Foras Taighde ar Oideachas
Educational Research Centre

PISA 2022

Non-response bias analysis for Ireland

This report was prepared in the Educational Research Centre by Brenda Donohue, Rachel Perkins, David Millar, Theresa Walsh and Mary Delaney, in consultation with Anastasios Karakolidis and Keisha Taylor.

Table of Contents

1. Introduction

PISA, the Programme for International Student Assessment, is a project of the Organisation for Economic Co-operation and Development (OECD) that assesses the skills of 15- and 16-year-olds in mathematics, reading, and science every three years. Ireland first participated in the PISA assessment in the year 2000. In Ireland, the Educational Research Centre (ERC) administers PISA on behalf of the Ministry of Education (known as the Department of Education in Ireland).

The study usually takes place every three years. However, the administration of the Main Study for the 2021 cycle was delayed internationally by one year due to the global COVID-19 pandemic. In addition to the delay to the Main Study caused by the pandemic, the testing window for PISA in Ireland moved from spring to autumn in 2022. This decision was the culmination of a process that began before PISA 2018 testing. Over various PISA cycles, schools and policy-makers in Ireland indicated that springtime was a crowded and busy time in the Irish school calendar, with practical tests, exam preparation and project work scheduled between March and May each year. In response to these concerns, the ERC conducted a study in autumn 2018 to assess the feasibility of autumn testing for PISA in Ireland (Denner, 2020). The results indicated that there would likely be no significant change in the overall performance of students on reading literacy, mathematics and science if testing moved to autumn relative to a spring administration. Based on these results, as well as a number of external considerations, the ERC in conjunction with the Department of Education requested to move to autumn testing for the PISA 2022 Main Study.

In order to provide a representative sample of 15- and 16-year old students in the Irish education system, a two-stage stratified sampling design was employed in the PISA 2022 study. The first stage utilised a systematic probability-proportionate-to-size technique. In total, 170 schools in Ireland were sampled, with each school having two replacement schools (similar in characteristics to the sampled school) also identified, should the original school not participate in PISA. In Ireland, the national PISA team was successful in recruiting all 170 of the originally-sampled schools. However, one school was later unable to participate due to a critical event that took place on campus, and the first replacement school was then recruited as a substitute. Data collection took place between 17th October and 2nd December 2022. All 170 schools met the minimum 33% student participation rate specified in the PISA Technical Standards, meaning that Ireland reached the international standards for school response, and a non-response bias analysis (NRBA) at school level is not required.

Although Ireland met the PISA Technical Standard thresholds for school-level sampling, it did not meet the threshold for student-level response (i.e. a minimum weighted response rate of 80%). Table 1 provides a breakdown of numbers of students in each participation category. Of the 7635 students sampled for the Main Study, 5569 students participated in the assessment. 1689 students were absent, including 512 students who did not participate due to parental refusal, as well as 14 students who had been present on the test day but were not included as participants due to non-engagement with the test. 266 students were excluded due to SEN, and 111 students were ineligible due to age, or because they were no longer enrolled in the sampled school.

Table 1. PISA 2022 participation in Ireland by participation category (unweighted)

Participation category	**N**
Participant	**5569**
Absent	1689
SEN exclusion	266
Ineligible	111
Total	7635

The final weighted student response rate for Ireland was 76.8%, meaning Ireland is required to undertake an analysis to explore potential bias at student level.

2. Methodology

This section describes the different approaches used to explore potential non-response bias at the student level in the achieved sample for Ireland in PISA 2022.

2.1 Samples used in the analysis

Two student samples (and associated subsamples) were used in the analysis.

1. The **full sample**, which includes all sampled PISA students that were selected and eligible to participate – 7258 students in total spread across Grades 8 to 12. It does not include students that were excluded from participating in PISA due to SEN, or who were not eligible for other reasons.[1] In the analyses described, base weights supplied by Westat were applied to the full sample.
   The full sample is compared to the **respondent-only subsample**. This includes only the 5569 students who participated in PISA, excluding 14 students who were present for the assessment but did not engage with the test. Non-response adjusted weights were applied to analyses involving this subsample.
2. The **achievement sample** includes PISA students that were eligible to participate with matching achievement data. This consists of 5040 students in Grades 10 and 11 only. As in the full sample, the achievement sample does not include students that were excluded or ineligible. In the analyses described, base weights supplied by Westat were applied to the achievement sample.
   The achievement sample is compared to the **respondent-only achievement subsample**. This includes only students who completed some parts of the PISA test and for whom achievement data are available (n = 3844). Non-response adjusted weights were applied to the achievement respondent-only subsample.

The analysis was conducted in three stages;

1. The first stage used descriptive statistics to compare the characteristics of the full sample, with base weights applied, to those of the respondent-only subsample, with adjusted weights applied.

2. The second stage compared the achievement sample to the respondent-only achievement subsample. The mean achievement scores of both groups were compared.

3. Finally, using the achievement sample, a binary logistic regression analysis was used to explore the relationship between students' participation status in the PISA test (i.e., 'participant' and 'non-participant'[2]) and a selected number of school- and student-level variables.

Specialist software was required to conduct this analysis, as the complexity of the PISA sample design needed to be accounted for. The SPSS Complex Sample module was used for this purpose. Student base weights (supplied by Westat) were applied to the full and achievement samples, while non-response adjusted weights were applied to the respondent-only subsamples for comparison.

[1] It was agreed with Westat to conduct the current analysis only on students that had either participated in PISA or who were eligible to participate but were absent. Therefore, of the original PISA sample of 7635 students, 7258 students were included in the analysis. 377 students were excluded, of these 266 had SEN exemptions, and 111 were age-ineligible or no longer attended the sampled schools.

[2] Non-participant in this analysis refers to students who did not take part due to refusal or absence on the test day.

2.2 Reference variables

The analysis in this report draws on a number of school- and student-level variables.

2.1.1 School- and student-level variables

Data on school-level characteristics, sourced from the Department of Education, and student-level characteristics, available from the PISA Student Tracking Forms, were used in this analysis and are described in Table 2.

Table 2. School- and student-level variables examined in the non-response bias analysis

	Variable	Description
School level	School gender	All boys, All girls, Mixed
	School type	Community, Comprehensive, Secondary, Vocational
	DEIS status*	DEIS, non DEIS
Student level	Gender	Male, Female
	Grade	International grades 8-12
	Study programme	Junior Cycle, Transition Year, Leaving Certificate Established, Leaving Certificate Applied, Leaving Certificate Vocational

*DEIS is the acronym used to refer to the Delivering Equality of Opportunity in Schools programme in Ireland. It is one measure of disadvantage used in the Irish school system. DEIS allocates resources to schools in line with the level of disadvantage of their student population ("https://assets.gov.ie/24451/ba1553e873864a559266d344b4c78660.pdf").

2.1.2 Achievement Scores

Students' performance in the Junior Cycle (JC) state examinations in 2022 was used as reference data and a proxy for achievement in our analysis. This compulsory state examination is usually taken by all students at the end of Grade 9 in Ireland, and students then choose to proceed to the optional Grade 10 (known as Transition Year), or proceed directly to Grade 11. Students can sit examinations in up to 13 different subject areas. All students are required to complete an examination in English and mathematics, but participation in the science examination is optional. Due to the COVID-19 pandemic, no Junior Cycle examinations were conducted in 2020 or 2021. For our analysis then, reference achievement data was available for only a subsample of the eligible PISA sample of 7258 students. These students were expected to be found primarily in Grade 10 (the grade with the highest concentration of PISA students) and Grade 11 in the PISA sample.

Student's performance in the JC examination in 2022 was used as a proxy of achievement for PISA students in Ireland. JC achievement data in English, mathematics and science were obtained from the national State Examinations Commission (SEC) to link with PISA student data for the expected subset of PISA students in Grades 10 and 11.[3]

These grades were converted to a composite Junior Cycle achievement scale for ease of comparison. Data for a maximum of three subjects for each student were received from the SEC. The mean achievement across subjects was computed for each student. Where students had data for only two subjects, the mean of two subjects was calculated.

[3] The data used in this analysis includes student scores on the state-certified final written examinations at the end of a three-year programme called the Junior Cycle. These scores are combined with classroom-based assessments and other short courses to produce a final award of the Junior Cycle Profile of Achievement. Only the examinations scores are used in this analysis.

The computation of the Junior Cycle composite score took into account the fact that the examination could be taken at different levels for two of the subjects concerned: English and mathematics. Science scores, instead, were on a common scale.[4] For the NRBA, a system was adopted that mirrored the approach taken by National Standardisation Group[5] in calculating grades for the Leaving Certificate state examination during the COVID-19 pandemic. As part of this approach for the non-response analysis, Junior Certificate achievement scores were mapped to a common scale (with a range 0-145), accounting for differences in grades as well as the level at which examinations were taken. More detail on this process can be found in Appendix A.

2.1.3 Matching Achievement Scores to PISA Sample

66,519 unique JC records were received from the SEC. As the ERC does not obtain national student unique identifiers when conducting the PISA assessment, identifiers were created to match cases in both the JC and PISA datasets using a combination of the national school identifier, date of birth, and gender of students. Because these identifiers were not unique, students' names were also included in the matching process (see Appendix B for further detail on the matching process). In total, 5048 JC records could be matched to students in the full sample. Table 3. shows the numbers of students in the full sample with and without matched JC achievement data at each grade level.

Table 3. Unweighted numbers of PISA students in the full sample by grade, with and without JC achievement data

International Grade	National Grade	JC Achievement Data		Total
		Available	Not Available	
8	**2 (Second Year)**	0	23	23
9	**3 (Third Year)**	4	1866	1870
10	**4 (Transition Year)**	3942	130*	4072
11	**5 (Fifth Year)**	1098	151	1249
12	**6 (Sixth Year)**	4	40	44
Total		**5048**	**2210**	**7258**

*All students at Grade 10 (Transition Year) and most students at Grade 11 (Fifth Year) were expected to have completed the Junior Certificate examination in 2022. PISA Student Lists and Tracking Forms were consulted to investigate why there were not JC achievement data for all students in Grade 10. Ninety-one out of 130 unmatched students were listed as either visiting foreign students, or new students that could not be linked to other schools in Ireland. No further information was available for the remaining 39 unmatched students.

As seen in Table 3 above, the majority of students with JC achievement data available were in Grades 10 and 11 (3942 and 1098, respectively). Small numbers of students with achievement data were also found in Grades 9 and 12. As students complete the JC examinations at the end of Grade 9, before proceeding to either Grade 10 or Grade 11, it is unclear why data would be available for students in these grades. Given the uncertainty around these students' status, data for these cases were judged to be unreliable and they were not included in subsequent achievement analyses. Overall, therefore, 5040 students of the 7258 students in the full sample had JC data associated with them, representing 69.6% of the PISA-eligible students selected to take part in the Main Study. Table 4 shows the composition of the full sample, and the achievement sample by grade level and participation status.

[4] For the Junior Cycle examination, English and Mathematics can be taken at Higher level, or at Ordinary level.
[5] The National Standardisation Group was responsible for overseeing the Leaving Certificate Calculated Grades process which was implemented for the state examination at Ireland's senior cycle level during the COVID-19 pandemic, when it was not possible or desirable to hold in-person examinations (DES, 2020)

Table 4. Unweighted numbers of PISA students in the full sample and achievement sample by grade and PISA participation status

International grade	National grade	Full sample			Achievement sample		
		PISA participant	PISA non participant	Total	PISA participant	PISA non participant	Total
8	2 (Second Year)	12	11	23	-	-	-
9	3 (Third Year)	1526	344	1870	-	-	-
10	4 (Transition Year)	3138	934	4072	3059	883	3942
11	5 (Fifth Year)	868	381	1249	785	313	1098
12	6 (Sixth Year)	25	19	44	-	-	-
Total		5569	1689	7258	3844	1196	5040

Note: Only students in Grades 10 and 11 with matched JC achievement data were included in the achievement sample.

3. Analysis

This section describes the non-response bias analysis, which was carried out in three stages. The first stage compared the characteristics of the full sample to those of the respondent-only subsample. Secondly, the mean achievement of the achievement sample was compared to that of the respondent-only achievement subsample. Finally, binary logistic regression analysis was employed to explore the relationship between students' participation status in PISA and a selected number of school- and student-level variables.

3.1 Comparison of the full and respondent-only samples on key variables

In this section, the school- and student-level characteristics of the full sample (with base weights applied) are compared to those of the respondent-only subsample (with adjusted weights applied).

A number of variables were selected at school- and student-level to describe the characteristics of students within the full sample, and the respondent-only subsample. The aim of this analysis was to examine differences in the two samples, when the appropriate weights were applied.

As mentioned above in section 2.1.1, the selected school variables were drawn from data provided to the ERC by the Department of Education in Ireland, and the student variables were gathered from the PISA Student Tracking Forms. At the school level, school gender, school type and disadvantaged status (referred to as DEIS status in Ireland) were employed in the analysis. At the student level, student gender, grade and study programme were analysed. Tables 5 and 6 show the outcomes of this analysis at school and student level.

In general, at the school level (Table 5), the percentages of students in each category in the respondent-only subsample (when the non-response adjusted weights were applied) are very close to the corresponding percentages for the full sample (when the base weights were applied). Some very small differences between the full sample, and the respondent-only subsample can be observed for school gender (where a difference of +0.1% was noted for all girls and mixed schools). Looking at disadvantaged school status, known as DEIS, we see that slightly fewer students from DEIS schools were included in the respondent-only subsample (-0.50%). Small differences between the full sample and respondent-only subsample are also seen for school type, where slightly fewer students from community schools (-0.20%) and slightly more students from comprehensive schools (+0.20%) were included in the respondent subsample.[6]

[6] For sampling purposes in Ireland, these two school types are collapsed into a single category (Community/Comprehensive) because both types of school fall under the same management structure. Small differences are evident here when the descriptive analysis is carried out with community and comprehensive schools treated separately. However, when the analysis is replicated with the two categories combined into a single category, the percentage difference between the full sample and the respondent-only subsample is reduced to zero.

Table 5. Weighted percentages of students in the full sample and the respondent-only subsample by school-level variables

	Full sample %	Respondent-only subsample %	Difference
School gender			
Boys	16.50	16.50	0.00
Girls	17.70	17.80	0.10
Mixed	65.80	65.70	-0.10
School type			
Community	13.80	13.60	-0.20
Comprehensive	2.40	2.60	0.20
Secondary	53.40	53.40	0.00
Vocational	30.40	30.40	0.00
DEIS status			
DEIS	78.50	79.00	0.50
non-DEIS	21.50	21.00	-0.50

Similarly, at student level (Table 6), the percentages of students in the full sample and respondent-only subsample are very similar. Some small differences in the percentages of students by grade level and study programme, which are closely related, are observed between the two samples. In particular, there is a slightly greater percentage of students at Grade 10 (+0.9%) and a slightly lower percentage of students at Grade 11 (-0.9%) in the respondent-only subsample when compared to the full sample. This echoes the pattern seen for the study programme variable, where there are slightly more students in the Transition Year programme (which corresponds to Grade 10 in the Irish system) and slightly fewer students in the Leaving Certificate Applied (LCA) programme amongst the respondent-only subsample when compared to the full sample. It should be noted that students at Grade 11 and 12 can follow this programme, and the number of PISA students in this programme is very small (162/7258 students in the full sample were recorded as following this study programme). Additionally, the LCA programme is likely to be followed by lower-achieving students. The LCA programme focuses on practical and vocational skills, and graduates are not eligible for direct entry to a degree-level course at Third Level.[7]

[7]The Leaving Certificate Applied has as "its primary objective the preparation of participants for adult and working life and the development of the participants' literacy and numeracy skills." (NCCA, 2023)

Table 6. Weighted percentages of students in the full sample and the respondent-only subsample by student-level variables

	Full sample %	Respondent-only subsample %	Difference
Gender			
Female	48.70	48.70	0.00
Male	51.30	51.30	0.00
Grade			
8	0.30	0.20	-0.10
9	25.90	26.10	0.20
10	56.10	57.00	0.90
11	17.10	16.20	-0.90
12	0.60	0.50	-0.10
Study Programme			
Junior Cycle	26.20	26.30	0.10
Transition Year	56.10	57.00	0.90
LC Applied	2.20	1.30	-0.90
LC Established	14.20	14.10	-0.10
LC Vocational	1.30	1.30	0.00

Note: Base weights were applied for the full sample and non-response adjusted weights were applied to the respondent-only subsample.

3.2 Comparison of the achievement outcomes of achievement and the respondent-only achievement samples

The achievement sample was used to conduct a bias analysis using the JC composite score as a proxy for achievement for the PISA students.Table 7 compares the JC achievement outcomes of the achievement sample and the respondent-only achievement subsample using the JC composite scores (described in section 2.1.2 and Appendix A), which range from 0 to 145. The mean score of the respondent-only achievement subsample is 2.35 marks higher than that of the achievement sample. This equates to approximately one tenth of a Standard Deviation.[8]

Table 7. Comparison of weighted mean scores in the achievement sample and respondent-only subsample

JC Composite Score	Achievement Sample	Respondent-only achievement subsample
N weighted	45291.57	45875.93
Mean	97.50	99.85
Std. Error of Mean	0.74	0.63
95% Confidence Interval Lower	96.02	98.59
95% Confidence Interval Upper	98.98	101.11
Std. Deviation	20.71	18.81

Note: Base weights were applied for the achievement sample and non-response adjusted weights were applied to the respondent-only achievement subsample.

[8] Tests for significance were not carried out due to the interdependence of the samples.

An additional analysis was then carried out, whereby the mean scores of students within the achievement sample and the respondent-only achievement subsample were compared to the same school- and student-level characteristics used in the descriptive analysis outlined in section 3.1. Tables 8 and 9 present comparisons of school- and student-level variables respectively.

At the school level, differences in achievement can be observed in some of categories. Across all subgroups examined, the mean achievement scores of students in the respondent-only achievement subsample were somewhat higher than those in the achievement sample. Looking at school gender, the difference in mean scores between the two samples was greatest for students in mixed schools (+2.74 points for the respondent-only achievement subsample). The largest difference in mean achievement between both samples when looking at school type can be seen in students attending community schools; they have a mean achievement score that is 4.16 marks above the mean of the achievement sample. Finally, looking at DEIS status, the difference between the two samples is larger for students in DEIS schools (+3.65 points for the respondent-only subsample) than students in non-DEIS schools (+1.92 points for the respondent-only subsample).[9]

Table 8. *Comparison of weighted mean scores by selected school-level variables in the achievement sample and the respondent-only achievement subsample*

	Achievement Sample		**Respondent-only achievement subsample**		
	Mean JC Composite Score	**S.E**	**Mean JC Composite Score**	**S.E.**	**Difference**
School gender					
Boys	102.82	1.88	103.66	1.61	0.84
Girls	99.44	2.25	101.86	1.83	2.42
Mixed	95.51	0.83	98.25	0.72	2.74
School type					
Community	92.49	2.34	97.09	1.97	4.60
Comprehensive	98.19	2.30	98.94	2.64	0.75
Secondary	100.20	1.07	102.10	0.91	1.90
Vocational	95.06	1.13	97.20	0.98	2.14
DEIS status					
DEIS	86.84	1.67	90.49	1.43	3.65
Non-DEIS	100.47	0.65	102.39	0.57	1.92

Note: Base weights were applied for the achievement sample and non-response adjusted weights were applied to the respondent-only achievement subsample.

At student level, again the mean scores for all subgroups are higher for students in the respondent-only achievement subsample than those in the achievement sample. With regard to students' gender, the difference in mean achievement between the two samples is greater for boys (+3.02 for the respondent-only achievement subsample) than for girls (+1.68 for the respondent-only achievement subsample). Students in the respondent-only achievement subsample at Grade 11 had a mean score that was 4.69 marks higher than the achievement sample, a bigger gap that the one observed at Grade 10 where there was a difference of 1.60 marks. Finally, looking at study programme, the largest differences can be seen in two programmes; Leaving Certificate Applied and Leaving Certificate Vocational (LCV). Students in the respondent-only subsample who were following the LCA programme have a mean score that is 4.8 marks higher than the mean of the achievement

[9] Tests for significance were not carried out due to the interdependence of the samples.

sample, while students in the respondent-only achievement subsample undertaking the LCV programme have a mean score that is 5.3 marks higher than the corresponding score for the achievement sample. As can be seen from the mean achievement scores for both samples, students in these study programmes perform at lower levels compared to those who follow the Transition Year and LC Established programme. It should be noted that the percentages of students following these programme are very low, at less than 2.2% (see Table 6).

Table 9. Comparison of weighted mean scores by selected student-level variables in the achievement sample and the respondent-only subsample

	Achievement Sample		**Respondents-only achievement subsample**		
	Mean JC Composite Score	**S.E**	**Mean JC Composite Score**	**S.E.**	**Difference**
Gender					
Female	99.78	0.91	101.46	0.79	1.68
Male	95.19	0.97	98.21	0.85	3.02
Grade					
10	100.44	0.68	102.04	0.61	1.60
11	86.90	1.23	91.59	1.21	4.69
Study Programme					
Transition Year	100.44	0.68	102.04	0.61	1.60
LC Applied	58.06	2.31	62.86	3.16	4.80
LC Established	91.62	1.19	94.23	1.18	2.61
LC Vocational	86.73	2.65	92.03	2.49	5.30

Note: Base weights were applied for the achievement sample and non-response adjusted weights were applied to the respondent-only achievement subsample.

3.3 Logistic regression

Finally, a binary logistic regression was conducted to investigate the relationship between students' non-participation in PISA and the school- and student-level characteristics (presented in Table 2), along with students' achievement on the JC examination. This analysis was carried out on the achievement sample without any weights applied, with participation status as the dependent variable.

In preparation for the logistic regression analysis, the association between each school- and student-level variable and participation status was tested using chi-square analysis. From this analysis, school type, DEIS status, grade and study programme were found to be statistically significantly associated with participation status, and therefore were chosen for inclusion in the logistic regression model. The remaining variables (school gender, student gender) were not shown to be significantly associated with participation, and therefore were not included in the model (see Appendix C for further details).

Table 10 outlines the variables that were included in the binary logistic regression model. Two of the categorical variables were recoded into dummy variables (study programme and school type). For study programme, Transition Year was used as the reference category, while for school type, secondary school was the reference category, as both of these groups contained the highest percentage of students across the variable categories.

Table 10. Variables included in the binary logistic regression model

Variables	Coding structure
Participation status	Participant Non-participant
Study programme	Transition Year *(reference category)* LC Applied LC Established LC Vocational
School type	Secondary schools *(reference category)* Comprehensive/Community* Vocational
DEIS status	DEIS school *(reference category)* Non-DEIS school
JC composite score	Scale (range: 0 to 145)

*For this analysis, the categories Community and Comprehensive were collapsed into a single variable as the percentages of students attending comprehensive schools was low (2.2% of the achievement sample n=113).

All variables were entered simultaneously into the logistic regression model (Table 11). Two variables were found to be significant predictors of students' non-participation in PISA (JC composite score and study programme), while for the remaining two variables (DEIS status and school type) there is no statistically significant difference in the likelihood of participation in PISA.

With regard to students' performance on the JC composite score, the model demonstrates that the higher students score on this scale, the less likely they are to be non-participants in PISA – that is that higher achieving students were significantly more likely to participate in the study (Odds ratio=0.977, suggesting a relatively small effect).

No significant differences were observed between Transition Year students and students following the LC Established programme, or between Transition Year students and those in the LC Vocational programme. However, students following the LCA programme were almost twice as likely as Transition Year students to be non-participants in PISA and this finding is statistically significant (p = .001). It should be noted that the LCA cohort only represents 1-2% of students (Table 6) so, although significant, this difference is unlikely to have a major impact on the overall sample.

Table 11. Binary logistic regression with student participation in PISA as the outcome variable, and one school-level and three student-level variables tested together (unweighted)

Model of response status		B	S.E	Wald	df	P	Exp (B) [Odds Ratio]
Study programme *(reference: Transition Year*	LC Applied	.618	.189	10.705	1	**.001**	1.854
	LC Established	-.161	.092	3.025	1	.082	.851
	LC Vocational	-.229	.281	.668	1	.414	.795
School type *(reference: Secondary schools)*	Comprehensive/Community	.065	.097	.450	1	.503	1.067
	Vocational	.114	.081	1.999	1	.157	1.121
DEIS status *(reference DEIS)*		.085	.085	.984	1	.321	1.088
JC composite score *(range 0 to 145)*		-.023	.002	179.137	1	**<0.001**	.977

Summary

This non-response bias analysis has examined response patterns in selected samples in Ireland from the PISA 2022 Main Study testing.

A comparison of the characteristics of the full PISA sample with those of the respondent-only subsample found that, overall, when the non-response adjusted weights were applied to the respondent-only subsample, the distribution of this subsample broadly matched the distribution of the full sample (with base weights applied). A few small differences in the distributions between the two groups were observed, however. At the school level, the percentage of students attending DEIS schools (an indicator of disadvantage) was very slightly higher in the respondent-only sample when compared to the full sample (a difference of 0.5 percentage points). At the student level, the percentage of students at Grade 10 (and therefore following the Transition Year programme) is also slightly higher (by 0.9 percentage points) among the respondent-only sample, while the percentage of students following the Leaving Certificate Applied programme was slightly lower (by 0.9 percentage points). However, these differences remain small and may not point to an important bias once adjusted weights have been applied.

A second analysis looked at student achievement. For this, a reduced sample of students for whom national state examination achievement data was available was used. A comparison of the achievement sample with base weights applied against a respondent-only achievement subsample with adjusted weights applied was carried out. A difference of 2.35 points in mean achievement between the two groups was found here, equating to approximately a tenth of a Standard Deviation, indicating a small upward bias in achievement among the PISA respondents.

Finally, a binary logistic regression was carried out. This analysis indicated that two variables were statistically significant in predicting participation status in PISA. Higher achievement scores on the national state examination was associated with a higher likelihood of participation in PISA, while students following the Leaving Certificate Applied programme (which is often undertaken by lower-achieving students) were found to be significantly less likely to participate in PISA than those in the modal grade of Transition Year.

In attempting to answer the central question of this analysis, what is the potential bias in the achieved sample in Ireland for the PISA 2022 Main Study, differences between samples, and between respondents and non-respondents have been examined. While some differences in the distributions and mean scores have been found in this analysis, they appear to be small.

However, a difference in the achievement scores between respondents and non-respondents was found (equating to approximately one tenth of a Standard Deviation), indicating that there is some evidence of an upwards bias within the respondent sample.

Appendix A: Description of process to convert JC grades to composite score

The Junior Cycle exam results for mathematics, English and science were received by the ERC from the State Examinations Commission. In the dataset received, each student had scores for a maximum of three subjects: English, mathematics and science. These marks appeared on different scales (as the English and mathematics examinations can be taken at Higher and Ordinary levels, while science is only available at a common level).

Data from these examinations were combined into a single composite score using a scaling procedure to properly reflect the distinct levels of achievement represented by examinations at different levels within each subject. Marks from each subject were mapped to a common scale, following the procedure already set out by the Junior Certificate Overall Performance Scale (OPS), which was developed in the ERC (Kellaghan & Millar, 2003; Martin & Hickey, 1993; Sofroniou, Cosgrove & Shiel, 2002; Weir & Kavanagh, 2018).

This mapping procedure produced a mark for each student on a single scale for each subject for which we had data. Finally, the Junior Cycle composite score was calculated by computing the mean of the scores across the subjects which were available for each student. For students who did not study science, the mean score was based on the average of their scores for English and mathematics examinations.

Appendix B: Description of matching processes

Table B.1 outlines four different types of matching based on student demographic variables in order to identify PISA students in the JC achievement dataset. Note that the total numbers of cases reported as matched here (5196) is higher than reported in the main report body (5040) as it includes students in the PISA sample who were ineligible to participate due to either SEN exclusions, being age-ineligible or because students no longer attended the sampled schools.

Table B.1. Description of matching processes employed

Match	Variables matched	Match type/process	No of cases matched
1	National school identifier (roll number), student date of birth, gender, name	All identifiers based on this combination of variables was unique in both datasets and therefore matches were found in a ratio of 1:1 for all records here.	4977
2	National school identifier, date of birth, gender	Those students not matched in Match 1 were matched in a "fuzzy match" (i.e. non unique match) using unique identifiers based on national school identifier, student date of birth and gender as some students were expected to have different variations of their names across the two datasets. The student names across both datasets were then scanned manually and matches were identified where there were small typographical or linguistic differences between names in both datasets.	102
3	National school identifier, student name	A second "fuzzy match" was run based on national school identifier and student name to scan cases where date of birth and gender almost matched e.g. the month of birth was different across both datasets but all else was constant.	82
4	Student date of birth, gender, name	A third "fuzzy match" was carried out to identify students who had the same demographic details in both datasets but different national school identifiers. Students were in matched in cases where the students were unique in each dataset in all details and they were identified in the PISA student lists as being new to the PISA school.	30
5	National school identifier, student date of birth, gender,	A final "fuzzy match" was carried out to identify students listed with different names in both datasets but where all else was constant. Students with name changes or gender changes that were identified by the student lists were matched in this final match.	4

Appendix C: Chi-square analyses

Table C1. Chi-square analysis conducted on the full sample with base weights applied

	Relationship between different variables		Chi Square	Adjusted F	df1	df2	Sig.
School-level variables	School gender and PISA participation	Pearson	19.49	3.16	1.98	170.52	0.05
		Likelihood Ratio	19.83	3.22	1.98	170.52	0.05
	School type and PISA participation	Pearson	41.23	5.85	2.86	245.77	<0.001
		Likelihood Ratio	41.26	5.86	2.86	245.77	<0.001
	DEIS and PISA participation	Pearson	59.44	18.30	1.00	86.00	<0.001
		Likelihood Ratio	56.82	17.50	1.00	86.00	<0.001
Student-level variables	Student gender and PISA participation	Pearson	1.99	1.11	1.00	86.00	0.30
		Likelihood Ratio	1.99	1.11	1.00	86.00	0.30
	Student grade and PISA participation	Pearson	80.16	15.73	3.88	333.76	<0.001
		Likelihood Ratio	76.89	15.09	3.88	333.76	<0.001
	Student study programme and PISA participation	Pearson	152.43	26.39	3.83	329.65	0.00
		Likelihood Ratio	130.67	22.62	3.83	329.65	<0.001

References

Denner, S. (2020). *PISA testing in Spring vs Autumn 2018: A Feasibility Study*. Unpublished report submitted to the Department of Education and Skills.

Department of Education and Skills (DES). (2017). *DEIS Plan 2017. Delivering Equality of Opportunity in Schools*. Author. https://assets.gov.ie/24451/ba1553e873864a559266d344b4c78660.pdf.

DES. (2020). *Calculated Grades for Leaving Certificate 2020: Report from the National Standardisation Group to the Independent Steering Committee and the Programme Board*. Author. https://www.gov.ie/pdf/?file=https://assets.gov.ie/86709/aacaf08c-a32b-4e74-817b-5a3ba1e2a7e8.pdf#page=null.

Kellaghan, T. & Millar, D. (2003). *Grading in the Leaving Certificate Examination: A Discussion Paper*. Dublin: Educational Research Centre.

National Council for Curriculum and Assessment (NCCA). (2023) *Leaving Certificate Applied Programme Statement*. Author. Revised-LCAPS.pdf (curriculumonline.ie).

Martin, M.O., & Hickey, B.L. (1993). *The 1992 Junior Certificate Examination: a Review of Results.* Dublin: National Council for Curriculum and Assessment.

Sofroniou, N., Cosgrove, J., & Shiel, G. (2002). Using PISA variables to explain performance on Junior Certificate examinations in mathematics and science. *Irish Journal of Education, 33*, 99-124.

Weir, S., & Kavanagh, L. (2018). *The evaluation of DEIS at post-primary level: Closing the achievement and attainment gaps*. Dublin: Educational Research Centre.