immigrant background as reference), $sex_i$ is the pupil's sex, $age_i$ is age in months, $\mu_s$ is a set of intercepts for the different sampling strata, $\varepsilon_{ij}$ is an individual error term clustered on school classes, and $\alpha$, $\beta$, $\gamma$, $\delta$, and $\lambda$ are (sets of) parameters to be estimated.

The model is estimated separately for each source of SES: child, parent, and register. Figure 3 shows the resulting gradients for parental education and immigrant background. As SES and immigrant background enter simultaneously into the model, their respective estimates are net of each other. What is allowed to vary are the sources of SES and the sample selection. The top panels use only cases with information from all three sources (complete cases), whereas bottom panels use all available cases for each source (full sample).

In the top left panel, we see no gradient according to parental education from pupil reports, whereas results

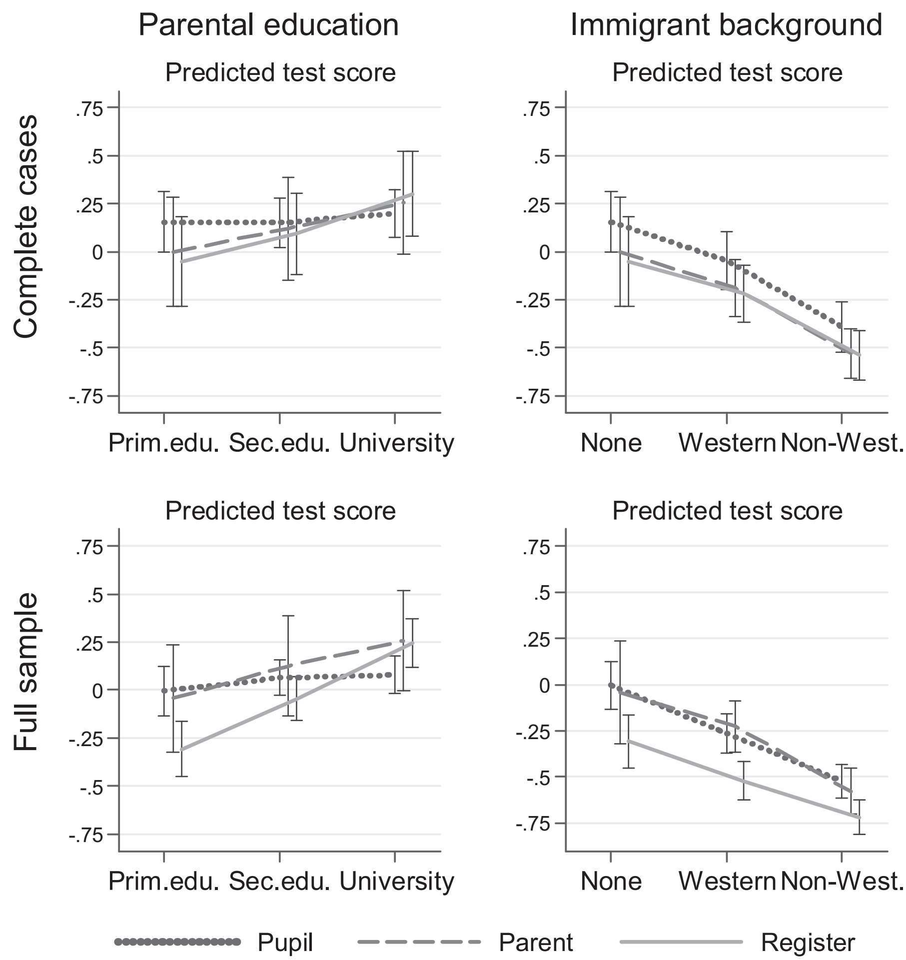


**Figure 3**. Sweden. Predicted mean z-standardized test scores with cluster robust 95 per cent confidence intervals from linear regression on parental education and immigrant background (each entered as three dummy variables), controlling for sex, age, and stratum. Graphs in left panels display estimated coefficients for parental education (holding immigrant background constant at 'none'), those in right panels the estimated immigrant–non-immigrant gap (holding parental education constant at 'primary'). Graphs in top panels use only cases with information from all three sources (complete cases), those in bottom panels use all available cases for each source (full sample). Top panel $N = 2,483$; bottom panel $N = 4,413$ (pupils), 2,607 (parents), 4,635 (register).

based on parent reports and registers both show a visible gradient in the expected direction (albeit statistically non-significant).

In the bottom left panel, we instead let the sample size vary by each source of SES (child, parent, and register) using all available data for each. The intercept drops dramatically for both pupil reports and register data, reflecting the selectivity of parental nonresponse shown in Table 1. Results from parent reports stay unchanged, as this analysis uses much the same individuals as the previous one. The zero association using pupil reports persists despite the increased sample. In contrast, the gradient from register data gains in size. Point estimates for secondary and university educated are 0.26 and 0.55 vis-à-vis the reference group of lower education, whereas corresponding figures from parent reports are 0.17 and 0.30. These are non-trivial differences. Note also the greater efficiency of register estimates; the wider confidence spans around parent estimates do not allow us to reject the null of no association between parental education and test scores, whereas the slope for register data is clearly significant.

Our next question is how misreporting and nonresponse in SES impact on the estimated immigrant–non-immigrant gap, as reported in the right-hand panels of Figure 3. Estimates for immigrant background appear surprisingly robust to both, despite the fact that nonresponse is substantial and larger among immigrant parents. At first this appears puzzling, as the conventional wisdom is that selection on the outcome should weaken all coefficients equally. Our analyses discard all cases where an SES report was not retained, so selectivity would seem to be just as much of a problem for the ethnic gradient.

However, nonresponse bias can be stronger for some regressors than others, namely, if based on an *interaction* of a regressor with the outcome, and this is in fact what we find. Not only is parental selection into the sample a positive function of the child's cognitive score; this selectivity is *stronger* among pupils with low-SES parents. Consequently low-SES pupils in the sample will be more positively selected, less representative of their sub-group of the population, and closer in achievement to high-SES pupils than if selection was uniform across SES. It is the availability of registers that allows us to draw this conclusion.[5] We find no such interaction by immigrant background.

## Bias from Misreporting—Comparative Results

How do the consequences of selectivity compare in magnitude with those of misreporting that most previous studies have focused on? To answer this, we compare results for all four countries. We fit a model similar to Equation 2, but to display results more economically, we let parental education be captured by a single coefficient by assigning number of years in education, and use a binary immigrant background variable. Figures 4 and 5 show estimates for SES and immigrant background from these models.

Figure 4, where the SES gradients represent 5 years of parent's education, shows that ME is a problem, with point estimates from child reports biased as expected in England and, especially, the Netherlands. Consequences of sample selection appear somewhat less dramatic for other countries compared with Sweden, but as we have seen, most of this bias goes undetected when child reports are the only comparison. Again, estimates for the immigrant–non-immigrant gap are hardly affected at all.

Next, we replace parental education with occupation, ISEI-08 (multiplying the coefficients by 40). Figure 5 demonstrates that the discrepancies in estimates are smaller than for parental education. In fact, with few exceptions, our estimates for parental occupation would be almost identical whether we used information from parents or children.

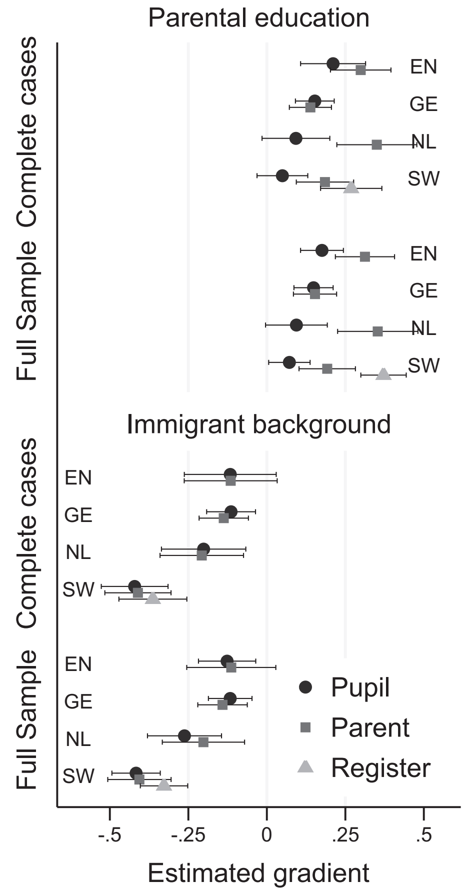


**Figure 4.** England, Germany, the Netherlands, and Sweden. Estimated gradients (with cluster robust 95 per cent confidence intervals): standardized cognitive test score regressed on parental education, immigrant background, and covariates (sex, age, and stratum). Parental education as reported by pupils, parents, and obtained from registers. SES estimates represent 5 years of education. *N* per country = 3,796–4,475 (full sample, pupil reports), 1,463–3,469 (full sample, parent reports), and 1,409–3,387 (complete cases).

The immigrant–non-immigrant gap is somewhat smaller when using parental occupation as compared with education. This could be due to the finer-grained measure or because immigrants hold jobs that do not match their education. But the weakening of the immigrant gradient is largely uniform across countries and so the ordering of countries is also here robust.

## Conclusions and Discussion

We asked what role error in measuring socio-economic background plays for analyses of social and ethnic inequality, as exemplified by between-group differences in

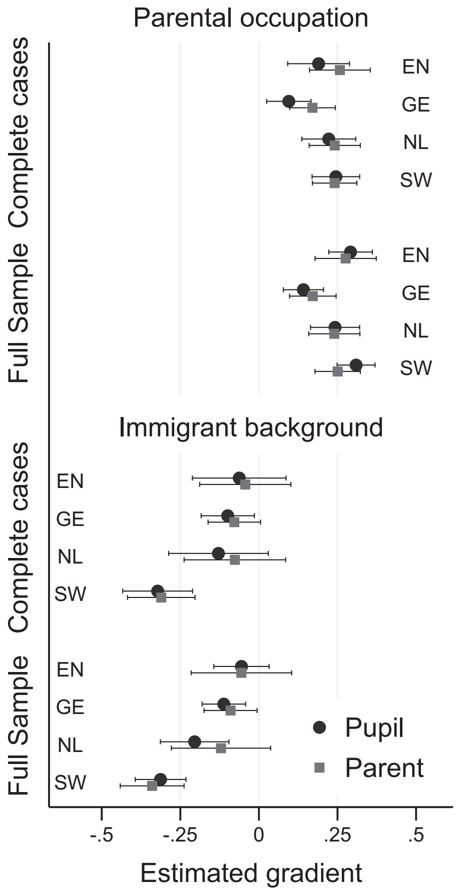


**Figure 5.** England, Germany, the Netherlands, and Sweden. Estimated gradients (with cluster robust 95 per cent confidence intervals): standardized cognitive test score regressed on parental occupation, immigrant background, and covariates (sex, age, and stratum). Parental occupation as reported by pupils and parents. SES estimates represent a move of 40 steps along the ISEI-08 scale. *N* per country = 3,650–4,432 (full sample, pupil reports), 1,311–2,980 (full sample, parent reports), and 1,247–2,905 (complete cases).

pupil's cognitive test scores. With the aid of data from pupils and parents from four European nations—England, Germany, the Netherlands, and Sweden—we estimate the difference between children's and parents' reports of parental education and parental occupation. For Sweden, we also have access to register data on parental education with almost total coverage, making it possible to address the question of nonresponse bias in SES gradients.

We find, in line with previous research, that children's reports are less aligned with parents' own reports for education than for occupation—suggesting that reports on occupations are more accurate and should be preferred. In fact, so good is the resemblance between parent and pupil reports that conclusions about SES gradients would be almost identical irrespective of who gave the information. Codeable pupil responses on parental occupation are also well over 90 per cent for non-immigrants (and in the high 80 s or above for those with immigrant background), making the more cost-effective pupil reporting an attractive option.

This good news is counterbalanced by some bad news: nonresponse from parents tends to be selective in a way that leads to a downward bias in SES estimates. When we 'thought away' the parent nonresponse in our Swedish data by using register data on parental education for all pupils, the regression coefficients got appreciably stronger—to a large extent because of the inclusion of nonresponding parents. This makes intuitive sense, as we know that nonresponse is rarely random, but it is a virtue of our analysis to show just how large this bias can be, in our case reducing coefficients by almost half their size. Importantly, we have shown that the bias due to parental nonresponse will go largely undetected when pupil reports serve as the only point of comparison.

We went further in our analyses, using SES background not only as a predictor of cognitive test results but also as a control variable. This is important in several analytical designs, perhaps particularly—as we used it here—in estimating the gradient of immigrant background. Here, we are back to good news. Even in the cases where we found downward bias—namely, in the gradient of parental education—this did not affect the regression coefficient for immigrant background much. That is, even when we rely on pupils' (what we believe are relatively inaccurate) reports on parental education in controlling for SES, our conclusion about immigrant–non-immigrant differences would not be severely affected. However, our estimate of the average level of cognitive ability would change. In our data, the selective nonresponse means that cognitive test scores are over-estimated—although to an equal extent for each immigrant-status group.

Is there some bigger lesson to be learnt from our analyses? We believe that the result that parental occupation is a relatively reliable measure when reported by 14-year-olds, is of great practical value—and, while this is a result supported by previous research, by showing this convincingly on new data for four countries we think our study brings to bear on the issue of how to construct school questionnaires. The flip side, but equally important, is the lesson that there is less correspondence in reports on parental education. Together, the results suggest that school-based surveys would benefit from including questions on parental occupation. But perhaps as important is the result that no matter our control variable, the estimate of immigrant background was not severely affected—this is something, then, that will strengthen analyses in that field of research (Heath and Brinbaum, 2014).

Finally, one important but not really resolved matter is what use we may have from parental information as collected by surveys, if, as commonly is the case, they come with low response rates. We have shown that such nonresponse will bias SES estimates, to an unknown but certainly troubling extent, and one way of seeing it is that it is probably not worth the while (or the cost) if the main information one wants is an indicator of parental SES. The nonresponse has the potential of also biasing international comparisons because it tends to vary across countries. Normally, of course, researchers elicit also other types of information from parents (such as their attitudes) that cannot be given by the responding child. But as long as this other information shows the same type of selectivity, even these reasons may not be enough for complicated, expensive, and unsuccessful attempts at surveying parents.

## Notes

1 In practice, some of these conditions cannot hold whenever the regressor is categorical or, more generally, bounded. Nevertheless, across a wide range of circumstances, the consequence of ME is still one of attenuation, so the classical model remains a useful tool for intuition.

2 In descriptives, we display unweighted statistics not accounting for sampling design. In OLS regression analyses, we account for the sampling design by including a set of dummy variables for our four sampling strata. This strategy is efficient and unbiased given that the model suffers from no omitted variables (Winship and Radbill, 1994). Alternative specifications indicate that our cross-country comparisons of the immigrant–non-immigrant gap are not entirely unaffected by the weighting strategy used, but the differences are slight enough that the methodological conclusions we draw are unaffected.

3 We take some care to ensure that, as far as possible, aggregated pupil and parent variables refer to the same person. The motivation is that we want discrepancies across sources to reflect underlying differences in reliability rather than being artefacts of survey design. For various reasons, some parents did not receive or answer a module about the other parent's characteristics (Kalter *et al.*, 2013). If so, we use only pupil reports that refer to the responding parent (most often the mother). Conversely, because pupils were asked about biological parents, we ignore parent reports in case they refer to a non-biological parent, such as a step-parent. These modifications turn out to be largely unimportant for our regression estimates, but allow us to gain greater precision in the calculation of inter-reporter reliabilities.

4 That we are restricted to pupils for whom we have a measure of the outcome (the test score) means that we cannot say anything about those who did not take part in the survey owing to unit nonresponse on the school, class, or pupil level. As success in school recruitment differed markedly between our four countries, with Sweden on top and England in the bottom (Kalter *et al.*, 2013), we put less trust in the substantive pattern of estimated coefficients across countries than in the pattern across different indicators *within* countries, which is our main focus.

5 Although it is not apparent from the figure, the gap in cognitive scores in Sweden depending on the status of the parent report (present or missing) is roughly 0.15 standard deviations wider for those whose parents *actually* have less than university education compared with the rest (assuming again that register information represents true values).

## Acknowledgements

For valuable comments on earlier versions of this article, the authors thank participants at the European Consortium for Sociological Research (ECSR) Spring School on the Analysis of Educational Inequalities, Moncalieri, March 2013, the NORFACE Interdisciplinary Conference on Migration, London, April 2013, the Spring meeting of Research Committee 28 on Social Stratification and Mobility (RC28) of the International Sociological Association, Trento, May 2013, as well as seminar participants at Stockholm University and the Swedish Institute for Social Research. The authors also thank two anonymous reviewers for comments that greatly improved the final draft.

## Funding

This work was supported by NORFACE research programme on Migration in Europe: Social, Economic, Cultural and Policy Dynamics; FAS, Swedish Council for Working Life and Social Research [2009-1320, 2012-1741]; and VR, the Swedish Research Council [421-2012-5598].

## References

Andersen, A. *et al.* (2008). High agreement on family affluence between children's and parents' reports: international study of 11-years-olds. *Journal of Epidemiology and Community Health*, **62**, 1092–1094.

Berk, R. A. and Ray, S. C. (1982). Selection biases in sociological data. *Social Science Research*, **11**, 352–398.

Black, D., Sanders, S. and Taylor, L. (2003). Measurement of higher education in the census and current population survey. *Journal of the American Statistical Association*, **98**, 545–554.

Bound, J., Brown, C. and Mathiowetz, N. (2001). Measurement error in survey data. In Heckman, J. J. and Leamer, E. (Eds.), *Handbook of Econometrics*, Vol. 5 Ch. 59. Amsterdam: Elsevier.

Breen, R. and Jonsson, J. O. (1997). How reliable are studies of social mobility? An investigation into the consequences of errors in measuring social class. *Research in Social Stratification and Mobility*, **15**, 91–112.

Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: recent research on educational attainment and social mobility. *Annual Review of Sociology*, **31**, 223–243.

Currie, C. *et al*. (1997). Indicators of socioeconomic status for adolescents: the WHO health behaviour in school-aged children survey. *Health Education, Research*, **12**, 385–397.

Currie, C. *et al*. (2008). Researching health inequalities in adolescents: the development of the health behaviour in school-aged children (HBSC) family affluence scale. *Social Science and Medicine*, **66**, 1429–1436.

Ensminger, M. E. *et al*. (2000). The validity of measures of socioeconomic status of adolescents. *Journal of Adolescent Research*, **15**, 392–419.

Feskens, R. *et al*. (2006). Collecting data among ethnic minorities in an international perspective. *Field Methods*, **18**, 284–304.

Ganzeboom, H. B. (2010). *International Standard Classification of Occupations ISCO-08 with ISEI-08 Scores*, Version of July 27 2010, available from: <http://www.harryganzeboom.nl/isco08/isco08_with_isei.pdf> [accessed 11 November 2013].

Ganzeboom, H. B. and Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, **25**, 201–239.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, **70**, 646–675.

Groves, R. M. and Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: a meta–analysis. *Public Opinion Quarterly*, **72**, 167–189.

Hanushek, E. A. and Woessmann, L. (2011). Sample selectivity and the validity of international student achievement tests in economic research. *Economics Letters*, **110**, 79–82.

Heath, A. F. and Brinbaum, Y. (Eds.) (2014). *Unequal Attainments. Ethnic Educational Inequalities in Ten Western Countries*. Oxford: Oxford University Press for the British Academy.

Jerrim, J. and Micklewright, J. (2014). Socioeconomic gradients in children's cognitive skills: are cross-country comparisons robust to who reports family background? *European Sociological Review*, **30**, 766–781.

Kalter, F. *et al*. (2013). *Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU)*. GESIS Data Archive, Cologne. doi: 10.4232/cils4eu.5353.1.0.0.

Klein, D. (2014). KALPHA: stata module to compute krippendorff's alpha-reliability. *Statistical Software Components*. http://EconPapers.repec.org/RePEc:boc:bocode:s457862.

Kreuter, F. *et al*. (2010). Pupil reports of parents' education level: does it matter whom you ask and what you ask about? *Survey Research Methods*, **4**, 127–138.

Krippendorff, K. (2010). Krippendorff's alpha. In Salkind, N. J. (Ed.), *Encyclopedia of Research Design*. Thousand Oaks: Sage, pp. 669–674. doi: 10.4135/9781412961288.n206.

Kvist, A. V. and Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: a test of Cattell's investment theory. *Intelligence*, **36**, 422–436.

Laganà, F. *et al*. (2013). National minorities and their representation in social surveys: which practices make a difference? *Quality and Quantity*, **47**, 1287–1314.

Le Donné, N. (2014). European variations in socioeconomic inequalities in students' cognitive achievement: the role of educational policies. *European Sociological Review*, **30**, 329–343.

Lien, N. *et al*. (2001). Adolescents' proxy reports of parents' socioeconomic status: how valid are they? *Journal of Epidemiology and Community Health*, **55**, 731–737.

Looker, E. D. (1989). Accuracy of proxy reports of parental status characteristics. *Sociology of Education*, **62**, 257–276.

Micklewright, J., Schnepf, S. V. and Skinner, C. (2012). Nonresponse biases in surveys of school children: the case of the English programme for international student assessment (PISA) samples. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, **175**, 915–938.

Nordahl, H. *et al*. (2011). Measurement of ethnic background in cross-national school surveys: agreement between students' and parents' responses. *Journal of Adolescent Health*, **49**, 272–277.

OECD. [The Organisation for Economic Co-operation and Development]. (2012). *PISA 2009 Technical Report*. OECD Publishing, available from: <http://dx.doi.org/10.1787/9789264167872-en>.

Parameshwaran, M. and Engzell, P. (2014). Ethnicity in England: What parents' country of birth can and can't tell us about their children's ethnic identification. *Journal of Ethnic and Migration Studies*, doi: 10.1080/1369183X.2014.920690.

Statistics Sweden. (2011). *SCBDOK 3.2: Befolkningens utbildning (Utbildningsregistret, UREG) [Education of the Population (The Register of Education, UREG)]. UF0506, UF0537*, available from: <http://www.scb.se/Statistik/UF/UF0506/_dokument/UF0506_DO_2011.pdf> [accessed 11 November 2013].

Traynor, A. and Raykov, T. (2013). Household possessions indices as wealth measures: a validity evaluation. *Comparative Education Review*, **57**, 662–688.

van de Werfhorst, H. G. and Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: a comparative perspective. *Annual Review of Sociology*, **36**, 407–428.

Vereecken, C. A. and Vandegehuchte, A. (2003). Measurement of parental occupation: agreement between parents and their children. *Archives of Public Health*, **61**, 141–149.

Wardle, J., Robb, K. and Johnson, F. (2002). Assessing socioeconomic status in adolescents: the validity of a home affluence scale. *Journal of Epidemiology and Community Health*, **56**, 595–599.
Weiß, R. H. (2006). *Grundintelligenztest Skala 2 Revision (CFT–20R)*. Göttingen, Germany: Hogrefe.
West, P., Sweeting, H. and Speed, E. (2001). We really do know what you do: a comparison of reports from 11 year olds and their parents in respect of parental economic activity and occupation. *Sociology*, **35**, 539–559.
Winship, C. and Radbill, L. (1994). Sampling weights and regression analysis. *Sociological Methods and Research*, **23**, 230–257.

## Appendix A

Mapping of ISCED level to years of education. Adapted from OECD (2012: pp. 364).

| | ISCED 1 | ISCED 2 | ISCED 3A or 4 | ISCED 5A or 6 |
|---|---|---|---|---|
| England | 6 | 9 | 13 | 16 |
| Germany | 4 | 10 | 13 | 18 |
| The Netherlands | 6 | 10 | 12 | 16 |
| Sweden | 6 | 9 | 12 | 15.5 |

## Appendix B

Average (countrywise z-standardized) cognitive test score for children with report on parental characteristic present, report missing, and the difference between the two; separately by characteristic (parental education, parental occupation) and source (pupil report, parent report).

| | Parental education | | | Parental occupation | | |
|---|---|---|---|---|---|---|
| | Present | Missing | Difference | Present | Missing | Difference |
| England | | | | | | |
| Pupil report | 0.03 | –0.46 | **–0.49** | 0.07 | –0.57 | **–0.64** |
| Parent report | 0.19 | –0.09 | **–0.28** | 0.20 | –0.08 | **–0.28** |
| Germany | | | | | | |
| Pupil report | 0.03 | –0.59 | **–0.62** | 0.03 | –0.43 | **–0.46** |
| Parent report | 0.08 | –0.19 | **–0.27** | 0.10 | –0.16 | **–0.26** |
| The Netherlands | | | | | | |
| Pupil report | 0.01 | –0.31 | **–0.32** | 0.02 | –0.46 | **–0.48** |
| Parent report | 0.12 | –0.30 | **–0.42** | 0.14 | –0.25 | **–0.39** |
| Sweden | | | | | | |
| Pupil report | 0.04 | –0.50 | **–0.54** | 0.05 | –0.60 | **–0.65** |
| Parent report | 0.19 | –0.23 | **–0.42** | 0.21 | –0.23 | **–0.44** |

## Appendix C

Krippendorff's alpha between pupil and parent reports of parental characteristics (cf. Tables 2–3).

In the main text, we used the simple linear correlation (Pearson's $r$) as an indicator of inter-reporter agreement. It has the virtue of being well known, easily interpretable by researchers in the field, and comparable to much previous literature (Looker, 1989). As a measure of agreement, however, it might not be ideal. Therefore, we here display inter-reporter agreement in an alternative metric, Krippendorff's alpha (Krippendorff, 2010).

Alpha is a generalization of several commonly used reliability indices, and as such, it accommodates data on any scale level. We are therefore able to assess agreement both for years of education and the ISCED level as an ordinal variable. In addition, we also display alpha on the nominal-scale level, and here we recode the occupation variable to five classes using the first digit of the ISCO-08 code, similarly to Jerrim and Micklewright (2014); our classes comprise digits 1, 2, 3, 4–6, and 7–9, respectively.

Krippendorff's alpha in most circumstances ranges from 0 to 1, where 0 represents no better than chance agreement, and 1 perfect agreement. It may also take on negative values if agreement is worse than that expected by chance. For interval data, it is comparable with Pearson's intraclass correlation (not the product-moment correlation $r$), for ordinal data with the Spearman rank correlation, and for nominal data with Scott's $\pi$ (Krippendorff, 2010).

Here we report alpha coefficients for parent and child reports on parental education and parental occupation, obtained using Stata code written by Klein (2014). The interval- and ordinal-level coefficients are similar to results in Table 2–3. The nominal coefficients for alpha are generally lower, but in most cases, higher for occupation than education. We conclude that the greater agreement on occupation is not an artefact of the detail of measurement, scale level, or metric used.

| | Parental education | | | Parental occupation | | |
|---|---|---|---|---|---|---|
| | Interval α | Ordinal α | Nominal α | Interval α | Ordinal α | Nominal α |
| England | | | | | | |
| Non-immigrant | 0.54 | 0.59 | 0.36 | 0.63 | 0.62 | 0.39 |
| Immigrant | 0.58 | 0.66 | 0.41 | 0.71 | 0.70 | 0.50 |
| Germany | | | | | | |
| Non-immigrant | 0.48 | 0.41 | 0.23 | 0.75 | 0.74 | 0.57 |
| Immigrant | 0.36 | 0.32 | 0.16 | 0.61 | 0.57 | 0.54 |
| The Netherlands | | | | | | |
| Non-immigrant | 0.59 | 0.57 | 0.53 | 0.67 | 0.67 | 0.44 |
| Immigrant | 0.41 | 0.38 | 0.24 | 0.47 | 0.48 | 0.36 |
| Sweden | | | | | | |
| Non-immigrant | 0.45 | 0.45 | 0.38 | 0.71 | 0.68 | 0.43 |
| Immigrant | 0.38 | 0.43 | 0.24 | 0.57 | 0.55 | 0.39 |