PISA 2009 bias analysis – England

5th March 2010

In this report we present the results of investigations of possible bias at school level in the PISA 2009 sample in England. In the first section the participating main sample schools are compared with the non-participating main sample schools. In the second section all participating schools (main and replacement) are compared with the original main sample schools.

The aim of this investigation was to identify any differences which existed at the 5% level of significance and to consider the implications of any which were found for possible sample bias.

1   *Comparison of main sample participants with main sample non-participants*

As a first stage, we compared the 131 participating main sample schools with the 59 non-participating main sample schools, first using chi-squared analysis to compare according to the stratification variables. This is reported in Table 1.1 below. We then carried out a logistic regression which again included the stratification variables along with additional variables which are generally considered to be possible indications of differences in the attainment of pupils in schools. The results of this are reported in Table 1.5.

1.1   Chi-squared analysis

Participating and non-participating main sample schools were first compared using the stratification variables, ie:

- School type
- Region
- School attainment band
- Gender

Tables 1.1 – 1.4 show these comparisons. There were no significant differences found on any of these analyses.

Table 1.1: School type

		main sample non-participants	main sample participants	total
Independent	n	5	8	13
	%	8.5%	6.1%	6.8%
Maintained non-selective	n	52	116	168
	%	88.1%	88.5%	88.4%
Maintained selective	n	2	7	9
	%	3.4%	5.3%	4.7%
Total		59	131	190

	Value	df	sig	
Pearson Chi-Square	0.662	2	0.718	

Table 1.2: Region

		main sample non-participants	main sample participants	total
Greater London	n	5	19	24
	%	8.5%	14.5%	12.6%
Midlands	n	22	39	61
	%	37.3%	29.8%	32.1%
North	n	20	37	57
	%	33.9%	28.2%	30.0%
South	n	12	36	48
	%	20.3%	27.5%	25.3%
Total		59	131	190

	Value	df	sig	
Pearson Chi-Square	3.141	3	0.370	

nfer

Table 1.3: School attainment (GCSE band)

		main sample non-participants	main sample participants	total
Lowest band	n	11	23	34
	%	18.6%	17.6%	17.9%
2nd lowest band	n	12	26	38
	%	20.3%	19.8%	20.0%
Middle band	n	12	27	39
	%	20.3%	20.6%	20.5%
2nd highest band	n	12	27	39
	%	20.3%	20.6%	20.5%
Highest band	n	12	28	40
	%	20.3%	21.4%	21.1%
Total		59	131	190

	Value	df	sig	
Pearson Chi-Square	0.055	4	1.000	

Table 1.4: Gender

		main sample non-participants	main sample participants	total
Boys	n	2	6	8
	%	3.4%	4.6%	4.2%
Girls	n	8	9	17
	%	13.6%	6.9%	8.9%
Mixed	n	49	116	165
	%	83.1%	88.5%	86.8%
Total		59	131	190

	Value	df	sig	
Pearson Chi-Square	2.313	2	0.315	

nfer

1.2 Logistic regression

The second stage of comparing participating and non-participating main sample schools was to conduct a logistic regression analysis. Included in this were the stratification variables plus additional variables which are generally considered to be connected with school attainment. These additional variables were:

- Number of full-time equivalent teachers in the school

- Pupil-teacher ratio

- Percentage of pupils eligible for free school meals (considered an indication of socio-economic status)

- Percentage of pupils with English as an additional language

- Average points score for General Certificate of Secondary Education

- Rural/urban school

The results of this analysis are shown in Table 1.5. The only differences that were found to be significant at the 5% level were small effects indicating that boys' schools and schools with high pupil-teacher ratios in the main sample were slightly less likely to respond than other similar schools.

Table 1.5   Results of logistic regression

	B	S.E.	df	Sig.
Full time equivalent teachers	-0.008	0.009	1	0.370
Pupil-teacher ratio	-0.161	0.080	1	0.045
% free school meals	-0.022	0.024	1	0.343
% with special educational needs	0.177	0.159	1	0.264
% with English as an additional language	0.024	0.017	1	0.150
average GCSE total points score	0.019	0.010	1	0.065
Non-selective school	0.930	1.262	1	0.461
Selective school	0.446	1.110	1	0.688
Greater London	0.677	0.714	1	0.343
North	0.194	0.448	1	0.664
South	0.790	0.464	1	0.089
Girl's school	-0.359	0.982	1	0.715
Boy's school	-1.587	0.732	1	0.030
Rural	1.283	1.344	1	0.340
School attainment - Lowest band	2.682	1.840	1	0.145
School attainment - 2nd Lowest band	2.157	1.375	1	0.117
School attainment - Middle band	1.335	1.076	1	0.215
School attainment - 2nd Highest band	0.774	0.837	1	0.355
Constant	-4.950	4.914	1	0.314

2 Comparison of main sample with all participating schools

The second stage of the bias analysis was to compare the 165 participating main sample and replacement schools with the original 190 schools in the main sample.

Again we first used chi-squared analysis to compare according to the stratification variables, and then used logistic regression with the inclusion of additional variables which are generally considered to be possible indications of differences in the attainment of pupils in schools. Since schools can be included in both the sample of respondents and the main sample the significance tests are not precise but are intended to give an indication of the severity of any differences.

2.1 Chi-squared analysis

They were first compared using the sampling variables, ie:

- School type
- Region
- School attainment band
- Gender

Tables 2.1 – 2.4 show this comparison. There were no significant differences at the 5% level found on any of these analyses.

Table 2.1: School type

		main sample	total participants
Independent	n	13	11
	%	6.8%	6.7%
Maintained non-selective	n	168	147
	%	88.4%	89.1%
Maintained selective	n	9	7
	%	4.7%	4.2%
Total		190	165

	Value	**df**	**sig**	
Pearson Chi-Square	0.056	2	0.972	

nfer

Table 2.2: Region

		main sample	total participants
Greater London	n	24	21
	%	12.6%	12.7%
Midlands	n	61	54
	%	32.1%	32.7%
North	n	57	46
	%	30.0%	27.9%
South	n	48	44
	%	25.3%	26.7%
Total		190	165
	Value	**df**	**sig**
Pearson Chi-Square	0.215	3	0.975

Table 2.3: School attainment (GCSE band)

		main sample	total participants
Lowest band	n	34	28
	%	17.9%	17.0%
2nd lowest band	n	38	31
	%	20.0%	18.8%
Middle band	n	39	36
	%	20.5%	21.8%
2nd highest band	n	39	37
	%	20.5%	22.4%
Highest band	n	40	33
	%	21.1%	20.0%
Total		190	165
	Value	**df**	**sig**
Pearson Chi-Square	0.376	4	0.984

nfer

Table 2.4: Gender

		main sample	total participants
Boys	n	8	8
	%	4.2%	4.8%
Girls	n	17	14
	%	8.9%	8.5%
Mixed	n	165	143
	%	86.8%	86.7%
Total		190	165

	Value	df	sig
Pearson Chi-Square	0.102	2	0.950

2.2   Logistic regression

The second stage was to conduct a logistic regression analysis. Included in this were the stratification variables plus additional variables which are generally considered to be connected with school attainment. These additional variables were:

- Number of full-time equivalent teachers in the school

- Pupil-teacher ratio

- Percentage of pupils eligible for free school meals (considered an indication of socio-economic status)

- Percentage of pupils with English as an additional language

- Average points score for General Certificate of Secondary Education

- Rural/urban school

The results of this analysis are shown in Table 2.5. Again no significant differences were found.

Table 2.5   Results of logistic regression

	B	S.E.	df	Sig.
Full time equivalent teachers	-0.004	0.006	1	0.441
Pupil-teacher ratio	-0.051	0.056	1	0.356
% free school meals	-0.009	0.015	1	0.554
% with special educational needs	0.039	0.085	1	0.647
% with English as an additional language	0.007	0.010	1	0.490
average GCSE total points score	0.004	0.005	1	0.417
Non-selective school	0.208	0.650	1	0.749
Selective school	-0.442	0.746	1	0.554
Greater London	-0.181	0.424	1	0.670
North	-0.041	0.302	1	0.891
South	0.095	0.290	1	0.742
Girl's school	0.218	0.600	1	0.716
Boy's school	-0.025	0.461	1	0.957
Rural	-0.117	0.702	1	0.867
School attainment - Lowest band	0.640	0.997	1	0.521
School attainment - 2nd Lowest band	0.502	0.757	1	0.507
School attainment - Middle band	0.402	0.616	1	0.514
School attainment - 2nd Highest band	0.322	0.491	1	0.512
Constant	-1.047	2.779	1	0.706

3   Conclusion

Few significant differences were found between participating and non-participating main sample schools. Furthermore none of these analyses have shown significant differences between the final participating schools after replacement and the main sample. We therefore feel confident that there is no evidence of bias at school level in the 2009 PISA sample in England. Indeed there is strong evidence that the participating sample after replacement is representative of the main sample schools selected.

nfer