DISCUSSION PAPER SERIES

IZA DP No. 8392

Socioeconomic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?

John Jerrim
John Micklewright

August 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Socioeconomic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?

John Jerrim
IOE, University of London

John Micklewright
IOE, University of London
and IZA

Discussion Paper No. 8392
August 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8392
August 2014

ABSTRACT

Socioeconomic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?*

The international surveys of pupil achievement – PISA, TIMSS, and PIRLS – have been widely used to compare socioeconomic gradients in children's cognitive abilities across countries. Socioeconomic status is typically measured drawing on children's reports of family or home characteristics rather than information provided by their parents. There is a well established literature based on other survey sources on the measurement error that may result from child reports. But there has been very little work on the implications for the estimation of socioeconomic gradients in test scores in the international surveys, and especially their variation across countries. We investigate this issue drawing on data from PISA and PIRLS, focusing on three socioeconomic indicators for which both child and parental reports are present for some countries: father's occupation, parental education, and the number of books in the family home. Our results suggest that children's reports of their father's occupation provide a reliable basis on which to base comparisons across countries in socioeconomic gradients in reading test scores. The same is not true, however, for children's reports of the number of books in the home – a measure commonly used – while results for parental education are rather mixed.

JEL Classification: C21, C81, I24

Keywords: educational inequality, socioeconomic status, measurement error, international comparisons, PISA, PIRLS

Corresponding author:

John Micklewright
Department of Quantitative Social Science
Institute of Education
University of London
20 Bedford Way
London WC1H 0AL
United Kingdom
E-mail: j.micklewright@ioe.ac.uk

* We thank the Russell Sage Foundation for funding on a related project which led to the idea for this paper. The work was partly funded by the European Union FP7 GINI project. Helpful comments were made at a GINI workshop at the LSE and by seminar participants at the Institute of Education. We thank Dick Wiggins for his comments on a draft. Micklewright thanks the Melbourne Institute at the University of Melbourne for hospitality during a stay as 2012 Downing Fellow when he worked on the paper.

1. Introduction

There has long been interest in how child outcomes are associated with family background. Parents with better education, higher occupations and greater income can invest more in their offspring and provide other stimuli to child development (e.g. Haveman and Wolfe 1995). The international studies of children's cognitive achievement – the Programme for International Student Assessment (PISA), the Progress in International Reading Literacy Study (PIRLS), and the Trends in International Mathematics and Science Study (TIMSS) – have led to increased interest in how and why these socioeconomic gradients in one important outcome vary across countries. Countries can be placed in a comparative context, showing whether the gradient in any one country is particularly large, and the impact of differences between countries in institutional structures and policies that may limit disadvantaged children's achievement can be explored. There has been research into variation in socioeconomic gradients at a given age (e.g. OECD 2001, Marks 2005, Schütz et al. 2008, Woessmann 2008, Jerrim 2012, Hermann and Horn 2011) and whether gradients steepen as children move through their schooling (e.g. Ammermueller 2006, Waldinger 2007, Jakubowski 2010, Jerrim and Micklewright 2012). See also the reviews in Hanushek and Woessmann (2011).

The information on family background that has been used by analysts of these surveys is typically collected from the children rather than from their parents. (The children are aged 9/10 in PIRLS, 9/10 and 13/14 in TIMSS, and 15 in PISA.) All three surveys are conducted via schools and parental questionnaires are either only an optional supplement that countries can include if they wish (and with their own problems of non-response), or, in the case of TIMSS and in early rounds of PISA and still in later rounds for many countries, are not part of the survey at all. There is a well-established literature drawing on other datasets which shows that children may report their parents' and other home characteristics with error. Looker (1989) suggests that children have particular problems reporting parental education. Lien et al. (2001) find that agreement between parental and child reports is reasonably good for occupation, but only 'fair' for education. West et al. (2001) take an optimistic view when considering occupation but emphasise the importance of the fieldwork conditions when collecting the data. Buchmann (2002), citing work by Koretz (1992), notes that the correlation between parent and child reports of home possessions can be low, and that discrepancies 'raise serious questions about the accuracy of these data' (p181). If children's reports of their family background are inaccurate, and this information is used to obtain

estimates of socioeconomic gradients in cognitive achievement, then bias in the estimates can be expected.

In this paper we consider the implication of using the children's reports to estimate the variation across countries in socioeconomic gradients. Research on measurement error in family background in the international achievement surveys has either used data from small scale field trial surveys of parents simply to illustrate the association between parent and child reports (e.g. Adams and Wu 2002, Schulz 2006), or in the important work by Kreuter et al. (2010) on the reporting of parental education, has analysed just one country (Germany). Although insightful, these investigations do not show the differences across countries and do not compare different SES measures. The purpose of this paper is to fill these gaps.

We consider three measures of socioeconomic status (SES): parental education, father's occupation, and the number of books in the family home. The first two are standard SES measures. The number of books is a less obvious measure, but has become an increasingly common SES proxy in cross-national research given its presence in all three of the international surveys of children's achievement, PISA, TIMSS and PIRLS, as well as other international surveys (e.g. Ammermueller 2006, Waldinger 2007, Schütz et al. 2008, Ammermueller and Pischke 2009, Machin 2009, Evans et al. 2010, Jakubowski 2010, Ermish and Del Bono 2010, Hermann and Horn 2011, Brunello et al. 2012, Jerrim and Micklewright 2012). We have some doubts about the validity of books in the home as a measure of SES but we do not pursue the issue in this paper, accepting that – rightly or wrongly – this is a widely used measure.

In Section 2 we discuss the circumstances under which measurement error in SES variables leads to bias in estimates of cross-national differences in SES gradients in test scores. Section 3 describes the data from the two international surveys we use, PIRLS and PISA. PIRLS has always included a parental questionnaire which provides information on books in the home that can be compared with child reports on this measure for most countries. In recent survey rounds PISA has collected reports from parents as well as children on parental education and occupation for a subset of countries.

Section 3 also describes our methods: we first assess the extent of parental and child agreement for the three SES measures. Then, for each measure, we compare the picture obtained of differences across countries in SES gradients in test scores from use in regression models of the parent reports with that obtained from use of the child reports. Since parent reports of SES may also be subject to error, we cannot be sure that the data reveal the extent of measurement error in the child reports or of its impact. Hence our emphasis is on

'robustness' of conclusions: does the switch from child to parent reports make a difference? Section 4 presents our results and Section 5 gives our conclusions and recommendations.

2. Measurement error and cross-national comparisons of SES gradients

What difficulties do measurement errors in SES cause in international comparisons of SES gradients in children's test scores? For ease of illustration, we assume a simple bivariate linear model:

$$Y_i = \alpha + \beta . X_i^* + \varepsilon_i \qquad (1)$$

where Y_i is the test score of child i, X_i^* is the child's true (perfectly measured) SES, and ε_i is a random term. The OLS estimate of β is given by:

$$\hat{\beta} = \frac{cov\,(X_i^*, Y_i)}{var\,(X_i^*)} \quad . \qquad (2)$$

where lower case 'cov' and 'var' refer to the sample values. Assume that we observe not X_i^*, but a misreported measure, X_i:

$$X_i = X_i^* + v_i$$

where v_i is the measurement error. The standard treatment of measurement error is to assume that it is 'classical' in form: X_i is just a noisy measure of SES, with v_i uncorrelated both with the true value X_i^*, and with ε_i (and hence Y_i). In this case (and suppressing the individual child sub-script i from now on), the OLS estimate becomes:

$$\hat{\beta} = \frac{cov\,(X^*, Y)}{var\,(X^*) + var(v)} \qquad (3)$$

The estimate of the SES gradient is biased downwards in absolute size – the textbook case of 'attenuation bias'. The proportional bias in the probability limit of the estimate of β depends

on the ratio of the population variances, $VAR(v)/VAR(X^*)$, the so-called noise-to-signal ratio (e.g. Cameron and Trivedi 2005:903).

If the child reports of SES in each country are subject to 'classical' error, then the estimates of the SES gradients in test scores based on these reports will all be biased downwards. Whether the estimated pattern of differences across countries in these gradients is biased depends on the cross-country variation in noise-to-signal ratios. Were these ratios all the same, estimates of the relative differences in the SES gradients would be unaffected. The estimate of the gradient in each country would be too low, but conclusions based on the data of the type 'the SES gradient in Country A is twice that in Country B' would be unaffected. In this case, one could assess appropriately the relative differences across countries despite the measurement error in the data.

However, it seems reasonable to expect some variation in the noise-to-signal ratios. For example, consider children's reports of parental education in England and Germany. There are comparatively few exit points from the school system in England, and a quite limited number of educational routes. Germany, on the other hand, has a greater number of different educational pathways and qualifications. As a result, there may be a higher probability of children misreporting parental education in Germany. Another example of variation in noise-to-signal ratios could be when different survey instruments are used to collect data in different countries. Despite the significant resources invested to ensure comparability in national instruments in international surveys, it is possible that national variations are not always perfectly captured. A more substantial problem may arise when using a set of country-specific surveys (e.g. the growing trend for comparisons between longitudinal datasets drawn from different countries that have not been designed to be comparable).

But a more fundamental objection is the reliance on the textbook assumption that measurement error in children's reports is 'classical', uncorrelated with the true value of SES or with the child's test score. In their major survey paper on measurement error, Bound et al. (2001) note that this assumption tends to 'reflect convenience rather than conviction' and argue that 'the possibility of non-classical measurement error should be taken much more seriously by those who analyze survey data'. Mason et al. (1976) discuss this issue of non-random measurement error with specific reference to children's reports of socio-economic characteristics. Kreuter et al. (2010) draw upon cognitive theory of response behaviour to argue that children who score highly on achievement tests are more likely to report parental SES measures correctly. Using a follow-up of the PISA 2000 sample in Germany which

collected SES information from parents, they find some evidence of this for parental education (assuming parents reports to be free of error). They then shed doubt on the assumption that reporting error is uncorrelated with the true SES value: children with the highest value of parental education as reported by their parents (a university degree), can only provide the same information or under-report, while those with the lowest values reported by parents (no qualifications) can only provide the same information or over-report. Hence Kreuter et al. argue that a more realistic assumption is that the higher the true value of SES, the more likely the reporting error to be negative.

What are the implications of these patterns of measurement error? Kreuter et al. (2010) note the expression for the OLS estimator in the presence of measurement error that is not restricted to be 'classical' in form:

$$\hat{\beta} = \frac{cov\ (X^*,Y) + cov(v,Y)}{var\ (X^*) + var(v) + cov\ (v,X^*)} \qquad (4)$$

Consider the single country setting. If, as found by Kreuter et al., more able children provide better reports, then $cov(v,Y) < 0$ and there is a downward impact on the estimate of β coming through the numerator of (4). And if, as just discussed, $cov(v,X^*) < 0$, an upward impact comes through the denominator. In this situation the estimated SES gradient in test scores may not be attenuated as in the classical case. Rather, it depends on the exact values of $cov(v,Y)$, $cov(v,X^*)$ and $var(v)$, with a downward bias in the estimate of β only ensured when $var(v) + cov(v,X^*) > 0$. The direction of bias in the estimated SES gradient becomes an empirical question.

What about the cross-national setting, when the goal is to compare SES gradients across countries? Unlike in the case of 'classical' measurement error, the form of the bias in the estimate of β for each country is not a multiplicative adjustment factor. So even in the very unlikely circumstances of all countries having the same values of $cov(v,Y)$, $cov(v,X^*)$, and $var(v)$, bias in the ratio of the SES estimates for any pair of countries would result. Even in this special case, a conclusion that 'the SES gradient in Country A is twice that in Country B' could not be drawn safely.

To this point we have considered the case of regression with a single explanatory variable. But in practice researchers often estimate SES gradients controlling for other individual or family characteristics or school-level variables. This further complicates cross-

country comparison. Returning to the case of 'classical' error, the attenuation bias increases, relative to the case with a single explanatory variable, the greater is the correlation between the true value of the mis-measured variable and the other variables (assumed free of measurement error) that are included in the regression (e.g. Bound et al. 2001: equation (5), Wooldridge 2002: 75). This correlation may well vary across countries. For example, the correlation between SES and school type will be stronger in countries with selective school systems. So even in this simpler case of 'classical' error and in the unlikely event that the noise-to-signal ratios are everywhere the same, the cross-country pattern of SES gradients in test scores estimated with a mis-measured SES variable is likely to be misleading.

In the rest of the paper we investigate empirically the problems that we have raised: we compare child and parent reports of SES, and how their differences correlate with test scores, and then compare estimates of SES gradients in test scores obtained with the child reports with those obtained with the parent reports. Although we have framed this investigation with discussion of children's reporting errors, we cannot rule out that the parents also report SES with error, rather than providing the true values, X^*. We would expect errors in parental reports to be smaller than those of the children, especially when the SES measures refer to their own individual characteristics. But the possibility of parental error means that the focus of our empirical investigation is more on the robustness of results: how do results differ when we switch between use of child and parent reports of SES?

3. Data and methods

Our data are drawn from the Progress in International Reading Literacy Study (PIRLS) and the Programme for International Student Assessment (PISA). PIRLS is a study of 4th grade (age 9/10) children while PISA focuses on 15 year olds. The surveys have two stage sample designs. First, schools are sampled with probability proportional to size. In PIRLS, one or two classes are then randomly selected from each institution, while in PISA a random sample of 35 children is taken in each school.[1] We use data from the 2001 and 2006 rounds of PIRLS and the 2006 and 2009 rounds of PISA. These data provide us with measures of SES reported separately by the study children and by their parents. By contrast, the Trends in International Mathematics and Science Study (TIMSS) collects information on SES from the study

[1] PIRLS is organised by the International Association for the Evaluation of Educational Achievement (IEA) and PISA by the Organisation for Economic Co-operation and Development (OECD). Average response rates for both surveys (and for both schools and pupils) are high (around 90%), although this varies moderately between countries.

children only so we cannot use this survey here. We restrict ourselves only to considering OECD member countries (it is these countries that form the core of PISA).

The purpose of both PIRLS and PISA is to test children's 'functional literacy' – how well they can use the skills examined in 'real life' situations. Children participating in PIRLS sit a one hour test in reading while respondents to PISA take a two hour test that covers reading, maths and science; we focus on results for reading.[2] Both surveys summarize children's answers to the test questions into a single score using an item-response model; the intuition is that true ability is unobserved and must be estimated from responses to the test. Five 'plausible values' of true reading proficiency are generated for each child. Scores are scaled by the organizers of both surveys to have a mean (across core participating countries) of 500 and a standard deviation of 100. In view of the large volume of data, we use the first of these plausible values throughout our analysis.[3] We report all our results in units of national z-scores: we subtract the country specific mean from the 'test score' (i.e. the first plausible value) and divide by the country specific standard deviation.

PIRLS has always had a parental questionnaire in addition to the questions asked of the study children. One or two countries however do not include the survey of parents – England in 2011 and the USA in all survey rounds – and response by parents has been notably low in several other countries, as we show later in this section. PISA administered no questions to parents in its first two rounds in 2000 and 2003. But ten OECD countries conducted a new 'home' questionnaire in 2006 to collect information from parents and nine countries did so in 2009, although there was again a substantial problem of non-response that we come to later. In both surveys, the most common pattern is for the child's mother to complete the parental questionnaire alone if parents do respond to the survey: this is true in around 70% of cases in most countries in PISA and 75% of cases in PIRLS (these figures refer to PISA in 2006 and PIRLS in 2001). It is unusual for both parents to complete the questionnaire together (typically about 10% in both surveys).

We consider in turn the information collected from parents and children on parental occupation and parental education (in PISA) and the number of books in the home (in PIRLS). Unfortunately, only children – but not parents – in PISA are asked about the number

[2] We re-ran a selection of our analyses using the maths scores. Our conclusions on the robustness of socio-economic gradients to who reports the family background characteristics were unchanged.

[3] We experimented by estimating models with each of the plausible values and then averaging the resulting parameter estimates, as recommended by the survey organisers. We found very little change in our results. (OECD (2009: 129) note that 'analysing one plausible value instead of five plausible values provides unbiased population estimates'.)

of books, and only parents – but not children – in PIRLS are asked about parental occupation and education. Hence we cannot compare the consistency of child and parent reports between the two surveys for the same SES measure. In both surveys, our understanding is that parents as well as children answer self-completion questionnaires – there are no interviewers to provide further guidance.

Parental occupation

Children in PISA are asked the following two questions about both their mothers' and fathers' occupations:

What is your [mother's / father's] main job? (e.g. school teacher, cook, sales manager)
- Please write in the job title

What does your [mother / father] do in his main job? (e.g. teaches high school students)
- Please use a sentence to describe the kind of work he does or did in that job

The 'home' questionnaire in 2006 asked identical questions to the children's parents. No information on occupation was asked of parents in 2009. We therefore have 10 countries where parent and child reports of parental occupation can be compared.

The descriptions given by respondents were converted by the survey organisers into four digit ISCO codes (the ILO categorization of occupations). [4] We recode the data in a similar manner to Marks (2005). This leads to five social class groupings: high professional, low professional, routine white collar, skilled manual, and semi-skilled or unskilled manual.[5] We also comment on results under possible alternatives, including the continuously measured occupational status index proposed by Ganzeboom and Treiman (1996). We focus on results for father's occupation, which has been widely used to study SES differences in the sociological literature. (We demonstrate in Appendix A that use of mother's occupation gives similar results.)

[4] Details can be found at www.ilo.org/public/english/bureau/stat/isco/isco88/publ4.htm. We assume that the parental and child reports are coded independently.

[5] These have been created using the Stata code written by Hendrickx (2004). Due to the small number of children with parents reported as farm workers, we include this category with the semi-skilled or unskilled manual category.

Parental education

Children in PISA were asked the following two questions about their mother's and father's educations. Country specific options were provided in the questionnaire (the phrases in brackets illustrate those for the USA):

What is the <highest level of schooling> completed by your [mother/father]?

- ISCED level 3A [She completed grade 12]
- ISCED level 3B/3C [N/A]
- ISCED level 2 [She completed grade 9]
- ISCED level 1 [She completed grade 6]
- She did not complete ISCED level 1 [She did not complete grade 6]

Does your [mother/father] have any of the following qualifications?

- ISCED level 6 [Masters, doctoral, or professional degree such as medicine]
- ISCED level 5A [Bachelor degree – a 4 year college degree]
- ISCED level 5B [Associate degree – 2 year degree from a community college]
- ISCED level 4 [Vocational or technical certificate/diploma after high school]

The 'home' questionnaire in 2006 and in 2009 asked parents directly about their own educational qualifications:

Does the child's [mother/father] have any of the following qualifications?

- ISCED level 5A/6
- ISCED level 5B
- ISCED level 4
- ISCED level 3A

Note that parents were not asked to discriminate between levels of education below ISCED 3A (a level that implies the parent completed high school in the USA). We therefore combine children's reports of parental education below ISCED level 3A into a single group so we have the same set of five categories for child and parent reports: ISCED level 5A/6, level 5B, level 4, level 3A, below level 3A. We then calculate a 'highest parental education' variable which is based upon the highest qualification achieved by either of the child's parents.[6] (Appendix B presents analysis where maternal and paternal education are considered separately – the two variables give similar results.)

[6] We focus on the highest education level obtained by either parent (rather than focus on mother's/father's education level separately) as this is the most widely used measure in the existing literature.

Books in the home

We have to switch to PIRLS data to investigate our third SES measure – the number of books in the family home.[7] Books in the home have been widely used in analyses of SES gradients in children's test scores based on PIRLS and TIMSS in particular, but also PISA. PIRLS is the only study among the three where information on books in the home is collected from both children and parents. Children in PIRLS are asked:

> *About how many books are there in your home? (Do not count magazines, newspapers, or your school books)*

- None or very few (0-10 books)
- Enough to fill one shelf (11-25 books)
- Enough to fill one bookcase (26-100 books)
- Enough to fill two bookcases (101-200 books)
- Enough to fill three or more bookcases (more than 200)

A diagram in the questionnaire accompanies each category of books. For example, the diagram for 26-100 books shows a full bookcase with four shelves.

The parents are invited to answer a very similar question:

> *About how many books are there in your home? (Do not count magazines, newspapers or children's books.)*

- 0-10 books
- 11-25 books
- 26-100 books
- 101-200 books
- More than 200 books

It is important to recognise that children in PIRLS reporting this information are younger than children reporting parental education and occupations in PISA. If we find that books in the homes is the least consistently reported SES measure when comparing child and parent reports, this could be due to younger children being less able to provide reports of home and

[7] There is debate over the extent to which books in the home is a measure of SES or of 'scholarly culture' (Evans et al. 2010) which we do not enter here.

family characteristics than are older children (see Looker 1989) rather than to the number of books in the home being particularly difficult for children to estimate.[8]

In all three cases – occupation, education, and books in the home – there are limits in practice to the assessment of measurement error in children's reports by comparing their responses to those of their parents. First, the information provided by parents may also contain error (including as a result of proxy reporting e.g. the mother reporting the father's education). We expect this to be especially important for the number of books and least important for occupation. In practice we can therefore only investigate the *consistency* of children's and parents' reports of our SES measures.

Second, there is the problem of complex family structures. The PISA questionnaire provides the following guidance to responding children:

> *Some of the following questions are about your mother and father or those persons who are like a mother or father to you — for example, guardians, step-parents, foster parents, etc. If you share your time with more than one set of parents or guardians, please answer the following questions for those parents/guardians you spend the most time with.*

No such guidance is given in the parental questionnaire.[9] This could result in child reports referring to one person (e.g. the biological father) while parental reports refer to someone else (e.g. a step father) thus weakening the observed association between the two variables.

Finally, there is the practical matter of missing reports of SES, either from children or, much more commonly, from parents – notably due to non-response to the parental questionnaire. We show the extent of this problem in the PISA and PIRLS data for 2006 in Table 1.

< Table 1 here >

For father's occupation and parental education (in PISA) there is at least one report missing for over a quarter of the sample in 6 countries (occupation) and 5 countries (education) out of the 10 where the parental questionnaire was conducted. The figure varies considerably so that

[8] There is also a small difference in the question on books posed to parents and children, although we assume this is not important in practice.

[9] Rather the questionnaire simply states: 'In this section we ask questions about the background of both the mother and the father of the student'.

information is complete for father's occupation in only 48% of cases in Denmark but 92% in South Korea. (The countries are ranked within each panel in declining order of the percentage of cases with both child and parent reports.) The problem is driven by the figures for missing parental reports. The extent of this problem further increases the incentive for analysts of PISA data to use the child reports of SES, an incentive that is already high given the relatively few countries administering the 'home' questionnaire. Missing information is less common for books in the homes (in PIRLS) – the modal country has complete information for 89% of cases – although there are very large numbers of missing parental reports in a small number of countries: in 30% or more cases in the Netherlands, Spain, New Zealand, Scotland, and England. Recall also that no parental questionnaire is issued in the USA.

Parents and children who fail to report SES may be systematically different to those who do provide this information. Table 2 presents the mean reading test scores for children with the parent SES report missing (left panel) or the child report missing (right panel) minus the mean score for children with complete information (both reports present), measured in national z-scores. (The figures again refer to 2006 data only.) All figures are negative: children with missing parent or child reports of SES have lower average test scores (the differences are almost always statistically significant).[10] The extent of this selectivity seems to vary across countries; for instance there is a stronger association between missing information on parental education and occupation data in Germany (0.7 to 0.8 standard deviations) than in Luxembourg (0.3 to 0.4 standard deviations).

< Table 2 here >

In principle, we could take into account the selectivity of the missing data problem. But the relevant methods rely on strong assumptions.[11] Consequently, we undertake a 'complete case' analysis, where we restrict our sample to those children where both parent

[10] PISA data for 2009 contain more information on family structure and show that (i) the great majority of children without a father present in the household do report their father's characteristics, and (ii) controlling for single parent status does not appreciably reduce the sizes of the differences in Table 2.

[11] If there were a variable within the dataset that could predict well non-response to the parental questionnaire but not the outcome of interest (children's reading achievement), one could use this as an instrument to correct for sample selectivity. However we do not see a credible candidate in either the PIRLS or PISA data. Alternatively, one could use multiple imputation to predict missing reports on the basis of other observable characteristics. This would hinge on the untestable 'Missing At Random' (MAR) assumption.

and child reports are available (cases in Table 1, column 1), recognising that this limits the generalisability of our results.[12]

Methods

We first document for each SES variable the frequency with which parents and children report the same category of the variable concerned. (Each variable has the same number of categories – five.) This measure of consistency does not take into account that parents and children may simply agree by chance. We therefore also present estimates of Cohen's Kappa, a measure of 'inter-rater reliability' that adjusts for chance agreement. Following Kreuter et al. (2010), we then consider whether children whose SES reports agree with those of their parents (i.e. same category reported) have test scores that differ on average from those of other children (giving an indication of whether *cov(v,Y) = 0* as discussed in Section 2).

Finally, we use a simple linear regression model to estimate the SES gradients in test scores for each country. The dependent variable is the child score on the PISA or PIRLS reading test, with SES – defined by one of our three measures – the covariate of interest. We include controls for children's age, gender, immigrant status, and the interaction of immigrant status with SES (dummy variables for missing information for the controls are also included – all cases included in the regressions have complete information on the SES variable in question). The estimated standard errors allow for the clustering of children within schools. This particular specification has been chosen for consistency with the existing literature (e.g. Schütz et al. 2008). Formally, we estimate the model:

$$T_{ij} = \alpha + \boldsymbol{\beta}'.\boldsymbol{SES_i} + \gamma_1.Gender_i + \gamma_2.Age_i + \gamma_3.Imm_i + \boldsymbol{\gamma_4'}.\boldsymbol{SES} * Imm_i + \epsilon_{ij} \qquad (5)$$

(i = child, j = school)

where *T* is the child's reading test score and *SES* is the measure of socio-economic status of interest. *Gender* is a dummy variable indicating the child's gender (reference = girl), *Age* is a continuous variable indicating child's age in months, and *Imm* is a dummy variable indicating the child is either a first or second generation immigrant. Since *SES* is a categorical variable it

[12] If the least able children are the most likely to mis-report measures of socio-economic status, then dropping these individuals from our analysis could lead us to overstate the extent to which estimates of the socio-economic gradient in children's reading achievement are robust.

is entered as a vector of dummy variables, with the most disadvantaged group as the reference group (e.g. semi/unskilled manual occupation). When using the books variable, we combine the bottom two categories (0-10 and 11-25 books) to form the reference group due to the sparse number of observations in the lowest category. The parameter estimate for the most advantaged group (e.g. high professional occupation) provides our estimate of the SES gradient in test scores (we check the robustness of our results to this choice).

We estimate the model twice for each country for each of our three SES measures of interest, once using child reports of the variable (education, occupation, or books) and once using parent reports. Comparison of the results from the two sets of regressions shows whether the cross-country pattern of SES gradients is robust to who reports SES – the child or the parent.

4. Results

The extent of agreement between parents and children

Table 3 shows the percentage agreement between parent and child reports for each of the three SES measures. Figure 1 gives box-and-whisker plots of Kappa statistics of inter-rater reliability. Kappa values can vary between 0, indicating no agreement, and 1, indicating perfect agreement. To aid interpretation of the estimated values, we follow the rules of thumb in Landis and Koch (1977), who suggested that 0.01 to 0.20 indicates 'slight' agreement, 0.21 to 0.40 'fair', 0.41 to 0.60 'moderate', 0.61 to 0.80 'substantial' and 0.81 to 0.99 'almost perfect' agreement. We are interested in two features of the results: (a) whether agreement between child and parent reports differs across the three SES measures, and (b) the extent of variation across countries for any given measure.

< Table 3 here >

< Figure 1 here >

For books in the home, the data points generally lie towards the left hand side of Figure 1, with most Kappa statistics less than 0.2, indicating only 'slight' agreement

according to the Landis and Koch classification.[13] Table 3 shows the percentage of children reporting the same category of number of books as their parents to average only 40%, and reaching 50% only in one country (Turkey). There are also signs of moderate variation across countries, with Kappa statistics ranging from 0.12 (the Netherlands in 2001) to 0.38 (Slovakia in 2006).

Parental education shows more consistency between parent and child reports. Most Kappa statistics exceed 0.40, suggesting 'moderate' agreement, and the percentage of children reporting the same category as their parents is above 50% everywhere, with an average of 63%. However, there is quite a lot of variation across countries.

The results for father's occupation are the most encouraging. Consistency between parent and child reports tends to be higher than for the other SES measures – Kappa statistics typically exceed 0.60, indicating 'substantial' agreement on the Landis and Koch rules of thumb, and percentage agreement averages 71%. There is only limited variation across countries. We also computed correlations between parent and child reports using an alternative variable for father's occupation – the continuous occupational status index developed by Ganzeboom and Treiman (1996). Values were reasonably high (averaging 0.69) and again showed only moderate variation across countries.

Our results show that agreement between parents and children is best for parental occupation and least good for books in the home.[14] The lower agreement for books may reflect the younger age of the children in PIRLS than in PISA (although Vereecken and Vandegehuchte (2003) find children only a little older than the PIRLS children giving reports on parental occupation that agreed well with those of their parents). Or it may reflect the difficulty people of all ages have in estimating numbers of books, whether a child or an adult, implying parental as well as child error is to be expected. By contrast, we would expect the least parental error – including error due to proxy responses by mothers – for reports of father's occupation.

Differential measurement error

[13] The graph presents estimates of Kappa statistics that do not account for the ordinal nature of the categorical SES variables. We also estimated Kappa values that give more weight to disagreement of increased gravity (cells in the cross-tabulation that are further away from the leading diagonal). The overall pattern in the graph was unchanged.

[14] We checked the sensitivity of this conclusion by restricting attention to the five countries (Italy, Iceland, Poland, Germany and New Zealand) with data on each of the three SES measures; books in the home still showed the lowest levels of consistency between parent and child reports, and father's occupation the highest.

Do children whose SES reports agree with those of their parents have better test scores? In the terminology of Section 2, this provides an indication of how the measurement error varies with the response, $cov(v, Y)$, indicating a departure from the assumptions of classical measurement error. (This interpretation assumes that the parents' reports are error free.) Figure 2 shows the difference in mean reading achievement (in national z-scores) between children whose reports agree with those of their parents and other children. (The two-letter country labels are given in Table 3.) We subtract the mean for children who disagree from the mean from those who agree: positive values show that children providing consistent reports have higher test scores. Countries are ranked by the size of the difference in means for the SES measure concerned.

< Figure 2 here >

We start with father's occupation. The 95% confidence interval includes zero for 9 of the 10 countries indicating no significant difference in mean reading scores. Moreover, the cross-national variation in the estimates is quite modest, with the values in most countries typically less than 0.15 national standard deviations. This again provides encouraging signs of the validity of children's reports of father's occupation, and the consistency of cross-national estimates based upon this measure.

The lengths of the bars are notably larger for parental education. The differences in means are all positive and in every case we easily reject at the 5% level the null that the differences in the population are zero. Children who agree with their parents score on average 0.2 to 0.3 of a national standard deviation higher on the PISA reading test. This extends the conclusion by Kreuter et al. (2010) based on PISA data for Germany of differential measurement error in children's reports of parental education – showing this is the case for a much larger set of countries. The 95% confidence intervals tend to overlap – there is relatively little variation across countries. Germany (DE) is the median country and we unable to reject the hypothesis that the difference in mean scores here is the same as that in other countries in our sample.

The pattern for books in the home is quite similar to that for parental education. Children who report the same category of number of books as their parents tend to have higher maths test scores (in only two countries does the 95% confidence interval include zero). Cross-national variation is again moderate, with the confidence intervals overlapping for most countries. The outliers include England (GB(E)), where children who agree with

their parents' reports of books score on average 0.35 of a national standard deviation more on the PIRLS reading test (the difference is significantly greater at the 5% level than in the majority of other countries).[15] This suggests that estimates of the SES gradient based on child reports in PIRLS may be subject to pronounced downward bias in England, resulting from negative covariance between the measurement error in the reports and the test score – the term $cov(v, Y)$ in the numerator of equation (4). However, this conclusion ignores any offsetting biases resulting from the negative covariance $cov(v, X^*)$ expected in the denominator of (4). A cross-tabulation of parent and child reports for England demonstrates this pattern in Table 4. When parents report the lowest category (0 to 10 books), 79% of children report a higher one. On the other hand, 46% of children report a lower figure when parents report the highest category (more than 200 books). We emphasised in Section 2 that the net result of these two effects is an empirical issue, to which we now turn.

< Table 4 here >

Robustness of the SES gradient to the switch between child and parent reports

We now investigate whether the pattern of differences in SES gradients in test scores across countries depends on the choice of child or parents reports on which to base the SES measure.

We start with father's occupation, the measure for which child and parent reports appear the most consistent. We estimate the regression model in equation (5), entering father's occupation as a series of dummy variables. We summarise the SES gradient by the estimated coefficient on highest level of occupation ('high professional'), which measures the average difference in reading score between children with fathers in this group and children with fathers in the base category, the lowest occupational group ('semi/unskilled manual'), controlling for other variables in the model.

Figure 3 shows estimates when using children's reports of father's occupation on the horizontal axis and when using parental reports on the vertical axis. The data points are all close to the 45 degree line (the correlation coefficient is 0.94). The size of the gradient varies greatly across countries, from less than 0.5 to over 1.0 national standard deviations. However, the close correspondence between the two sets of estimates indicates that the pattern and size

[15] Our restriction to cases with complete data means that we are excluding half the sample in England (see Table 2).

of the differences in the SES gradient across countries is robust to the choice of child or parental reports on which to base the measure – in line with our finding of a high degree of consistency between the reports. We checked the sensitivity of our results to use of other occupational measures – the Ganzeboom and Treiman (1996) index and the ISCO nine major groups – based on the same underlying data collected from parents and children. The correlation between SES gradients based on child and parent reports remained around 0.9.[16]

< Figure 3 here >

We now turn to parental education. We enter a series of dummy variables for parental education in the regression models (in place of those for father's occupation). The SES gradient is taken as the average difference in points scored between those who reported a low level of parental education (ISCED below level 3A – did not complete high school) and those who reported a high level (ISCED 5A and above – university degree), controlling for the other variables in the model. Estimates using children's reports are on the horizontal axis of Figure 4 and those using parental reports on the vertical axis. Since we are now using data for both 2006 and 2009, we append the last digit of the survey year to the two-letter country code (e.g. 'DE6' refers to Germany in 2006).

< Figure 4 here >

A striking feature of Figure 4 is that almost all the data points (17 out of 19) sit above the 45 degree line; the estimates are larger when using the parental reports. If the goal is to estimate the size of the SES gradient within any one country, then use of data on parental education reported by children will lead to an underestimate (assuming the parent reports are error-free). In terms of the direction of bias, this is consistent with measurement error in the child reports being classical in form although our analysis in Section 3 suggests this not to be the case. The magnitude of the bias is typically about 0.1 of a standard deviation but is larger in some countries, notably New Zealand. However, the similarity in the apparent bias across most countries means that that the correlation between the two sets of estimates of the SES gradient is quite high (r = 0.83). Consequently, conclusions on the differences across

[16] When using ISCO codes, we assigned children into one of three groups: 'high' (ISCO major groups 1 and 2), 'medium' (ISCO major groups 3, 4 and 5) and 'low' (ISCO major groups 6, 7, 8 and 9). The SES gradient was then defined as the test score difference between the 'low' and the 'high' group (again with controls).

countries in the SES gradient are moderately robust to whether child or parent reports of parental education are used.

Finally we consider the number of books in the home (Figure 5). The SES gradient is estimated as the average difference in the test scores between children with many books in the home (more than 200) and those with few books (0 to 25), controlling for the other variables in the regression. There are many more data points than in Figures 3 and 4 as information for books in the home is drawn from the PIRLS surveys where parental questionnaires were administered in a much larger number of countries than in the PISA surveys.

< Figure 5 here >

The correlation between the estimates based on parent reports and estimates based on child reports is clearly lower than for the other two SES measures (r=0.37). Moreover, there is no obvious overall pattern of an upward or downward change in the estimates based on the parent reports compared to those based on the child reports: there is a fair amount of scatter around the 45 degree line. There are a number of outlier countries where there is a large change in results, e.g. Italy and Greece. The cross-country pattern of differences in SES gradients is not that robust to the switch between child reports and parent reports. It is difficult to have confidence in the finer details of the pattern of estimates when using books in the home, particularly when trying to identify the countries where SES gradients in test scores are atypically large or small.

To illustrate this point, we highlight the case of England. Using TIMSS data, Schütz et al. (2008) found England to have a particularly strong association between books in the home and children's test scores, a finding that has been widely cited (e.g. Chevalier et al. 2009, Machin 2009, Crawford et al. 2011, Ermisch and Del Bono 2010, Jerrim 2012). Yet England is a notable outlier in the Figure 5 results based on PIRLS; it is only when using children's reports that SES differences stand out as large in England compared to those elsewhere. The estimate for England based on parental reports is about 40% smaller, and below the international average.[17]

[17] If we estimate the model using all children in the PIRLS sample reporting books in the home (rather than the 'complete case' sample) England still stands out as a country where the SES gradient is particularly large. For example, the difference in test scores between children with few (0 – 25) and many (more than 200) books is bigger in England than in 18 of the 21 other counties included in our analysis from the PIRLS 2006 wave.

The results in Figures 3-5 refer to differences in average test scores (controlling for other characteristics in the regression model) between children in the top and bottom categories of each of the SES variables, i.e. the most and the least advantaged backgrounds. We now check the robustness of our results to the selection of other pairs of categories for each SES measure. For example, we can redefine the SES gradient as the average difference (conditional on the control variables) between test scores in the top category of a variable and an intermediate category, rather than the bottom category. In each case we re-compute the correlation coefficient for the set of country SES gradients based on child reports and the set based on parent reports. We do this for all possible pairs of categories for each SES measures (Table 5). The circled figures refer to the correlations reported earlier for Figures 3-5.

< Table 5 near here >

For father's occupation, most of the correlations are high. (The lowest figures, 0.75 and 0.77, refer to SES gradients estimated using adjacent categories.) Largely speaking, the pattern of variation across countries in the SES gradients is robust both to whether the child or the parent reports the information on family background and to the selection of the occupational groups to compare.

There is more sensitivity of the results to the selection of comparison groups in the case of parental education. Low correlations are found when comparing children in the ISCED level 3 group to the other categories – the correlation is below 0.45 in each case and is only just positive for the level 3/below 3A comparison. One possible explanation is that the questions asked to parents and children were not identical, with the main discrepancy occurring around the ISCED level 3A category (see Section 3). Alternatively, it could just be that this is a problematic category for children to report. Otherwise, the correlations are all in the range 0.62 to 0.84, suggesting that our results in Figure 4 are reasonably robust. However, in most cases, the data points in analogous graphs (not shown) no longer clearly sit above the 45 degree line: the general pattern of attenuation in estimates of SES gradients based on the children's reports that we saw in Figure 4 when comparing top and bottom categories is no longer found.

For books in the home, the correlations range from 0.51 to a low of 0.07, with all but one of the six figures below a modest 0.40. This re-iterates the point made earlier: cross-national comparisons of SES gradients in test score based on books in the home seem to be

quite sensitive to whether the child or the parent provides the information on the number of books.

5. Conclusions

We have investigated whether estimates of differences across countries in SES gradients in test scores in international surveys of children's reading achievement are robust to who reports the SES information, parent or child. Using data drawn from PISA and PIRLS, we have considered SES measures based on parental occupation, parental education, and the number of books in the home. This is an important exercise given the reliance on the child reports in much research with the international surveys. Analysts of these surveys either must use the child reports to construct SES variables (TIMSS or early rounds of PISA) or may choose to do so in order maximise the number of countries under investigation (PIRLS and, even more the case, later rounds of PISA) and the number of cases with complete information (we have shown there to be substantial non-response to the parental questionnaires, especially in PISA).

There is substantial agreement between parental reports of father's occupation and those of their 15 year old children in the PISA data for the 10 countries that we have been able to analyse. In only one country did average reading test scores differ significantly between children reporting the same category of father's occupation (out of five possible categories) and those reporting a different category, suggesting a lack of any systematic reporting bias linked to reading ability. The differences across countries in SES gradients in test scores seem robust to whether the child or the parent reports are used as the basis for the SES measure. Although we would like to see these findings replicated for a larger number of countries, we think these results imply that children's reports of parental occupation provides a reliable basis with which to explore cross-national differences in SES test score gradients.

We are less sanguine about books in the home. There is much lower agreement between parent and child reports over the number of books in the home than for parental occupation, although we cannot tell to what extent this is due to the data relating to a younger group of children, 9/10 year olds. In all but one of the 24 countries in our PIRLS data, less than 50% of children report the same category of books in the home as their parents (out of five possible categories). Children reporting a different category from their parents have average reading scores that are almost always significantly lower than those reporting the same category, suggesting that reporting error – assuming the parental reports to be closer to

the truth – is not ‘classical’ in form. Estimates of cross-national differences in the SES gradient of reading test scores are quite sensitive to who provides the information, the parent or the child, with some countries moving dramatically in the international rankings e.g. England. Hence we advise caution with the use of this variable when estimating SES differences in children’s test scores. (This is without questioning its validity as an SES measure.)

The appropriate conclusion regarding parental education is not quite so obvious. Although the consistency of the information given by parents and children was reasonable, reports given by children in some countries appear better than in others. The pattern of differences across countries in the SES gradients was fairly robust to the choice of parent or child reports when comparing the most and least advantaged groups, yet not all such comparisons based on the reported data were quite so encouraging. The evidence available does not seem strong enough for us to either clearly reject or support the use of this SES measure.

Information on family background is a key component in any survey of school children. Collecting that information from children only is likely to remain an attractive low-cost option. Our recommendation to the PISA, PIRLS and TIMSS survey organisers is that further efforts be made to validate the quality of children’ reports of SES measures, encompassing all participating countries. Researchers using data from these surveys to investigate SES differences in test scores should take heart from some of our findings while proceeding with caution given others. At least some of the information available on family background seems to be quite well reported by children, judging by the yardstick of their parents’ reports of the same information.

References

Adams, R. and Wu, M. (2002) *PISA 2000 Technical Report*, Paris: OECD

Ammermueller, A. (2006) 'Educational opportunities and the role of institutions', ZEW Discussion Paper, Number 05-44

Ammermueller A. and Pischke J. (2009) 'Peer effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study', *Journal of Labor Economics*, 27 (3): 315-48.

Bound, J., Brown, C., and Mathiowetz, N. (2001) 'Measurement error in survey data' in J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, vol. 5, Amsterdam: Elsevier.

Brunello, G., Weber, G., and Weiss, C. (2012), 'Books are forever: early life conditions, Education and Lifetime Income'. IZA Discussion Paper 6386.

Buchmann, C. (2002) 'Measuring family background in international studies of education: conceptual issues and methodological challenges' in A. Porter and A. Gamoran (eds.) *Methodological Advances in Cross-National Surveys of Educational Achievement*, Washington DC: National Academy Press.

Cameron, C. and Trivedi, P. (2005) *Microeconometrics*, Cambridge: Cambridge University Press.

Chevalier, A., Denny, K. and McMahon, D. (2009) 'Intergenerational mobility and education equality', in P. Dolton, R. Asplund and E. Barth (eds), *Education and Inequality across Europe*, London: Edward Elgar

Crawford, C., Johnson, P., Machin, S. and Vignoles, A. (2011) 'Social Mobility: A Literature Review', Department for Business, Innovation and Skills research report March 2011

Ermish, J. and Del Bono, E. (2010) *Education Mobility in England – The link between the education levels of parents and the educational outcomes of teenagers*, Sutton Trust research report

Evans, M., Kelley J., Sikora J., and Treiman, D. (2010) 'Family scholarly culture and educational success: books and schooling in 27 nations', *Research in Social Stratification and Mobility*, 28 (2): 171-97.

Ganzeboom, H. and Treiman, D. (1996) 'Internationally comparable measures of occupational status for the 1988 international standard classification of occupations', *Social Science Research*, 25: 201–239.

Haveman, R. and Wolfe, B. (1995). 'The determinants of children's attainments: a review of methods and findings.' *Journal of Economic Literature* 33(4): 1829-78.

Hanushek, E. and Woessmann L. (2011) 'The economics of international differences in educational achievement' in E. Hanushek, S. Machin and L. Woessmann (eds.) *Handbook of the Economics of Education*, volume 3, North-Holland.

Hendrickx, J. (2004) 'ISKO: Stata module to recode 4 digit ISCO-88 occupational codes', Boston College Department of Economics.

Hermann, Z. and Horn, D. (2011) 'How inequality of opportunity and mean student performance are related? A quantile regression approach using PISA data', IEHAS Discussion Papers 1124.

Jakubowski, M. (2010) 'Effects of tracking on achievement growth: exploring difference-in-differences approach to PIRLS, TIMSS and PISA data' in J. Dronkers (ed) *Quality and Inequality of Education*. Springer: Netherlands.

Jerrim, J. (2012). 'The socio-economic gradient in teenagers' reading skills: how does England compare to other countries?', *Fiscal Studies*, 33 (2) 2: 159-84.

Jerrim, J. and Micklewright, J. (2012). 'Parental socio-economic status and children's cognitive achievement at ages 9 and 15: how do the links vary across countries?' in J. Ermisch, M. Jantti and R. Erikson (eds) *From Parents to Children: The Intergenerational Transmission of Advantage*, New York: Russell Sage Foundation.

Koretz, D. (1992) *Evaluating and validating indicators of mathematics and science education*, RAND note N-2900-NSF. Santa Monica, CA: RAND.

Kreuter, F. Eckman, S. Maaz, K. and Watermann, R. (2010) 'Children's reports of parents' education level: does it matter whom you ask and what you ask about?', *Survey Research Methods*, 4(3): 127-138.

Landis, J. and Koch, G. (1977) 'The measurement of observer agreement for categorical data', *Biometrics*, 33 (1): 159-174.

Looker, D. (1989) 'Accuracy of proxy reports of parental status characteristics', *Sociology of Education*, 62 (4): 257-276.

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

Machin, S. (2009) 'Inequality and education' in W. Salverda, B. Nolan, and T. Smeeding (eds.) *The Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press.

Marks, G. (2005) 'Cross-national differences and accounting for social class inequalities in education', *International Sociology*, 20 (4): 483–505.

Marks, G. (2006) 'Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries', *Educational Research*, 48 (1): 21-40.

Mason, W., Hauser, R. M., Kerckhoff, A. C., Poss, S. S., & Manton, K. (1976). 'Models of response error in student reports of parental socioeco-nomic characteristics' in W. H. Sewell,

R. M. Hauser, & D. L. Featherman (eds.), *Schooling and achievement in American society* (pp. 443-519). New York: Academic Press.

OECD (2001) *Knowledge and Skills for Life – First results from PISA 2000*, OECD, Paris.

OECD (2009) *PISA data analysis manual: SPSS second edition*, Paris: OECD

Schulz, W. (2006) 'Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003', Annual Meetings of the American Educational Research Association. San Francisco (April 7-11, 2005).

Schütz G., Ursprung, H. and Woessmann, L. (2008) 'Education policy and equality of opportunity' *Kyklos*, 61 (2): 279–308.

Vereecken C., Vandegehuchte A. (2003) 'Measurement of parental occupation: Agreement between parents and their children' *Archives of Public Health*, 61: 141-9.

Waldinger, F. (2007) 'Does ability tracking exacerbate the role of family background for students' test scores?', unpublished manuscript, London School of Economics.

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 : 539-59

Woessmann, L. (2008) 'How equal are educational opportunities? Family background and student achievement in Europe and the United States', *Zeitschrift für Betriebswirtschaft*, 78 (1): 45-70.

Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge MA: The MIT Press.

Table 1. Parental and child reports of family background (PIRLS and PISA, 2006)

			1	2	3	4	5
	Country	n	Parent and child report present (%)	Child report missing (%)	Parent report missing (%)	Parent and child report missing (%)	Total (%)
Father's Occupation	Korea	5176	92	1	4	2	100
	Poland	5547	86	5	5	5	100
	Turkey	4942	84	1	6	8	100
	Italy	2177 3	79	2	17	2	100
	Portugal	5109	74	4	16	6	100
	Germany	4891	64	8	18	9	100
	Luxembourg	4567	62	5	25	7	100
	New Zealand	4823	58	4	29	8	100
	Iceland	3789	57	4	35	4	100
	Denmark	4532	48	8	33	11	100
Parental Education	Korea	5176	93	2	4	1	100
	Poland	5547	93	3	3	2	100
	Turkey	4942	88	2	10	1	100
	Portugal	5109	78	3	17	2	100
	Italy	2177 3	77	2	20	2	100
	Germany	4891	69	7	18	7	100
	Iceland	3789	59	3	34	4	100
	Luxembourg	4567	59	7	28	6	100
	New Zealand	4823	57	7	27	9	100
	Denmark	4532	55	3	37	5	100
Books in the home	Slovakia	5337	95	1	3	0	100
	Belgium (Flemish)	4552	95	2	4	0	100
	Italy	3581	94	2	4	0	100
	Poland	4854	93	4	3	0	100
	Slovenia	5337	93	2	5	0	100
	Austria	5067	92	3	5	0	100
	Denmark	4001	91	2	6	0	100
	Sweden	4394	90	3	7	1	100
	Luxembourg	5101	89	1	10	0	100
	Hungary	4086	89	2	9	1	100
	France	4404	87	5	7	1	100
	Norway	3837	85	7	7	1	100
	Belgium (French)	4552	84	6	9	1	100
	Canada	2056 5	81	4	14	1	100
	Germany	7899	80	7	10	4	100
	Iceland	3673	73	2	23	2	100
	Netherlands	4156	67	1	31	1	100
	New Zealand	6256	60	3	33	3	100
	Spain	4094	60	2	35	2	100
	Scotland	3775	51	1	47	2	100
	England	4036	46	1	52	2	100

Note: unweighted data.

Table 2. Difference in mean reading scores between children with missing SES information and those with complete information (PIRLS and PISA 2006)

		Parent reports		Child reports	
	Country	difference	S.E.	difference	S.E.
Father's Occupation	Germany	-0.68	0.06	-0.68	0.06
	Korea	-0.62	0.09	-0.52	0.09
	New Zealand	-0.57	0.04	-0.64	0.06
	Italy	-0.46	0.05	-0.47	0.09
	Iceland	-0.43	0.04	-0.49	0.07
	Denmark	-0.40	0.04	-0.36	0.04
	Luxembourg	-0.38	0.08	-0.39	0.06
	Portugal	-0.29	0.07	-0.33	0.05
	Poland	-0.22	0.05	-0.31	0.06
	Turkey	-0.22	0.07	-0.06	0.07
Parental Education	Germany	-0.71	0.06	-0.88	0.07
	New Zealand	-0.54	0.04	-0.72	0.05
	Iceland	-0.48	0.04	-0.70	0.08
	Korea	-0.42	0.09	-0.29	0.09
	Italy	-0.42	0.04	-0.58	0.12
	Denmark	-0.41	0.04	-0.84	0.09
	Portugal	-0.41	0.07	-0.96	0.09
	Poland	-0.36	0.11	-0.62	0.12
	Luxembourg	-0.32	0.08	-0.29	0.08
	Turkey	-0.18	0.06	-0.82	0.16
Books in the home	Slovakia	-0.83	0.19	-1.14	0.17
	Belgium (Flemish)	-0.61	0.11	-0.60	0.13
	Italy	-0.60	0.09	-0.92	0.19
	Germany	-0.58	0.12	-1.01	0.16
	Poland	-0.55	0.09	-0.75	0.09
	Sweden	-0.54	0.08	-0.67	0.12
	New Zealand	-0.53	0.04	-0.88	0.10
	France	-0.50	0.06	-0.40	0.08
	Norway	-0.50	0.09	-0.61	0.16
	Austria	-0.48	0.07	-0.94	0.10
	England	-0.46	0.05	-1.22	0.14
	Netherlands	-0.46	0.05	-0.88	0.15
	Denmark	-0.43	0.08	-0.44	0.14
	Luxembourg	-0.42	0.05	-0.66	0.13
	Slovenia	-0.40	0.08	-0.99	0.11
	Scotland	-0.39	0.04	-0.92	0.15
	Iceland	-0.34	0.04	-0.88	0.08
	Canada	-0.30	0.04	-0.77	0.06
	Spain	-0.29	0.06	-1.10	0.13
	Belgium (French)	-0.26	0.09	-0.41	0.11
	Hungary	-0.24	0.09	-0.80	0.14

Note: we subtract the mean score for children with complete information from the mean for children with missing information. Figures reported in national z-scores. Final student weights applied.

Table 3. Percentage agreement between parent and child reports (PIRLS and PISA)

		Books	Education	Occupation
		%	%	%
Portugal	PT	-	64	76
Denmark	DK	42	53	75
New Zealand	NZ	36	59	72
Germany	DE	36	52	71
Poland	PL	40	79	70
Italy	IT	37	58	69
Iceland	IS	36	56	69
Luxembourg	LU	45	51	68
Korea	KR	-	75	58
Turkey	TR	53	85	85
Chile	CL	-	67	-
Hungary	HU	42	62	-
Slovakia	SK	47	-	-
Czech	CZ	44	-	-
Slovenia	SI	44	-	-
Sweden	SE	43	-	-
France	FR	41	-	-
Norway	NO	40	-	-
England	GB(E)	39	-	-
Spain	ES	39	-	-
Canada	CA	37	-	-
Greece	GR	37	-	-
Austria	AT	37	-	-
Flemish Belgium	BE(Fl)	36	-	-
Scotland	GB(S)	35	-	-
French Belgium	BE(Fr)	35	-	-
Netherlands	NL	34	-	-
Average		40	63	71

Note: if data are available from more than one year (books in 2001 and 2006 and parental education in 2006 and 2009) we show the unweighted average of the figures for the two years. Calculations are based on unweighted data. The averages across all countries are unweighted figures.

Table 4. Parent and child reports of books in the home for England, PIRLS (row %)

		Child report						
	number of books	0 - 10	11 - 25	26 - 100	101 - 200	> 200	total	n
Parent report	0 - 10	21	30	34	10	5	100	198
	11 - 25	17	28	37	12	6	100	468
	26 - 100	9	13	42	19	17	100	1,133
	101 - 200	3	7	30	29	31	100	892
	> 200	2	4	20	20	54	100	888

Notes: the data are the pooled PIRLS 2001 and 2006 samples. Unweighted data.

Table 5. Correlations between country estimates of SES gradients in reading test scores based on parent and child reports

(a) Father's occupation

	High professional	Low professional	Routine white collar	Skilled manual	Semi/unskilled manual
High professional	-				
Low professional	0.86	-			
Routine white collar	0.80	0.77	-		
Skilled manual	0.92	0.94	0.91	-	
Semi/unskilled manual	0.94	0.95	0.90	0.75	-

(b) Parental education

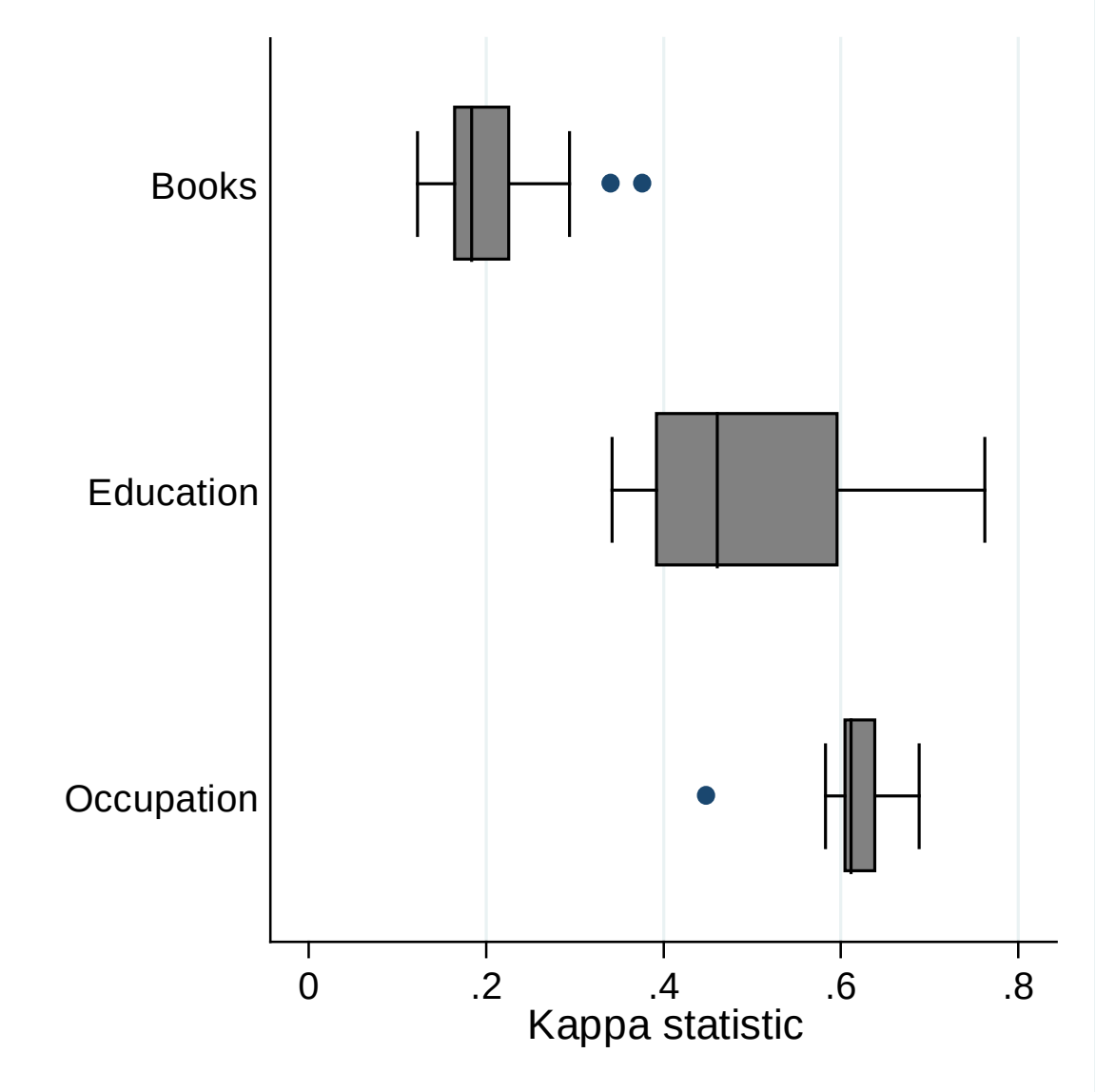
	ISCED 5A	ISCED 5B	ISCED 4	ISCED 3A	Below 3A
5A	-				
5B	0.68	-			
4	0.75	0.83	-		
3A	0.39	0.32	0.43	-	
Below 3A	0.83	0.62	0.84	0.05	-

(c) Books in the home

	>200 books	101-200 books	26-100 books	0-25 books
>200	-			
101-200	0.33	-		
26-100	0.51	0.07	-	
0-25	0.37	0.27	0.39	-

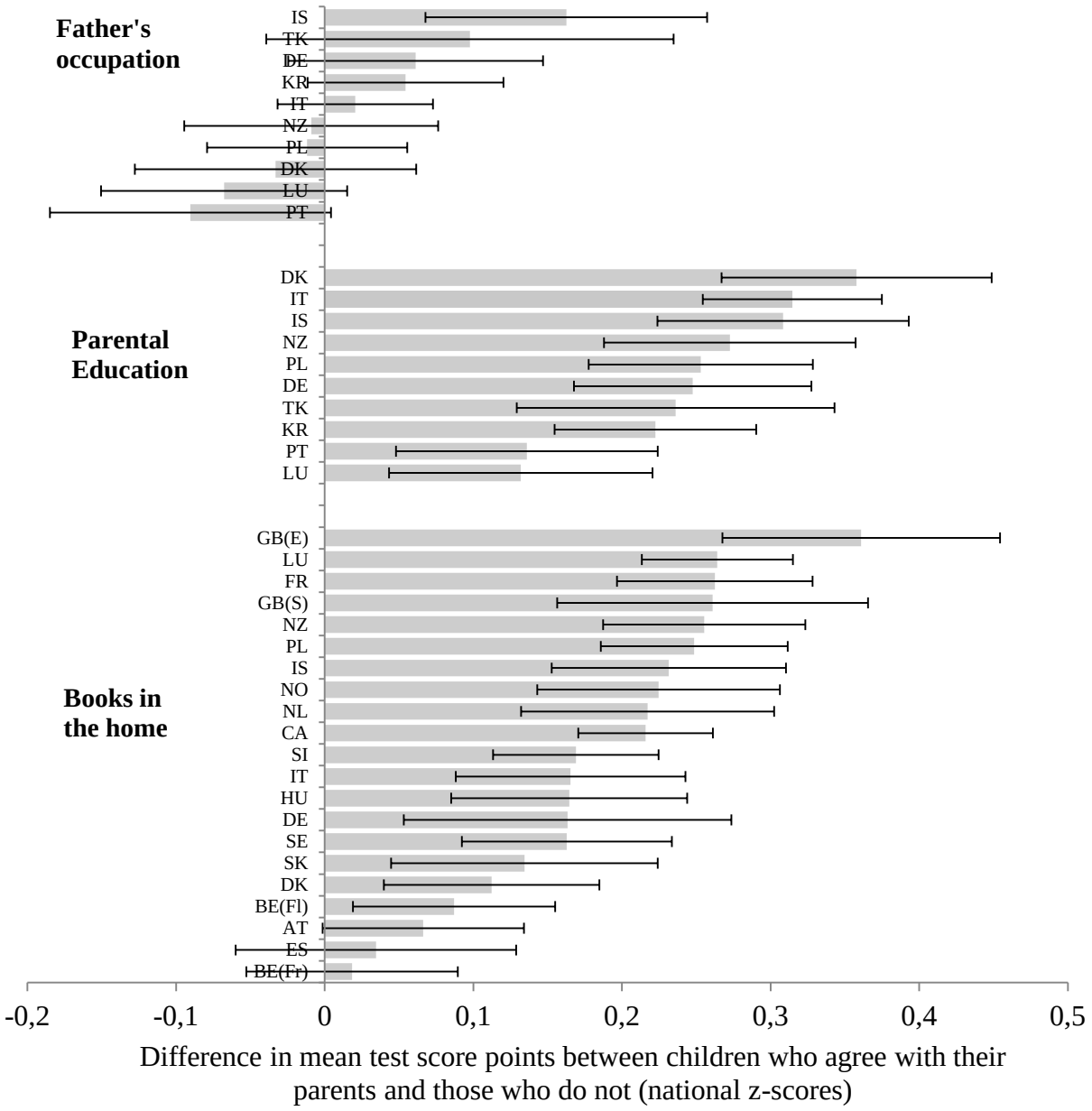
Notes: SES gradients in reading test score have been estimated using every possible pair of categories within each of the measures. For each SES measure, the figures give the correlations between the two sets of estimates of gradients, one based on child reports of SES and one on parent reports. For example, the circled figure in panel (a) refers to the gradients estimated using the 'high professional' and 'semi/unskilled manual' groups (that is the difference in average reading test scores between these two categories, controlling for other variables in the regression model). The circled figures in each panel give the correlations for the scattergrams in Figures 3 to 5. Final students weights applied.

Figure 1. The extent of agreement between parent and child reports for different SES measures (PIRLS and PISA): Kappa statistics of inter-rate reliability



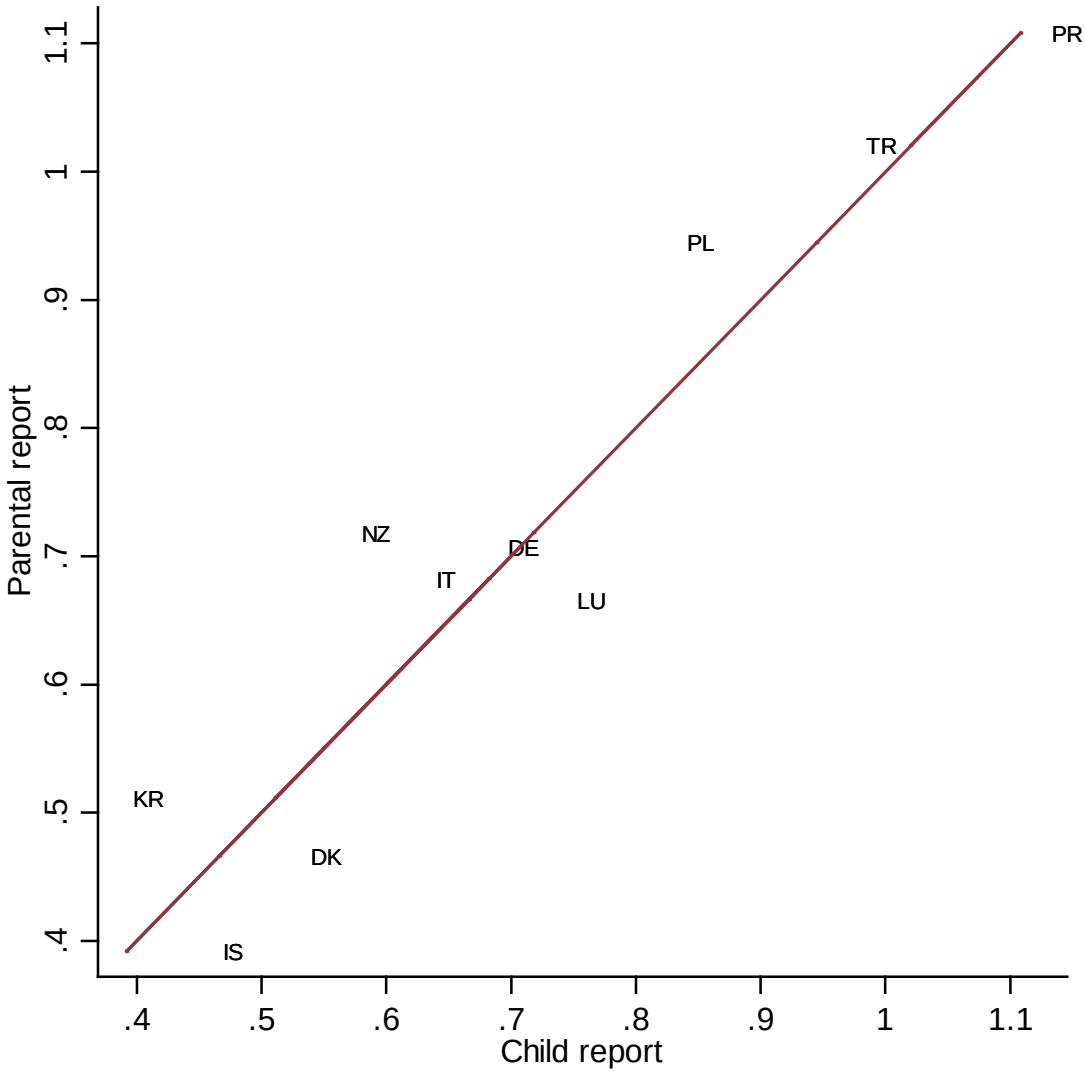
Note: data on books in the home are from PIRLS 2001 and 2006, data on education are from PISA 2006 and 2009, data on occupation from PISA 2006. Kappa statistics are estimated for each country-year. Unweighted data.

Figure 2. Difference in average test scores between children reporting the same SES category as their parents and those who do not (PISA and PIRLS 2006)



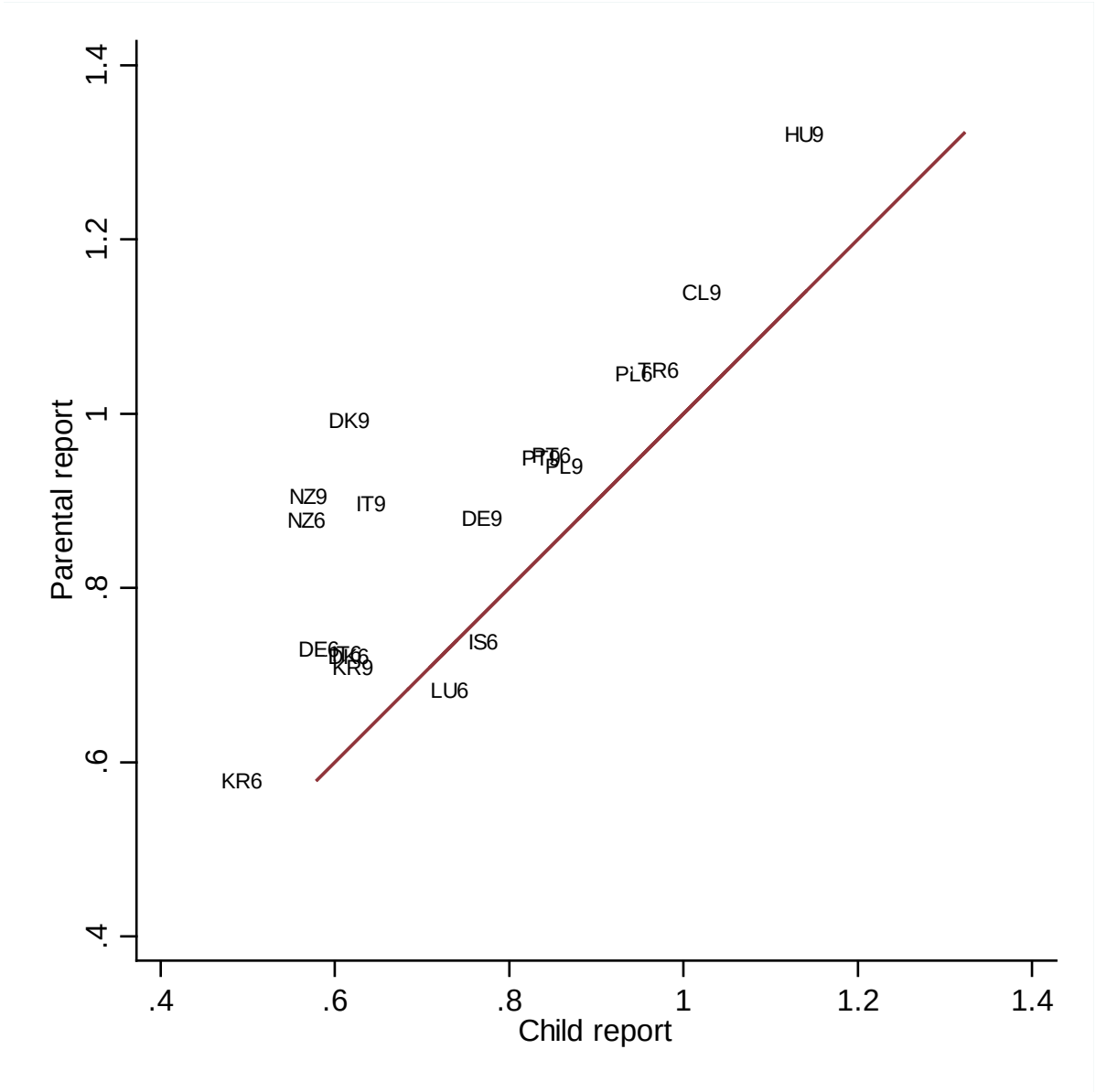
Notes:
The graph shows the difference in average reading test scores between children who report the same SES category as their parents and those who do not (subtracting the latter from the former). All figures refer to national standard deviations. The line running through the centre of each bar shows the estimated 95% confidence interval. Final students weights applied.

Figure 3. Estimated SES gradient in child test scores – comparison of results based on parent reports and child reports of father's occupation (PISA)



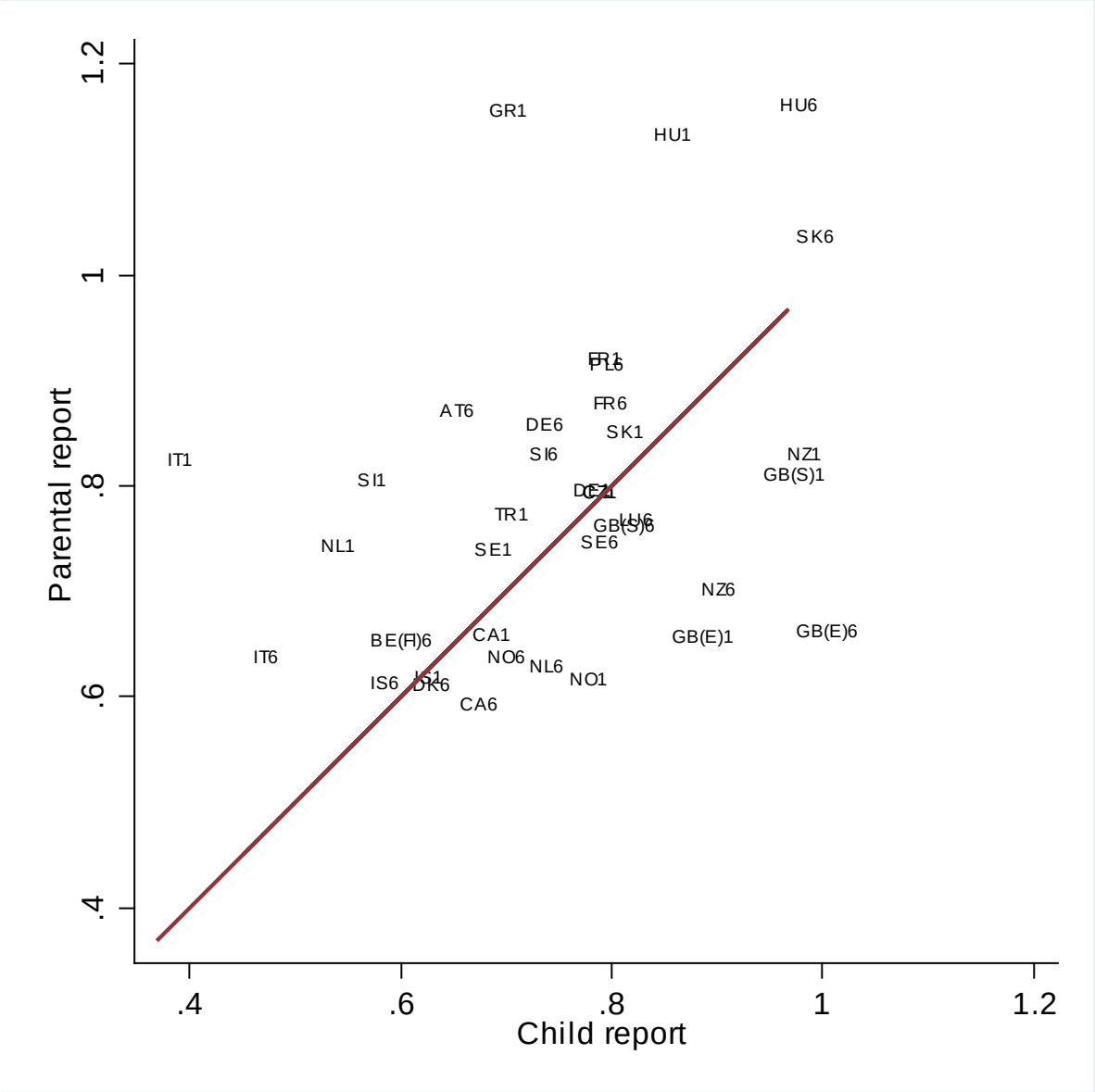
Notes: the graph shows the difference in average reading test scores, controlling for other variables included in the regression model, between children with the highest ('high controller') and lowest ('semi/unskilled manual') levels of father's occupation. The units are national standard deviations. The correlation between the two sets of estimates is 0.94. Two letter country abbreviations are given in Table 3. The data refer to 2006. Final students weights applied.

Figure 4. Estimated SES gradient in child test scores – comparison of results based on parent reports and child reports of parental education (PISA)



Notes: the graph shows the difference in average reading test scores, controlling for other variables included in the regression model, between children with the highest (ISCED 5A+) and lowest (ISCED below 3A) levels of parental education (father's or mother's education, whichever is higher). The units are national standard deviations. The correlation between the two sets of estimates is 0.83. Two letter country abbreviations are given in Table 3; the digit ending the label shows the year to which the data point refers ('6' = 2006, '9' = 2009). Final students weights applied.

Figure 5. Estimated SES gradient in child test scores – comparison of results based on parent reports and child reports of books in the home (PIRLS)



Notes:

Notes: the graph shows the difference in average reading test scores, controlling for other variables included in the regression model, between children with the highest (more than 200 books) and lowest (0-25 books) reported levels of books in the home. The units are national standard deviations. The correlation between the two sets of estimates is 0.37. Two letter country abbreviations are given in Table 3; the digit ending the label shows the year to which the data point refers ('1' = 2001, '6' = 2006). Final students weights applied.

Appendix A. Consistency between parent and child reports of mother's occupation

The main body of the paper investigated the consistency between parent and child reports of father's occupation. We extend this analysis by considering mother's occupation, using the same PISA 2006 data (recall that the parental questionnaire included questions about mother's and father's occupation only in this year). Table A.1 provides details on missing information. As noted in Section 3, much non-response is due to the entire parental questionnaire not being completed. However, comparing with Table 1 in the main paper, it seems missing information is a bigger problem for mother's occupation than for father's occupation in some countries due to greater item non-response, e.g. Italy, Korea, and most notably Turkey (the reporting of occupation for mothers not in work may be an issue). Additional analysis, not presented for brevity, suggests that children with missing data achieved lower PISA reading test scores on average. All results presented in this appendix are based upon the subset of children with both parent and child reports available.

Appendix Table A.2 presents kappa statistics for the consistency between parent and child reports, together with the percentage agreement (parents and children reporting the same occupational group). These are based upon the same five occupational groupings described in Section 3 (high professional, low professional, routine white collar, skilled manual, and semi-skilled or unskilled manual). It seems that information on maternal occupation provided by children is reasonably consistent with that provided by their parents. Kappa statistics in most countries sit around 0.65, with relatively modest cross-national variation (though estimates are notably smaller in South Korea, 0.48, and larger in Turkey, 0.84, than other countries). These findings are similar to those for father's occupation in Section 4.

In Figure A.1 we investigate whether children with higher PISA reading test scores are more likely to agree with parental reports of their mother's occupation (i.e. an indication of how the measurement error varies with the response, $cov(v, Y)$). As described in Section 3, we create a binary variable that takes a value of 1 if parent and child reports agree, and 0 otherwise. Average reading test scores are then compared across the two groups. Recall that for father's occupation, differences in average reading test scores were not statistically different from 0 at the 5% level in nine of the ten countries considered (see Figure 2). Yet the same does not seem to hold true for maternal occupation. In five out of the ten countries (Poland, South Korea, New Zealand, Iceland, Germany) average reading test scores are significantly higher amongst children who report the same maternal occupation category as

their parents. However the estimated 95% confidence intervals continue to overlap for most countries, suggesting only limited evidence of cross-national variation in this result.

In Figure A.2 we re-estimate the regression model described in Section 3 using mother's occupation as the chosen measure of socio-economic status. The SES gradient is taken to be the average difference in points scored between the semi-skilled/unskilled manual group and the high professional group. Estimates when using children's reports are on the x-axis, and those using parental reports on the y-axis. The data points are generally quite close to the 45 degree line, with a correlation of 0.86 between the two sets of results. Although this correlation is slightly weaker than for father's occupation (0.94 – see Figure 3), it nevertheless suggests that cross-national comparisons of SES test score gradients based upon maternal occupation are quite robust to who reports this family background characteristic.

The results in Figure A.2 refer to differences in average test scores between children from the most (high professional) and the least (semi-skilled/unskilled) advantaged backgrounds. We now check the robustness of our results to the selection of other pairs of maternal occupation categories. In each case we re-compute the correlation coefficient for the set of country SES gradients based on child reports and the set based on parent reports. We do this for all possible pairs of categories. (The procedure is as for father's occupation described in Section 4.) Results are given in Table A.3. Most of the correlations are high, standing at 0.85 or above in all but one case. The exception (where the correlation is 0.56) refers to the comparison between the two lowest socio-economic groups (skilled manual and semi/unskilled manual). Thus the SES test score gradient seems robust to who provides information on mother's occupation, regardless of the categories that are compared. And overall, our results in this appendix for mother's occupation do not depart notably from those for father's occupation in the main paper.

Table A.1. Missing parental and child reports of mother's occupation

		1	2	3	4	5
	n	Parent and child report present (%)	Child report missing (%)	Parent report missing (%)	Parent and child report missing (%)	Total (%)
Poland	5,547	82	6	6	6	100
Portugal	5,109	64	5	16	15	100
Germany	4,891	62	8	19	11	100
South Korea	5,176	58	7	14	21	100
New Zealand	4,823	56	5	29	10	100
Iceland	3,789	56	4	35	6	100
Italy	21,773	54	6	15	25	100
Luxemburg	4,567	54	8	25	13	100
Denmark	4,532	50	7	33	11	100
Turkey	4,942	13	2	3	83	100

Table A.2 Percentage agreement and kappa statistics for the consistency between parent and child reports of mother's occupation

	Mother's Occupation	
	% agreement	Kappa
Turkey	88	0.84
Luxemburg	77	0.67
Poland	75	0.67
Italy	75	0.66
Germany	75	0.66
Denmark	76	0.66
Portugal	75	0.66
New Zealand	74	0.64
Iceland	72	0.62
Korea	63	0.48
Average	75	0.66

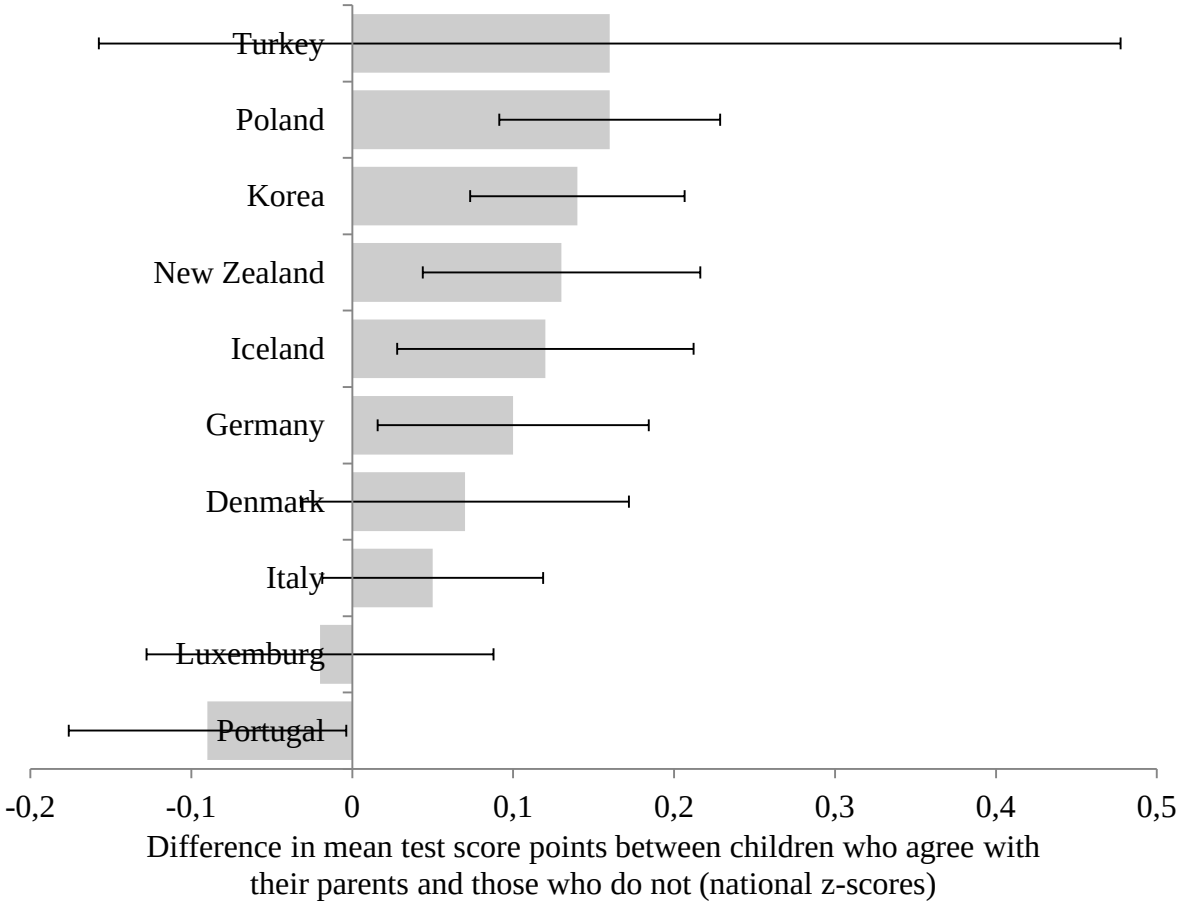
Note: the average is the unweighted.

Table A.3 Correlation matrices showing the consistency of the estimated socio-economic gradient between parent and child reports

	High professional	Low professional	Routine white collar	Skilled manual	Semi/unsk. manual
High professional	-				
Low professional	0.91	-			
Routine white collar	0.94	0.91	-		
Skilled manual	0.94	0.86	0.92	-	
Semi/unskilled manual	0.86	0.96	0.84	0.56	-

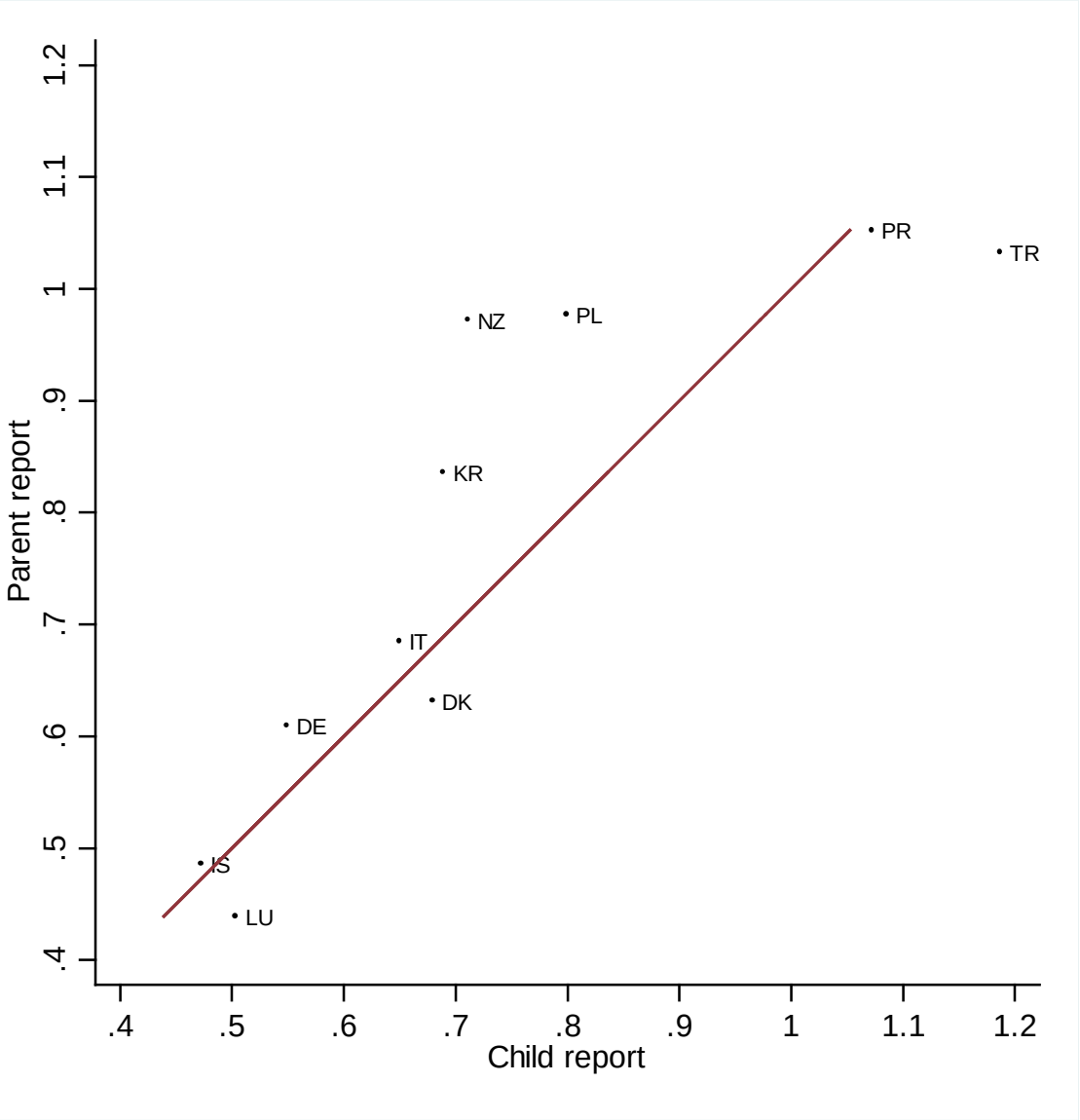
Note: see text for explanation.

Figure A.1 Difference in average test scores between children reporting the same maternal occupation category as their parents and those who do not (PISA 2006)



Notes: The graph shows the difference in average reading test scores between children who report the same maternal occupation category as their parents and those who do not (subtracting the latter from the former). All figures refer to national standard deviations. The line running through the centre of each bar shows the estimated 95% confidence interval.

Figure A.2 Estimated SES gradient in child test scores – comparison of results based on parent reports and child reports of mother's occupation (PISA)



Notes: the graph shows the difference in average reading test scores, controlling for other variables included in the regression model, between children with the highest ('high controller') and lowest ('semi/unskilled manual') levels of mother's occupation. The units are national standard deviations. The correlation between the two sets of estimates is 0.86. Two letter country abbreviations are given in Table 3. The data refer to 2006.

Appendix B. Consistency between parent and child reports of mother's and father's education

In the main body of the paper we investigated the consistency of parental and child reports for the highest level of parental education. This appendix provides separate results for the education level of the mother and the education level of the father. All analyses are based on the subset of children with complete information – the extent and pattern of missing data is very similar to that reported elsewhere in the paper (see Table 1 and Table A.1). To begin, we consider the consistency between parent and child reports, measured by Kappa statistics and percentage agreement (same level of education reported by parent and child) – see Table B.1. The consistency between parent and child reports of maternal education is reasonable, although with moderate variation across countries. Kappa statistics range from 0.31 in Germany to 0.69 in Turkey, with the median percentage agreement across the countries standing at 57%. Results for father's education are very similar.

Figure B.1 illustrates how average reading test scores differ between the group of children who report the same education category as the parent and the group who do not. For both mother's and father's education, the difference in test scores is statistically significant at the 5 percent level in most of the countries considered. In other words, children with higher PISA reading test scores are more likely to report the same category as their parents. This suggests that the measurement error in children's reports may be differential, under the assumption that parental reports are closer to the 'truth'. This is consistent with results for highest level of parental education presented in the main body of the paper (see Figure 2).

The regression model described in Section 3 is now estimated using either mother's education (left hand panel) or father's education (right hand panel) as the chosen measure of socio-economic status. The SES gradient is taken to be the average difference in points scored between those who reported a low level of mother/father education (below ISCED level 3A) and those who reported a high level (ISCED level 5A and above). Results can be found in Figure B.1. Estimates when using children's reports can be found on the x-axis, with parental reports on the y-axis. One can see that almost every data point (for both maternal and paternal education) sits above the 45 degree line. Thus the general pattern of attenuation reported in the main body of the paper (Figure 4) continues to hold. However, the correlation between the two sets of estimates is slightly stronger for father's education ($r = 0.88$) than it is for mother's ($r = 0.81$).

Finally, rather than simply defining the socio-economic gradient as the difference in test scores between the top and bottom SES categories, we now look at the consistency of results when drawing comparisons between all possible combinations of socio-economic groups. For instance, with parental education we no longer solely focus on the test score difference between the ISCED 5A/6 category and the ISCED below 3A groups, and whether this particular estimate varies between parent and child reports. Rather, we also look at difference between (say) the ISCED 5B and ISCED 3A categories as well. Our results are presented as Pearson correlation coefficients (analogous to those in Table 5 discussed in the main text) and can be found in Table B.3. Results are somewhat mixed, with particularly low correlations for comparisons involving the ISCED level 3A group. Otherwise, the correlations are generally satisfactory (around 0.70 to 0.80). This holds for both mother's and father's education.

These results are consistent with those for highest level of parental education presented in the main body of the paper. Indeed, the same conclusion regarding the robustness of SES test score gradients hold whether one focuses on mother's education, father's education, or the higher of the two.

Table B.1. Percentage agreement and Kappa statistics for the consistency between parent and child reports of mother's/father's education (PISA 2006 data)

	Mother's Education		Father's Education	
Country	% agreement	Kappa	% agreement	Kappa
Turkey	86	69	85	76
Poland	80	68	81	62
Chile	70	60	68	57
South Korea	69	53	69	56
Portugal	66	49	64	44
Hungary	64	54	59	46
Iceland	58	46	57	43
Italy	58	43	60	47
New Zealand	57	43	60	48
Denmark	53	34	52	36
Luxemburg	52	36	53	41
Germany	49	31	52	38
Average	63	49	63	49

Notes: where data are available from more than one year (both 2006 and 2009) we show the unweighted average of the figures for the two years. The averages across all countries are unweighted figures.

Table B.2. Correlation matrices illustrating the consistency of the estimated socio-economic gradient between parent and child reports

(a) Mother's education

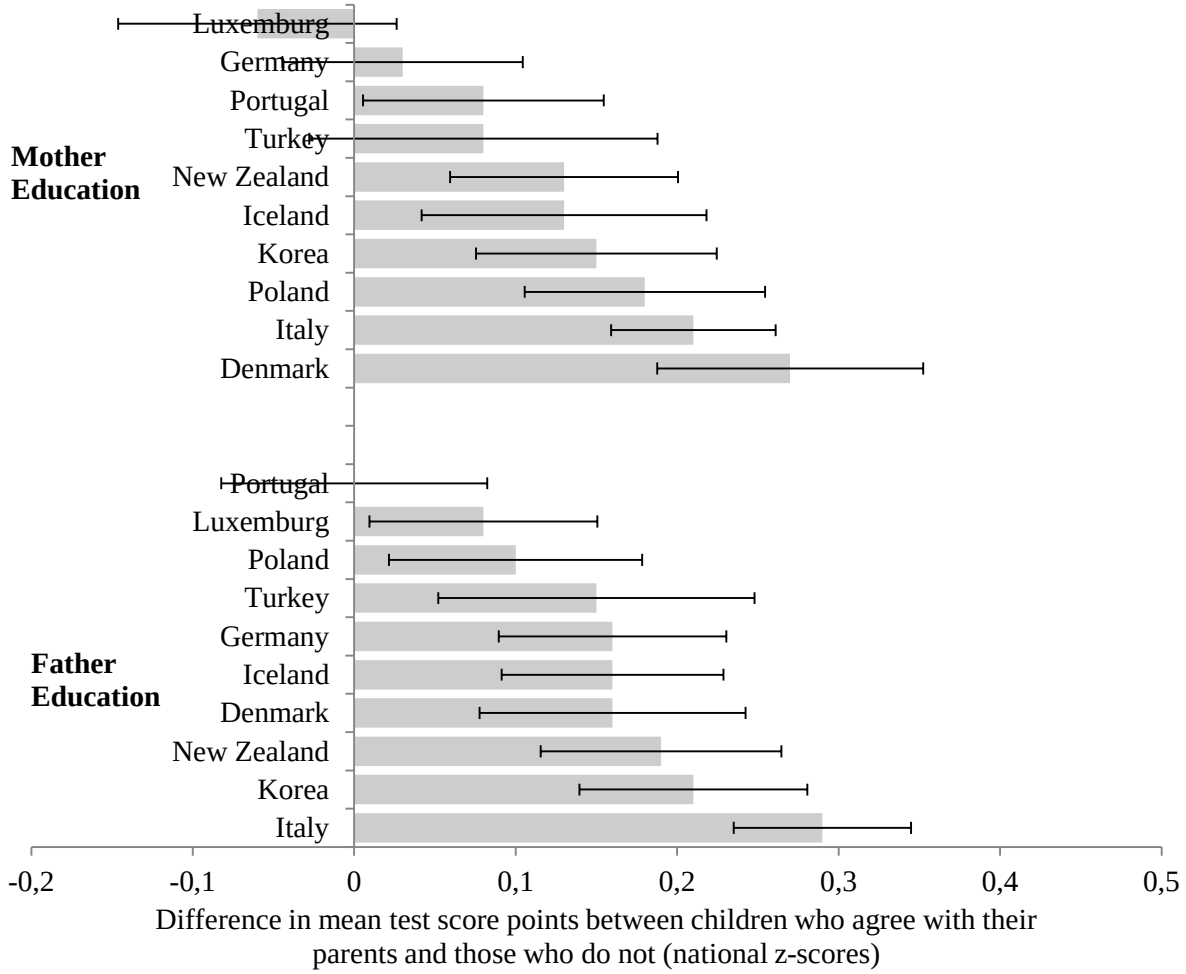
	ISCED 5A	ISCED 5B	ISCED 4	ISCED 3A	Below 3A
5A	-				
5B	0.60	-			
4	0.71	0.77	-		
3A	0.47	0.24	0.49	-	
Below 3A	0.81	0.64	0.93	0.18	-

(b) Father's education

	ISCED 5A	ISCED 5B	ISCED 4	ISCED 3A	Below 3A
5A	-				
5B	0.73	-			
4	0.79	0.70	-		
3A	0.47	0.24	0.38	-	
Below 3A	0.88	0.61	0.86	0.24	-

Notes: SES gradients in reading test score have been estimated using every possible pair of categories within each measure. For each SES measure, the figures give the correlations between the two sets of estimates of gradients, one based on child reports of maternal/paternal education and one on parent reports. For example, the circled figure in panel (a) refers to the gradients estimated using father's education and the 'ISCED 5A' and 'ISCED Below 3A' groups (that is the difference in average reading test scores between these two categories, controlling for other variables in the regression model described in Section 3).

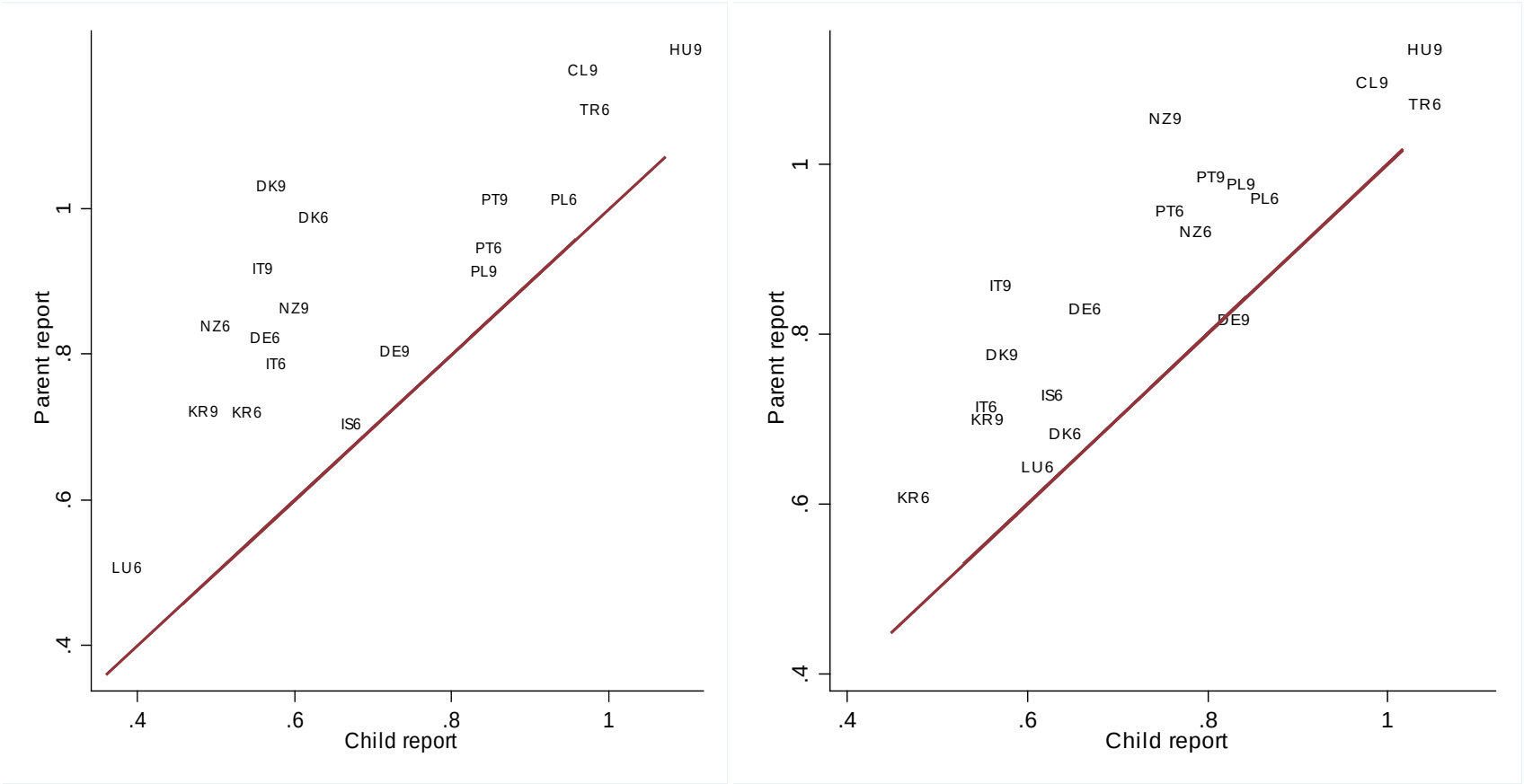
Figure B.1 Difference in average test scores between children reporting the same mother / father education category as their parents and those who do not (PISA 2006)



Notes: The graph shows the difference in average reading test scores between children who report the same maternal/paternal education category as their parents and those who do not (subtracting the latter from the former). All figures refer to national standard deviations. The line running through the centre of each bar shows the estimated 95% confidence interval.

Figure B.2. Estimated socio-economic gap in children's test scores between those with 'high' and 'low' levels of education – a comparison of results between using parent and child reports (PISA 2006 data)

(a) Mother's education **(b) Father's education**



Notes: In the figure above we investigate the difference in test scores between children with the highest (ISCED 5A+) and lowest (ISCED below 3A) levels of mother's/father's education. Estimates along the x-axis are based upon children's reports, with the y-axis referring to estimates using parental reports. Figures refer to national standard deviations. The correlations are 0.81 in the left hand panel and 0.88 in the right hand panel.