

METHODOLOGY

Open Access



Evaluating the risk of nonresponse bias in educational large-scale assessments with school nonresponse questionnaires: a theoretical study

Sabine Meinck^{1*} , Diego Cortes² and Sabine Tieck¹

*Correspondence:
sabine.meinck@iea-dpc.de
¹ IEA Data Processing
and Research Center,
Hamburg, Germany
Full list of author information
is available at the end of the
article
This manuscript is intended
to be published in the
conference special issue of
the 6th IEA International
Research Conference, 24–26
June 2015, Cape Town, South
Africa

Abstract

Survey participation rates can have a direct impact on the validity of the data collected since nonresponse always holds the risk of bias. Therefore, the International Association for the Evaluation of Educational Achievement (IEA) has set very high standards for minimum survey participation rates. Nonresponse in IEA studies varies between studies and cycles. School participation is at a higher risk relative to within-school participation; school students are more likely to cooperate than adults (i.e., university students or school teachers). Across all studies conducted by the IEA during the last decade, between 7 and 33% of participating countries failed to meet the minimum participation rates at the school level. Quantifying the bias introduced by nonresponse is practically impossible with the currently implemented design. During the last decade social researchers have introduced and developed the concept of nonresponse questionnaires. These are shortened instruments applied to nonrespondents, and aim to capture information that correlates with both: survey's main outcome variable(s), and respondent's propensity of participation. We suggest in this paper a method to develop such questionnaires for nonresponding schools in IEA studies. By these means, we investigated school characteristics that are associated with students' average achievement scores using correlational and multivariate regression analysis in three recent IEA studies. We developed regression models that explain with only 11 school questionnaire variables or less up to 77% of the variance of the school mean achievement score. On average across all countries, the R^2 of these models was 0.24 (PIRLS), 0.34 (TIMSS, grade 4) and 0.36 (TIMSS grade 8), using 6–11 variables. We suggest that data from such questionnaires can help to evaluate bias risks in an effective way. Further, we argue that for countries with low participation rates a change in the approach of computing nonresponse adjustment factors to a system where school's participation propensity determines the nonresponse adjustment factor should be considered.

Keywords: Nonresponse bias, Multistage sample surveys, Weighting adjustments

Background

In order to ensure the validity and reliability of cross-country comparative large-scale assessments, the IEA sets high quality standards for its survey instruments, as well as sampling and data collection procedures. All these quality indicators are regarded when

results of a study are reported and the data is made publicly available, and are meant to ensure a high quality and validity of the survey results.

Among other measures, the IEA outlines minimum participation rates. This is due to the fact that usually no or very little information is available for nonresponding units or individuals, which is why nonresponse always holds the risk of bias. Therefore, the general goal of any survey researcher is to achieve a 100% response rate. However, IEA studies acknowledge the difficulties in achieving this goal. Instead, they determine specific minimum participation rates to reduce the risk of bias due to nonresponse. As a standard rule, 85% of the sampled schools within a country as well as 85% of the sampled individuals must participate in the survey in order to accept the data and results for a final release. Participation rates in IEA studies vary among educational systems (further referred to as “countries”), target populations and surveys. Notably, highly developed western economies are facing increasing problems to comply with IEA’s response rate standards. As a general rule, data from participating countries that fail to meet these standards get annotated in the international reports or are even reported in separated report sections, highlighting the possibly reduced validity of the results to the readers. Interested readers are referred to the TIMSS International Report Appendix C.8 (Mullis et al. 2012a) for details on participation rates and guidelines for annotations.

A common approach to mitigate the risk of nonresponse bias in survey estimates is through adjustment cell reweighting, where participating units (schools, students, teachers etc.) carry the weight of nonresponding units. This technique is based on the assumption of a non-informative response model, that is, nonresponse occurs completely at random within each adjustment cell. This weighting adjustment method is used in all IEA studies, as no—or very limited—information is available about nonrespondent units. Explicit strata constitute, in most cases, the adjustment cells for school and class level nonresponse, while schools or classes constitute usually the adjustment cells for individual level nonresponse (Martin and Mullis 2013; Schulz et al. 2011; Meinck 2015). Since, there is no way to prove that the units’ nonresponse is completely at random within an adjustment cell, the IEA standards are very strict on response rate thresholds as pointed out above. This paper will propose a novel approach on how to evaluate the risk of bias due to nonresponse at the school level.

IEA surveys usually implement a two-stage stratified cluster sampling design. Normally, schools are selected first, and then individuals (or classes) are randomly selected within the sampled schools (hence, nonresponse can occur at both sampling stages). In order to validate our approach, we first provide evidence in this paper that school level participation is at a higher risk, relative to within-school participation. This implies that the highest burden for survey administrators is to convince schools to participate in these assessments, while high rates of within-school participation are usually easy to achieve. Hence, understanding nonresponse at the school level is of great importance, and adjusting for the bias introduced by any systematic nonresponse pattern recommended.

The current state of nonresponse bias analysis in LSA

Encouraging participating countries to achieve the highest response rate possible in order to maximize data quality is not unique to the IEA, but is rather a common feature of all international comparative large-scale assessments in education. The minimum

thresholds set for participation, though, vary substantially among studies as there is no universal consensus of what is the minimum participation rate acceptable. Increasing nonresponse rates motivate study centers to develop further strategies to ensure high data quality besides setting minimum requirements. However, no general standards are extant that help countries facing low participation rates to analyze their data to verify the bias risk due to poor response rates. To our knowledge there are three international comparative surveys in education which have systematically conducted nonresponse bias analysis to evaluate the risk of bias due to poor participation. In what follows we briefly summarize the different approaches implemented by these studies.

All participating countries in the Programme for the International Assessment of Adult Competences (PIAAC) (OECD 2013) were required to carry out a “basic nonresponse bias analysis”. This consisted in comparing survey respondents and nonrespondents on individual characteristics which were assumed to be associated with the main outcome variable of the survey. All countries had to include in this analysis at least the following variables: age, gender, education, employment and region. When participating countries were not able to achieve an overall participation rate of 70%, they were required to perform a more in-depth nonresponse bias analysis (Mohadjer et al. 2013). Examples of such an analysis are: comparing survey total estimates with census totals, comparison of responding rates by demographic characteristics, and correlation analysis of weighting adjustment variables with proficiency measures (outcome variables). To name one exemplary outcome of such analysis, in Germany, Helmschrott and Martin (2014) found that age, citizenship, the level of education, the type of house the sampled persons live in, and municipality size were the main factors influencing response to PIAAC.

The Teaching and Learning International Study (TALIS) (OECD 2014) is a comparative international large-scale survey on teacher competences. The international survey and sampling design of TALIS coincides, to a larger extent, with the design of most other IEA studies. The primary sampling units are schools and the responses are also at risk at both sampling stages (in the case of TALIS schools and the teachers within sampled schools). The TALIS International Consortium invited those countries facing participation problems at any sampling stage to conduct a nonresponse bias analysis to evaluate the risk of bias.

The first step proposed was to compare the weighted estimates of characteristics from the school and teacher surveys with official statistics. This was done to show that (non) response propensity is independent of teacher or school characteristics. Establishing the impact of response propensities on teachers’ characteristics was analyzed as a second step. This analysis consisted of comparing teachers’ and/or schools’ characteristics between participating schools having different within-school participation rates. The aim was to show that survey results from schools with high participation rates can be compared with those from schools with low rates of participation. Analysis results of affected countries are not publicly available.

ICILS was the first IEA study to systematically conduct a nonresponse analysis in order to evaluate the risk of bias due to systematic non-participation (Meinck and Cortes 2015). ICILS aims to infer on two populations: students and teachers, and the nonresponse analysis was performed at the student and teacher levels (i.e., within participating

schools). At the student level, associations between response propensities, gender and students' *information and computer literacy* (ICILS' main outcome variable) were explored. At the teacher level, distributions of respondents and nonrespondents were compared with respect to age, subject domain and gender. These were the only available individual characteristics for respondents and nonrespondents that ICILS collected. The analysis showed that different response patterns between boys and girls were negligible, but significant for gender, age and main subject domains of teachers (Meinck and Cortes 2015).

The approaches presented above vary significantly in the way the common goal—evaluating potential bias introduced by nonresponse—was addressed. The common feature between PIAAC and TALIS is that they use auxiliary variables for the nonresponse analysis which might have not been present on the sampling frame, therefore allowing country-specific variables on the analysis. ICILS, on the other hand, exploited the very little information available for respondents and nonrespondents for all countries in a standardized way. The restraints of both applied approaches are obvious: (1) the availability and reliability of auxiliary statistical information varies substantially across countries, and (2) restrictions in the array of available information on nonresponding units limit the explanatory power of the analyses. From the authors' point of view, the approach followed by ICILS is more consistent in a cross-country comparative framework, but very limited in terms of available information.

Another approach to evaluate bias was developed in non-educational social surveys. So-called nonresponse or basic questionnaires are handed out to individuals who refuse participation, or who could not be contacted in the main data collection (e.g., Bethlehem and Kersten 1985; Lynn 2003; Stoop 2004; Matsuo et al. 2010). These questionnaires contain a significantly reduced number of survey questions. The items in the questionnaires are assumed to be highly associated to survey's main outcome variables and with unit's participation propensity. This allows researchers to evaluate the risk of bias arising from nonresponse, determine methods of nonresponse adjustments (e.g., weight adjustments related with the features of nonrespondents), or identify missing data imputation models. Recent research provided evidence that it is possible to achieve high participation rates in nonresponse questionnaires, which is the precondition for a meaningful use of the collected data (Lynn 2003; Stoop 2004; Matsuo et al. 2010). To our knowledge, nonresponse questionnaires have yet to be used in any large cross-national comparative assessment in education.

Research focus, methods and data sources

There is extensive evidence on the literature that the main outcome variables in IEA assessments (usually achievement scores in specific subject domains) are highly associated with background characteristics of the participants (Caldas and Bankston 1997; Fuller 1987; Grace and Thompson 2003), suggesting that school context explains an important portion of the variability of student achievement scores (e.g., Koretz et al. 2001; Lamb and Fullarton 2001; Baker et al. 2002; Mullis et al. 2012a, b).

In a first step, this paper will evaluate the scope of nonresponse in IEA surveys. All IEA studies conducted within the last ten years will be reviewed with respect to nonresponse levels at the different sampling stages.

We will then focus on the methodological feasibility of the development of a school-level nonresponse questionnaire by identifying items that serve as good predictors of school average achievement. We will thereby address also operational constraints by trying to keep the number of items at a minimum. Note, since the practical implementation of such questionnaires is pending, we cannot yet evaluate whether the items do also correlate with response propensities. The potential content of these questionnaires will be determined through analyzing the association of school-level variables with student-level results using data of TIMSS and PIRLS 2011. Regression analysis, using only school-level characteristics, will be applied to identify the best-fitting model in predicting averaged student achievement scores. We will compare cross-country standardized models with country-specific models.

We accounted for the complex sample design (i.e., stratification and unequal selection probabilities of schools) by applying sampling weights for the estimation of population parameters and jackknife repeated replication for the estimation of standard errors.

Between and within-school nonresponse rates across IEA studies

Table 1 summarizes nonresponse rates of all IEA studies within the last decade. It can be seen that the amount of nonresponse varies between studies and cycles. Overall, about 17% of the countries failed to meet the minimum participation standards at the school level when the target population was school students. In ICCS 2009 and ICILS 2013 however, even every third country could not convince at least 85% of the sampled schools to participate in the study. In contrast, countries hardly ever struggle to reach the minimum participation rates for the sampled students within participating schools. Looking through the technical documentation of IEA studies, one will find that in the majority of all countries, the student participation rates are well above 90%. Hence, even if non-participants deviate systematically from participants, the risk of bias is very low. When adults comprise the target population, achieving high participation rates at both sampling stages becomes even more challenging, as shown in the lower part of Table 1. On average, 40% of the countries failed to meet the minimum participation requirements for the sampled schools, and more than 30% failed to meet these requirements within participating schools.

Replacing sampled schools that refuse to participate with predefined (replacement) schools is a common strategy to support countries facing school participation problems. In most student surveys the use of replacement schools has helped countries to achieve survey's minimum participation rates. However, there might be a risk of bias due to the use of replacement schools. Specific methods are used to determine replacement schools in all IEA studies in order to keep this risk as low as possible: replacements are assigned in a way to ensure that they share similar features with the originally sampled school (i.e., they belong to the same stratum and have a similar size). However, since information on the originally sampled schools is very limited, one cannot be certain that there are no systematic differences between the sampled and their replacement schools that could cause nonresponse on one side but not on the other. Therefore, the bias risk is not quantifiable; this is why the use of replacement schools is strictly limited in IEA studies. Countries that meet the minimum participation requirements only after including replacement schools get annotated in the international reports.

Table 1 Percentages of countries failing the participation rate requirements in IEA studies (last 10 years)

| Study, cycle and target population | Number of participating countries | Countries failing participation rate requirements* at | | | | | |
|---|-----------------------------------|---|-------|----------------------------------|-------|-----------------------------------|-------|
| | | School level (before replacement) | | School level (after replacement) | | Within schools (individual level) | |
| | | % | Count | % | Count | % | Count |
| Studies with school students comprising the target population | | | | | | | |
| ICCS 2009—grade 8 students | 42 | 33.3 | 14 | 14.3 | 6 | 0.0 | 0 |
| ICILS 2013—grade 8 students | 21 | 33.3 | 7 | 23.8 | 5 | 9.5 | 2 |
| PIRLS 2006—grade 4 students | 47 | 12.8 | 6 | 4.3 | 2 | 4.3 | 2 |
| PIRLS 2011—grade 4 students | 58 | 13.8 | 8 | 1.7 | 1 | 1.7 | 1 |
| TIMSS 2007—grade 4 students | 43 | 20.9 | 9 | 7.0 | 3 | 0.0 | 0 |
| TIMSS 2007—grade 8 students | 56 | 14.3 | 8 | 7.1 | 4 | 0.0 | 0 |
| TIMSS 2011—grade 4 students | 50 | 20.0 | 10 | 6.0 | 3 | 0.0 | 0 |
| TIMSS 2011—grade 8 students | 59 | 6.8 | 4 | 3.4 | 2 | 0.0 | 0 |
| TIMSS advanced 2008 | | | | | | | |
| Advanced mathematics students | 10 | 10.0 | 1 | 10.0 | 1 | 0.0 | 0 |
| TIMSS advanced 2008 | | | | | | | |
| Physics students | 9 | 22.2 | 2 | 0.0 | 0 | 11.1 | 1 |
| Overall | 395 | 17.5 | 69 | 6.8 | 27 | 1.5 | 6 |
| Studies with adults comprising the target population | | | | | | | |
| ICCS 2009—grade 8 teachers | 37 | 37.8 | 14 | 29.7 | 11 | 21.6 | 8 |
| ICILS 2013—grade 8 teachers | 21 | 42.9 | 9 | 38.1 | 8 | 28.6 | 6 |
| SITES 2006—math teachers | 22 | 59.1 | 13 | 40.9 | 9 | 22.7 | 5 |
| SITES 2006—science teachers | 22 | 59.1 | 13 | 40.9 | 9 | 31.8 | 7 |
| TEDS-M 2008—future primary | | | | | | | |
| Math teachers** | 16 | 6.3 | 1 | 6.3 | 1 | 37.5 | 6 |
| Secondary math teachers** | | | | | | | |
| TEDS-M 2008—future lower | 16 | 31.3 | 5 | 31.3 | 5 | 56.3 | 9 |
| TEDS-M 2008 | | | | | | | |
| University educators** | 16 | 37.5 | 6 | 37.5 | 6 | 37.5 | 6 |
| Overall | 150 | 40.7 | 61 | 32.7 | 49 | 31.3 | 47 |

* Minimum participation rates are 85% for schools; 85% for individuals within schools (computed across all participating schools)

** In TEDS-M, “schools” were institutions that offer education programs for future primary or lower-secondary mathematics teachers

In conclusion, IEA studies face a non-negligible amount of nonresponse, which occurs especially at school level in student surveys and at both sampling stages when adults are the target population. Therefore, enhancing methods of analyzing and addressing non-response is of general importance in order to attain evidence that study results remain unaffected by nonresponse.

Results

Association of school-level variables with student-level results using selected IEA survey data

The analyses and procedural steps explicated in this section were carried out with the goal to develop a shortened school questionnaire. This questionnaire would have variables that could comprise a regression model with a high explanatory power on the

school's average achievement score. Analysis was conducted first with data of TIMSS 2011, grade 4, and repeated with data of TIMSS 2011, grade 8 and PIRLS 2011.

As the first step, we calculated mathematics or reading score averages by school (across students and plausible values) and merged these with school level data. Then, we determined the relationship between each variable from the school questionnaire with the average student achievement by running a correlation analysis for each participating country, weighted by the school level weight (SCHWGT).

Standardized Questionnaire

In an effort to develop a questionnaire that may work in a standardized format for any participating country, we considered now all variables with cross-country average correlation coefficients $r \geq \pm 0.2$ for further analysis. Table 2 shows which variables fulfilled this condition in the considered studies. As can be seen, some variables fulfill the criterion in all studies; others only in one or two. In TIMSS grade 4, only six variables fulfilled the criterion while ten and eleven variables respectively were kept for TIMSS grade 8 and PIRLS. Then, we ran regression models separately for each country and study as

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

with y being the students' achievement score averaged at school level, α being the intercept of the regression equation, β comprising the regression coefficients (assuming linear effects on the school mean scores), x the relevant school questionnaire variables, and subscript n denoting the number of variables included into the model. We estimated and reported the adjusted R^2 of each model, which is the portion of the average achievement scores' variance explained by the model. For any given country and study, we started with a model with only one variable and added then step by step the next considered variable to the model in order to monitor the increase in R^2 . As expected, the explained variance portion varied significantly between countries as shown in the Tables 3, 4, 5. The standard model explained as much as 77% of the achievement scores' variance in Chinese Taipei (PIRLS), 67% in Korea (TIMSS grade 8) and 66% again in Chinese Taipei (TIMSS grade 4). To get an overview on the effectiveness of the models across countries, we computed the cross-country average of R^2 for each model and study (Table 6). On average across countries, the explained variance was 34% for PIRLS (model with 11 variables), 24% for TIMSS grade 4 (model with 6 variables) and 36% for TIMSS grade 8 (model with 10 variables).

Country-specific questionnaires

Often times, the standardized models were able to explain a relatively high level of variation between the school's student achievement averages in some countries but not always in others. Therefore, we instead considered applying tailor-cut models for specific countries. We conducted respective analyses exemplarily for the five countries with the lowest participation rates in PIRLS 2011—Belgium (French), England, Netherlands, Northern Ireland and Norway. In order to determine the best fitting model for each country, we fitted regression models with stepwise in-/exclusion of the variables according to specific model parameters (probability of F for entry = 0.05 and of 0.1 for removal). We selected the model solution with 11 variables in order to be able to

Table 2 School questionnaire variables with cross-country average correlation coefficients $r \geq \pm 0.2$ with the students' achievement scores averaged at school level

| Label | Name | Used in model* |
|---|----------|-------------------------|
| Approximately what percentages of students in your school have the following backgrounds? (Come from economically disadvantaged homes) | BCBG03A | T8—Model 1 to Model 10 |
| | ACBG03A | P4—Model 11 |
| | | T4—Model 1 to Model 6 |
| Approximately what percentages of students in your school have the following backgrounds? (Come from economically affluent homes) | BCBG03B | T8—Model 2 to Model 10 |
| | ACBG03B | P4—Model 10 to Model 11 |
| | | T4—Model 2 to Model 6 |
| How many people live in the city, town, or area where your school is located? | BCBG05A | T8—Model 8 to Model 10 |
| | ACBG05A | P4—Model 3 to Model 11 |
| Which best describes the immediate area in which your school is located? | BCBG05B | T8—Model 7 to Model 10 |
| Which best characterizes the average income level of the school's immediate area? | BCBG05C | T8—Model 10 |
| | ACBG05C | P4—Model 1 to Model 11 |
| | | T4—Model 3 to Model 6 |
| How would you characterize each of the following within your school? (Teachers' expectations for student achievement) | BCBG11D | T8—Model 3 to Model 10 |
| | ACBG12D | P4—Model 5 to Model 11 |
| How would you characterize each of the following within your school? (Parental support for student achievement) | BCBG11E | T8—Model 5 to Model 10 |
| | ACBG12E | P4—Model 7 to Model 11 |
| | | T4—Model 4 to Model 6 |
| How would you characterize each of the following within your school? (Parental involvement in school activities) | BCBG11F | T8—Model 6 to Model 10 |
| | ACBG12F | P4—Model 9 to Model 11 |
| | | T4—Model 5 to Model 6 |
| How would you characterize each of the following within your school? (Students' desire to do well at school) | BCBG11H | T8—Model 4 to Model 10 |
| | ACBG12H | P4—Model 8 to Model 11 |
| | | T4—Model 6 |
| To what degree is each of the following a problem among <fourth/eighth-grade> students in your school? (Unjustified absenteeism) | BCBG12AB | T8—Model 9 to Model 10 |
| | ACBG13AB | P4—Model 2 to Model 11 |
| About how many of the students in your school can do the following when they begin primary/elementary school? (Read some words) | ACBG16B | P4—Model 4 to Model 11 |
| About how many of the students in your school can do the following when they begin primary/elementary school? (Recognize most of the letters of the alphabet) | ACBG16A | P4—Model 5 to Model 11 |

* T4—TIMSS grade 4; T8—TIMSS grade 8; P4—PIRLS

compare the country-specific models with the standard model. As shown in Table 7 the standard model was as good as the tailor-cut model in Belgium (French) and England, while R^2 of the country-specific model was higher in Northern Ireland, Netherlands and Norway. The variables included in the country specific models are presented in Table 8.

Discussion and conclusions

We showed in this article that a significant portion of the variance of the school averaged student achievement scores could be explained based on relatively few variables from TIMSS and PIRLS school questionnaires. Therefore, the risk of bias due to nonresponse could be evaluated in effective and efficient ways when collecting this information from nonresponding schools. With the information at hand, one could compare the school characteristics of responding and nonresponding schools, bearing in mind that the compared characteristics are associated with the main outcome variables. Further,

Table 3 TIMSS grade 4—explained variance of school-averaged mathematics score by model and country

| Country | Model 1* R ² (adj.) | Model 2* R ² (adj.) | Model 3* R ² (adj.) | Model 4* R ² (adj.) | Model 5* R ² (adj.) | Model 6* R ² (adj.) |
|--------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| Armenia | 0.00 (0.01) | 0.01 (0.02) | 0.01 (0.03) | 0.08 (0.06) | 0.11 (0.07) | 0.11 (0.08) |
| Australia | 0.31 (0.06) | 0.33 (0.07) | 0.33 (0.08) | 0.41 (0.06) | 0.42 (0.06) | 0.42 (0.06) |
| Austria | 0.09 (0.06) | 0.15 (0.06) | 0.16 (0.06) | 0.19 (0.07) | 0.19 (0.07) | 0.20 (0.07) |
| Republic of Azerbaijan | 0.02 (0.03) | 0.05 (0.05) | 0.06 (0.06) | 0.05 (0.06) | 0.07 (0.05) | 0.09 (0.07) |
| Bahrain | 0.12 (0.05) | 0.14 (0.06) | 0.15 (0.06) | 0.21 (0.07) | 0.22 (0.07) | 0.23 (0.07) |
| Belgium (Flemish) | 0.21 (0.08) | 0.21 (0.08) | 0.23 (0.09) | 0.27 (0.08) | 0.28 (0.08) | 0.30 (0.08) |
| Botswana | 0.25 (0.09) | 0.29 (0.09) | 0.32 (0.09) | 0.35 (0.10) | 0.37 (0.10) | 0.37 (0.10) |
| Canada (Alberta) | 0.26 (0.09) | 0.35 (0.09) | 0.36 (0.10) | 0.36 (0.10) | 0.37 (0.10) | 0.37 (0.10) |
| Canada (Ontario) | 0.19 (0.06) | 0.22 (0.07) | 0.25 (0.07) | 0.28 (0.06) | 0.28 (0.06) | 0.28 (0.06) |
| Canada (Quebec) | 0.07 (0.05) | 0.07 (0.06) | 0.07 (0.07) | 0.14 (0.08) | 0.15 (0.08) | 0.16 (0.09) |
| Chile | 0.37 (0.05) | 0.38 (0.06) | 0.39 (0.06) | 0.39 (0.06) | 0.39 (0.06) | 0.47 (0.07) |
| Chinese Taipei | 0.50 (0.09) | 0.60 (0.06) | 0.65 (0.06) | 0.66 (0.06) | 0.66 (0.07) | 0.66 (0.07) |
| Croatia | 0.15 (0.09) | 0.17 (0.10) | 0.17 (0.10) | 0.17 (0.10) | 0.17 (0.09) | 0.20 (0.09) |
| Czech Republic | 0.03 (0.04) | 0.03 (0.04) | 0.11 (0.08) | 0.14 (0.07) | 0.15 (0.08) | 0.21 (0.06) |
| Denmark | 0.12 (0.08) | 0.09 (0.07) | 0.09 (0.07) | 0.10 (0.07) | 0.10 (0.06) | 0.10 (0.06) |
| England | 0.24 (0.07) | 0.30 (0.08) | 0.31 (0.08) | 0.35 (0.07) | 0.36 (0.07) | 0.36 (0.07) |
| Finland | 0.12 (0.06) | 0.12 (0.06) | 0.12 (0.07) | 0.17 (0.06) | 0.18 (0.06) | 0.21 (0.08) |
| Georgia | 0.05 (0.05) | 0.03 (0.05) | 0.04 (0.05) | 0.05 (0.06) | 0.05 (0.06) | 0.06 (0.07) |
| Germany | 0.44 (0.06) | 0.44 (0.06) | 0.44 (0.07) | 0.49 (0.07) | 0.49 (0.07) | 0.51 (0.07) |
| Republic of Honduras | 0.04 (0.05) | 0.13 (0.09) | 0.13 (0.08) | 0.13 (0.08) | 0.13 (0.09) | 0.15 (0.08) |
| Hong Kong, SAR | 0.01 (0.06) | 0.00 (0.12) | 0.01 (0.12) | 0.01 (0.12) | 0.03 (0.11) | 0.05 (0.09) |
| Hungary | 0.37 (0.06) | 0.39 (0.06) | 0.40 (0.06) | 0.40 (0.06) | 0.40 (0.06) | 0.41 (0.06) |
| Islamic Republic of Iran | 0.14 (0.05) | 0.15 (0.06) | 0.16 (0.06) | 0.21 (0.06) | 0.21 (0.06) | 0.21 (0.06) |
| Ireland | 0.19 (0.05) | 0.28 (0.07) | 0.29 (0.08) | 0.32 (0.08) | 0.32 (0.08) | 0.33 (0.09) |
| Italy | 0.00 (0.01) | 0.01 (0.01) | 0.01 (0.03) | 0.02 (0.04) | 0.03 (0.04) | 0.03 (0.04) |
| Japan | 0.04 (0.05) | 0.07 (0.06) | 0.11 (0.08) | 0.15 (0.07) | 0.18 (0.07) | 0.28 (0.07) |
| Kazakhstan | 0.01 (0.03) | 0.01 (0.04) | 0.12 (0.06) | 0.12 (0.06) | 0.13 (0.06) | 0.13 (0.06) |
| Republic of Korea | 0.31 (0.11) | 0.39 (0.11) | 0.53 (0.08) | 0.56 (0.08) | 0.56 (0.08) | 0.56 (0.08) |
| Kuwait | 0.03 (0.04) | 0.08 (0.05) | 0.11 (0.05) | 0.17 (0.07) | 0.22 (0.06) | 0.22 (0.07) |
| Lithuania | 0.20 (0.08) | 0.21 (0.08) | 0.22 (0.08) | 0.22 (0.09) | 0.22 (0.08) | 0.23 (0.09) |
| Malta | 0.19 (0.07) | 0.21 (0.08) | 0.22 (0.08) | 0.21 (0.07) | 0.20 (0.07) | 0.20 (0.07) |
| Morocco | 0.00 (0.01) | 0.05 (0.04) | 0.06 (0.05) | 0.08 (0.06) | 0.08 (0.05) | 0.15 (0.08) |
| Netherlands | 0.23 (0.08) | 0.22 (0.10) | 0.29 (0.11) | 0.29 (0.12) | 0.29 (0.12) | 0.27 (0.12) |
| New Zealand | 0.39 (0.05) | 0.43 (0.06) | 0.43 (0.06) | 0.47 (0.06) | 0.47 (0.06) | 0.49 (0.06) |
| Northern Ireland | 0.40 (0.06) | 0.42 (0.06) | 0.40 (0.06) | 0.41 (0.07) | 0.44 (0.08) | 0.50 (0.07) |
| Norway | 0.03 (0.03) | 0.03 (0.03) | 0.04 (0.04) | 0.09 (0.06) | 0.13 (0.05) | 0.13 (0.06) |
| Oman | 0.04 (0.03) | 0.04 (0.03) | 0.08 (0.04) | 0.08 (0.04) | 0.10 (0.04) | 0.11 (0.04) |
| Poland | 0.07 (0.07) | 0.08 (0.07) | 0.21 (0.07) | 0.22 (0.07) | 0.22 (0.07) | 0.22 (0.07) |
| Portugal | 0.07 (0.08) | 0.04 (0.08) | 0.04 (0.08) | 0.03 (0.07) | 0.03 (0.08) | 0.05 (0.07) |
| Qatar | 0.04 (0.03) | 0.06 (0.03) | 0.07 (0.04) | 0.15 (0.06) | 0.15 (0.06) | 0.16 (0.06) |
| Romania | 0.00 (0.01) | 0.01 (0.03) | 0.01 (0.03) | 0.03 (0.04) | 0.03 (0.05) | 0.08 (0.06) |
| Russian Federation | 0.07 (0.06) | 0.07 (0.07) | 0.09 (0.07) | 0.09 (0.07) | 0.10 (0.08) | 0.10 (0.08) |
| Saudi Arabia | 0.01 (0.02) | 0.03 (0.05) | 0.05 (0.05) | 0.05 (0.05) | 0.05 (0.05) | 0.05 (0.05) |
| Serbia | 0.00 (0.01) | 0.07 (0.04) | 0.08 (0.05) | 0.08 (0.05) | 0.08 (0.06) | 0.10 (0.06) |
| Singapore | 0.17 (0.05) | 0.25 (0.06) | 0.25 (0.06) | 0.29 (0.06) | 0.29 (0.06) | 0.30 (0.06) |
| Slovak Republic | 0.06 (0.05) | 0.10 (0.08) | 0.10 (0.08) | 0.11 (0.08) | 0.18 (0.09) | 0.19 (0.09) |
| Slovenia | 0.04 (0.04) | 0.05 (0.05) | 0.05 (0.05) | 0.09 (0.07) | 0.09 (0.08) | 0.12 (0.09) |

Table 3 continued

| Country | Model 1* R ² (adj.) | Model 2* R ² (adj.) | Model 3* R ² (adj.) | Model 4* R ² (adj.) | Model 5* R ² (adj.) | Model 6* R ² (adj.) |
|----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| Spain | 0.20 (0.09) | 0.19 (0.09) | 0.27 (0.09) | 0.32 (0.09) | 0.33 (0.09) | 0.33 (0.09) |
| Sweden | 0.25 (0.07) | 0.26 (0.08) | 0.35 (0.09) | 0.35 (0.08) | 0.35 (0.09) | 0.36 (0.08) |
| Thailand | 0.01 (0.01) | 0.08 (0.05) | 0.07 (0.06) | 0.08 (0.06) | 0.14 (0.08) | 0.15 (0.07) |
| Tunisia | 0.11 (0.05) | 0.17 (0.06) | 0.17 (0.06) | 0.22 (0.05) | 0.22 (0.06) | 0.23 (0.05) |
| Turkey | 0.14 (0.04) | 0.15 (0.05) | 0.18 (0.05) | 0.28 (0.07) | 0.33 (0.08) | 0.33 (0.07) |
| United Arab Emirates | 0.04 (0.02) | 0.06 (0.03) | 0.07 (0.03) | 0.13 (0.04) | 0.13 (0.04) | 0.15 (0.04) |
| United Arab Emirates (Abu Dhabi) | 0.02 (0.02) | 0.03 (0.04) | 0.04 (0.05) | 0.07 (0.06) | 0.08 (0.06) | 0.13 (0.06) |
| United Arab Emirates (Dubai) | 0.15 (0.07) | 0.12 (0.07) | 0.22 (0.07) | 0.33 (0.07) | 0.32 (0.07) | 0.32 (0.07) |
| United States | 0.09 (0.17) | 0.08 (0.18) | 0.09 (0.23) | 0.10 (0.22) | 0.10 (0.22) | 0.16 (0.22) |
| Yemen | 0.07 (0.08) | 0.13 (0.11) | 0.14 (0.11) | 0.20 (0.12) | 0.29 (0.13) | 0.29 (0.13) |
| Yemen (grade 6) | 0.01 (0.02) | 0.01 (0.03) | 0.03 (0.04) | 0.03 (0.05) | 0.11 (0.07) | 0.11 (0.07) |
| Cross-country average | 0.13 (0.01) | 0.16 (0.01) | 0.18 (0.01) | 0.21 (0.01) | 0.22 (0.01) | 0.24 (0.01) |

Standard errors appear in parenthesis

* Variables included in model: Model 1: ACBG03A; Model 2: ACBG03A, ACBG03B; Model 3: ACBG03A, ACBG03B, ACBG05C; Model 4: ACBG03A, ACBG03B, ACBG05C, ACBG12E; Model 5: ACBG03A, ACBG03B, ACBG05C, ACBG12E, ACBG12F; Model 6: ACBG03A, ACBG03B, ACBG05C, ACBG12E, ACBG12F, ACBG12H

using the regression coefficients, one may estimate average achievement scores of the nonresponding schools and compare them (i.e., means, distributions) with those of the responding schools. In this case, country-specific models are preferable because they have fewer multicollinearity problems. The results of these analyses could be presented in the studies' technical documentations and may inform sample adjudication.

A more conclusive and consequent step would be to replace the non-informative response model for nonresponse adjustments by a model that uses the information collected from nonresponding schools. One possibility would be to estimate the response propensities of respondents by logistic regression models and compute the weight adjustment factors based on these models (e.g., Lepidus Carlson and Williams 2001; Watson 2012). However, this approach can result in rather unstable adjustment coefficients (Joncas 2015, personal communication). A more robust method would be to use the results of the logistic regression analysis to define more effective adjustment cells than those used by default, since propensity rank strata can render the nonresponse adjustment more 'stable'. To date, the only information used for school-level nonresponse adjustment in IEA studies is schools' allocation to explicit strata. In TIMSS, PIRLS and ICCS, the variance of the achievement scores explained by the explicit stratification is however only about 5% on average (source: own computations); this is why the models presented in this paper explain five to seven times higher portions of this variance.

While the current standard approach of adjusting for non-response is acceptable and valid in all countries with high participation rates, the current adjustment methods can be improved by the use of nonresponse questionnaires to lower the risk of bias. Therefore, school nonresponse questionnaires may be applied in future studies in countries experiencing low participation rates in past assessments or that foresee such problems. We believe that high response rates could be achieved for such questionnaires, because the burden of completing them is considerably lower compared to a full study participation of the school. However, great care is needed to develop procedures on how to

Table 4 TIMSS grade 8—explained variance of school-averaged mathematics score by model and country

| Country | Model 1* R ² (adj.) | Model 2* R ² (adj.) | Model 3* R ² (adj.) | Model 4* R ² (adj.) | Model 5* R ² (adj.) | Model 6* R ² (adj.) | Model 7* R ² (adj.) | Model 8* R ² (adj.) | Model 9* R ² (adj.) | Model 10* R ² (adj.) |
|--------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|
| Armenia | 0.06 (0.07) | 0.07 (0.07) | 0.14 (0.06) | 0.14 (0.06) | 0.14 (0.07) | 0.15 (0.07) | 0.21 (0.09) | 0.21 (0.09) | 0.26 (0.09) | 0.27 (0.09) |
| Australia | 0.20 (0.06) | 0.22 (0.06) | 0.28 (0.06) | 0.33 (0.06) | 0.33 (0.06) | 0.33 (0.06) | 0.35 (0.07) | 0.36 (0.08) | 0.39 (0.08) | 0.40 (0.08) |
| Bahrain | 0.11 (0.06) | 0.14 (0.08) | 0.26 (0.08) | 0.39 (0.11) | 0.41 (0.11) | 0.43 (0.11) | 0.44 (0.11) | 0.43 (0.12) | 0.43 (0.13) | 0.43 (0.13) |
| Botswana | 0.32 (0.10) | 0.37 (0.10) | 0.46 (0.10) | 0.50 (0.11) | 0.52 (0.11) | 0.51 (0.11) | 0.56 (0.08) | 0.60 (0.07) | 0.60 (0.07) | 0.64 (0.07) |
| Canada (Alberta) | 0.20 (0.09) | 0.27 (0.09) | 0.27 (0.09) | 0.29 (0.10) | 0.30 (0.10) | 0.32 (0.10) | 0.32 (0.10) | 0.36 (0.11) | 0.36 (0.11) | 0.37 (0.11) |
| Canada (Ontario) | 0.07 (0.04) | 0.13 (0.05) | 0.15 (0.05) | 0.25 (0.05) | 0.26 (0.06) | 0.27 (0.06) | 0.29 (0.08) | 0.29 (0.08) | 0.29 (0.08) | 0.29 (0.08) |
| Canada (Quebec) | 0.01 (0.08) | 0.03 (0.08) | 0.07 (0.13) | 0.07 (0.14) | 0.15 (0.08) | 0.15 (0.08) | 0.16 (0.09) | 0.16 (0.10) | 0.19 (0.09) | 0.20 (0.10) |
| Chile | 0.26 (0.08) | 0.25 (0.10) | 0.36 (0.09) | 0.46 (0.06) | 0.47 (0.07) | 0.47 (0.06) | 0.48 (0.06) | 0.51 (0.06) | 0.53 (0.06) | 0.54 (0.06) |
| Chinese Taipei | 0.26 (0.14) | 0.38 (0.12) | 0.38 (0.12) | 0.44 (0.13) | 0.44 (0.13) | 0.48 (0.13) | 0.52 (0.12) | 0.52 (0.12) | 0.52 (0.12) | 0.54 (0.11) |
| England | 0.07 (0.05) | 0.13 (0.08) | 0.15 (0.09) | 0.15 (0.08) | 0.16 (0.09) | 0.16 (0.09) | 0.17 (0.09) | 0.18 (0.09) | 0.26 (0.10) | 0.26 (0.10) |
| Finland | 0.05 (0.05) | 0.04 (0.04) | 0.11 (0.06) | 0.14 (0.07) | 0.16 (0.08) | 0.16 (0.08) | 0.17 (0.07) | 0.20 (0.07) | 0.24 (0.07) | 0.26 (0.07) |
| Georgia | 0.04 (0.03) | 0.09 (0.08) | 0.14 (0.09) | 0.14 (0.09) | 0.15 (0.09) | 0.18 (0.09) | 0.18 (0.09) | 0.23 (0.09) | 0.23 (0.09) | 0.24 (0.10) |
| Ghana | 0.04 (0.04) | 0.08 (0.05) | 0.15 (0.06) | 0.19 (0.07) | 0.19 (0.07) | 0.19 (0.08) | 0.27 (0.08) | 0.29 (0.09) | 0.31 (0.10) | 0.36 (0.09) |
| Republic of Honduras | 0.11 (0.09) | 0.14 (0.11) | 0.17 (0.11) | 0.18 (0.11) | 0.18 (0.11) | 0.22 (0.13) | 0.26 (0.12) | 0.28 (0.18) | 0.40 (0.14) | 0.41 (0.14) |
| Hong Kong, SAR | 0.19 (0.10) | 0.21 (0.10) | 0.35 (0.08) | 0.39 (0.08) | 0.39 (0.08) | 0.39 (0.08) | 0.43 (0.08) | 0.45 (0.08) | 0.57 (0.07) | 0.57 (0.07) |
| Hungary | 0.30 (0.07) | 0.34 (0.06) | 0.41 (0.07) | 0.42 (0.07) | 0.42 (0.07) | 0.42 (0.06) | 0.46 (0.08) | 0.48 (0.08) | 0.53 (0.09) | 0.54 (0.09) |
| Indonesia | 0.11 (0.08) | 0.10 (0.08) | 0.12 (0.08) | 0.12 (0.08) | 0.12 (0.09) | 0.13 (0.09) | 0.16 (0.09) | 0.17 (0.10) | 0.17 (0.10) | 0.18 (0.09) |
| Islamic Republic of Iran | 0.23 (0.07) | 0.22 (0.07) | 0.23 (0.08) | 0.26 (0.09) | 0.27 (0.09) | 0.27 (0.09) | 0.35 (0.09) | 0.41 (0.08) | 0.41 (0.08) | 0.42 (0.09) |
| Israel | 0.33 (0.09) | 0.36 (0.09) | 0.36 (0.10) | 0.38 (0.10) | 0.43 (0.09) | 0.43 (0.09) | 0.44 (0.09) | 0.45 (0.09) | 0.48 (0.08) | 0.51 (0.08) |
| Italy | 0.10 (0.05) | 0.14 (0.08) | 0.20 (0.08) | 0.22 (0.08) | 0.23 (0.08) | 0.26 (0.08) | 0.26 (0.08) | 0.26 (0.08) | 0.30 (0.09) | 0.30 (0.09) |
| Japan | 0.22 (0.10) | 0.27 (0.11) | 0.30 (0.10) | 0.31 (0.10) | 0.31 (0.10) | 0.32 (0.10) | 0.42 (0.08) | 0.42 (0.08) | 0.44 (0.08) | 0.45 (0.08) |
| Jordan | 0.14 (0.08) | 0.14 (0.10) | 0.22 (0.06) | 0.28 (0.07) | 0.31 (0.07) | 0.32 (0.07) | 0.37 (0.07) | 0.36 (0.07) | 0.37 (0.06) | 0.37 (0.06) |
| Kazakhstan | 0.02 (0.03) | 0.04 (0.04) | 0.04 (0.04) | 0.05 (0.04) | 0.07 (0.05) | 0.08 (0.05) | 0.11 (0.04) | 0.12 (0.04) | 0.18 (0.06) | 0.32 (0.07) |
| Republic of Korea | 0.21 (0.05) | 0.37 (0.08) | 0.37 (0.10) | 0.47 (0.10) | 0.55 (0.10) | 0.56 (0.10) | 0.60 (0.10) | 0.60 (0.10) | 0.65 (0.08) | 0.67 (0.08) |
| Lebanon | 0.23 (0.09) | 0.19 (0.09) | 0.26 (0.09) | 0.31 (0.09) | 0.33 (0.09) | 0.33 (0.09) | 0.32 (0.09) | 0.31 (0.09) | 0.33 (0.07) | 0.35 (0.08) |
| Lithuania | 0.08 (0.08) | 0.17 (0.06) | 0.20 (0.06) | 0.25 (0.07) | 0.27 (0.08) | 0.27 (0.08) | 0.38 (0.07) | 0.40 (0.07) | 0.41 (0.07) | 0.41 (0.07) |
| Macedonia | 0.17 (0.06) | 0.20 (0.07) | 0.21 (0.06) | 0.21 (0.06) | 0.28 (0.08) | 0.31 (0.09) | 0.43 (0.08) | 0.43 (0.08) | 0.46 (0.08) | 0.46 (0.08) |

Table 4 continued

| Country | Model 1* R ² (adj.) | Model 2* R ² (adj.) | Model 3* R ² (adj.) | Model 4* R ² (adj.) | Model 5* R ² (adj.) | Model 6* R ² (adj.) | Model 7* R ² (adj.) | Model 8* R ² (adj.) | Model 9* R ² (adj.) | Model 10* R ² (adj.) |
|----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|
| Malaysia | 0.04 (0.03) | 0.04 (0.03) | 0.04 (0.04) | 0.06 (0.05) | 0.06 (0.05) | 0.08 (0.04) | 0.14 (0.06) | 0.17 (0.06) | 0.30 (0.07) | 0.31 (0.07) |
| Morocco | 0.25 (0.10) | 0.36 (0.11) | 0.45 (0.11) | 0.48 (0.10) | 0.49 (0.10) | 0.49 (0.10) | 0.53 (0.09) | 0.55 (0.09) | 0.55 (0.09) | 0.55 (0.09) |
| New Zealand | 0.24 (0.07) | 0.30 (0.06) | 0.30 (0.06) | 0.33 (0.06) | 0.37 (0.07) | 0.37 (0.07) | 0.39 (0.06) | 0.40 (0.06) | 0.40 (0.06) | 0.42 (0.06) |
| Oman | 0.05 (0.03) | 0.09 (0.05) | 0.11 (0.05) | 0.14 (0.07) | 0.17 (0.06) | 0.21 (0.06) | 0.21 (0.06) | 0.25 (0.07) | 0.25 (0.07) | 0.25 (0.07) |
| Palestinian National Authority | 0.03 (0.04) | 0.04 (0.04) | 0.05 (0.04) | 0.10 (0.07) | 0.11 (0.07) | 0.13 (0.07) | 0.13 (0.07) | 0.13 (0.07) | 0.14 (0.07) | 0.14 (0.07) |
| Qatar | −0.01 (0.00) | −0.02 (0.01) | 0.02 (0.05) | 0.02 (0.05) | 0.01 (0.05) | 0.03 (0.06) | 0.02 (0.07) | 0.19 (0.12) | 0.19 (0.12) | 0.20 (0.11) |
| Romania | 0.06 (0.04) | 0.07 (0.04) | 0.14 (0.07) | 0.20 (0.06) | 0.20 (0.06) | 0.20 (0.06) | 0.28 (0.06) | 0.30 (0.06) | 0.34 (0.06) | 0.39 (0.07) |
| Russian Federation | 0.08 (0.06) | 0.11 (0.09) | 0.14 (0.09) | 0.14 (0.09) | 0.14 (0.10) | 0.14 (0.10) | 0.15 (0.10) | 0.15 (0.11) | 0.16 (0.11) | 0.16 (0.11) |
| Saudi Arabia | 0.02 (0.04) | 0.04 (0.04) | 0.11 (0.05) | 0.15 (0.09) | 0.19 (0.09) | 0.19 (0.10) | 0.24 (0.09) | 0.25 (0.09) | 0.29 (0.09) | 0.30 (0.09) |
| Singapore | 0.21 (0.06) | 0.24 (0.08) | 0.25 (0.08) | 0.30 (0.08) | 0.31 (0.08) | 0.30 (0.08) | 0.30 (0.08) | 0.30 (0.08) | 0.33 (0.08) | 0.33 (0.08) |
| Slovenia | 0.02 (0.02) | 0.06 (0.04) | 0.06 (0.04) | 0.06 (0.04) | 0.07 (0.04) | 0.09 (0.05) | 0.10 (0.05) | 0.10 (0.05) | 0.10 (0.06) | 0.12 (0.06) |
| South Africa | 0.14 (0.09) | 0.22 (0.08) | 0.24 (0.08) | 0.24 (0.09) | 0.27 (0.10) | 0.37 (0.09) | 0.52 (0.06) | 0.52 (0.07) | 0.52 (0.07) | 0.61 (0.06) |
| Sweden | 0.12 (0.05) | 0.08 (0.05) | 0.13 (0.08) | 0.15 (0.08) | 0.16 (0.10) | 0.16 (0.10) | 0.16 (0.10) | 0.16 (0.10) | 0.44 (0.10) | 0.47 (0.10) |
| Arab Republic of Syria | 0.01 (0.03) | 0.00 (0.02) | 0.01 (0.03) | 0.01 (0.03) | 0.02 (0.03) | 0.03 (0.05) | 0.05 (0.06) | 0.05 (0.05) | 0.09 (0.07) | 0.15 (0.09) |
| Thailand | 0.01 (0.02) | 0.09 (0.10) | 0.09 (0.10) | 0.09 (0.10) | 0.10 (0.10) | 0.10 (0.10) | 0.10 (0.10) | 0.11 (0.10) | 0.12 (0.11) | 0.12 (0.11) |
| Tunisia | 0.05 (0.04) | 0.05 (0.04) | 0.08 (0.06) | 0.08 (0.06) | 0.10 (0.06) | 0.09 (0.06) | 0.10 (0.06) | 0.10 (0.06) | 0.13 (0.06) | 0.16 (0.05) |
| Turkey | 0.21 (0.06) | 0.23 (0.07) | 0.35 (0.08) | 0.41 (0.08) | 0.42 (0.07) | 0.43 (0.07) | 0.46 (0.07) | 0.49 (0.07) | 0.50 (0.07) | 0.50 (0.07) |
| Ukraine | 0.00 (0.01) | 0.01 (0.02) | 0.06 (0.05) | 0.09 (0.07) | 0.09 (0.07) | 0.09 (0.07) | 0.11 (0.09) | 0.18 (0.08) | 0.23 (0.07) | 0.26 (0.06) |
| United Arab Emirates | 0.02 (0.01) | 0.02 (0.01) | 0.17 (0.04) | 0.24 (0.04) | 0.24 (0.04) | 0.25 (0.04) | 0.28 (0.04) | 0.28 (0.04) | 0.31 (0.04) | 0.31 (0.04) |
| United Arab Emirates (Abu Dhabi) | 0.02 (0.03) | 0.01 (0.03) | 0.10 (0.06) | 0.21 (0.07) | 0.21 (0.07) | 0.22 (0.07) | 0.34 (0.07) | 0.37 (0.07) | 0.38 (0.07) | 0.39 (0.06) |
| United Arab Emirates (Dubai) | 0.08 (0.05) | 0.12 (0.05) | 0.38 (0.07) | 0.41 (0.07) | 0.41 (0.06) | 0.42 (0.07) | 0.41 (0.07) | 0.42 (0.06) | 0.44 (0.06) | 0.45 (0.06) |
| United States | 0.27 (0.10) | 0.28 (0.10) | 0.28 (0.10) | 0.38 (0.20) | 0.41 (0.22) | 0.41 (0.22) | 0.41 (0.22) | 0.41 (0.22) | 0.42 (0.23) | 0.42 (0.23) |
| Cross-country average | 0.12 (0.01) | 0.15 (0.01) | 0.20 (0.01) | 0.24 (0.01) | 0.25 (0.01) | 0.26 (0.01) | 0.30 (0.01) | 0.31 (0.01) | 0.34 (0.01) | 0.36 (0.01) |

Standard errors appear in parenthesis

* Variables included in model: Model 1: BCBG03A; Model 2: BCBG03A, BCBG03B; Model 3: BCBG03A, BCBG03B, BCBG11D; Model 4: BCBG03A, BCBG03B, BCBG11D, BCBG11H; Model 5: BCBG03A, BCBG03B, BCBG11D, BCBG11H, BCBG11E; Model 6: BCBG03A, BCBG03B, BCBG11D, BCBG11H, BCBG11E, BCBG11F; Model 7: BCBG03A, BCBG03B, BCBG11D, BCBG11E, BCBG11F; Model 8: BCBG03A, BCBG03B, BCBG11D, BCBG11H, BCBG11E, BCBG11F; Model 9: BCBG03A, BCBG03B, BCBG11D, BCBG11H, BCBG11E, BCBG11F; Model 10: BCBG03A, BCBG03B, BCBG11D, BCBG11H, BCBG11E, BCBG11F, BCBG05B, BCBG05A, BCBG12AB, BCBG05C

Table 5 PIRLS grade 4—explained variance of school-averaged reading score by model and country

| Country | Model 1* R ² (adj.) | Model 2* R ² (adj.) | Model 3* R ² (adj.) | Model 4* R ² (adj.) | Model 5* R ² (adj.) | Model 6* R ² (adj.) | Model 7* R ² (adj.) | Model 8* R ² (adj.) | Model 9* R ² (adj.) | Model 10* R ² (adj.) | Model 11* R ² (adj.) |
|------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|------------------------------------|
| Australia | 0.32 (0.09) | 0.40 (0.07) | 0.40 (0.08) | 0.41 (0.08) | 0.41 (0.08) | 0.42 (0.09) | 0.52 (0.06) | 0.52 (0.06) | 0.52 (0.06) | 0.51 (0.07) | 0.52 (0.06) |
| Austria | 0.08 (0.05) | 0.08 (0.05) | 0.10 (0.05) | 0.14 (0.05) | 0.14 (0.06) | 0.14 (0.05) | 0.14 (0.06) | 0.14 (0.06) | 0.18 (0.06) | 0.30 (0.06) | 0.37 (0.05) |
| Republic of Azerbaijan | 0.02 (0.05) | 0.02 (0.07) | 0.05 (0.08) | 0.07 (0.08) | 0.12 (0.06) | 0.14 (0.08) | 0.14 (0.08) | 0.15 (0.09) | 0.18 (0.08) | 0.17 (0.15) | 0.18 (0.15) |
| Belgium (French) | 0.28 (0.07) | 0.30 (0.05) | 0.32 (0.05) | 0.32 (0.05) | 0.32 (0.06) | 0.32 (0.06) | 0.35 (0.06) | 0.36 (0.06) | 0.40 (0.06) | 0.50 (0.05) | 0.55 (0.06) |
| Botswana | 0.28 (0.08) | 0.40 (0.07) | 0.47 (0.07) | 0.57 (0.08) | 0.58 (0.08) | 0.58 (0.08) | 0.60 (0.08) | 0.61 (0.08) | 0.62 (0.08) | 0.62 (0.08) | 0.58 (0.08) |
| Bulgaria | 0.06 (0.07) | 0.09 (0.09) | 0.15 (0.07) | 0.25 (0.07) | 0.27 (0.08) | 0.28 (0.09) | 0.28 (0.09) | 0.28 (0.10) | 0.30 (0.09) | 0.37 (0.08) | 0.37 (0.08) |
| Canada | 0.11 (0.03) | 0.11 (0.03) | 0.13 (0.03) | 0.14 (0.03) | 0.15 (0.03) | 0.19 (0.04) | 0.22 (0.04) | 0.22 (0.04) | 0.22 (0.04) | 0.23 (0.04) | 0.23 (0.05) |
| Canada (Alberta) | 0.27 (0.08) | 0.30 (0.09) | 0.31 (0.09) | 0.33 (0.09) | 0.34 (0.09) | 0.34 (0.09) | 0.35 (0.09) | 0.36 (0.09) | 0.37 (0.10) | 0.38 (0.10) | 0.37 (0.10) |
| Canada (Ontario) | 0.11 (0.05) | 0.11 (0.05) | 0.12 (0.05) | 0.14 (0.07) | 0.15 (0.07) | 0.27 (0.07) | 0.28 (0.08) | 0.28 (0.08) | 0.28 (0.08) | 0.27 (0.08) | 0.25 (0.08) |
| Canada (Quebec) | 0.06 (0.06) | 0.05 (0.06) | 0.07 (0.08) | 0.08 (0.10) | 0.08 (0.10) | 0.12 (0.10) | 0.13 (0.11) | 0.14 (0.12) | 0.14 (0.14) | 0.12 (0.16) | 0.11 (0.16) |
| Chinese Taipei | 0.40 (0.08) | 0.40 (0.08) | 0.54 (0.09) | 0.61 (0.09) | 0.61 (0.09) | 0.62 (0.10) | 0.73 (0.08) | 0.73 (0.08) | 0.73 (0.08) | 0.75 (0.08) | 0.77 (0.08) |
| Colombia | 0.00 (0.02) | 0.02 (0.04) | 0.22 (0.09) | 0.22 (0.09) | 0.24 (0.09) | 0.31 (0.12) | 0.38 (0.11) | 0.40 (0.10) | 0.40 (0.10) | 0.48 (0.12) | 0.51 (0.12) |
| Croatia | 0.10 (0.06) | 0.09 (0.06) | 0.27 (0.09) | 0.30 (0.10) | 0.30 (0.10) | 0.30 (0.10) | 0.30 (0.10) | 0.31 (0.09) | 0.33 (0.09) | 0.35 (0.09) | 0.36 (0.10) |
| Czech Republic | 0.11 (0.07) | 0.11 (0.08) | 0.14 (0.07) | 0.16 (0.07) | 0.16 (0.07) | 0.17 (0.07) | 0.17 (0.07) | 0.23 (0.07) | 0.23 (0.07) | 0.24 (0.07) | 0.27 (0.08) |
| Denmark | 0.14 (0.05) | 0.15 (0.05) | 0.16 (0.06) | 0.18 (0.06) | 0.18 (0.06) | 0.18 (0.06) | 0.21 (0.05) | 0.21 (0.05) | 0.22 (0.06) | 0.24 (0.06) | 0.25 (0.06) |
| England | 0.27 (0.07) | 0.32 (0.07) | 0.31 (0.08) | 0.35 (0.08) | 0.35 (0.08) | 0.39 (0.09) | 0.39 (0.09) | 0.39 (0.09) | 0.40 (0.09) | 0.50 (0.08) | 0.55 (0.07) |
| Finland | 0.07 (0.05) | 0.08 (0.05) | 0.09 (0.05) | 0.14 (0.06) | 0.14 (0.06) | 0.15 (0.07) | 0.16 (0.07) | 0.21 (0.09) | 0.21 (0.09) | 0.24 (0.09) | 0.28 (0.10) |
| France | 0.25 (0.06) | 0.25 (0.06) | 0.26 (0.06) | 0.25 (0.06) | 0.25 (0.06) | 0.26 (0.06) | 0.30 (0.06) | 0.36 (0.06) | 0.37 (0.06) | 0.37 (0.07) | 0.38 (0.07) |
| Georgia | 0.00 (0.01) | 0.00 (0.01) | 0.08 (0.05) | 0.08 (0.05) | 0.08 (0.05) | 0.15 (0.05) | 0.15 (0.06) | 0.16 (0.06) | 0.17 (0.06) | 0.16 (0.05) | 0.17 (0.06) |
| Germany | 0.08 (0.05) | 0.19 (0.10) | 0.19 (0.10) | 0.19 (0.10) | 0.21 (0.09) | 0.32 (0.10) | 0.45 (0.06) | 0.46 (0.06) | 0.46 (0.06) | 0.47 (0.06) | 0.54 (0.07) |
| Republic of Honduras | 0.13 (0.07) | 0.23 (0.08) | 0.28 (0.09) | 0.28 (0.10) | 0.28 (0.10) | 0.36 (0.08) | 0.36 (0.08) | 0.40 (0.07) | 0.41 (0.08) | 0.43 (0.10) | 0.43 (0.10) |
| Hong Kong, SAR | 0.00 (0.00) | 0.03 (0.05) | 0.02 (0.06) | 0.02 (0.06) | 0.05 (0.07) | 0.05 (0.08) | 0.04 (0.08) | 0.04 (0.09) | 0.05 (0.08) | 0.10 (0.10) | 0.12 (0.10) |
| Hungary | 0.26 (0.06) | 0.48 (0.08) | 0.55 (0.08) | 0.55 (0.08) | 0.56 (0.07) | 0.57 (0.07) | 0.58 (0.06) | 0.58 (0.07) | 0.59 (0.07) | 0.57 (0.08) | 0.60 (0.07) |
| Indonesia | 0.06 (0.04) | 0.08 (0.05) | 0.12 (0.05) | 0.12 (0.05) | 0.13 (0.05) | 0.13 (0.05) | 0.13 (0.06) | 0.14 (0.06) | 0.14 (0.06) | 0.26 (0.08) | 0.26 (0.08) |

Table 5 continued

| Country | Model 1* R ² (adj.) | Model 2* R ² (adj.) | Model 3* R ² (adj.) | Model 4* R ² (adj.) | Model 5* R ² (adj.) | Model 6* R ² (adj.) | Model 7* R ² (adj.) | Model 8* R ² (adj.) | Model 9* R ² (adj.) | Model 10* R ² (adj.) | Model 11* R ² (adj.) |
|--------------------------|-----------------------------------|-----------------------------------|-----------------------------------|--|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|------------------------------------|
| Islamic Republic of Iran | 0.06 (0.04) | 0.08 (0.04) | 0.32 (0.05) | 0.32 (0.05) | 0.33 (0.05) | 0.36 (0.05) | 0.39 (0.06) | 0.40 (0.06) | 0.42 (0.06) | 0.43 (0.06) | 0.43 (0.06) |
| Ireland | 0.06 (0.05) | 0.09 (0.07) | 0.12 (0.08) | Models were not computed because questions ACBG16A and ACBG16B were not administered | | | | | | | |
| Israel | 0.21 (0.06) | 0.33 (0.06) | 0.35 (0.05) | 0.35 (0.05) | 0.35 (0.05) | 0.38 (0.06) | 0.40 (0.05) | 0.41 (0.05) | 0.41 (0.05) | 0.41 (0.06) | 0.45 (0.06) |
| Italy | 0.02 (0.04) | 0.01 (0.04) | 0.01 (0.04) | 0.01 (0.04) | 0.03 (0.05) | 0.03 (0.04) | 0.04 (0.05) | 0.04 (0.05) | 0.04 (0.05) | 0.04 (0.06) | 0.09 (0.06) |
| Kuwait | 0.01 (0.02) | 0.00 (0.02) | 0.00 (0.03) | −0.01 (0.03) | 0.01 (0.05) | 0.02 (0.05) | 0.02 (0.05) | 0.16 (0.09) | 0.18 (0.09) | 0.22 (0.11) | 0.23 (0.12) |
| Lithuania | 0.17 (0.06) | 0.17 (0.05) | 0.30 (0.06) | 0.28 (0.06) | 0.29 (0.06) | 0.30 (0.06) | 0.32 (0.06) | 0.31 (0.06) | 0.34 (0.08) | 0.37 (0.07) | 0.44 (0.08) |
| Malta | 0.09 (0.07) | 0.15 (0.07) | 0.15 (0.09) | 0.16 (0.09) | 0.16 (0.09) | 0.19 (0.08) | 0.24 (0.08) | 0.23 (0.08) | 0.23 (0.07) | 0.23 (0.08) | 0.28 (0.08) |
| Malta (Maltese) | 0.18 (0.06) | 0.21 (0.07) | 0.20 (0.07) | 0.19 (0.07) | 0.19 (0.08) | 0.19 (0.08) | 0.18 (0.08) | 0.17 (0.08) | 0.16 (0.08) | 0.16 (0.09) | 0.21 (0.10) |
| Morocco | 0.10 (0.05) | 0.10 (0.05) | 0.20 (0.06) | 0.20 (0.07) | 0.20 (0.07) | 0.20 (0.07) | 0.20 (0.07) | 0.24 (0.06) | 0.24 (0.06) | 0.36 (0.07) | 0.35 (0.07) |
| Morocco (Grade 6) | 0.14 (0.06) | 0.14 (0.06) | 0.26 (0.05) | 0.27 (0.06) | 0.27 (0.06) | 0.27 (0.06) | 0.32 (0.06) | 0.35 (0.05) | 0.35 (0.05) | 0.38 (0.06) | 0.35 (0.07) |
| Netherlands | 0.22 (0.06) | 0.23 (0.06) | 0.26 (0.07) | 0.27 (0.07) | 0.28 (0.08) | 0.28 (0.07) | 0.28 (0.08) | 0.29 (0.07) | 0.30 (0.07) | 0.32 (0.09) | 0.32 (0.09) |
| New Zealand | 0.19 (0.08) | 0.26 (0.07) | 0.26 (0.07) | 0.29 (0.06) | 0.31 (0.06) | 0.31 (0.06) | 0.31 (0.06) | 0.32 (0.07) | 0.32 (0.07) | 0.33 (0.09) | 0.33 (0.09) |
| Northern Ireland | 0.24 (0.07) | 0.25 (0.06) | 0.24 (0.08) | 0.24 (0.08) | 0.25 (0.08) | 0.29 (0.10) | 0.29 (0.11) | 0.30 (0.11) | 0.31 (0.11) | 0.36 (0.14) | 0.43 (0.12) |
| Norway | 0.03 (0.03) | 0.05 (0.05) | 0.15 (0.08) | 0.18 (0.09) | 0.20 (0.08) | 0.23 (0.08) | 0.27 (0.07) | 0.27 (0.07) | 0.28 (0.07) | 0.29 (0.07) | 0.28 (0.06) |
| Oman | 0.07 (0.03) | 0.14 (0.05) | 0.15 (0.05) | 0.15 (0.05) | 0.20 (0.07) | 0.24 (0.06) | 0.24 (0.07) | 0.24 (0.06) | 0.26 (0.07) | 0.25 (0.07) | 0.29 (0.07) |
| Poland | 0.17 (0.07) | 0.17 (0.07) | 0.24 (0.07) | 0.26 (0.07) | 0.28 (0.07) | 0.29 (0.07) | 0.29 (0.07) | 0.29 (0.07) | 0.32 (0.07) | 0.33 (0.07) | 0.34 (0.07) |
| Portugal | 0.20 (0.08) | 0.20 (0.08) | 0.22 (0.08) | 0.22 (0.08) | 0.23 (0.08) | 0.28 (0.08) | 0.28 (0.08) | 0.29 (0.08) | 0.31 (0.09) | 0.26 (0.10) | 0.30 (0.12) |
| Qatar | 0.03 (0.03) | 0.08 (0.05) | 0.21 (0.08) | 0.24 (0.08) | 0.27 (0.08) | 0.27 (0.08) | 0.30 (0.08) | 0.31 (0.07) | 0.32 (0.07) | 0.31 (0.07) | 0.24 (0.07) |
| Romania | 0.03 (0.03) | 0.06 (0.06) | 0.17 (0.07) | 0.20 (0.07) | 0.21 (0.08) | 0.21 (0.08) | 0.22 (0.08) | 0.25 (0.08) | 0.25 (0.09) | 0.37 (0.07) | 0.36 (0.07) |
| Russian Federation | 0.04 (0.04) | 0.04 (0.04) | 0.14 (0.06) | 0.14 (0.06) | 0.14 (0.06) | 0.15 (0.06) | 0.15 (0.06) | 0.16 (0.06) | 0.19 (0.07) | 0.18 (0.07) | 0.19 (0.09) |
| Saudi Arabia | 0.08 (0.06) | 0.11 (0.07) | 0.11 (0.06) | 0.14 (0.06) | 0.16 (0.06) | 0.20 (0.06) | 0.21 (0.06) | 0.29 (0.05) | 0.29 (0.05) | 0.28 (0.06) | 0.30 (0.07) |
| Singapore | 0.15 (0.06) | 0.17 (0.06) | 0.17 (0.06) | 0.17 (0.06) | 0.17 (0.06) | 0.21 (0.06) | 0.26 (0.06) | 0.26 (0.06) | 0.26 (0.06) | 0.29 (0.06) | 0.32 (0.06) |
| Slovak Republic | 0.03 (0.02) | 0.16 (0.06) | 0.24 (0.06) | 0.25 (0.06) | 0.25 (0.06) | 0.27 (0.07) | 0.28 (0.07) | 0.29 (0.07) | 0.29 (0.07) | 0.33 (0.08) | 0.35 (0.08) |
| Slovenia | 0.02 (0.04) | 0.01 (0.04) | 0.03 (0.05) | 0.04 (0.05) | 0.04 (0.05) | 0.05 (0.05) | 0.09 (0.05) | 0.10 (0.05) | 0.13 (0.05) | 0.14 (0.06) | 0.19 (0.10) |
| South Africa | 0.26 (0.08) | 0.39 (0.09) | 0.48 (0.10) | 0.51 (0.08) | 0.56 (0.08) | 0.57 (0.09) | 0.62 (0.07) | 0.62 (0.07) | 0.62 (0.07) | 0.62 (0.06) | 0.63 (0.07) |

Table 5 continued

| Country | Model 1* | Model 2* | Model 3* | Model 4* | Model 5* | Model 6* | Model 7* | Model 8* | Model 9* | Model 10* | Model 11* |
|----------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) | R ² (adj.) |
| Spain | 0.29 (0.06) | 0.30 (0.07) | 0.31 (0.07) | 0.34 (0.07) | 0.34 (0.07) | 0.34 (0.07) | 0.35 (0.06) | 0.36 (0.06) | 0.36 (0.06) | 0.39 (0.07) | 0.39 (0.08) |
| Spain (Andalucía) | 0.31 (0.07) | 0.31 (0.07) | 0.30 (0.07) | 0.32 (0.07) | 0.34 (0.08) | 0.37 (0.08) | 0.38 (0.08) | 0.40 (0.08) | 0.40 (0.08) | 0.38 (0.08) | 0.43 (0.08) |
| Sweden | 0.10 (0.08) | 0.10 (0.08) | 0.10 (0.08) | 0.13 (0.06) | 0.14 (0.07) | 0.15 (0.07) | 0.16 (0.07) | 0.19 (0.07) | 0.21 (0.08) | 0.20 (0.08) | 0.25 (0.08) |
| Trinidad And Tobago | 0.12 (0.06) | 0.14 (0.07) | 0.17 (0.08) | 0.17 (0.08) | 0.17 (0.08) | 0.18 (0.08) | 0.21 (0.08) | 0.22 (0.08) | 0.21 (0.07) | 0.23 (0.08) | 0.21 (0.08) |
| United Arab Emirates | 0.03 (0.02) | 0.07 (0.03) | 0.10 (0.04) | 0.13 (0.05) | 0.13 (0.05) | 0.13 (0.04) | 0.17 (0.04) | 0.20 (0.05) | 0.20 (0.05) | 0.19 (0.05) | 0.21 (0.04) |
| United Arab Emirates (Abu Dhabi) | 0.01 (0.02) | 0.04 (0.03) | 0.08 (0.04) | 0.17 (0.05) | 0.16 (0.05) | 0.17 (0.05) | 0.17 (0.05) | 0.18 (0.05) | 0.18 (0.05) | 0.17 (0.07) | 0.26 (0.08) |
| United Arab Emirates (Dubai) | 0.19 (0.06) | 0.20 (0.06) | 0.22 (0.07) | 0.22 (0.07) | 0.22 (0.07) | 0.23 (0.07) | 0.26 (0.07) | 0.29 (0.07) | 0.28 (0.08) | 0.27 (0.08) | 0.30 (0.08) |
| United States | 0.33 (0.04) | 0.37 (0.04) | 0.38 (0.05) | 0.37 (0.05) | 0.37 (0.06) | 0.41 (0.05) | 0.41 (0.05) | 0.41 (0.06) | 0.41 (0.06) | 0.46 (0.05) | 0.52 (0.06) |
| Cross-country average | 0.13 (0.01) | 0.17 (0.01) | 0.21 (0.01) | 0.23 (0.01) | 0.24 (0.01) | 0.26 (0.01) | 0.28 (0.01) | 0.29 (0.01) | 0.30 (0.01) | 0.32 (0.01) | 0.34 (0.01) |

Standard errors appear in parenthesis

[illegible]

Table 6 Descriptive statistics of R^2 (explained variance of achievement score) across countries by model and study

| PIRLS | Model | | | | | | | | | | |
|-----------------------|-------|-------|------|-------|------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Minimum | 0.00 | 0.00 | 0.00 | −0.01 | 0.01 | 0.02 | 0.02 | 0.04 | 0.04 | 0.04 | 0.09 |
| Maximum | 0.40 | 0.48 | 0.55 | 0.61 | 0.61 | 0.62 | 0.73 | 0.73 | 0.73 | 0.75 | 0.77 |
| Standard Deviation | 0.10 | 0.12 | 0.13 | 0.13 | 0.13 | 0.13 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 |
| Cross-country average | 0.13 | 0.17 | 0.21 | 0.23 | 0.24 | 0.26 | 0.28 | 0.29 | 0.30 | 0.32 | 0.34 |
| TIMSS grade 4 | Model | | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | | | | | |
| Minimum | 0.00 | 0.00 | 0.01 | 0.01 | 0.03 | 0.03 | | | | | |
| Maximum | 0.50 | 0.60 | 0.65 | 0.66 | 0.66 | 0.66 | | | | | |
| Standard deviation | 0.13 | 0.14 | 0.14 | 0.15 | 0.14 | 0.14 | | | | | |
| Cross-country average | 0.13 | 0.16 | 0.18 | 0.21 | 0.22 | 0.24 | | | | | |
| TIMSS grade 8 | Model | | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| Minimum | −0.01 | −0.02 | 0.01 | 0.01 | 0.01 | 0.03 | 0.02 | 0.05 | 0.09 | 0.12 | |
| Maximum | 0.33 | 0.38 | 0.46 | 0.50 | 0.55 | 0.56 | 0.60 | 0.60 | 0.65 | 0.67 | |
| Standard deviation | 0.10 | 0.11 | 0.12 | 0.14 | 0.14 | 0.14 | 0.15 | 0.15 | 0.14 | 0.14 | |
| Cross-country average | 0.12 | 0.15 | 0.20 | 0.24 | 0.25 | 0.26 | 0.30 | 0.31 | 0.34 | 0.36 | |

Table 7 R^2 (explained variance of achievement score) by country (PIRLS)

| Country | Participation rate of schools (before replacement) (%) | Standard model R^2 (adj.) | Country-specific model R^2 (adj.) |
|------------------|--|-----------------------------|-------------------------------------|
| Belgium (French) | 77 | 0.55 (0.06) | 0.51 (0.08) |
| England | 75 | 0.55 (0.07) | 0.55 (0.07) |
| Netherlands | 68 | 0.32 (0.09) | 0.47 (0.07) |
| Northern Ireland | 62 | 0.43 (0.09) | 0.57 (0.17) |
| Norway | 57 | 0.28 (0.06) | 0.42 (0.12) |

Standard errors appear in parenthesis

administer these questionnaires, ensuring that the participation in the actual survey is not jeopardized. Methodological and financial considerations will determine whether a standard approach (one standardized questionnaire for all affected countries) or a tailored approach (country-specific questionnaires) is more efficient. Further investigations are needed to show whether the presented approach of developing nonresponse questionnaires is also applicable to other large-scale assessments and if nonresponse questionnaires for individuals could be developed in similar ways. Moreover, a study on the feasibility of the practical application is pending. Careful consideration is needed to optimally integrate the administration of such questionnaires in the tight schedule of large-scale assessments. High participation rates would be needed to ensure the usability of this instrument. In this sense, short questionnaires might be favorable, while another option would be to administer full school questionnaires. The latter would simplify data processing and operations, but also be beneficial regarding the quality of the nonresponse bias analysis.

Table 8 PIRLS—variables included in the standard and the country-specific models (5 countries)

| Name | Label | Variables used in standard model | Variables used in country-specific models | | | | |
|----------|---|----------------------------------|---|---------|-------------|------------------|--------|
| | | | Belgium (French) | England | Netherlands | Northern Ireland | Norway |
| ACBG02 | What is the total enrollment of <fourth-grade> students in your school as of <first day of month PIRLS/TIMSS testing begins, 2010/2011>? | | X | | | | |
| ACBG03A | Approximately what percentages of students in your school have the following backgrounds? (Come from economically affluent homes) | X | | X | | X | |
| ACBG03B | Approximately what percentages of students in your school have the following backgrounds? (Come from economically disadvantaged homes) | X | X | X | | | |
| ACBG04 | Approximately what percentages of students in your school have <language of test> as their native language? | | X | | | | |
| ACBG05A | How many people live in the city, town, or area where your school is located? | X | | | X | | |
| ACBG05C | Which best characterizes the average income level of the school's immediate area? | X | | | X | | |
| ACBG06A | How many days per year is your school open for instruction? | | | X | | X | |
| ACBG06BA | What is the total instructional time, excluding breaks, in a typical day? | | | | X | | |
| ACBG09B | Approximately how many titles of magazines and other periodicals does your school library have? | | | | | X | |
| ACBG10AD | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (General School Resources: Heating/cooling and lighting systems) | | X | | | | |
| ACBG10AF | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (General School Resources: Technologically competent staff) | | | | | | X |
| ACBG10AG | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (General School Resources: Computers for instruction) | | | | | | X |
| ACBG10BD | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (Audio-visual resources for reading instruction) | | | X | | | |
| ACBG10CA | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (Teachers with a specialization in mathematics) | | | | | | X |
| ACBG10CB | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (Computer software for mathematics instruction) | | | | | X | |
| ACBG10CC | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (Library materials relevant to mathematics instruction) | | X | | | | |

Table 8 continued

| Name | Label | Variables used in standard model | Variables used in country-specific models | | | | |
|----------|---|----------------------------------|---|---------|-------------|------------------|--------|
| | | | Belgium (French) | England | Netherlands | Northern Ireland | Norway |
| ACBG10CE | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (Calculators for mathematics instruction) | | | | | X | |
| ACBG10DA | How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? (Teachers with a specialization in science) | | | X | | | |
| ACBG11AA | How often does your school do the following for parents concerning individual students? (Inform parents about their child's learning progress) | | | X | | | X |
| ACBG11AB | How often does your school do the following for parents concerning individual students? (Inform parents about the behavior and well-being of their child at school) | | X | | | | |
| ACBG11AC | How often does your school do the following for parents concerning individual students? (Discuss parents' concerns or wishes about their child's learning) | | X | | | | |
| ACBG11BA | How often does your school ask parents to do the following? (Volunteer for school projects, programs, and trips) | | X | | | X | |
| ACBG11BB | How often does your school ask parents to do the following? (Serve on school committees) | | X | | | | |
| ACBG11CB | How often does your school do the following for parents in general? (Inform parents about school accomplishments (e.g., tournament results, facility improvements)) | | X | | | | X |
| ACBG11CC | How often does your school do the following for parents in general? (Inform parents about the educational goals and pedagogic principles of the school) | | | X | | | |
| ACBG12B | How would you characterize each of the following within your school? (Teachers' understanding of the school's curricular goals) | | | | | X | |
| ACBG12D | How would you characterize each of the following within your school? (Teachers' expectations for student achievement) | X | | | | | |
| ACBG12E | How would you characterize each of the following within your school? (Parental support for student achievement) | X | | | | | |
| ACBG12F | How would you characterize each of the following within your school? (Parental involvement in school activities) | X | | | | | |
| ACBG12G | How would you characterize each of the following within your school? (Students' regard for school property) | | | | | X | X |

Table 8 continued

| Name | Label | Variables used in standard model | Variables used in country-specific models | | | | |
|----------|---|----------------------------------|---|---------|-------------|------------------|--------|
| | | | Belgium (French) | England | Netherlands | Northern Ireland | Norway |
| ACBG12H | How would you characterize each of the following within your school? (Students' desire to do well at school) | X | X | | | | X |
| ACBG13AA | To what degree is each of the following a problem among <fourth-grade> students in your school? (Arriving late at school) | | | X | | | |
| ACBG13AB | To what degree is each of the following a problem among <fourth-grade> students in your school? (Unjustified absenteeism) | X | | | | | |
| ACBG13AC | To what degree is each of the following a problem among <fourth-grade> students in your school? (Classroom disturbance) | | | | | X | |
| ACBG13AJ | To what degree is each of the following a problem among <fourth-grade> students in your school? (Intimidation or verbal abuse of teachers or staff (including texting, emailing, etc.)) | | X | | | | |
| ACBG13BB | To what degree is each of the following a problem among teachers in your school? (Absenteeism) | | | X | | | |
| ACBG14A | In your school, are any of the following used to evaluate the practice of <fourth-grade> teachers? (Observations by the principal or senior staff) | | | | X | X | X |
| ACBG14C | In your school, are any of the following used to evaluate the practice of <fourth-grade> teachers? (Student achievement) | | | X | | | |
| ACBG15C | During the past year, approximately how much time have you spent on the following school leadership activities in your role as a school principal? (Monitoring teachers' implementation of the school's educational goals in their teaching) | | | | | | X |
| ACBG15D | During the past year, approximately how much time have you spent on the following school leadership activities in your role as a school principal? (Monitoring students' learning progress to ensure that the school's educational goals are reached) | | | X | | | |
| ACBG15G | During the past year, approximately how much time have you spent on the following school leadership activities in your role as a school principal? (Addressing disruptive student behavior) | | | | X | | |
| ACBG15H | During the past year, approximately how much time have you spent on the following school leadership activities in your role as a school principal? (Creating a climate of trust among teachers) | | | | X | | X |
| ACBG15K | During the past year, approximately how much time have you spent on the following school leadership activities in your role as a school principal? (Visiting other schools or attending educational conferences for new ideas) | | | X | X | | |

Table 8 continued

| Name | Label | Variables used in standard model | Variables used in country-specific models | | | | |
|---------|---|----------------------------------|---|---------|-------------|------------------|--------|
| | | | Belgium (French) | England | Netherlands | Northern Ireland | Norway |
| ACBG16A | About how many of the students in your school can do the following when they begin primary/elementary school? (Recognize most of the letters of the alphabet) | X | | | | | |
| ACBG16B | About how many of the students in your school can do the following when they begin primary/elementary school? (Read some words) | X | | | | | |
| ACBG17G | At which grade do the following reading skills and strategies first receive a major emphasis in instruction in your school? (Identifying the main idea of a text) | | X | | | | |
| ACBG17N | At which grade do the following reading skills and strategies first receive a major emphasis in instruction in your school? (Determining the author's perspective or intention) | | | | X | | |
| ACBG18C | Compared with other areas of the curriculum (e.g., mathematics and science), how much emphasis does your school place on teaching the following language and literacy skills to students in <grades 1–4>? (Speaking/listening (oral language) | | | | | | X |
| ACBG19 | For students in <fourth grade> and below, does your school make provisions for reading instruction in mother tongue for students whose mother tongue is not <language of test>? | | X | X | | | |

Authors' contributions

SM developed the research questions and design, supervised data compilation, conducted major parts of the statistical analysis and interpretation of results and drafted major parts of the manuscript. DC conducted parts of the statistical analysis, drafted minor parts of the manuscript and critically revised all other parts of the manuscript. ST was responsible for data compilation and merging, preparation of data analysis, drafting all tables, and manuscript revision. All authors have given final approval of the manuscript version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Author details

¹ IEA Data Processing and Research Center, Hamburg, Germany. ² Johannes-Gutenberg-University, Mainz, Germany.

Acknowledgements

The authors are thankful to Marc Joncas and Plamen Mirazchiyski and two peer reviewers for their very useful comments.

Competing interests

The authors declare to have no competing interests.

Received: 21 September 2015 Accepted: 16 January 2017

Published online: 31 January 2017

References

- Baker, D. P., Goesling, B., & Letendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the "Heyneman-Loxley effect" on mathematics and science achievement. *Comparative Education Review*, 46(3), 291–312.
- Bethlehem, J. G., & Kersten, H. M. P. (1985). On the treatment of nonresponse in sample surveys. *Journal of Official Statistics*, 1(3), 287–300.
- Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, 90(5), 269–277.
- Fuller, Bruce. (1987). What school factors raise achievement in the Third World. *Review of Educational Research*, 37, 255–293.
- Grace, K., & Thompson, J. S. (2003). Racial and ethnic stratification in Educational Achievement and Attainment. *Annual Review of Sociology*, 29, 417–442.
- Helmschrott, S., & Martin, S. (2014). Nonresponse in PIAAC Germany. *Methods, Data, Analysis I*, 8(2), 243–266.
- Koretz, D., McCaffrey, D., & Sullivan, T. (2001). Predicting variations in mathematics performance in four countries using TIMSS. *Education Policy Analysis Archives*, 9(34). Retrieved on 08/27/2009 from <http://epaa.asu.edu/epaa/v9n34/>.
- Lamb, S., & Fullarton, S. (2001). *Classroom and school factors affecting mathematics achievement: A comparative study of the US and Australia using TIMSS*. Trends in International Mathematics and Science Study (TIMSS), TIMSS Australia Monograph Series, Australian Council for Educational Research.
- Lepidus Carlson, B. & Williams, S. (2001). A comparison of two methods to adjust weights for non-response: propensity modelling and weighting class adjustments. In *Proceedings of the annual meeting of the American Statistical Association*, August 5–9, 2001.
- Lynn, P. (2003). PEDAKSI: Methodology for collecting data about survey nonrespondents. *Quality & Quantity*, 37, 239–261.
- Martin, M. O., & Mullis, I. V. S. (Eds.). (2013). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: Lynch School of Education, Boston College.
- Matsuo, H., Billiet, J., Loosvelt, G., & Kleven, O. (2010). Measurement and adjustment of nonresponse bias based on nonresponse surveys: The case of Belgium and Norway in the European Social Survey Round 3. *Survey Research Methods*, 4(3), 165–178.
- Meinck, S. (2015). *Computing sampling weights in large-scale assessments in education. Survey insights: Methods from the field, weighting: Practical issues and 'how to' approach*. Retrieved from <http://surveyinsights.org/?p=5353>.
- Meinck, S., & Cortes, D. (2015). Sampling weights, nonresponse adjustments and participation rates. In: Fraillon, J., Schulz, W., Friedman, T., Ainley, J., & Gebhardt, E. (Eds.), *International computer and information literacy study 2013 technical report*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).
- Mohadjer, L., Krenzke, T., & van de Kerckhove, W. (2013). Indicators of the quality of the sample data. In: Kirsch I., & Thorn, W. (Eds.), *Technical report of the survey of adult skills (PIAAC)*. Paris, France: Organization for Economic Co-operation and Development (OECD).
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012a). *TIMSS 2011 international results in mathematics*. Boston: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012b). *PIRLS 2011 international results in reading*. Boston: TIMSS & PIRLS International Study Center, Boston College.
- OECD (2013). *OECD skills outlook 2013: First results from the survey of adult skills*. New York: OECD Publishing. doi:10.1787/9789264204256-en
- OECD (2014). *TALIS 2013 results: An international perspective on teaching and learning*. New York: OECD Publishing.
- Schulz, W., Ainley, J., & Fraillon, J. (Eds.). (2011). *ICCS 2009 technical report*. Amsterdam: The International Association for the Evaluation of Educational Achievement.
- Stoop, I. A. L. (2004). Surveying nonrespondents. *Field Methods*, 16(1), 23–54.
- Watson, N. (2012). *Longitudinal and cross-sectional weighting methodology for the HILDA survey*. HILDA Project Technical Paper Series, 2/12.