# School Size and Student Achievement in TIMSS 2003

Draft paper prepared for the 2nd IEA International Research Conference

November 9-11,

2006-10-20 Washington, DC

Gabriela Schütz

Ifo Institute for Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
Phone: (+49) 89-9224 1698
E-mail: schuetz@ifo.de

November 5th, 2006

Preliminary and incomplete.
Do not quote nor circulate without author's explicit permission.
Comments welcome.

## 1. Introduction

A large strand of literature within different social sciences of education deals with the effects of school size on student achievement. Much of the existing research seems to indicate that up to a point, small school size is associated with higher student achievement, with optimal school sizes of some 600-900 students. However, evidence is mostly confined to the United States. Using TIMSS data, this paper will look at the association between school size and student achievement within different countries to identify whether internationally there is such a thing as an optimal school size. Due to data limitations, however, the aim of this paper cannot be to identify causal effects of school size on student performance, but rather to draw a first picture of the kind of relationship found between school size and student performance in the different countries. Although TIMSS 2003 targeted and yielded representative samples of students both at the eighth grade and at the fourth grade, we will restrict our analysis to eighth grade students for the purpose of this paper.

A second important issue with respect to school size that the paper will look at is the question: If there is a benefit attached to a certain school size, who profits most? Here, we specifically focus on groups of students that differ with respect to socioeconomic status and with respect to immigrant status.

## 2. Relation to the Literature

## 3. Data

To analyze at the association between school size and student achievement within different countries, we employ student-level micro data from the Trends in International Mathematics and Science Study (TIMSS) 2003. TIMSS 2003 is an extensive international student achievement tests conducted by the International Association for the Evaluation of Educational Achievement (IEA), an independent cooperation of national research institutes and governmental research agencies. All participating countries received the same test items, so that the ensuing measures of the educational performance of students in math and science are directly comparable across countries.[1] Furthermore, by using representative sampling methods to draw random samples of schools, the test provides representative samples of students in each participating country.

The target populations of TIMSS 2003 were all students at the end of their eighth and fourth years of formal schooling in the participating countries. TIMSS 2003 was constructed as a curriculum-valid test, which ensures greater validity of international comparisons of student achievement. To ensure comparability with the former TIMSS achievement tests, the formal definition for the grade eight student populations was the upper of the two adjacent grades with the largest share of 13-year-olds; this was usually the eighth grade in most countries.[2] TIMSS 2003 yielded internationally comparable data for representative samples of students in 47 countries and four benchmarking participants at the eighth grade and in 25 countries and three benchmarking participants at the fourth grade.[3]

TIMSS applied IRT scaling and calculated plausible values for the proficiency in mathematics and science for each participating student. The scores were placed on the metric used to report the results from previous assessments. At the eighth grade, the metric was established by setting the average of the mean scores of the countries that participated also in TIMSS 1995 at the eighth grade to 500 and the standard deviation to 100. Fourth grade dta were placed on the fourth grade scale of TIMSS 1995.[4] In this paper, we use the mean of the students' plausible values for mathematics as our measure of educational performance.

TIMSS provides not only achievement data for representative samples of students in the participating countries but also a rich array of background information for each

---

1 The development of the test contents was a cooperative process involving national research coordinators from all participating countries, and the curriculum framework was endorsed by all participating countries. The study also performed a test-curriculum matching analysis that restricted the analysis to items definitely covered in each country's curriculum, which made little difference for the overall achievement patterns. For more detailed information, see Mullis et al. (2004).

2 See Foy and Joncas (2004).

3 The benchmarking participants in the eighth grade include the Basque Country in Spain, the US State of Indiana in the United States and the Ontario and Quebec Provinces in Canada. In the fourth grade the benchmarking participants were again Indiana State in the United States and the Ontario and Quebec Province in Canada. With the exception of Yemen, all countries that participated in the fourth grade participated in the eighth grade as well. England and Scotland count as individual countries here as they have separate school systems which participated separately in the tests. In Belgium only the Flemish school system participated in TIMSS 2003.

4 See Gonzalez et al. (2004).

student, his mathematic and science teachers and his school. In separate background questionnaires, students were asked to provide information on various features of their family background. These features include the number of books in their home, their parents' education, whether they themselves, their mother and their father were born in the country, their gender and age. Background questionnaires were also distributed to and answered by the students' teachers and the head masters of the sampled school. Thus, we have also information on the students' learning environment and the general background of their schools. Data from these questionnaires where spliced together using identifiers for the students, teachers and schools. Since in this paper we use the students' test performance in mathematics, we included in our dataset only information on the mathematics teachers. In the cases where more than one teacher could be linked to a student, we chose information only from that teacher that taught the most hours in the TIMSS class. If this did not uniquely identify a teacher (e.g. due to missing information), we chose the teacher who had the most teaching experience. The reason behind this is that if two teachers teach the same student, it might most likely be because one of the teachers is still in a training period and teaches only occasionally. In the cases where there was still more than one teacher to a student we choose one teacher arbitrarily to be included in the dataset.

Our dataset for the eighth grade contains data from 51 countries or regions. On average, the sample contains 141 schools in each country. Although schools were chosen at random, the design of the school sample in TIMSS was such that it provides an optimal sample of students – not of schools (Martin 2003). The sampling probability of a school was proportional to their sizes, which lead to larger schools being more frequently sampled than if a representative school sample had been drawn. Even though we analyze school size only within the complete student sample, where we correct for the sampling probability of each student, we should bear in mind that the school sample is rather small and not representative for the school population. In some countries or regions the number of schools sampled in the eighth grade is especially small. In Bahrain only 67 schools were sampled, in Cyprus 58, in England 57, in Morocco 88, in Scotland 88 and in the US State of Indiana only 50.[5] We think that making inferences from such small school samples is even more problematic. Results for those countries will be also presented but the size of the school sample should be born in mind.

[5] Without these countries, the average number of schools sampled in each country is 151.

# 4. Framework of the Analysis

### *4.1 Including School Size in Educational Production Functions*

The easiest way to include size in an educational production function is of course by including it as a continuous variable and estimating a linear relationship. However, even at first sight it is not clear why enrolment size – be it school or grade enrolment – should impact linearly on student achievement. Increasing enrolment from 100 to 200 has probably a different impact on the social interactions within a school than increasing enrolment from 1000 to 1100. Therefore, experimenting with non-linear specifications with respect to school size is more advisable. Moreover, if we indeed suspect that an optimal school size range exists, we cannot approach this question by including a linear term. Instead, to account for the suspected non-linear relationship between size and achievement we entered size in the regressions together with its quadratic term. The results for the complete sample indicated that the relationship between enrolment and test scores followes in fact an inversely u-shaped relationship which suggests the existence of an optimal size range.

Another possibility to account for non-linearity in the relationship between size and performance is to construct different size categories and to include dummies for these categories in the regressions. Most of the literature in the field of school size (e.g. Lee and Smith 1997) has proceeded that way and it is a very convenient tool since the restrictions imposed on the model are fewer than with a linear or a quadratic specification. Using size categories relationships can be much more flexibly estimated. With our dataset, however, this approach was not possible, since in most cases, the coefficients for the size dummies were not statistically significantly estimated. This problem was probably due to the small sample of schools.

*Identification of Causal Effects and Unobserved Heterogeneity*

It is not clear that the regression coefficients on size in an educational production function should reflect a true causal effect of size on student achievement. The problem encountered here as in any kind of similar analysis is whether it can be assumed that children are randomly assigned to schools of different sizes.

If parents and students have the possibility to choose between different schools, they will most certainly do so. In most countries in our international dataset, parents are free in chosing the school for their child. The practice that residence determines which school the child will attend is an US phenomenon that is not so much the case in most other countries. Also that housing prices reflect the quality of schooling is a phenomenon almost entirely restricted to the US. But even where choice is restricted, parents will in most cases find ways to excert the best possible choice for their child. This is even truer for parents with high socio-economic status than for parents with socio-economic status.

Thus, selection of students into schools is an issue in our international dataset. However, little is known on how parents and children choose their schools. One could argue that size is not a parameter in making the choice and that parents might choose the smaller school only between schools of equal quality but even this is not clear.

Even if we do not know how school size enters into the choice function we can see the problems that arise from the selection of students into schools. For example, if schools that are smaller are for some reason also held to be better in certin aspects of quality, all parents would try to send their children to the smaller (and suspectedly better) schools. In countries where choice is restricted and also where capacity is a problem, well-off families are probably most successful in sending their children to the preferred schools. Since children from these families at the same time also perform better in student achievement tests, the results will be biased towards finding that smaller schools are better.

One way of evading this problem of non-random selection, is to restrict the sample only to such schools and children that did not have a possibility to choose (either the school the students or the students the school).[6] Unfortunately, TIMSS 2003 does not provide such an opportunity. Another possibility would be at least to account for private or public schools or to exclude private schools from the sample. The problem here is that TIMSS provides this information only for a few countries.

### *4.2 Sampling Design*

TIMSS sampled students using a two-stage stratified cluster design within each country. At the first stage, schools were sampled, and at the second stage a sample of students – mostly from one class - from the target grade in the sampled schools was drawn (Foy and Joncas 2004). Thus, the primary sampling unit (PSU) in TIMSS was the school. The

[6] However, PISA provides this information in its dataset. Students were asked how they came to visit exactly the school they are in. Similarly, school principals were asked how students are selected into this school. This information of students and principals could give the opportunity to evaluate the effects of school size where selection does not play a role.

performance of students within the same school, however, may not be independent from one another (cf. Moulton 1986 for this problem of hierarchical data structure). This suggests that the independence assumption usually made with respect to individual observations in standard econometric methods should be relaxed in favor of the assumption that only the variation between schools (PSUs) provides independent variation. This is implemented by the clustering-robust linear regression (CRLR) method, which allows any given amount of correlation of the error terms within PSUs and requires only that observations be independent across PSUs (cf. White 1984; Deaton 1997). Since TIMSS used a stratified sampling design within each country, sampling probabilities vary for different students (Martin and Kelly 1996; Martin et al. 2000). We obtain nationally representative coefficient estimates by employing weighted least squares (WLS) regressions, using the sampling probabilities as weights. WLS estimation guarantees that the proportional contribution to the parameter estimates of each stratum in the sample is the same as if a complete census had been obtained (cf. DuMouchel and Duncan 1983; Wooldridge 2001).

### *4.3 Structure of the Analysis*

Our analysis will proceed in three steps. In the first step we aim at obtaining a first picture of the relationship between school size and student performance in the data. We will therefore in a first regression in each country only enter school size as explanatory variable. Because of the suspected non-linearity of the relationship we will depict school size both as a linear and a quadratic term in these regressions. In the second step, we will add control variables to capture the effects of school, teacher and student background variables and calculate the size effects net of these other influencing factors. In the third step we will analyze whether the effects of school size differ over different groups of student populations.

All regressions take into account the stratified sampling design by employing student weights and by using clustering-robust linear regression.

## 5. Analysis

### *5.1 First Picture*

Research on the effects of school size has often found a kind of optimal size range where performance was highest. In contrast, schools that enroll fewer students perform worse - as do schools with higher enrolment.[7] This kind of association would imply an inversely u-shaped relationship between school size and student performance.

We approached the question of whether there exists an optimal school size by including size both as a linear and as a quadratic term in the regressions. First, to obtain an idea of the general pattern in the data, student performance was regressed solely on school size and it's squared. Table 1 displays the results of this regression for each participating country as well as for the complete sample and the sub-sample of OECD countries. The table contains the number of student observations available for the regression, the number of primary sampling units (i.e. schools), and the estimated coefficients and standard errors for the linear and the quadratic term of school size. It also depicts the pattern of the relationship between test scores and enrolment as indicated by the estimated coefficients, i.e. whether the relationship is u-shaped, inversely u-shaped, or rather linear, and the "optimal" or "non-optimal" school size calculated from the coefficients.[8]

Within the complete sample and the sub-sample of OECD countries, both the coefficients for the linear and for the quadratic term of school size are statistically significant. The signs of the coefficients indicate an inversely u-shaped relationship between school size and student performance within both regressions. Applying simple algebra, the estimates suggest an optimal school size of 3373 students in the complete sample and of 1624 in the sub-sample of the OECD countries.

However, when looking at the results of the regressions for each of the different countries contained in our dataset, the finding of an inversely u-shaped relationship is not easily verified for most countries. In 27 out of the 51 participating countries and regions, the estimated coefficients for the linear and the quadratic term of school size are not statistically significant. For those countries, it is probably most reasonably assumed

---

[7] See Andrews et al. (2002) for a review.

[8] With the estimated coefficients obtained for the linear and the quadratic term, the school size where performance would be highest or lowest according to the assumed relationship can be calculated. This is done by calculating the level of enrollment for which the slope of the function between school size and student performance is zero. Depending on the pattern of the effects this enrollment level is either a maximum or a minimum.

that there is no relationship – neither linear nor quadratic - between student performance and the size of the schools.

In 9 of the countries, only the coefficient on linear term is statistically significantly estimated.[9] In these countries, overall, the relationship is positive, meaning that on average, students in bigger schools perform better. The only exception is the US State of Indiana, where the relationship seems to be negative.

However, for the remaining 15 participating countries and regions, both the coefficient for the linear and for the quadratic term of school size are statistically significant, indicating the existence of a non-linear relationship between school size and student performance.[10] In most of these countries (i.e. in 13 cases), the signs of the estimated coefficients indicate an inversely u-shaped relationship between size and test score. This supports the hypothesis that there might be an optimal size range within which student performance is best, while performance is worse in schools that are smaller and in schools that are larger than this optimal range. Yet in two countries, Bahrain and Singapore, the signs of the estimated coefficients of school size and it's squared indicate a u-shaped relationship between school size and student performance. This suggests a non-optimal size rather than an optimal size, meaning that at first test performance decreases while size increases until the trend is reversed and the relationship becomes positive.

For the countries where the estimated coefficients were statistically significant, the implied "optimal" or "non-optimal" size can be calculated from the coefficients. Comparing the calculated optima for the single countries in Table 1, we see that the "optimal" sizes vary widely between the countries in our sample. The optimal size is calculated for Cyprus at 490 students, while in Chile an enrolment of 4010 students seems to be optimal.

Thus, from a first view at the results, we can already conclude that even if a relationship exists between student performance and school size, it is not the same across all countries. In some countries the relationship seems to be positive over the whole range of school size observed, while in other countries it seems to be rather negative. We also find incidence of quadratic relationships, both of inversely u-shaped

---

[9] These countries are: Ghana, Korea, Latvia, Lithuania, the Netherlands, the Russian Federation, Slovak Republic, Syrian Arab Republic, and the US State of Indiana.

[10] These countries are: Australia, Bahrain, Belgium (Flemish), Chile, Chinese Taipei, Cyprus, Hungary, Indonesia, Iran, Lebanon, Romania, Serbia, Singapore, South Africa, and the Basque Country in Spain.

and u-shaped relationships. Even if we compare only those countries where an inversely u-shaped relationship between size and performance is found, we see that the predicted "optimal" enrolment levels vary considerably across countries.

This is not really surprising though, since the school size range observed in the data does also vary widely between the countries. Table 2 depicts again the number of observations and primary sampling units from the first regression as well as the smallest and highest school size observed in the respective country. The difference in enrolment between the largest and the smallest school is comparably low in Cyprus and in Botswana, where these differences amount to only 649 and 797 students. On the other hand, in countries like the Philippines and Indonesia, the difference in enrolment between the largest and the smallest school is rather high, with differences of 9661 and 6358 students.

Table 2 takes up again the calculated "optimal" or "non-optimal" size from the first regression. Comparing these "optima" with the observed smallest and highest enrolment levels in the respective countries, gives the impression that the calculated "optimal" or "non-optimal" school sizes are located roughly at the middle of the school size distribution in the respective country. This is important insofar as it implies that the results of our estimations are reasonably well nested within the data which gives confidence in the findings.

However, it is necessary to take a closer look at the distribution of school size within the student populations of the respective countries to judge whether the estimated relationship between size and performance is really plausible. A quadratic relationship might be fitted to the data while the result is in fact driven by just a few schools at either end of the size distribution. We therefore looked at the percentage of students within the respective countries that actually attended schools with enrolment levels higher than the suggested "optimal" or "non-optimal" size. Thus, we obtained an idea of how much support for the quadratic specification there is in the data.

In roughly half of the countries where a quadratic relationship could be significantly estimated, as well as in the complete sample and the sub sample of OECD countries, the percentage of students attending schools larger than the estimated optimum is below ten.[11] In Belgium, Chinese Taipei, Hungary and South Africa, the percentage of students enrolled at larger schools lies between 10 and 20. Only in Bahrain (43.94), Cyprus

[11] These countries are: Australia, Chile, Indonesia, Iran, Lebanon, Romania and the Basque Country in Spain.

(64.47), Serbia (24.99) and Singapore (90.08), more than 20 percent of the students in our sample attend schools with an enrolment level that lies beyond the calculated “optimal” or “non-optimal” size.

The last column in Table 2 gives the percentage of students attending schools that enroll between 100 students less and 100 students more than the suggested “optimum”. These numbers illustrate how dense the school size distribution is around the “optimal” or “non-optimal” size. We see that in most countries only a minor part of the student population attends schools with a size that falls within this range. In Chile, Indonesia, Iran, and Lebanon, the percentage is even zero. But these are also countries were the differences in enrolment between the largest and the smallest school is rather high and where school size therefore varies widely.

Altogether the first look at the data suggests that only in less than one third of the countries, a quadratic relationship between student performance and school size exists. The results from Table 2 would lead one to assume that in some countries, even if the coefficients of the linear and the quadratic term of school size are significantly estimated, the quadratic relationship might still not be appropriate because only very few observations can be found at schools that are larger than the suggested optimum.

### *5.2 Regressions with control variables*

The first regression described above was conducted as a kind of descriptive analysis to obtain an idea of whether there is a relationship between test performance and school size and of how the relationship might look like. However, the method employed was little more than a bivariate analysis in that we included only one explanatory variable, its squared and a constant. The general problem encountered with bivariate analysis is that they do not consider other factors that might have an influence on the dependent variable (here: test score) while being correlated with the explanatory variable of interest (school size). Not taking such factors into account will lead to biased estimates of the relationship between size and performance. If for example bigger schools reveal better student performance, it might be because of better educated teachers or better resources in these schools. Also, schools of a certain size might attract a different student population which is then reflected in differences in test scores. It is possible for example, that large schools in cities attract students with better educated parents than small schools in rural areas. We then might find that larger schools are better while in fact they are only better because their students come from well educated families. Also, if larger schools are better because they are better equipped, our results will be biased in favor of finding that bigger is better if we do not control for resources at the same time.

Therefore, to reduce the bias in the relationship between school size and student performance we will in the following subsection include school size in an education production function and look at the relationship between size and test performance net of other potentially influencing factors.

### *5.2.1 Control variables*

Since TIMSS collected background information about the students, their families, teachers and schools, we can control for influencing factors at these different levels. We controlled for student characteristics and family background by including student age, a dummy for student gender, three dummies for whether the father, the mother and the student was born in the country of test, a dummy for whether the language of test was sometimes or never spoken at home, the number of books at home in five categories and a set of dummies for the highest level of education attained by the parents[12]. At the teacher and school level, the regression controlled for teacher experience and its squared and for a set of dummies describing the teacher's level of education. Since our dependent variable is the student's mean of the plausible values in mathematics, the teacher variables refer to the students' teachers for this subject. Further, the regression included the size of the TIMSS class and a dummy for whether more than fifty percent of the students enrolled at this specific school come from a disadvantaged background. It also included a set of dummies indicating whether the school principal reported that instruction in the school was affected by a lot or some shortage of several kinds of resources: instructional material, budget for supplies, school buildings and grounds, or teachers. Moreover, we included two dummies for the size of the community where the school was located. One dummy for whether more than 100,000 people live in the community and one for less than 15,000 people.

Since missing values had been imputed were possible, the regressions also include imputation dummies to control for imputation effects.[13]

---

[12] We only considered the highest education level of either mother or father

[13] We estimated the same regression also without using imputed values. In most cases the results were not much different than those reported here except for the standard errors which were larger in the sample without imputed values.

Descriptive statistics of the variables describing the students, families, teachers and schools in our dataset can be found in the appendix.

*5.2.2 Results*

Controlling for family background, and teacher and school effects, the pattern of countries for which a relationship between school size and student performance is found changes slightly. Table 3 contrasts the results obtained in the first descriptive regression with the results of the regression that includes control variables. Of the 15 countries where significant coefficients were estimated for both the linear and the quadratic term of school size in the first regression, only in eight the same coefficients remain significant after including control variables. This is the case for Bahrain, Belgium, Chile, Indonesia, Lebanon, Serbia, Singapore, and South Africa as well as for the results obtained from the complete sample regression. In Bahrain, however, the pattern of the relationship between size and performance changes from a u-shaped to an inversely u-shaped relationship. But since only 67 schools were sampled in Bahrain, and the school sample was further reduced to 50 by including control variables, results obtained for this country would not seem very reliable in any specification. In most countries the optimal school size calculated from both regressions does not differ very much. The differences are however rather pronounced in Chile, Indonesia and South Africa. In Chile, the newly calculated optimal school size lays approximately 200 students below the optima calculated from the first regression, and in Indonesia the optimal size is increased by 400. But since the differences in size between the smallest and the largest school are very high in those countries, those changes do not seem very dramatic. In South Africa, the new optimum is at a level that is increased by approximately 300 students. Since the sample did not change at all due to the inclusion of control variables, we can attribute this change entirely to the enhanced specification.

With the new specification, statistically significant coefficients for both the linear and the quadratic term of school size were now estimated also in Ghana, Hong Kong and Norway. In Syria and the United States only the coefficient on the quadratic term of school size could be significantly estimated. Repeating the same regression while including only the quadratic term, produced no significant estimate for the coefficient in Syria and we will therefore not discuss the results for this country further below. In the United States, the coefficient on the quadratic term was indeed also positive and significantly estimated in the regression that excluded the linear term. Because including a positive quadratic term of school size in the education production function without a linear term implies that student achievement increases as schools get larger, there seems to be no evidence for an optimal school size in the United States. In the Basque Country in Spain, the coefficient on the quadratic term is no longer statistically significantly estimated, and only the coefficient on the linear term remains significant once the control variables are included in the regression. But at the same time, also the number of schools within the sample decreases from 111 to 87. We can therefore not be certain whether the observed change in the pattern of the relationship between school size and student performance does not rather stem from a change in the pool of available student observations.

Unfortunately, including additional control variables also reduced the regression sample in most of the other countries. In order to check whether the difference in results between the first descriptive regression and the regression which includes controls could have been driven by differences in the samples available, we repeated the first regression for the same observations that were used in the regression with control variables. In most countries the results were not sensitive to those changes in the sample.

We are therefore confident that including control variables produces more reliable results even if the sample size is smaller and that at least some of the biases in the estimated relationship between school size and student achievement is reduced.

Table 4 depicts the estimated coefficients and standard errors for the linear and the quadratic term of school size as it results from the more detailed model specification discussed above. In all but one of the countries, where the coefficients of linear and the quadratic term of school size were statistically significantly estimated, the signs of the coefficients indicate an inversely u-shaped relationship between school size and student performance. The exception is again Singapore, where a u-shaped relationship is suggested implying that small schools and large schools perform best.

Table 5 gives again more detailed information on the distribution of schools within the regression sample. We repeated from Table 4 the number of observations and schools (primary sampling units) contained in the regression and the “optimal” or “non-optimal” school size calculated from the estimated coefficients. To obtain an idea approximately where within the respective school size distribution this “optimum” lies, we included information on the smallest and largest school in the regression sample, the percentage of students enrolled at schools that are larger than the calculated “optimum”, and the percentage of students enrolled at schools where enrolment is within plus or minus 100 students difference from the calculated optimum.

The calculated "optimal" school size lies at the far right end of the students' population in Chile, Ghana, Hong Kong, Indonesia, Lebanon and South Africa. In these countries less than 5 percent of the student population is enrolled at schools that are larger than the optimal size. There, also the percentage of students attending schools that are larger than the optimal size minus hundred and smaller than the optimal size plus hundred, is rather small.

In the other countries, the percentage of students' enrolled at schools larger than the "optimal" or "non-optimal" size in above 10. In Bahrain 65.15 percent of the students are enrolled in larger schools, in Belgium 10.28 percent, in Norway 14.92, in Serbia 28.09 and in Singapore 89.47 percent. Only in Norway and in Serbia, is the school size distribution rather dense around the calculated optimum. In Norway, 29.40 percent of the students attend schools which are larger than the optimal size minus hundred and at the same time also smaller than the optimal size plus hundred, and in Serbia 14.59 percent of the students.

### *5.3 Graphical Analysis*

Using control variables, we were able to calculate the relationship between student achievement and school size net of other influencing factors. Even after controlling for other determinants of student achievement, a quadratic relationship seems to exist in eleven of the 51 countries and regions that participated in TIMSS 2003.

However, we also noted that in some of the countries the support in the data for an inversely u-shaped relationship between school size and test performance might be rather week in so far as only a minor part of the students were unrolled in schools that are larger than the optimum calculated from the regression results.

We will therefore now look in more detail at the estimated model and the distribution of students across schools of different sizes for those countries where a significant quadratic relationship between size and performance was found. From what we observed in Table 5 we would not expect to find proof for the existence of an optimal school size in Chile, Ghana, Hong Kong, Indonesia, Lebanon, South Africa and the United States. In Bahrain, Belgium, Norway, Serbia and Singapore, however, there seems to be more support in the student observation for a quadratic relationship.

For each country, we plotted the fitted test scores from the previous model that included control variables against the size of the schools that the students attended. In the same diagram we also drew the quadratic function between school size and student performance estimated by the model. Thereby we were able to see a bit clearer how well the model fitted the actual data and also to discern possible outliers that might be driving the results. The graphs discussed in the following can be found in the Appendix A.1.

*Complete Sample*

Plotting the fitted test scores against school size together with the predicted relationship between school size and student performance for the complete sample revealed nothing peculiar. The student observations are more concentrated at the left side of the school size distribution, but overall, there are still several school and student observations to the right of the optimal school size. At the same time, however, we also note that after an enrolment level of approximately 6000 students the school size distribution becomes much less dense. Schools that enroll more than 6000 students can be found exclusively in Chile, Indonesia and the Philippines and schools with more than 7000 students only in the Philippines. From the estimated coefficients we calculated that on average, students enrolled at the smallest schools (enrolment of 21 students) achieve 20 test scores less than students in schools that enroll 500 students more. In line with the quadratic relationship estimated, the function becomes flatter as school size increases. Students in schools enrolling 1000 students perform on average by 17 points better than comparable students enrolled at schools of size 500. We can compare this difference in test scores to the average difference in math achievement between boys and girl which is approximately 15 test score points.

*Bahrain*

In Bahrain, the quadratic specification seems to fit the data reasonably well. The optimal enrolment level is estimated at 623 students and lies well in the middle of the school size distribution. No outliers are discernible that might be driving the results. However, the differences in test scores that might be related to differences in school size are rather small overall. The smallest school observed in Bahrain enrolls 170 students. Students in this school perform on average 12 test score points below students enrolled at schools with an optimal size. From the optimal school size to the highest school size observed in the regression sample (1034 students), average test scores fall again by 12 test score points. Differences in students' performance between students from different family backgrounds are usually much higher. Therefore and also because the regression sample consists only of 50 schools, we would rather conclude that no relationship exists between school size and student performance in this country.

*Belgium*

In Belgium, the diagram reveals a big gap in the school size distribution. The two largest schools enroll 2691 and 2786 students while the third largest school with 1596 enrolled students is substantially smaller. It appears that the fitted quadratic relationship between school size and performance might be driven by those two largest schools. If we perform the same underlying regression analysis without including the two largest schools into our sample, the coefficients for the quadratic and the linear term of school size are no longer statistically significant. However, since the graph suggested that a linear relationship might exist within the restricted sample, we repeated the same regression for the restricted sample while including only the linear term of school size. Without the two largest schools, the linear term of school size became statistically significant. The estimate suggested that the average difference in test scores between schools that have a difference in size of 100 students amounts to 2.3 test score points. Over the range of schools in the restricted sample, where the smallest school enrolls 39 and the largest school enrols1596 students, this implies a difference of 37 test score points. As a comparison, girls in Belgium perform on average 15 test score points lower than boys, and the difference between children whose parents went to university and children whose parents completed only lower secondary education is 21 test score points on average. Altogether, there seems to be little support for an inversely u-shaped relationship between school size and student performance in Belgium. The relationship in Belgium seems to be better described by "bigger is better".

*Chile*

Looking at the diagram for Chile, it seems that the finding of a quadratic relationship between size and performance is again driven by an outlier. But in fact, this outlier consists of three schools that all enroll 6410 students each. Between those very large school and the rest of the school size distribution spans a wide gap: the next largest school in the sample enrolls only 2800 students. Repeating the regression without these three largest schools, we find no more evidence for an inversely u-shaped relationship since only the linear coefficient is now significantly estimated. The second diagram for Chile therefore depicts the estimated relationship between size and performance from a regression that includes only the linear term of school size and is restricted to schools that enroll less than 3000 students. The estimated coefficient for school size predicts that increasing school size by 100 is accompanied by an increase in test scores of 1.7. The difference in test scores between students enrolled at schools that differ with respect to size by 1000 is therefore as high as the average difference in performance between boys and girls, which is 17 test score points in Chile. Since the previously observed quadratic relationship is driven by only three outliers, the conclusion for Chile is that on average, bigger schools perform better than smaller schools.

*Ghana*

The first diagram for Ghana plots the quadratic relationship as it was found in the previously discussed regression. From Table 5 we already know that in Ghana, the optimal school size was estimated at 1430 students and that the largest school size observed in the regression sample enrolls 1500 students. The graph depicts this situation rather nicely. Because the slope is so smooth, we do not even observe that the function in fact reaches its maximum before the end of the school size distribution. Increasing enrolment from 25, the smallest enrolment level observed at Ghana, to 500 translates into an increase in test scores by 37 points on average. The difference in student performance between a school that enrolls 500 students and a school that enrolls 1000 students is only 23 test scores. Thereafter, the slope of the relationship becomes really flat. By increasing enrolment further to the estimated optimal of 1430 students, the test scores rise only by 6 on average.

Because of the gap in the school size distribution between the largest (1500) and the second largest school (enrolment of 993 students), we also estimated the quadratic specification excluding the largest school from our sample. Consequently, the coefficient for the quadratic term was no longer significantly estimated. We then estimated a model including only the linear term of school size for the restricted and then also for the complete sample. The coefficients were significant in both cases but the difference between them was negligible. Since overall, the fit of the quadratic model specification was better, our conclusion for Ghana is that the relationship between school size and student performance is one of decreasing returns to scale. Bigger schools are better on average but the gain in student performance becomes smaller as school size increases over the range of the distribution.

*Hong Kong*

In Hong Kong, the optimal school size was estimated at 1519 students. Again, this is almost an out of sample prediction, since only two schools have an enrolment that is actually higher: those two schools enroll 1521 and 1750 students. From the scatter plot and the function describing the relationship between size and performance as predicted by the model, it does not become clear why a quadratic specification should in fact provide the best fit for the data. Moreover, if only the linear term of school size is included in the model, the fit of the model is not significantly reduced. Our analysis for Hong Kong therefore suggests that the hypothesis of the existence of an optimal school size should be discharged. Student performance increases over the entire observed school size range even if it is not clear whether this increase should be best described by a linear relationship or a quadratic relationship with decreasing returns to size.

*Indonesia*

In Indonesia, the quadratic specification using control variables revealed evidence for an inversely u-shaped relationship between size and performance. The optimal school size was estimate to be at 3535 students. However, looking at the diagram depicting the function between school size and performance together with the test scores predicted from the model, we see that this result is probably driven by three outliers. These three schools enroll 2735, 4660 and 6549 students. In comparson, the fourth largest school enrolls only 1695 students. Repeating the same regression as before without the three largest schools also yields significant effects for both the linear and the quadratic term of school size. The optimal school size is now estimated at 1047 students. Since no more outliers are discernible that might be driving the results and 17.41 percent of the student population is enrolled at schools that have a higher enrolment level, we are convinced that the quadratic specification fits the data reasonably well. Although we would not take the calculated optimal size literally, we would conclude that test scores are on average higher in bigger schools and that this trend is reversed once a school enrolls significantly more than 1000 students.

*Lebanon*

Also for Lebanon, the regression using a quadratic and a linear term of school size suggested an inversely u-shaped relationship between school size and student performance. The optimal enrollment level was estimated at 3005 students while only 1.39 percent of the student population attends schools larger than this estimated optimum. Looking at the diagram that combines the predicted test scores from the model with the estimated relationship between size and performance, we see that only two schools enroll more students than the suggested optimal school. These two schools enroll 3850 and 4916 students. Compared to the third largest school enrolling 2898 students, the two largest schools appear to be outliers rather than typical schools. Repeating the same regression as before excluding the two largest schools, still finds evidence for a quadratic relationship. This time, however, the estimated optimal size lays at 2062. Test scores are on average 26 points higher in schools that enroll 542 students than in schools that enroll only 43 students - the smallest school size observed in Lebanon. Increasing school size by additional 500 students is on average accompanied by an increase in test scores by 19 test scores. We are still not very much convinced by this quadratic relationship because also the newly estimated optimum is situated quite far to the right side of the school size distribution in Lebanon. Altogether, test scores are generally higher in larger schools while the relationship becomes weaker as schools become larger.

*Norway*

In Norway, the previously estimated optimal school size lies at 466 students while 14.92 of the student population are enrolled in schools larger than this optimum. However, looking at the diagram leads one to conclude that really no relationship between school size and student performance exists. The function is relatively flat and it is not clear whether the estimated quadratic relationship is in fact not rather driven by the two comparably large schools enrolling 820 and 999 students. In contrast, the third largest school enrolls 600 students. Repeating the same regression for the sub sample of schools that are smaller that 601 students, the coefficients for the linear and the quadratic term of school size are not statistically significantly estimated. Also iincluding only the linear term in the model, it is no longer significantly estimated. We therefore reject the idea of an optimal school size in Norway. Student performance does not seem to be correlated at all with school size.

*Serbia*

In Serbia, the optimal school size was estimated at 1107 students in the regression that used control variables. The diagram again depicts the predicted test score from this model and the estimated function between school size and student performance. Increasing school size from 123 (smallest school in Serbia) to 623, translates into an average rise in test scores by 19. The difference in test scores between a school that enrolls the estimated optimal number of students and a school that enrolls 500 students less is approximately seven test scores. The diagram also shows one outlier: the largest school enrolls 2282 students, the second largest only 1825 students. To judge whether this outlier might drive our findings, we repeated the regression while excluding the largest school from our sample. In the new diagram, the function between size and performance seems only a little bit flatter than before and the estimated optimum now lies at an enrollment level of 1167 students. Plotting the functions estimated from the complete and the sub sample within the same diagram indeed shows that both functions are very similar. The difference in the optimal sizes estimated from the two regression samples is at 60 really very small. Therefore, we conclude that for Serbia the quadratic specification fits the data very well. No outlier seems to be driving the results. Instead we indeed observe an optimal school size which lies between 1100 and 1200 students.

*Singapore*

Singapore is the only country where in the enhanced specification a u-shaped relationship was found between school size and performance. The relationship seems to be negative for small schools and becomes positive for schools that enroll more than 862 students. The graphical evidence supports this finding. In the diagram that displays the fitted test scores together with the estimated function between size and performance, no outliers can be identified. Average test scores are decreasing as the school size increases from the smallest school (197 students) until a school size of approximately 862 students and are increasing in school size thereafter. The difference in test scores associated with increasing enrolment from 197 to 697 students is 55 test scores. As enrolment increases from the non-optimal school size of 862 by 500 students, the average test scores of the students increase by 33 points and remain increasing over the rest of the school size range. The differences in student performance associated with differences in school size are rather large considering that the average difference in student performance is 40 test scores between children of parents whose highest education is primary education or below and children of parents who received university education. Overall; schools that enroll more than 1500 students perform best.

*South Africa*

In South Africa the quadratic specification using control variables suggested an optimal enrolment level of 1519 students. However, we learned from Table 5, that only 4.73 percent of the student population is enrolled at schools larger than this estimated optimum. So the question is whether this estimated quadratic relationship is supported by the graphical analysis. The average difference in test scores between students enrolled in the smallest school and students enrolled in schools that have 500 students more is 21. By increasing enrollment from 500 students to 1000 students the average test scores still increase by 16. But we also see from the diagram that there is only one school that enrolls more than 2000 students (2017 students) and that the second largest school enrolls only 1647 students. We therefore repeated the same regression for the sample of schools where enrolment levels were below 2000. The estimated relationship between size and performance does not change much due to this change in the sample. The optimal enrolment level is now smaller and is estimated at 1286 students. All in all, there is some support for the inversely u-shaped relationship between size and performance in South Africa but the function is rather flat and the peak not very pronounced.

*United States*

In the United States only the quadratic term of school size was significantly estimated in the regression that included control variables. For the diagram we therefore repeated the regression analysis excluding the linear term from the equation. Since the estimated coefficient is positive this implies that student performance is increasing over the entire range of school size.[14] Moving from the smallest to the largest school observed in the sample (from 38 students to 2489 students) leads to an increase in test scores by 57 on average. Excluding the largest school from the sample since the second largest school enrolled only 1752 students, the linear coefficient was again not significantly estimated

[14] We also tried one specification that included only the linear term of school size. However, in this setting the coefficient on the linear term was not estimated significantly neither.

and therefore excluded from the regression.[15] Without the largest school in the sample the slope for the estimated relationship between size and performance becomes steeper. Over the entire range of school size observed in this regression sample (38 - 1752), test scores rise on average by 97. As a comparison, with the previously estimated coefficient for the complete sample, the rise would only be 28 over the same range. All in all, even if we cannot settle on the exact relationship, there seems to be evidence that in the United States bigger schools perform better.

In this subsection we looked further into the relationship between school size and student performance by plotting the fitted test scores into a diagram together with the predicted function between size and performance. From what we had observed previously in Tables 4 and 5 we had suspected to find evidence for a quadratic relationship in Bahrain, Belgium, Norway, Serbia and Singapore but not in the other countries. Indeed, the finding of a quadratic relationship was supported only in Bahrain, Serbia and Singapore. In most of the other countries we found evidence for the fact that bigger schools performed better, albeit sometimes with strongly decreasing returns to size. Only for Norway we concluded that school size and student performance are entirely uncorrelated.

### *5.4 Are size effects different for different groups of students?*

Providing equal opportunities for all students irrespective of their socio-economic and socio-cultural background is seen as one major goal of education systems by most people. First of all, simple considerations of justice of distribution can lead one to argue in favor of equality of educational opportunities. But equally often it is argued that it is a waste of human resources if people are not educated to their full potential and that therefore also for the sake of the economic prosperity of the country, education systems should seek to promote equality of opportunity. An additional goal of education systems especially in traditional countries of immigration is to act as an integrating tool for people from different countries of origin.

How well the education system of a country succeeds in these goals, is therefore of vital interest to the society. Differences in student achievement due to differences in their

[15] Again, the linear coefficient was also not significant if it was included on its own without the lquadratic term.

socio-economic or socio-cultural backgrounds are best kept as low as possible without negatively affecting other important goals, such as high average performance for example. Thus, we addressed the question whether school size influenced the degree to which equality of educational opportunities is achieved. We approached this issue by looking at the differences in the effects of being in schools that fall into different size categories for different groups of students.

We looked at equality of educational opportunities along two dimensions: along socio-cultural and along socio- economic backgrounds of students. Along the socio-cultural dimension we considered two characteristics as being informative on the background of the student: whether the student was born in the country of test and whether the language of test was always or almost always spoken at home. The socio-economic background of the students was proxied by the number of books on their home. We looked at each of these characteristics in turn while our control variables remained as before. We constructed dummy categories for each combination between the socio-cultural or socio-economic characteristic of interest and five school size categories. The school size categories where constructed separately for each country and such that each category contained the same number of weighted student observations. Thus, we ended up with 9 (5x2 -1) dummy coded variables to analyze the effects of differences between students from different backgrounds within schools of different size. From the regression results we calculated for each size category the difference in the effect between students with a favorable background characteristic and students from less favorable backgrounds. We also tested whether the differences in the effects between the two groups of students was statistically significant within each of the size categories. Table 6 gives the results. Our discussion will describe only Australia, Belgium and England, but the results for all countries are listed in Table 6.

In Australia, the difference between a student that speaks the language of test always or almost always at home and a student that does not or does only sometimes speak the language of test at home is high (between 36 and 78 test score points) and statistically significant only in the three largest schools size categories. The difference between a student that was born in Australia and a student that was born in a foreign country is statistically significant only in middle sized schools and amounts to 28 test score points. A student that has more than 100 books at home performs significantly better than a student with fewer books in all size categories but category four. The difference is highest (28) in category 1 and lowest (18) in category 2.

In Belgium, the difference between a student that speaks the language of test always or almost always at home and a student that does not or does only sometimes speak the language of test at home is almost 23 test score points on average in schools within size category 2 (330-483 students). In the other school size categories, the difference is much lower and not significant. The difference between a student that was born in Belgium and a student that was born in a foreign country is not significant in any of the size categories. A student that has more than 100 books at home performs on average 8 test score points better in size category 2 than a student with fewer books. The biggest difference (20 test scores) is found for students in size category four schools.

In England, the differences in the performance of students that belong to different socio-economic or socio-cultural backgrounds are almost always very high and statistically significant. However, the differences in theses differences across the different school size categories are not statistically significant

Overall, the differences in the effects of school size between groups of different socio-cultural or socio-economic backgrounds show no clear pattern across countries. What can be said is that the differences in school size effects between students that differ with respect to their family background in terms of number of books at home, is more often statistically significant than the differences found for students that differ with respect to the language spoken at home or to whether they are born in the country of test. At the same time, however, these differences are most often not statistically significant from each other across the school size categories. This implies that with respect to providing equal educational opportunities for children from different socio-economic family backgrounds no such thing as an optimal school size exists in most countries.

## 7. Conclusion

The aim of this paper was to look at the effects of school size on student performance in the countries that participated in TIMSS 2003. Much of the previous research on the effects of school size had been confined to the United States. There, evidence suggested a kind of optimal size range with schools between 600 and 900 students performing best.

Our research was lead by the question whether also in other countries an optimal school size range existed. First, we used a simple regression of test scores on school size and its squared to obtain a first picture of the relationship between those variables in our dataset. We found both estimated coefficients to be statistically significant in 15 of the 51 countries and regions. In most cases (i.e. 13) the relationship between size and performance was inversely u-shaped

Because school size can be hypothesized to be correlated with other factors that also impact on student performance, in a second step, we included several control variables. Our control variables accounted for the family background of the student and for teacher and schools effects. After including control variables, the set of countries for which a significant effect was changed and reduced to eleven.

However, considering the calculated optimal size and its location within the school size distribution in the respective country raised concerns about the appropriateness of the model specification. The optimal size calculated from the estimates was to the far right end of the school size distribution in most countries, which suggested that a quadratic specification did probably not fit the data very well.

Therefore, we proceeded by graphically analyzing the relationship between school size and student performance in those countries where the quadratic relationship was significantly estimated. The fitted test scores from the enhanced regression were plotted into a diagram together with the predicted function between size and performance. The finding of a quadratic relationship was supported only in Bahrain, Serbia and Singapore. In most of the other countries we found evidence for the fact that bigger schools performed better, albeit sometimes with strongly decreasing returns to size. In the United States, contrary to most other findings in the literature, the evidence suggested that bigger schools perform better. Only for Norway we concluded that school size and student performance are entirely uncorrelated.

Overall, it seems as if student performance is uncorrelated with school size in most countries, but if a significant relationship is estimated, it mostly implies that bigger is better.

A further question was wether the relationship between school size and student performance differed between different groups of soico-economic and socio-cultural status. We constructed dummy categories for each combination between the socio-cultural or socio-economic characteristic of interest and five school size categories. From the regression results we calculated for each size category the difference in the effect between students with a favorable background characteristic and students from less favorable backgrounds. Again, no clear pattern that applies to all countries emerges. It seems that mostly, the differences between the different groups are not significantly different from each other across schools of different sizes.

An interesting issue that could not be addressed within this paper is the question why schools of a certain size mihgt perform better than other schools that are smaller or larger. This question should be looked at using information on the extent of autonomy that the school has and information on how students and teachers describe the atmosphere within the school. We have tried to address the problem of selection bias due to non-random selection of students to schools but the dataset is not very rich in this respect.

## References


Andrews, Matthew, William Duncombe and John Yinger (2002). Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus? Economics of Education Review, 21: 245-262.

Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: The Johns Hopkins University Press.

DuMouchel, William H., Greg J. Duncan (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association* 78 (383): 535-543.

Foy, Pierre and Marc Joncas (2004): TIMSS 2003 Sampling Design. In: Michael O. Martin, Ina V.S. Mullis and Steven J. Chrostowski (eds.), TIMSS 2003 Technical Report, TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College, pp.108-123.

Gonzalez, Eugenio J., Joseph Galia and Isaac Li (2004). Scaling Methods and Procedures for the TIMSS 2003 Mathematics and Science Scales. In: Michael O. Martin, Ina V.S. Mullis and Steven J. Chrostowski (eds.), TIMSS 2003 Technical Report, TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College, pp. 252-273.

Lee, Valerie E., Julia B. Smith (1997): School Size: Which Works Best and for Whom? *Educational Evaluation and Policy Analysis* 19 (3), pp. 205-227.

Martin, Michael O. (eds.) (2003): *TIMSS 2003 User Guide for the International Database*. International Association for the Evaluation of Educational Achievement & TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

Martin, Michael O., Dana L. Kelly (eds.) (1996). *Third International Mathematics and Science Study Technical Report, Volume I: Design and Development*. Chestnut Hill, MA: Boston College.

Martin, Michael O., Kelvin D. Gregory, Steven E. Stemler (eds.) (2000). *TIMSS 1999 Technical Report*. Chestnut Hill, MA: Boston College.

Mullis, Ina V.S., M.O. Martin, E.J. Gonzalez and S.J. Chrostowski (2004). *Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: Boston College.

White, Halbert (1984). *Asymptotic Theory for Econometricians*. Orlando: Academic Press.

Wooldridge, Jeffrey M. (2001). Asymptotic Properties of Weighted *M*-Estimators for Standard Stratified Samples. *Econometric Theory* 17 (2): 451-470.

## Tables

Table 1: Regression of student performance on a linear and a quadratic term of school size

| Country | # of obser-vations | number of PSUs | coefficient on linear term | std error of linear term | coefficient on quadratic term | std error of quadratic term | Pattern | "optimal" size |
|---|---|---|---|---|---|---|---|---|
| Complete sample | 221392 | 7212 | 0.061 | (0.003691)*** | -0.000009 | (0.000001)*** | inverse u-shape | 3373 |
| OECD sample | 66867 | 2402 | 0.051 | (0.008391)*** | -0.000016 | (0.000004)*** | inverse u-shape | 1624 |
| Armenia | 4736 | 121 | -0.018 | (0.034129) | 0.00003 | (0.000022) | | |
| **Australia** | 4331 | 185 | 0.072 | (0.032798)** | -0.000024 | (0.000014)* | inverse u-shape | 1483 |
| Bahrain* | 4199 | 67 | -0.180 | (0.081058)** | 0.000113 | (0.000058)* | u-shape | 794 |
| **Belgium (Flemish)** | 4864 | 141 | 0.088 | (0.027340)*** | -0.000037 | (0.000010)*** | inverse u-shape | 1179 |
| Botswana | 4662 | 132 | 0.080 | (0.075905) | -0.00007 | (0.000061) | | |
| Bulgaria | 3984 | 160 | 0.064 | (0.046018) | -0.000036 | (0.000033) | | |
| Chile | 6309 | 193 | 0.043 | (0.011914)*** | -0.000005 | (0.000002)*** | inverse u-shape | 4010 |
| Chinese Taipei | 5314 | 148 | 0.042 | (0.016264)** | -0.000007 | (0.000003)** | inverse u-shape | 3193 |
| Cyprus | 3941 | 58 | 0.211 | (0.110319)* | -0.000216 | (0.000103)** | inverse u-shape | 490 |
| Egypt | 6916 | 211 | -0.006 | (0.013565) | 0.000004 | (0.000004) | | |
| **England** | 1845 | 57 | -0.201 | (0.121357) | 0.000069 | (0.000053) | | |
| Estonia | 3778 | 141 | 0.030 | (0.032454) | 0.000005 | (0.000025) | | |
| Ghana | 4661 | 139 | 0.147 | (0.042192)*** | -0.000029 | (0.000027) | linear effect significant | |
| Hong Kong, SAR | 4649 | 117 | 0.238 | (0.165496) | -0.000061 | (0.000071) | | |
| **Hungary** | 3112 | 147 | 0.218 | (0.060034)*** | -0.000159 | (0.000055)*** | inverse u-shape | 686 |
| Indonesia | 5671 | 148 | 0.073 | (0.019316)*** | -0.000011 | (0.000003)*** | inverse u-shape | 3186 |
| Iran, Islamic Republic | 4885 | 179 | 0.066 | (0.015007)*** | -0.000012 | (0.000003)*** | inverse u-shape | 2703 |
| Israel | 4032 | 136 | -0.055 | (0.050205) | 0.000034 | (0.000030) | | |
| **Italy** | 4256 | 170 | -0.036 | (0.087733) | 0.000032 | (0.000062) | | |
| **Japan** | 4835 | 145 | 0.017 | (0.050825) | 0.000022 | (0.000054) | | |
| Jordan | 4420 | 137 | 0.027 | (0.018393) | 0.000004 | (0.000005) | | |
| **Korea, Rep. of** | 5309 | 149 | 0.059 | (0.020719)*** | -0.00001 | (0.000010) | linear effect significant | |
| Latvia | 3325 | 128 | 0.065 | (0.033228)* | -0.000025 | (0.000024) | linear effect significant | |

| Country | # of obser-vations | number of PSUs | coefficient on linear term | std error of linear term | coefficient on quadratic term | std error of quadratic term | Pattern | "optimal" size |
|---|---|---|---|---|---|---|---|---|
| Lebanon | 3721 | 147 | 0.061 | (0.009088)*** | -0.00001 | (0.000002)*** | inverse u-shape | 3075 |
| Lithuania | 4323 | 127 | 0.079 | (0.024884)*** | -0.000024 | (0.000015) | linear effect significant | |
| Macedonia, Rep. of | 3728 | 142 | 0.044 | (0.043964) | -0.000027 | (0.000021) | | |
| Malaysia | 5314 | 150 | 0.010 | (0.025224) | 0.000003 | (0.000008) | | |
| Moldova, Rep. of | 3080 | 115 | 0.011 | (0.025337) | -0.000001 | (0.000010) | | |
| Morocco | 1986 | 88 | 0.013 | (0.019945) | -0.000005 | (0.000006) | | |
| **Netherlands** | 2895 | 122 | 0.072 | (0.029876)** | -0.000017 | (0.000012) | linear effect significant | |
| **New Zealand** | 3604 | 160 | 0.011 | (0.037933) | 0.000005 | (0.000016) | | |
| **Norway** | 4041 | 135 | 0.021 | (0.055698) | -0.000027 | (0.000067) | | |
| Palestinian Nat'l Auth. | 5308 | 144 | -0.019 | (0.028550) | 0.000015 | (0.000014) | | |
| Philippines | 6553 | 129 | 0.003 | (0.008310) | 0 | (0.000001) | | |
| Romania | 4013 | 145 | 0.084 | (0.019568)*** | -0.000023 | (0.000007)*** | inverse u-shape | 1791 |
| Russian Federation | 4595 | 211 | 0.046 | (0.023362)* | -0.000012 | (0.000014) | linear effect significant | |
| Saudi Arabia | 4087 | 150 | -0.040 | (0.064845) | 0.000067 | (0.000080) | | |
| **Scotland** | 2387 | 88 | 0.002 | (0.052755) | -0.000001 | (0.000029) | | |
| Serbia | 4137 | 144 | 0.121 | (0.022886)*** | -0.000053 | (0.000011)*** | inverse u-shape | 1141 |
| Singapore* | 5902 | 161 | -0.225 | (0.046656)*** | 0.000139 | (0.000020)*** | u-shape | 807 |
| **Slovak Republic** | 4190 | 178 | 0.115 | (0.050007)** | -0.00006 | (0.000041) | linear effect significant | |
| Slovenia | 3178 | 154 | 0.031 | (0.053303) | -0.000002 | (0.000043) | | |
| South Africa | 7405 | 212 | 0.193 | (0.044022)*** | -0.000079 | (0.000024)*** | inverse u-shape | 1224 |
| **Sweden** | 3971 | 149 | 0.078 | (0.067473) | -0.00004 | (0.000055) | | |
| Syrian Arab Republic | 4190 | 113 | 0.074 | (0.035275)** | -0.000025 | (0.000017) | linear effect significant | |
| Tunisia | 4628 | 141 | -0.004 | (0.019973) | 0.000005 | (0.000010) | | |
| **United States** | 7547 | 197 | -0.029 | (0.025225) | 0.000021 | (0.000014) | | |
| **Basque Country, Spain** | 2431 | 111 | 0.063 | (0.016410)*** | -0.00002 | (0.000006)*** | inverse u-shape | 1608 |
| **Indiana State, US** | 2028 | 50 | -0.117 | (0.050169)** | 0.000037 | (0.000025) | linear effect significant | |
| **Ontario Province, Can.** | 3981 | 176 | 0.020 | (0.044301) | 0 | (0.000036) | | |
| **Quebec Province, Can.** | 4135 | 164 | 0.021 | (0.015465) | -0.000005 | (0.000006) | | |

Notes:

* - The pattern of the coefficients in the countries marked with an asterisk predicts not an optimal school size, but a non-optimal size where test scores are lowest.
OECD member states marked in bold.
The number of strata is 51 in the complete sample and 17 in the OECD sample.
Less than 100 schools were sampled in Bahrain, Cyprus, England, Morocco, Scotland and in the US State of Indiana. Inferences from such small school samples seem problematic.

Table 2: School size in countries with significant coefficients for the linear and quadratic term of school size

| Country | # of observations | number of PSUs | smallest school size | highest school size | "optimal" school size | percentage of students at schools with higher enrolment than the "optimal" size | percentage of students at schools with enrolment +/- 100 of "optimal" size |
|---|---|---|---|---|---|---|---|
| Complete sample | 221392 | 7212 | 21 | 9960 | 3373 | 0.80 | 0.10 |
| OECD sample | 66867 | 2402 | 28 | 2925 | 1624 | 4.08 | 2.41 |
| Armenia | 4736 | 121 | 21 | 1637 | | | |
| **Australia** | 4331 | 185 | 94 | 2300 | 1483 | 5.67 | 3.00 |
| Bahrain* | 4199 | 67 | 161 | 1324 | 794 | 43.94 | 30.06 |
| **Belgium (Flemish)** | 4864 | 141 | 39 | 2786 | 1179 | 10.28 | 3.23 |
| Botswana | 4662 | 132 | 171 | 968 | | | |
| Bulgaria | 3984 | 160 | 60 | 1725 | | | |
| Chile | 6309 | 193 | 52 | 6410 | 4010 | 0.68 | 0.00 |
| Chinese Taipei | 5314 | 148 | 177 | 4960 | 3193 | 13.73 | 1.85 |
| Cyprus | 3941 | 58 | 143 | 792 | 490 | 64.47 | 56.13 |
| Egypt | 6916 | 211 | 83 | 5425 | | | |
| **England** | 1845 | 57 | 416 | 1807 | | | |
| Estonia | 3778 | 141 | 82 | 1359 | | | |
| Ghana | 4661 | 139 | 24 | 1500 | | | |
| Hong Kong, SAR | 4649 | 117 | 480 | 1750 | | | |
| **Hungary** | 3112 | 147 | 103 | 1087 | 686 | 16.90 | 21.43 |
| Indonesia | 5671 | 148 | 56 | 6549 | 3186 | 2.24 | 0.00 |
| Iran, Islamic Republic | 4885 | 179 | 49 | 5020 | 2703 | 0.47 | 0.00 |
| Israel | 4032 | 136 | 35 | 1650 | | | |
| **Italy** | 4256 | 170 | 232 | 1215 | | | |
| **Japan** | 4835 | 145 | 30 | 1097 | | | |
| Jordan | 4420 | 137 | 72 | 3187 | | | |
| **Korea, Rep. of** | 5309 | 149 | 75 | 2158 | | | |
| Latvia | 3325 | 128 | 47 | 1621 | | | |
| Lebanon | 3721 | 147 | 42 | 4916 | 3075 | 1.39 | 0.00 |
| Lithuania | 4323 | 127 | 51 | 1992 | | | |

| Country | # of observations | number of PSUs | smallest school size | highest school size | "optimal" school size | percentage of students at schools with higher enrolment than the "optimal" size | percentage of students at schools with enrolment +/- 100 of "optimal" size |
|---|---|---|---|---|---|---|---|
| Macedonia, Rep. of | 3728 | 142 | 60 | 2300 | | | |
| Malaysia | 5314 | 150 | 220 | 3296 | | | |
| Moldova, Rep. of | 3080 | 115 | 106 | 2980 | | | |
| Morocco | 1986 | 88 | 151 | 3359 | | | |
| **Netherlands** | 2895 | 122 | 175 | 2839 | | | |
| **New Zealand** | 3604 | 160 | 35 | 2358 | | | |
| **Norway** | 4041 | 135 | 59 | 999 | | | |
| Palestinian Nat'l Auth. | 5308 | 144 | 91 | 2010 | | | |
| Philippines | 6553 | 129 | 299 | 9960 | | | |
| Romania | 4013 | 145 | 66 | 3073 | 1791 | 2.03 | 0.64 |
| Russian Federation | 4595 | 211 | 23 | 2007 | | | |
| Saudi Arabia | 4087 | 150 | 24 | 904 | | | |
| **Scotland** | 2387 | 88 | 106 | 1851 | | | |
| Serbia | 4137 | 144 | 123 | 2282 | 1141 | 24.99 | 15.53 |
| Singapore* | 5902 | 161 | 197 | 1852 | 807 | 90.08 | 7.69 |
| **Slovak Republic** | 4190 | 178 | 53 | 1245 | | | |
| Slovenia | 3178 | 154 | 91 | 1076 | | | |
| South Africa | 7405 | 212 | 68 | 2017 | 1224 | 13.68 | 8.63 |
| **Sweden** | 3971 | 149 | 28 | 1126 | | | |
| Syrian Arab Republic | 4190 | 113 | 45 | 2355 | | | |
| Tunisia | 4628 | 141 | 33 | 2483 | | | |
| **United States** | 7547 | 197 | 38 | 2489 | | | |
| **Basque Country, Spain** | 2431 | 111 | 90 | 2761 | 1608 | 7.72 | 4.18 |
| **Indiana State, US** | 2028 | 50 | 49 | 1800 | | | |
| **Ontario Province, Can.** | 3981 | 176 | 61 | 1232 | | | |
| **Quebec Province, Can.** | 4135 | 164 | 50 | 2925 | | | |

Notes:

* - The pattern of the coefficients in the countries marked with an asterisk predicts not an optimal school size, but a non-optimal size where test scores are lowest.
OECD member states marked in bold.
The number of strata is 51 in the complete sample and 17 in the OECD sample.
Less than 100 schools were sampled in Bahrain, Cyprus, England, Morocco, Scotland and in the US State of Indiana. Inferences from such small school samples seem problematic.

Table 3: Contrast of the results from the regression without control variables and the regression with control variables

| Country | Regression without control variables | | | | Regression with control variables | | | |
|---|---|---|---|---|---|---|---|---|
| | number of observations | number of PSUs | Pattern | "optimal" size | number of observations | number of PSUs | Pattern | "optimal" size |
| Complete sample | 221392 | 7212 | inverse u-shape | 3373 | 217137 | 7145 | inverse u-shape | 3974 |
| OECD sample | 66867 | 2402 | inverse u-shape | 1624 | 66867 | 2402 | | |
| Armenia | 4736 | 121 | | | 3042 | 91 | | |
| **Australia** | 4331 | 185 | inverse u-shape | 1483 | 3288 | 143 | | |
| Bahrain* | 4199 | 67 | u-shape | 794 | 2866 | 50 | inverse u-shape | 623 |
| **Belgium (Flemish)** | 4864 | 141 | inverse u-shape | 1179 | 4864 | 141 | inverse u-shape | 1188 |
| Botswana | 4662 | 132 | | | 4662 | 132 | | |
| Bulgaria | 3984 | 160 | | | 3124 | 132 | | |
| Chile | 6309 | 193 | inverse u-shape | 4010 | 6309 | 193 | inverse u-shape | 3822 |
| Chinese Taipei | 5314 | 148 | inverse u-shape | 3193 | 4930 | 138 | | |
| Cyprus | 3941 | 58 | inverse u-shape | 490 | 3291 | 50 | | |
| Egypt | 6916 | 211 | | | 5120 | 157 | | |
| **England** | 1845 | 57 | | | 1285 | 44 | | |
| Estonia | 3778 | 141 | | | 3778 | 141 | | |
| Ghana | 4661 | 139 | linear effect sign | | 2860 | 114 | inverse u-shape | 1430 |
| Hong Kong, SAR | 4649 | 117 | | | 3539 | 93 | inverse u-shape | 1519 |
| **Hungary** | 3112 | 147 | inverse u-shape | 686 | 2872 | 136 | | |
| Indonesia | 5671 | 148 | inverse u-shape | 3186 | 5671 | 148 | inverse u-shape | 3535 |
| Iran, Islamic Republic | 4885 | 179 | inverse u-shape | 2703 | 4885 | 179 | | |
| Israel | 4032 | 136 | | | 2672 | 105 | | |
| **Italy** | 4256 | 170 | | | 3690 | 151 | | |
| **Japan** | 4835 | 145 | | | 4835 | 145 | | |
| Jordan | 4420 | 137 | | | 4004 | 126 | | |
| **Korea, Rep. of** | 5309 | 149 | linear effect sign | | 3322 | 125 | | |
| Latvia | 3325 | 128 | linear effect sign | | 2734 | 106 | | |
| Lebanon | 3721 | 147 | inverse u-shape | 3075 | 3721 | 147 | inverse u-shape | 3005 |

| Country | Regression without control variables | | | | Regression with control variables | | | |
|---|---|---|---|---|---|---|---|---|
| | number of observations | number of PSUs | Pattern | "optimal" size | number of observations | number of PSUs | Pattern | "optimal" size |
| Lithuania | 4323 | 127 | linear effect sign | | 3620 | 114 | | |
| Macedonia, Rep. of | 3728 | 142 | | | 3028 | 116 | | |
| Malaysia | 5314 | 150 | | | 5117 | 144 | | |
| Moldova, Rep. of | 3080 | 115 | | | 2180 | 82 | | |
| Morocco | 1986 | 88 | | | 1986 | 88 | | |
| **Netherlands** | 2895 | 122 | linear effect sign | | 2322 | 99 | | |
| **New Zealand** | 3604 | 160 | | | 2738 | 130 | | |
| **Norway** | 4041 | 135 | | | 4041 | 135 | inverse u-shape | 466 |
| Palestinian Nat'l Auth. | 5308 | 144 | | | 3757 | 118 | | |
| Philippines | 6553 | 129 | | | 5891 | 116 | | |
| Romania | 4013 | 145 | inverse u-shape | 1791 | 3332 | 124 | | |
| Russian Federation | 4595 | 211 | linear effect sign | | 4595 | 211 | | |
| Saudi Arabia | 4087 | 150 | | | 2925 | 118 | | |
| **Scotland** | 2387 | 88 | | | 1629 | 65 | | |
| Serbia | 4137 | 144 | inverse u-shape | 1141 | 3234 | 128 | inverse u-shape | 1107 |
| Singapore* | 5902 | 161 | u-shape | 807 | 5902 | 161 | u-shaped | 862 |
| **Slovak Republic** | 4190 | 178 | linear effect sign | | 3558 | 152 | | |
| Slovenia | 3178 | 154 | | | 3178 | 154 | | |
| South Africa | 7405 | 212 | inverse u-shape | 1224 | 7405 | 212 | inverse u-shape | 1519 |
| **Sweden** | 3971 | 149 | | | 3172 | 123 | | |
| Syrian Arab Republic | 4190 | 113 | linear effect sign | | 4190 | 113 | quadratic effect sign | |
| Tunisia | 4628 | 141 | | | 4628 | 141 | | |
| **United States** | 7547 | 197 | | | 6170 | 174 | quadratic effect sign | |
| **Basque Country, Spain** | 2431 | 111 | inverse u-shape | 1608 | 1902 | 87 | linear effect sign | |
| **Indiana State, US** | 2028 | 50 | linear effect sign | | 2028 | 50 | | |
| **Ontario Province, Can.** | 3981 | 176 | | | 3293 | 148 | | |
| **Quebec Province, Can.** | 4135 | 164 | | | 3202 | 128 | | |

Table 4

| Country | | # of observations | number of PSUs | coefficient on linear term | std error of linear term | coefficient on quadratic term | std error of quadratic term | Pattern | "optimal" size |
|---|---|---|---|---|---|---|---|---|---|
| Complete sample | * | 217137 | 7145 | 0.042997 | (0.002842)*** | -0.000005 | (0.000001)*** | inverse u-shape | 3974 |
| OECD sample | * | 66867 | 2402 | 0.008799 | (0.007153) | -0.000003 | (0.000003) | | |
| Armenia | | 3042 | 91 | -0.010711 | (0.039278) | 0.000005 | (0.000023) | | |
| **Australia** | | 3288 | 143 | 0.026459 | (0.035294) | -0.00001 | (0.000014) | | |
| Bahrain | | 2866 | 50 | 0.075391 | (0.035046)** | -0.00006 | (0.000030)** | inverse u-shape | 623 |
| **Belgium (Flemish)** | * | 4864 | 141 | 0.061209 | (0.018334)*** | -0.000026 | (0.000006)*** | inverse u-shape | 1188 |
| Botswana | * | 4662 | 132 | 0.039181 | (0.048138) | -0.00004 | (0.000044) | | |
| Bulgaria | | 3124 | 132 | -0.016957 | (0.039170) | 0.000003 | (0.000024) | | |
| Chile | * | 6309 | 193 | 0.025894 | (0.008154)*** | -0.000003 | (0.000001)*** | inverse u-shape | 3822 |
| Chinese Taipei | | 4930 | 138 | -0.003789 | (0.011658) | 0.000001 | (0.000002) | | |
| Cyprus | | 3291 | 50 | 0.091213 | (0.079444) | -0.000109 | (0.000072) | | |
| Egypt | | 5120 | 157 | -0.003786 | (0.011766) | 0.000002 | (0.000003) | | |
| **England** | | 1285 | 44 | -0.05912 | (0.133316) | -0.000005 | (0.000062) | | |
| Estonia | * | 3778 | 141 | 0.001279 | (0.024166) | 0.000009 | (0.000017) | | |
| Ghana | * | 2860 | 114 | 0.096266 | (0.027845)*** | -0.000034 | (0.000018)* | inverse u-shape | 1430 |
| Hong Kong, SAR | | 3539 | 93 | 0.261518 | (0.130732)** | -0.000086 | (0.000051)* | inverse u-shape | 1519 |
| **Hungary** | | 2872 | 136 | 0.030579 | (0.048052) | -0.000029 | (0.000042) | | |
| Indonesia | * | 5671 | 148 | 0.039164 | (0.017556)** | -0.000006 | (0.000003)** | inverse u-shape | 3535 |
| Iran, Islamic Republic | * | 4885 | 179 | -0.008018 | (0.014243) | 0.000002 | (0.000003) | | |
| Israel | | 2672 | 105 | 0.013318 | (0.055219) | -0.000003 | (0.000032) | | |
| **Italy** | | 3690 | 151 | 0.008868 | (.0814978) | 0.000003 | (0.0000549) | | |
| **Japan** | * | 4835 | 145 | -0.061448 | (0.070545) | 0.000057 | (0.000063) | | |
| Jordan | | 4004 | 126 | 0.031275 | (0.019513) | -0.000002 | (0.000005) | | |
| **Korea, Rep. of** | | 3322 | 125 | -0.00008 | (0.025656) | 0.000008 | (0.000010) | | |
| Latvia | | 2734 | 106 | 0.018929 | (0.031501) | -0.000009 | (0.000021) | | |
| Lebanon | * | 3721 | 147 | 0.043623 | (0.007821)*** | -0.000007 | (0.000002)*** | inverse u-shape | 3005 |
| Lithuania | | 3620 | 114 | 0.029365 | (0.025962) | -0.000005 | (0.000015) | | |

| Country | | # of observations | number of PSUs | coefficient on linear term | std error of linear term | coefficient on quadratic term | std error of quadratic term | Pattern | "optimal" size |
|---|---|---|---|---|---|---|---|---|---|
| Macedonia, Rep. of | | 3028 | 116 | -0.005365 | (0.038156) | -0.000009 | (0.000017) | | |
| Malaysia | | 5117 | 144 | 0.007189 | (0.020558) | 0.000002 | (0.000006) | | |
| Moldova, Rep. of | | 2180 | 82 | -0.026134 | (0.024591) | 0.000005 | (0.000008) | | |
| Morocco | * | 1986 | 88 | 0.007944 | (0.013314) | -0.000003 | (0.000004) | | |
| **Netherlands** | | 2322 | 99 | 0.022387 | (0.027477) | -0.000003 | (0.000010) | | |
| **New Zealand** | | 2738 | 130 | 0.007163 | (0.026484) | -0.000003 | (0.000011) | | |
| **Norway** | ***** | 4041 | 135 | 0.094742 | (0.046819)** | -0.000102 | (0.000048)** | inverse u-shape | 466 |
| Palestinian Nat'l Auth. | * | 3757 | 118 | -0.021791 | (0.027598) | 0.000017 | (0.000012) | | |
| Philippines | | 5891 | 116 | 0.006152 | (0.006948) | -0.000001 | (0.000001) | | |
| Romania | | 3332 | 124 | 0.033482 | (0.025315) | -0.00001 | (0.000008) | | |
| Russian Federation | * | 4595 | 211 | -0.00502 | (0.032123) | 0.000005 | (0.000016) | | |
| Saudi Arabia | | 2925 | 118 | 0.002851 | (0.046362) | 0.000009 | (0.000059) | | |
| **Scotland** | | 1629 | 65 | -0.051757 | (0.038934) | 0.000021 | (0.000019) | | |
| Serbia | ***** | 3234 | 128 | 0.058131 | (0.020738)*** | -0.000026 | (0.000009)*** | inverse u-shape | 1107 |
| Singapore | ***** | 5902 | 161 | -0.228356 | (0.048745)*** | 0.000133 | (0.000021)*** | u-shaped | 862 |
| **Slovak Republic** | | 3558 | 152 | -0.033175 | (0.045390) | 0.000026 | (0.000036) | | |
| Slovenia | * | 3178 | 154 | -0.029997 | (0.046831) | 0.000042 | (0.000039) | | |
| South Africa | ***** | 7405 | 212 | 0.064215 | (0.020990)*** | -0.000021 | (0.000013)* | inverse u-shape | 1519 |
| **Sweden** | | 3172 | 123 | 0.035509 | (0.048308) | -0.000035 | (0.000038) | | |
| Syrian Arab Republic | ***** | 4190 | 113 | 0.033846 | (0.025759) | -0.000019 | (0.000011)* | quadratic effect sign | |
| Tunisia | * | 4628 | 141 | 0.009145 | (0.017484) | -0.000002 | (0.000009) | | |
| **United States** | | 6170 | 174 | -0.029178 | (0.018668) | 0.000024 | (0.000009)** | quadratic effect sign | |
| **Basque Country, Spain** | | 1902 | 87 | 0.033071 | (0.015887)** | -0.000009 | (0.000006) | linear effect sign | |
| **Indiana State, US** | * | 2028 | 50 | -0.033255 | (0.0540501) | 0.0000004 | (0.0000261) | | |
| **Ontario Province, Can.** | | 3293 | 148 | 0.029617 | (0.042709) | -0.000011 | (0.000032) | | |
| **Quebec Province, Can.** | | 3202 | 128 | -0.009585 | (0.016559) | 0.000008 | (0.000006) | | |

Note:

- In countries marked with an asterisk, the regression includes imputed values both at the student and at the school level. In countries not marked, imputations at school level were not possible and thus the regression includes imputed values at the student level only.
- OECD member states marked in bold.
- The pattern of the coefficients in Singapore predicts not an optimal school size, but rather a non-optimal size where test scores are lowest.
- Less than 100 schools were sampled in Bahrain, Cyprus, England, Morocco, Scotland and in the US State of Indiana. Inferences from such small school samples seem problematic.
- The number of strata is 51 in the complete sample and 17 in the OECD sample.
- In the US state of Indiana, no information was provided on the experience of the teachers. The regression therefore does not control for this variable.
- In Norway, no information was provided on the percentage of students from disadvantaged family backgrounds. The regression therefore does not control for this variable.
- In Lebanon, no information was provided on the education of the teachers. The regression therefore does not control for this variable.

Table 5

| Country | | # of observations | number of PSUs | smallest school size in the regression sample | largest school size in the regression sample | "optimal" school size | percentage of students at schools with higher enrolment than the "optimal" size | percentage of students at schools with enrolment +/- 100 of "optimal" size |
|---|---|---|---|---|---|---|---|---|
| Complete sample | * | 217137 | 7145 | 21 | 9960 | 3974 | 1.64 | 0.11 |
| OECD sample | * | 66867 | 2402 | 28 | 2925 | | | |
| Armenia | | 3042 | 91 | 21 | 1637 | | | |
| **Australia** | | 3288 | 143 | 94 | 2300 | | | |
| Bahrain | | 2866 | 50 | 170 | 1034 | 623 | 65.15 | 27.48 |
| **Belgium (Flemish)** | ***** | 4864 | 141 | 39 | 2786 | 1188 | 10.28 | 3.23 |
| Botswana | * | 4662 | 132 | 171 | 968 | | | |
| Bulgaria | | 3124 | 132 | 70 | 1725 | | | |
| Chile | ***** | 6309 | 193 | 52 | 6410 | 3822 | 0.68 | 0.00 |
| Chinese Taipei | | 4930 | 138 | 177 | 4960 | | | |
| Cyprus | | 3291 | 50 | 143 | 792 | | | |
| Egypt | | 5120 | 157 | 118 | 5425 | | | |
| **England** | | 1285 | 44 | 416 | 1807 | | | |
| Estonia | * | 3778 | 141 | 82 | 1359 | | | |
| Ghana | ***** | 2860 | 114 | 24 | 1500 | 1430 | 1.54 | 1.54 |
| Hong Kong, SAR | | 3539 | 93 | 530 | 1750 | 1519 | 2.38 | 1.13 |
| **Hungary** | | 2872 | 136 | 103 | 1087 | | | |
| Indonesia | ***** | 5671 | 148 | 56 | 6549 | 3535 | 2.24 | 0.00 |
| Iran, Islamic Republic | * | 4885 | 179 | 49 | 5020 | | | |
| Israel | | 2672 | 105 | 118 | 1650 | | | |
| **Italy** | | 3690 | 151 | 232 | 1215 | | | |
| **Japan** | * | 4835 | 145 | 30 | 1097 | | | |
| Jordan | | 4004 | 126 | 72 | 72 | | | |
| **Korea, Rep. of** | | 3322 | 125 | 75 | 2046 | | | |
| Latvia | | 2734 | 106 | 49 | 1621 | | | |
| Lebanon | ***** | 3721 | 147 | 42 | 4916 | 3005 | 1.39 | 0.00 |

| Country | # of observations | number of PSUs | smallest school size in the regression sample | largest school size in the regression sample | "optimal" school size | percentage of students at schools with higher enrolment than the "optimal" size | percentage of students at schools with enrolment +/- 100 of "optimal" size |
|---|---|---|---|---|---|---|---|
| Lithuania | 3620 | 114 | 51 | 1992 | | | |
| Macedonia, Rep. of | 3028 | 116 | 60 | 2300 | | | |
| Malaysia | 5117 | 144 | 220 | 3296 | | | |
| Moldova, Rep. of | 2180 | 82 | 106 | 2980 | | | |
| Morocco * | 1986 | 88 | 151 | 3359 | | | |
| **Netherlands** | 2322 | 99 | 175 | 2839 | | | |
| **New Zealand** | 2738 | 130 | 35 | 2358 | | | |
| **Norway** * | 4041 | 135 | 59 | 999 | 466 | 14.92 | 29.40 |
| Palestinian Nat'l Auth. * | 3757 | 118 | 91 | 2010 | | | |
| Philippines | 5891 | 116 | 299 | 9544 | | | |
| Romania | 3332 | 124 | 66 | 3073 | | | |
| Russian Federation * | 4595 | 211 | 23 | 2007 | | | |
| Saudi Arabia | 2925 | 118 | 24 | 904 | | | |
| **Scotland** | 1629 | 65 | 106 | 1851 | | | |
| Serbia * | 3234 | 128 | 123 | 2282 | 1107 | 28.09 | 14.59 |
| Singapore * | 5902 | 161 | 197 | 1852 | 862 | 89.47 | 4.58 |
| **Slovak Republic** | 3558 | 152 | 122 | 1245 | | | |
| Slovenia * | 3178 | 154 | 91 | 1076 | | | |
| South Africa * | 7405 | 212 | 68 | 2017 | 1519 | 4.73 | 6.84 |
| **Sweden** | 3172 | 123 | 170 | 1126 | | | |
| Syrian Arab Republic * | 4190 | 113 | 45 | 2355 | | | |
| Tunisia * | 4628 | 141 | 33 | 2483 | | | |
| **United States** | 6170 | 174 | 38 | 2489 | | | |
| **Basque Country, Spain** | 1902 | 87 | 90 | 2761 | | | |
| **Indiana State, US** * | 2028 | 50 | 49 | 1800 | | | |
| **Ontario Province, Can.** | 3293 | 148 | 61 | 1232 | | | |
| **Quebec Province, Can.** | 3202 | 128 | 77 | 2925 | | | |

Note:

- In countries marked with an asterisk, the regression includes imputed values both at the student and at the school level. In countries not marked, imputations at school level were not possible and thus the regression includes imputed values at the student level only.
- OECD member states marked in bold.
- The pattern of the coefficients in Singapore predicts not an optimal school size, but rather a non-optimal size where test scores are lowest.
- Less than 100 schools were sampled in Bahrain, Cyprus, England, Morocco, Scotland and in the US State of Indiana. Inferences from such small school samples seem problematic.
- The number of strata is 51 in the complete sample and 17 in the OECD sample.
- In the US state of Indiana, no information was provided on the experience of the teachers. The regression therefore does not control for this variable.
- In Norway, no information was provided on the percentage of students from disadvantaged family backgrounds. The regression therefore does not control for this variable.
- In Lebanon, no information was provided on the education of the teachers. The regression therefore does not control for this variable.

Table 6

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| **Armenia** | | | | | |
| language of test seldom spoken | 1.65 | 9.69 | 28.86 | 54.93 | 34.66 |
| significant | no | no | yes | yes | yes |
| not born in country of test | 22.12 | 21.44 | 7.09 | 32.08 | 7.50 |
| significant | no | no | no | no | no |
| few books at home | 5.75 | 23.54 | 20.31 | 6.79 | 30.48 |
| significant | no | yes | yes | no | yes |
| **Australia** | | | | | |
| language of test seldom spoken | 2.79 | 14.17 | 59.67 | 36.30 | 78.37 |
| significant | no | no | yes | yes | yes |
| not born in country of test | 32.70 | 1.81 | 28.11 | 1.69 | 4.74 |
| significant | no | no | yes | no | no |
| few books at home | 27.81 | 18.24 | 22.01 | 6.53 | 22.65 |
| significant | yes | yes | yes | no | yes |
| **Bahrain** | | | | | |
| language of test seldom spoken | 10.60 | 10.94 | 10.04 | 1.89 | 1.20 |
| significant | no | yes | no | no | no |
| not born in country of test | 23.23 | 2.40 | 26.99 | 24.21 | 17.08 |
| significant | yes | no | yes | yes | no |
| few books at home | 19.71 | 8.17 | 21.39 | 21.14 | 9.08 |
| significant | yes | no | yes | yes | yes |
| **Belgium (Flemish)** | | | | | |
| language of test seldom spoken | 7.31 | 22.77 | 14.82 | 1.24 | 1.38 |
| significant | no | yes | no | no | no |
| not born in country of test | 9.98 | 6.72 | 16.91 | 10.97 | 8.39 |
| significant | no | no | no | no | no |
| few books at home | 9.29 | 8.05 | 7.82 | 20.36 | 6.68 |
| significant | no | yes | no | yes | no |
| **Botswana** | | | | | |
| language of test seldom spoken | 6.90 | 3.56 | 10.31 | 13.09 | 3.39 |
| significant | no | no | no | no | no |
| not born in country of test | 8.55 | 31.96 | 34.65 | 9.40 | 43.95 |
| significant | no | no | yes | no | yes |
| few books at home | 8.07 | 1.17 | 21.52 | 3.13 | 12.17 |
| significant | no | no | yes | no | no |
| **Bulgaria** | | | | | |
| language of test seldom spoken | 16.84 | 1.75 | 1.18 | 10.37 | 53.71 |
| significant | no | no | no | no | yes |
| not born in country of test | 20.21 | 38.77 | 41.85 | 43.18 | 77.95 |
| significant | no | no | yes | yes | yes |
| few books at home | 2.14 | 18.65 | 16.43 | 39.01 | 33.91 |
| significant | no | yes | yes | yes | yes |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| Chile | | | | | |
| language of test seldom spoken | 45.56 | 17.01 | 7.77 | 29.97 | 29.71 |
| significant | yes | yes | no | yes | yes |
| not born in country of test | 24.90 | 29.91 | 22.77 | 45.53 | 18.83 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 18.80 | 21.44 | 23.70 | 30.87 | 28.49 |
| significant | no | yes | yes | yes | yes |
| Chinese Taipei | | | | | |
| language of test seldom spoken | 41.72 | 10.23 | 42.29 | 32.59 | 36.13 |
| significant | yes | no | yes | yes | yes |
| not born in country of test | 64.13 | 72.23 | 78.60 | 51.67 | 20.91 |
| significant | yes | yes | yes | yes | no |
| few books at home | 43.76 | 29.57 | 32.71 | 35.96 | 31.59 |
| significant | yes | yes | yes | yes | yes |
| Cyprus | | | | | |
| language of test seldom spoken | 8.23 | 6.86 | 19.07 | 18.32 | 3.25 |
| significant | no | no | yes | yes | no |
| not born in country of test | 32.73 | 6.15 | 0.95 | 22.63 | 18.99 |
| significant | yes | no | no | yes | no |
| few books at home | 21.22 | 22.68 | 22.27 | 14.96 | 33.94 |
| significant | yes | yes | yes | yes | yes |
| Egypt | | | | | |
| language of test seldom spoken | 8.79 | 7.29 | 9.99 | 0.66 | 20.39 |
| significant | no | no | no | no | yes |
| not born in country of test | 35.64 | 38.02 | 29.82 | 13.27 | 32.63 |
| significant | yes | yes | yes | no | yes |
| few books at home | 6.95 | 6.72 | 11.77 | 13.44 | 9.25 |
| significant | no | no | no | no | no |
| England | | | | | |
| language of test seldom spoken | n.a. | n.a. | n.a. | n.a. | n.a. |
| significant | | | | | |
| not born in country of test | 35.69 | 58.73 | 28.93 | 67.00 | 0.49 |
| significant | yes | yes | yes | yes | no |
| few books at home | 2.31 | 47.49 | 50.20 | 43.49 | 51.28 |
| significant | no | yes | yes | yes | yes |
| Estonia | | | | | |
| language of test seldom spoken | 42.94 | 22.10 | 3.81 | 19.44 | 43.53 |
| significant | yes | no | no | yes | yes |
| not born in country of test | 7.64 | 36.40 | 2.05 | 7.07 | 19.27 |
| significant | no | yes | no | no | yes |
| few books at home | 21.96 | 25.36 | 28.17 | 21.20 | 6.97 |
| significant | yes | yes | yes | yes | no |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| Ghana | | | | | |
| language of test seldom spoken | 1.37 | 12.10 | 4.01 | 0.98 | 6.23 |
| significant | no | no | no | no | no |
| not born in country of test | 25.28 | 33.99 | 39.27 | 47.05 | 52.76 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 8.67 | 8.06 | 23.25 | 4.57 | 20.30 |
| significant | no | no | yes | no | yes |
| Hong Kong, SAR | | | | | |
| language of test seldom spoken | 52.71 | 22.52 | 42.57 | 38.41 | 44.02 |
| significant | yes | yes | yes | yes | yes |
| not born in country of test | 10.93 | 2.49 | 10.31 | 6.28 | 18.07 |
| significant | no | no | yes | no | yes |
| few books at home | 21.30 | 15.09 | 8.00 | 0.83 | 14.50 |
| significant | yes | yes | no | no | yes |
| Hungary | | | | | |
| language of test seldom spoken | 6.45 | 12.18 | 6.16 | 94.80 | 12.86 |
| significant | no | no | no | yes | no |
| not born in country of test | 22.36 | 17.38 | 47.48 | 9.23 | 35.99 |
| significant | no | no | no | no | yes |
| few books at home | 29.93 | 38.22 | 35.55 | 40.08 | 30.25 |
| significant | yes | yes | yes | yes | yes |
| Indonesia | | | | | |
| language of test seldom spoken | 19.66 | 18.71 | 39.63 | 23.40 | 21.30 |
| significant | no | yes | yes | yes | no |
| not born in country of test | 42.92 | 39.87 | 50.24 | 61.88 | 59.47 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 9.73 | 0.32 | 17.52 | 22.10 | 37.87 |
| significant | no | no | no | no | yes |
| Iran, Islamic Republic | | | | | |
| language of test seldom spoken | 5.41 | 27.40 | 17.76 | 4.40 | 13.09 |
| significant | no | yes | yes | no | yes |
| not born in country of test | 17.92 | 17.91 | 27.38 | 2.91 | 1.59 |
| significant | no | no | yes | no | no |
| few books at home | 15.98 | 35.60 | 33.06 | 14.70 | 12.91 |
| significant | yes | yes | yes | yes | yes |
| Israel | | | | | |
| language of test seldom spoken | 3.26 | 20.37 | 23.11 | 6.30 | 9.54 |
| significant | no | no | no | yes | no |
| not born in country of test | 22.54 | 14.11 | 3.60 | 35.09 | 12.60 |
| significant | no | no | no | yes | no |
| few books at home | 18.55 | 19.57 | 9.76 | 36.75 | 2.78 |
| significant | yes | yes | no | yes | no |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| Italy | | | | | |
| language of test seldom spoken | 40.46 | 3.14 | 23.69 | 22.40 | 12.85 |
| significant | yes | no | no | no | no |
| not born in country of test | 10.36 | 6.30 | 14.88 | 15.26 | 0.06 |
| significant | no | no | no | no | no |
| few books at home | 41.79 | 21.48 | 27.15 | 15.07 | 31.49 |
| significant | yes | yes | yes | yes | yes |
| Japan | | | | | |
| language of test seldom spoken | 7.17 | 21.19 | 0.42 | 67.60 | 41.07 |
| significant | no | no | no | yes | no |
| not born in country of test | 39.93 | 8.17 | 11.91 | 13.02 | 5.05 |
| significant | yes | no | no | no | no |
| few books at home | 20.92 | 30.34 | 27.81 | 31.60 | 34.94 |
| significant | yes | yes | yes | yes | yes |
| Jordan | | | | | |
| language of test seldom spoken | 5.89 | 41.62 | 3.34 | 17.33 | 1.97 |
| significant | no | yes | no | no | no |
| not born in country of test | 42.39 | 42.37 | 28.43 | 30.77 | 21.26 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 15.04 | 15.56 | 17.11 | 32.22 | 18.95 |
| significant | no | yes | yes | yes | yes |
| Korea, Rep. of | | | | | |
| language of test seldom spoken | 47.93 | 53.54 | 54.01 | 4.16 | 52.05 |
| significant | no | yes | no | no | yes |
| not born in country of test | 83.83 | 105.07 | 6.51 | 82.64 | 31.66 |
| significant | yes | yes | no | yes | no |
| few books at home | 44.63 | 31.91 | 42.87 | 35.73 | 48.00 |
| significant | yes | yes | yes | yes | yes |
| Latvia | | | | | |
| language of test seldom spoken | 29.85 | 17.07 | 9.65 | 12.23 | 9.21 |
| significant | no | no | no | no | no |
| not born in country of test | 71.81 | 5.72 | 1.56 | 34.05 | 4.02 |
| significant | yes | no | no | yes | no |
| few books at home | 16.61 | 28.00 | 14.89 | 12.68 | 26.83 |
| significant | yes | yes | yes | yes | yes |
| Lebanon | | | | | |
| language of test seldom spoken | 1.62 | 5.96 | 6.93 | 12.44 | 11.96 |
| significant | no | no | no | no | yes |
| not born in country of test | 0.23 | 2.13 | 8.23 | 0.29 | 9.63 |
| significant | no | no | no | no | no |
| few books at home | 11.67 | 19.34 | 14.44 | 30.84 | 19.92 |
| significant | no | yes | yes | yes | yes |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| Lithuania | | | | | |
| language of test seldom spoken | 5.58 | 21.00 | 24.07 | 18.37 | 38.65 |
| significant | no | no | no | no | yes |
| not born in country of test | 32.61 | 31.49 | 44.25 | 14.04 | 20.65 |
| significant | yes | yes | yes | no | no |
| few books at home | 20.72 | 26.22 | 26.80 | 27.47 | 25.83 |
| significant | yes | yes | yes | yes | yes |
| Macedonia, Rep. of | | | | | |
| language of test seldom spoken | 46.96 | 9.77 | 9.08 | 38.95 | 6.01 |
| significant | yes | no | no | no | no |
| not born in country of test | 28.73 | 32.15 | 38.86 | 42.33 | 52.38 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 7.42 | 10.48 | 0.77 | 8.24 | 9.86 |
| significant | no | no | no | no | no |
| Malaysia | | | | | |
| language of test seldom spoken | 28.66 | 46.01 | 26.87 | 39.23 | 48.83 |
| significant | yes | yes | yes | yes | yes |
| not born in country of test | 27.19 | 23.32 | 28.58 | 4.10 | 11.99 |
| significant | yes | yes | yes | no | no |
| few books at home | 22.91 | 37.10 | 15.13 | 14.80 | 22.12 |
| significant | yes | yes | yes | yes | yes |
| Moldova, Rep. of | | | | | |
| language of test seldom spoken | 12.22 | 16.34 | 24.43 | 14.86 | 24.61 |
| significant | no | no | no | no | yes |
| not born in country of test | 34.44 | 19.59 | 15.53 | 16.45 | 5.65 |
| significant | yes | no | no | yes | no |
| few books at home | 18.55 | 19.24 | 29.05 | 8.15 | 19.29 |
| significant | no | no | yes | no | yes |
| Morocco | | | | | |
| language of test seldom spoken | 5.02 | 7.11 | 10.02 | 7.57 | 8.68 |
| significant | no | no | no | no | no |
| not born in country of test | 54.08 | 17.86 | 36.23 | 17.07 | 29.67 |
| significant | no | no | yes | no | yes |
| few books at home | 11.85 | 2.06 | 2.83 | 5.10 | 7.92 |
| significant | no | no | no | no | no |
| Netherlands | | | | | |
| language of test seldom spoken | 0.89 | 12.08 | 19.47 | 7.07 | 32.58 |
| significant | no | no | yes | no | no |
| not born in country of test | 2.13 | 10.16 | 11.53 | 10.22 | 15.78 |
| significant | no | no | no | no | no |
| few books at home | 15.46 | 33.47 | 9.49 | 6.47 | 15.89 |
| significant | yes | yes | no | no | no |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| New Zealand | | | | | |
| language of test seldom spoken | 7.74 | 19.71 | 11.16 | 16.63 | 63.72 |
| significant | no | no | no | no | yes |
| not born in country of test | 17.75 | 4.12 | 10.13 | 11.78 | 8.23 |
| significant | no | no | no | no | no |
| few books at home | 29.32 | 23.45 | 20.58 | 26.15 | 29.14 |
| significant | yes | yes | no | yes | yes |
| Norway | | | | | |
| language of test seldom spoken | 15.91 | 34.81 | 10.90 | 15.51 | 15.13 |
| significant | no | yes | no | no | no |
| not born in country of test | 16.33 | 2.85 | 34.23 | 24.39 | 0.51 |
| significant | no | no | yes | yes | no |
| few books at home | 28.30 | 37.86 | 41.69 | 39.85 | 39.74 |
| significant | yes | yes | yes | yes | yes |
| Palestinian Nat'l Auth. | | | | | |
| language of test seldom spoken | 5.95 | 0.23 | 20.32 | 19.45 | 17.57 |
| significant | no | no | yes | no | yes |
| not born in country of test | 23.41 | 27.54 | 24.81 | 24.18 | 31.25 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 5.39 | 18.72 | 10.24 | 4.38 | 14.70 |
| significant | no | no | no | no | no |
| Philippines | | | | | |
| language of test seldom spoken | 25.34 | 7.61 | 21.61 | 12.04 | 10.46 |
| significant | yes | no | yes | no | no |
| not born in country of test | 20.78 | 21.50 | 63.13 | 26.59 | 18.26 |
| significant | no | no | yes | yes | no |
| few books at home | 14.96 | 5.09 | 8.45 | 11.97 | 12.88 |
| significant | no | no | no | no | no |
| Romania | | | | | |
| language of test seldom spoken | 42.00 | 3.56 | 3.78 | 24.37 | 13.42 |
| significant | yes | no | no | yes | no |
| not born in country of test | 18.58 | 19.71 | 16.68 | 7.34 | 37.24 |
| significant | no | no | no | no | no |
| few books at home | 28.17 | 38.95 | 22.28 | 23.89 | 36.47 |
| significant | yes | yes | yes | yes | yes |
| Russian Federation | | | | | |
| language of test seldom spoken | 2.80 | 27.54 | 15.64 | 31.60 | 20.04 |
| significant | no | yes | no | no | no |
| not born in country of test | 1.18 | 14.94 | 15.71 | 14.21 | 30.87 |
| significant | no | no | no | no | yes |
| few books at home | 9.44 | 20.88 | 25.78 | 19.33 | 11.32 |
| significant | no | yes | yes | yes | yes |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| Saudi Arabia | | | | | |
| language of test seldom spoken | | | | | |
| significant | | | | | |
| not born in country of test | 29.54 | 33.43 | 15.30 | 6.01 | 12.40 |
| significant | no | yes | no | no | no |
| few books at home | 3.30 | 15.48 | 3.31 | 31.10 | 3.36 |
| significant | no | no | no | yes | no |
| Scotland | | | | | |
| language of test seldom spoken | 85.89 | 6.64 | 35.51 | 8.96 | 22.96 |
| significant | yes | no | no | no | no |
| not born in country of test | 39.91 | 13.22 | 24.81 | 52.91 | 13.67 |
| significant | no | no | yes | yes | no |
| few books at home | 38.40 | 38.49 | 27.60 | 35.93 | 20.70 |
| significant | yes | yes | yes | yes | yes |
| Serbia | | | | | |
| language of test seldom spoken | 81.95 | 5.28 | 43.24 | 11.53 | 8.23 |
| significant | no | no | yes | no | no |
| not born in country of test | 41.63 | 29.59 | 22.45 | 34.87 | 37.44 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 16.58 | 51.55 | 37.96 | 25.69 | 23.98 |
| significant | no | yes | yes | yes | yes |
| Singapore | | | | | |
| language of test seldom spoken | 3.67 | 5.10 | 4.05 | 4.18 | 11.44 |
| significant | no | no | no | no | no |
| not born in country of test | 36.38 | 17.18 | 23.46 | 26.72 | 5.14 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 21.13 | 0.41 | 27.31 | 21.01 | 17.20 |
| significant | yes | no | yes | yes | yes |
| Slovak Republic | | | | | |
| language of test seldom spoken | 11.27 | 19.99 | 40.88 | 5.52 | 13.05 |
| significant | no | no | no | no | no |
| not born in country of test | 14.13 | 58.65 | 50.83 | 57.59 | 19.08 |
| significant | no | yes | yes | yes | no |
| few books at home | 42.07 | 34.49 | 43.78 | 44.44 | 24.22 |
| significant | yes | yes | yes | yes | yes |
| Slovenia | | | | | |
| language of test seldom spoken | 7.09 | 33.15 | 17.52 | 9.96 | 6.89 |
| significant | no | yes | no | no | no |
| not born in country of test | 1.67 | 18.78 | 8.47 | 31.46 | 0.11 |
| significant | no | no | no | yes | no |
| few books at home | 18.82 | 12.73 | 30.62 | 18.93 | 16.07 |
| significant | yes | yes | yes | yes | yes |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| South Africa | | | | | |
| language of test seldom spoken | 7.74 | 23.60 | 43.33 | 79.53 | 66.22 |
| significant | no | yes | yes | yes | yes |
| not born in country of test | 30.13 | 49.62 | 47.62 | 52.70 | 38.10 |
| significant | yes | yes | yes | yes | yes |
| few books at home | 3.17 | 11.58 | 10.50 | 48.49 | 8.44 |
| significant | no | no | no | yes | no |
| Sweden | | | | | |
| language of test seldom spoken | 1.83 | 0.84 | 20.88 | 4.98 | 0.32 |
| significant | no | no | no | no | no |
| not born in country of test | 5.75 | 8.77 | 28.31 | 16.73 | 17.95 |
| significant | no | no | yes | no | no |
| few books at home | 36.54 | 41.24 | 46.16 | 37.41 | 35.37 |
| significant | yes | yes | yes | yes | yes |
| Syrian Arab Republic | | | | | |
| language of test seldom spoken | 2.91 | 0.54 | 14.94 | 5.43 | 12.71 |
| significant | no | no | no | no | no |
| not born in country of test | 9.14 | 111.37 | 46.10 | 13.44 | 4.29 |
| significant | no | yes | yes | no | no |
| few books at home | 1.70 | 4.32 | 20.95 | 0.36 | 13.12 |
| significant | no | no | no | no | no |
| Tunisia | | | | | |
| language of test seldom spoken | 4.39 | 5.32 | 8.96 | 5.85 | 9.10 |
| significant | no | no | no | no | no |
| not born in country of test | n.a. | n.a. | n.a. | n.a. | n.a. |
| significant | | | | | |
| few books at home | 25.61 | 16.04 | 16.29 | 18.29 | 21.07 |
| significant | yes | no | yes | no | yes |
| United States | | | | | |
| language of test seldom spoken | 4.34 | 1.90 | 20.04 | 12.98 | 2.56 |
| significant | no | no | yes | no | no |
| not born in country of test | 46.20 | 15.64 | 23.85 | 10.21 | 21.91 |
| significant | yes | no | yes | no | yes |
| few books at home | 30.78 | 36.14 | 37.75 | 22.29 | 30.44 |
| significant | yes | yes | yes | yes | yes |
| Basque Country, Spain | | | | | |
| language of test seldom spoken | 10.90 | 12.61 | 5.20 | 6.34 | 15.58 |
| significant | yes | no | no | no | no |
| not born in country of test | 7.50 | 26.46 | 23.10 | 11.45 | 42.27 |
| significant | no | no | no | no | yes |
| few books at home | 9.28 | 11.78 | 17.19 | 31.58 | 20.00 |
| significant | yes | yes | yes | yes | yes |

Table 6 – continued

| | difference between groups within the same size category | | | | |
|---|---|---|---|---|---|
| | size category 1 | size category 2 | size category 3 | size category 4 | size category 5 |
| Indiana State, US | | | | | |
| language of test seldom spoken | 28.91 | 1.53 | 11.72 | 3.60 | 47.00 |
| significant | no | no | no | no | yes |
| not born in country of test | 0.80 | 19.96 | 10.46 | 34.10 | 38.77 |
| significant | no | no | no | no | yes |
| few books at home | 23.30 | 28.89 | 23.60 | 30.31 | 16.86 |
| significant | yes | yes | yes | yes | yes |
| Ontario Province, Can. | | | | | |
| language of test seldom spoken | 11.09 | 32.09 | 2.83 | 7.60 | 26.09 |
| significant | no | yes | no | no | yes |
| not born in country of test | 25.67 | 7.54 | 1.04 | 11.47 | 10.78 |
| significant | no | no | no | no | no |
| few books at home | 22.29 | 32.08 | 15.81 | 18.89 | 27.28 |
| significant | yes | yes | yes | yes | yes |
| Quebec Province, Can. | | | | | |
| language of test seldom spoken | 4.92 | 5.63 | 9.53 | 13.73 | 18.76 |
| significant | no | no | no | no | yes |
| not born in country of test | 18.69 | 14.95 | 25.23 | 37.60 | 10.85 |
| significant | yes | no | yes | yes | no |
| few books at home | 5.53 | 9.91 | 20.67 | 30.35 | 15.46 |
| significant | no | no | yes | yes | yes |