Recommendation on the Austrian PISA analysis

Huub Verstralen & Anton Béguin
Dept. POK, Cito, Arnhem, the Netherlands

February 15, 2010

Abstract

This report contains our findings on the impact of the student boycott on the Austrian PISA results. We conclude that it has been sufficiently shown that the reported results can be accepted as reflecting the achievements as if no boycott had been taken place.

1 Introduction

Following a conflict between government and student unions, the latter called on students to boycott the PISA survey. The call for this boycott was valid only from April 20 till 23, where the boycott call was withdrawn. During the boycott "test administrators were asked to postpone a test session if participation of students was not guaranteed" (pg 5 of the report). The period during which the PISA survey was held reached from April 15 to May 20. However, also after the boycott call was withdrawn there remained students who refused to participate. In the report the Austrian team has show to use various ways to identify the boycotting students and to quantify the effect on the results of the analyses. Our conclusions are that the number of boycotting students is small and their effect on the results most probably negligible.

2 Identification and numbers of boycotting students

To put the following numbers in their proper perspective it is important to note that the size of the complete sample, including boycotting students equals 6915. The easiest to identify are those students who refused to respond. These are the code 6 students numbered 158 (2.2%). Of them only their sex is known and males are overrepresented in this group (61.4%) against fifty-fifty in the sample. Then there are the students who gave silly answers or wrote something else in their booklet indicating that they did not seriously respond to the test questions. These are the code 8 students numbered 61 (0.9%). Analyses of
missing response frequencies showed that 65% or more missing or invalid responses were clearly more frequent (2.3%) in 2009 than in other years, where these frequencies in other years where from 0.2% to 0.7%. Additionally to the already code 6 and 8 students this gives 82 (1.2%) more students identified as boycotting, code 89. The last method to identify students whose PISA administration cannot be considered regular is by identifying schools where the administration of the test was so much tainted by boycotting students that the others cannot be expected to have been able to show their ability. This pertains to nine schools which add another 61 (0.9%) students to be excluded. In total we have 158+61+82+61=362 (5%) of the sample.

3 The effect of omitting boycott students

The next problem is to get an idea of the effect of these missing data in our sample. Can they be considered Missing Completely At Random (MCAR). First look at what we know of the students. Of all the boycotting students their sex is known. I have asked the bife team to reweigh the sample for overrepresentation of males in the boycotting students. They reported that the estimated mean ability did not change with reweighing. This must be due by small or negligible differences in mean ability among boys or girls, and/or by the small size of the missing data. So as checked via sex the missing data mechanism can be considered MCAR.

Another objection that was considered refers to the exclusion of the code 89 students, although there were much more students in the 2009 sample with 65% or more missing or invalid responses than in the other years. In these other years these numbers did not vanish, however. They ran from 0.2% through 0.7%. So on my request the team has drawn some random samples from the code 89 students with sample sizes that correspond to 0.2..0.7% of the total sample, and has put them back to recalculate the average ability. Again they did not report any difference.

The team itself did some analyses also on migration background in the total sample and the omitted students if recorded. First the numbers of students with a migration background in the omitted sample are very small, and the percentages do not deviate impressively from those in the total sample. Although reweighing the sample was not carried out I do not expect any real changes in the estimated mean.

4 Conclusion

The estimated mean from the sample with boycotting students omitted can be considered a statistically sound estimate of the population mean.