



**Is Canada really an education superpower? The impact of exclusions and non-response on results from PISA 2015**

**Jake Anders  
Silvan Has  
John Jerrim  
Nikki Shure  
Laura Zieger**

---

**Department of Quantitative Social Science**

Working Paper No. 19-11

December 2019

## **Disclaimer**

Any opinions expressed here are those of the author(s) and not those of the UCL Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

DoQSS Workings Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Department of Quantitative Social Science, UCL Institute of Education, University College London,  
20 Bedford Way, London WC1H 0AL, UK

# Is Canada really an education superpower? The impact of exclusions and non-response on results from PISA 2015

Jake Anders <sup>1</sup>, Silvan Has<sup>2</sup>, John Jerrim<sup>3</sup>, Nikki Shure<sup>4</sup> and Laura Zieger <sup>5</sup>

## **Abstract**

The purpose of large-scale international assessments is to compare educational achievement across countries. For such cross-national comparisons to be meaningful, the students who take the test must be representative of the whole population of interest. In this paper we consider whether this is the case for Canada, a country widely recognised as high-performing in the Programme for International Student Assessment (PISA). Our analysis illustrates how the PISA 2015 data for Canada suffers from a much higher rate of student exclusions, school non-response and pupil non-response than other high-performing countries such as Finland, Estonia, Japan and South Korea. We discuss how this emerges from differences in how children with Special Educational Needs are defined and rules for their inclusion in the study, variation in school response rates and the comparatively high rates of pupil test absence in Canada. The paper concludes by investigating how Canada's PISA 2015 rank would change under different assumptions about how the non-participating students would have performed were they to have taken the PISA test.

**JEL Codes:**

**Keywords:**

Note: All joint first-authors. Equal contribution made by all.

---

<sup>1</sup> University College London

<sup>2</sup> University College London

<sup>3</sup> University College London

<sup>4</sup> University College London

<sup>5</sup> University College London

## 1. Introduction

The Programme for International Student Assessment (PISA) is an important international study of 15-year-olds achievement in reading, science and mathematics. It is conducted every three years by the Organisation for Economic Cooperation and Development (OECD) and receives substantial attention from policymakers, the media, academics and the wider education community. Particular attention is often paid to the top-performing nations in PISA and these often inspire policy development in other countries (Raffe, 2011). Although Finland (Hendrickson 2012; Takayama, Waldow and Sung 2013) and the high-performing East Asian nations (Feniger and Lefstein 2014; Jerrim 2015) have often taken the limelight, a North American country, Canada, has also received significant attention. Indeed, despite its cultural, linguistic and historical similarities to many other Western nations, Canada achieves much higher average PISA scores than most OECD countries, while also apparently having a more equitable distribution of educational achievement. This is illustrated by Table 1, which benchmarks Canada's PISA 2015 reading scores against key comparators. Based upon these results, Canada has consequently been described as an 'education super-power' (Coughlan 2017), with Andreas Schleicher – the man who developed the OECD's PISA programme – suggesting that this is driven by its strong commitment to equity.

### << Table 1 >>

Such international comparisons of countries – of the type routinely undertaken through PISA– requires strict criteria to ensure one is comparing like-with-like. A long and extensive literature has discussed the importance of translation (e.g. Masri, Baird and Graesser 2016), cross-cultural comparability of the test instruments (e.g. Kankaraš and Moors 2014) and the importance of establishing measurement invariance across countries (Rutkowski and Rutkowski 2016). Yet issues surrounding population definitions, school enrolment rates, sample exclusions, school participation and pupil participation are also important. For instance, if country A systematically excludes many of its low-achieving students (e.g. by deeming them ineligible for the study or because they are absent on the day of the test) then the data and results generated may not be comparable with country B (where a truly representative cross-section of the student population participated). Consequently, comparisons of educational achievement across these two countries will not be meaningful. As this paper will describe, the PISA 2015 data for Canada are likely to suffer from such comparability issues, which clearly have the potential to undermine its apparently strong performance in the PISA study in terms of both equity and efficiency.

This is, of course, not the first paper to discuss issues of population coverage and non-response bias in the context of PISA. Similar concerns have previously been raised about the quality of data available from other countries. For instance, using trends in the PISA scores of Turkey as an example, Spaul (2018) highlights how limitations with the eligibility criteria used in PISA can lead to underestimation of both levels of academic achievement and of educational inequality. A similar analysis conducted by Education Datalab (2017) highlights how issues with differential school enrolment rates across countries can partially explain the strong PISA performance of Vietnam. Pereira (2011) focuses upon changes to the PISA sampling method used in Portugal over time, suggesting that this can help explain recent trends in this country's performance. Furthermore, Micklewright, Schnepf and Skinner (2012) and Durrant and Schnepf (2018) tackle the issue of school and student non-response in England. They find that low-achieving schools, and schools with a large proportion of disadvantaged pupils, are more likely to refuse to take part in PISA, which may bias estimates of educational achievement. Similarly, Jerrim (2013) illustrates how a combination of non-response bias, changes to the target population and test month led policymakers to reach erroneous conclusions about changes to PISA test scores in England.

In this paper, we add to this literature by explaining how data from one of the top PISA performers, Canada, potentially suffers from similar issues. We begin by discussing the rules that the OECD set for inclusion in the PISA study and investigate whether Canada meets each of these criteria. We find that it either fails to meet them, or meets them only marginally, in all cases. It is then demonstrated how this has a significant cumulative impact upon the PISA 2015 Canadian sample. Our empirical analysis then moves on to apply a sensitivity analysis of the Canadian PISA results, focusing upon how it compares to two genuinely high-performing countries (Japan and South Korea) where student exclusion and school/student non-response rates are much lower. These sensitivity analyses estimate the scores that excluded and non-responding students would need to have achieved in order to 'disturb' a finding (Gorard and Gorard 2018); in other words, to make the difference between countries disappear. We argue that this is a more important reflection of uncertainty in the Canadian PISA results than the standard forms of statistical inference (confidence intervals and statistical significance tests) that are routinely reported by the OECD (as it captures different forms of bias rather than just sampling variation alone). Our results illustrate how Canada's PISA results could change in non-trivial ways relative to other countries, under plausible assumptions about how excluded/non-responding students would have performed on the test. It is hence concluded that

the OECD should do more to communicate the uncertainty in PISA results due to sample exclusions and missing data.

The paper now proceeds as follows. Section 2 describes the criteria set by the OECD to try to ensure the PISA data are of high quality and illustrate how the data for Canada performs relative to these benchmarks. Our empirical approach is set out in section 3, with the robustness tests we have conducted around Canada's PISA results provided in section 4. Conclusions and implications for communication and interpretation of the PISA results follow in section 5.

## **2. Key elements of the design of the PISA study**

### Target population and exclusions

The target population of PISA is 15-year-olds attending educational institutions in seventh grade<sup>6</sup> or higher (OECD 2017: chapter 4). This definition has some subtle, but important, implications. In particular, note that young people not enrolled in education (due to, for instance, permanent exclusion, home schooling or having surpassed the minimum school leaving age) are excluded. As previous research has suggested, this definition means many 15-year-olds are excluded from PISA in low and middle-income countries (Spaul 2017; Education Datalab 2017) upwardly biasing results compared to those that would be expected if the target population were all 15-year-olds. Yet, as Table 2 illustrates, it is also not a trivial issue in some OECD countries. In Canada, around 4% of 15-year-olds are excluded from PISA due to non-enrolment at school. This is greater than in the other high-performing OECD nations of Estonia, Finland, Japan and South Korea, where between 98% and 100% of 15-year-olds are enrolled in an educational institution. Yet there are also some other OECD countries where this is clearly a very important issue, most notable Turkey (83% of 15-year-olds are enrolled in schools) and Mexico (62% of 15-year-olds enrolled in school).

### **<< Table 2 >>**

Countries are also allowed to exclude some schools or students from the PISA study. This is usually due to severe Special Educational Needs (SEN) limiting the opportunity for some young people to take part. The criteria set by the OECD is that a maximum of 5% of students can be excluded from PISA within any given country. As noted by Rutkowski and Rutkowski (2016), this maximum of 5% should “*ensure that any distortions in national mean scores due to omitted schools or students would be no more than  $\pm 5$  score points on the PISA scale*”.

---

<sup>6</sup> Based upon the education system in the United States.

Yet the second column of Table 2 illustrates how several countries breached this 5% threshold for exclusions in PISA 2015, but were still included within the study. This includes Canada, which has one of the highest rates of student exclusions (7.5%) – double the OECD average (3.7%). Further inspection of the PISA 2015 national report for Canada (O’Grady et al. 2016) indicates that the excluded students were mainly those with intellectual disabilities (5%), with a further 1.5% of students removed due to limited language skills and 0.5% for physical disabilities. As Table 2 illustrates, the percentage of excluded students differs across countries – with many more excluded in Canada than in some of the other high-performing OECD countries (e.g. Japan and South Korea).<sup>7</sup> This has the potential to bias comparisons between these nations, if certain groups we would not expect to perform well on the PISA test are routinely excluded in some nations (e.g. students with intellectual disabilities in Canada) but not in others (e.g. Japan and South Korea).

### Sample design

PISA utilises a stratified, clustered sample size. The purpose of stratification is to boost the efficiency of the sample (i.e. increase power to narrow confidence intervals) and to ensure there is adequate representation of important sub-groups.

To begin, each country selects a set of ‘explicit stratification’ variables, which should be closely related to PISA achievement<sup>8</sup>. These are essentially used to form different sub-groups (strata). Although these differ across countries, geographic region and school type are common choices. In Canada, province, language and school size are used. Within each of these explicit strata, schools are then ranked by a variable (or set of variables) that are likely to be strongly associated with PISA test scores. This is known as implicit stratification, with the ideal variable being some measure of prior academic achievement amongst pupils within the school. Unfortunately, the implicit stratification variables used in Canada (level of urbanisation, source of school funding and the ISCED level taught) are likely to only be relatively weakly associated with academic achievement. This creates a potential issue if replacement schools need to be targeted, which we discuss below.

---

<sup>7</sup> This in part demonstrates an inherent challenge in cross-national comparative research; the definition, identification and treatment of students with Special Educational Needs is likely to differ significantly across different national settings.

<sup>8</sup> The ideal stratification variables are usually based upon school/student performance in national examinations. Although standardised tests are conducted within Canada, they differ across provinces, potentially explaining why they are not used.

Schools are then randomly selected, with probability proportional to size, from within each of the explicit strata. The minimum number of schools to be selected is 150, although some countries oversample in order to be able to produce sub-group estimates. This is the case in Canada, where results are reported nationally and at the province level. Hence, in total, 1,008 Canadian schools were approached to take part.

Not all schools approached agree to participate in the PISA study. In Canada, 305 (30%) of schools initially approached refused to participate in PISA 2015. In this situation, PISA allows countries to approach two ‘replacement schools’ to take the place of the originally sampled schools. These replacement schools are those that are adjacent to the originally sampled school on the sampling frame. The intuition behind this approach is that the replacement schools will be ‘similar’ to the originally sampled school that they replaced. It is hence a form of ‘near neighbour’ donor imputation; however, this is only effective at reducing non-response bias if the stratification variables used in the sampling are strongly correlated with the outcome of interest (PISA scores). As noted above, this is questionable in the case of Canada, where only weak predictors of academic achievement were used as stratification variables.

After including these replacement schools, a total of 726 Canadian schools (72%) took part. This is much lower than in many other OECD countries (OECD average = 95%), including the other high-performing nations of Estonia (100%), Finland (100%), Japan (99%) and South Korea (99%), as illustrated by Table 1.

#### School response rate criteria

To encourage countries to achieve adequate school response rates, the OECD have set criteria to determine inclusion in the PISA study. This is depicted by Figure 1 and can be summarised as follows:

- Not acceptable (dark-blue region). Less than 65% of the originally sampled schools participated in the study.<sup>9</sup> Countries that fall in this category should be automatically excluded from the PISA results.
- Acceptable 1 (light-blue region). More than 85% of originally sampled schools participated. The PISA sample for countries in this category is assumed to be unbiased and automatically included in the results.

---

<sup>9</sup> One way to think of this criterion is that a maximum of 35% of the PISA data is allowed to be ‘imputed’ from the donor (substitute) schools.



- Acceptable 2 (light-blue region). Between 65% and 85% of originally sampled schools participated, with this percentage increasing substantially<sup>10</sup> once replacement schools are added. The PISA sample for countries in this category is assumed to be unbiased and automatically included in the results.
- Intermediate (blue region). Between 65% and 85% of originally sampled schools participated, with this percentage not increasing sufficiently even when replacement schools are added. Countries that fall into this category are required to undertake a Non-Response Bias Analysis (NRBA) as discussed in the following sub-section.

### << Figure 1 >>

Figure 1 illustrates how, in PISA 2015, four OECD countries (Italy, New Zealand, United States and Canada) fell into the intermediate category where a NRBA was required – with Canada the furthest from the ‘acceptable’ zone after replacement. The data for one other OECD country (the Netherlands) appears in the ‘not acceptable’ zone and, as such, should have been automatically excluded.<sup>11</sup> Nevertheless, all five countries were included in the final PISA 2015 rankings without any explicit warning given about their results.

#### Non-response bias analyses

Survey non-response does not necessarily introduce bias into the sample – it only does so if non-response is not random. One way of investigating whether certain ‘types’ of schools choose not to participate in PISA is to compare the observable characteristics of participating and non-participating schools.<sup>12</sup> Ideally, the variables used to compare responding and non-responding schools should be strongly associated with the outcome of interest (i.e. PISA scores) – such as national measures of school achievement. The intuition behind this approach is that, if responding and non-responding school differ in terms of (for instance) national measures of achievement (e.g. scores on a national mathematics exam) then they are also likely to differ in terms of their likely performance on PISA.

---

<sup>10</sup> A sliding scale is used, where a higher ‘after-replacement’ response rate is required if a lower ‘before-replacement’ response rate was obtained. For instance, a country that had a 65% response rate amongst initially sampled schools would require this to increase to around 95% once replacement schools are added.

<sup>11</sup> The PISA 2015 technical report (OECD 2017) notes that the Netherlands was permitted to conduct a non-response bias analysis (NRBA) rather than face automatic exclusion. Based upon the findings from the NRBA, the OECD deemed the Dutch sample to be of acceptable quality and hence could be included in the PISA 2015 results. We have obtained a copy of the NRBA for New Zealand using freedom of information laws and provide this for interested readers in online Appendix H.

<sup>12</sup> Borrowing terminology from the literature on missing data, such investigations attempt to establish whether the missing data are Missing at Random (MAR) rather than Missing Completely at Random (MCAR).

Unfortunately, few details about what a NRBA entails are published by the OECD. The only details available come from a small section in the technical report (OECD 2017: chapter 14). However, some more description of what is required is provided by some countries where NRBA have previously been conducted, such as Kastberg et al. (2017) for the United States. In summary, the characteristics of responding schools are compared to non-responding schools in terms of a small set of observable characteristics (usually the stratification variables included on the sampling frame plus, occasionally, some additional auxiliary variables).

The key criterion then used to determine evidence of bias seems to be whether or not the difference between participating and non-participating schools, in terms of the observable school-level characteristics available, was statistically significant. If there are no statistically significant differences this seems to be treated as an indication of a lack of bias and, hence, reason for inclusion in the PISA results. Critically, full results from these NRBA are *not* routinely published by the OECD (bar a nebulous paragraph included in the depths of the technical report – OECD 2017: Chapter 14), with the information eventually provided largely left to the discretion of individual countries within their national reports.

The only publicly available details about the NRBA conducted for the PISA 2015 Canada sample are provided within O’Grady et al. (2017: Appendix A). This explains that a NRBA was not conducted for Canada as a whole, but for just three provinces where school response rates were particularly low (Québec, Ontario and Alberta). The report goes on to explain how the characteristics of participating schools was compared to the characteristics of all originally sampled schools (i.e. both participating and non-participating) in terms of school funding source, language, size and recent results in provincial assessments. All analyses were conducted separately for the three provinces with schools (rather than students) being the unit of analysis.

Unfortunately, very little detail is provided about the specific analyses undertaken within the information made publicly available. Likewise, little formal detail is provided about the results (e.g. there are no tables illustrating the results of the NRBA conducted). Instead, it is simply offered that “*non-response analysis revealed no potential bias*” in Ontario and that “*very few statistically significant differences were observed between the non-response adjusted estimates and the population parameter estimates*” in Alberta (O’Grady et al. 2017: Appendix A). On the other hand, in Québec (a reasonably large province that accounts for approximately a fifth of Canada’s population) statistically significant differences were observed and it is reported

that the NRBA “*revealed potential bias*”. Yet, despite this, it was concluded that “*the PISA international consortium judged that the Canadian data overall were of suitable quality to be included fully in the PISA data sets without restrictions*”.

There are, however, at least two significant problems with the current approach to NRBA used within PISA, and the analysis for Canada shows:

1. Only a limited selection of variables is investigated, with the choice of these variables to some extent at the discretion of individual countries. Lack of evidence of a difference on this small handful of variables is then taken as indicating a lack of bias. Yet it could simply reflect that countries had not looked for bias very hard.
2. Whether there is a *statistically significant* difference between responding and non-responding schools seems to be the main criterion for evidence of bias. Yet with only a very limited number of schools (e.g. just 114 in the case of the NRBA conducted in Alberta) – and with the NRBA conducted at school-level – such significance tests are likely to be woefully underpowered. In other words, the small sample size will make it extremely difficult to detect ‘significant’ differences between participating and non-participating schools. In fact, the *magnitude* (effect size or standardised difference) of the differences are of much more use and interest (Imbens & Rubin, 2015). Yet, as in the example of Canada, such crucial information is not generally made publicly available.

The main consequence of the discussion above is that it is not clear that the PISA data for countries that ‘passed’ the NRBA is indeed unbiased. Not enough detail has been published by the OECD and countries themselves (including Canada) to allow proper independent scrutiny of the matter. It is for this reason that we have used freedom of information laws to publish – for the first time – the full school-level NRBA that was conducted for Canada in PISA 2015. This is provided in Appendix F. The limited information available within such NRBA’s are not particularly extensive, and we believe have been designed to support the inclusion of a country’s data wherever possible. Indeed, in Appendix B we document all occasions where a country has been required to complete a NRBA since 2000, noting that on 21 out of 24 occasions (88%) they have come through the process unscathed.

### Pupil response rates

The OECD stipulates that at least 80% of students from within participating schools complete the PISA assessment. Pupils who are selected to participate may end up not participating if

they are absent from school on the day of the test, they (or their parents) do not consent to participation in the study or there were issues with how the study was conducted (e.g. as a computer-based assessment, non-participation could have been the result of computers crashing). In 2015, Canada narrowly met this threshold (81%) but, as Table 2 illustrates, this is one of the lowest rates of student response across the OECD (OECD average = 89%). Yet, as the official student response rate criteria was met, no further evidence is available about the characteristics of these non-participants. This is despite analysis within previous PISA cycles suggesting students who were absent from the PISA test tending to achieve lower scores on Canadian provincial assessments (Knighton, Brochu and Gluszynski 2010) and that low student participation rates might be more problematic (in terms of introducing bias into the sample) than low school participation rates (Durrant and Schnepf 2018). The fact that almost a fifth of sampled Canadian pupils within participating schools did not take the PISA test is therefore a concern. Figure 2 and Appendix D illustrate how this is not a new problem facing the PISA sample for Canada; it has historically had both high-rates of student exclusions and low student participation rates relative to other high-performing countries.

### << Figure 2 >>

#### Weighting for non-response

The PISA database includes a set of response weights which attempt to adjust estimates for non-response (amongst other functions). These weights are only as effective in reducing non-response bias as the variables used in their construction. In an ideal world, they would be both (a) predictive of non-participation and (b) strongly associated with the outcome of interest (PISA scores). This is likely to be the case in, for example, the PISA data for England, where prior school achievement in high-stakes national examinations is included in the non-response adjustment at school level.

Unfortunately, this is unlikely to hold true in the case of Canada (and potentially many other countries as well). Only the implicit and explicit stratification variables are used to adjust for non-participation at the school level. These are level of urbanisation, source of school funding and the ISCED level taught (OECD 2017: chapter 4) which are only likely to be modestly related to PISA outcomes. Then, at the student level, essentially no correction for non-response has been made; as noted within the PISA technical report: *‘in most cases, this student non-response factor reduces to the ratio of the number of students who should have been assessed to the number who were assessed’* (OECD 2017: 122). This means that the fact a fifth of

Canadian students skipped the PISA test (as discussed in the sub-section above) has essentially been ignored in creating the weights. The main implication of this is that the application of the weights supplied as part of the Canadian PISA data should do little to mollify concerns over school and student non-response.

### Summary

Table 3 provides a summary of the combined impact of these issues upon the Canadian PISA sample, with further discussion following in Appendix I. In total, the OECD estimated there to be almost 400,000 15-year-olds in Canada. Yet, through a combination of some young people not being enrolled in school, exclusions from the sample, schools refusing to participate and student absence, the final (weighted) number of students assessed according to the technical report is 210,476 (OECD 2016) – around 53% (see Appendix I for further details about how this figure is calculated). This is quite some distance below the OECD average (77%) and especially far from the other high-performing OECD nations of Finland (91%), Estonia (86%), Japan (91%) and South Korea (90%).

### << Table 3 >>

Returning to Table 1, the other key point to note is that, despite the issues discussed in this section, the mean score for Canada has one of the narrowest confidence intervals. This is, of course, due to the Canadian PISA 2015 sample continuing to have a very large sample size (726 schools and 19,604 students). Yet, as this is the only measure of uncertainty routinely reported within PISA, it would be easy for non-academic audiences to conclude that the Canadian PISA results are amongst the most secure and robust. The reality is, of course, rather different – with uncertainty due to missing data particularly acute. This in turn highlights the need for more sensitivity analyses of the PISA results and for there to be clearer articulation by the OECD about the various different uncertainties that surround them (see Schnepf 2018). We turn to this issue in the following section.

### **3. Sensitivity analyses conducted for the Canadian results**

The issues raised above mean it is important to consider the potential cumulative effect upon Canada's PISA results. Our approach to doing so can be summarised as follows. We assume that students not enrolled in schools, students excluded from the study (due to, for instance, special educational needs), students in schools not-responding and non-responding pupils (within responding schools) have a different distribution of PISA achievement scores than those covered in the PISA data. As we know little about the characteristics about these students

(i.e. we have no micro data about them) we make some assumptions about the distribution of the likely PISA scores for these individuals.

Our starting point is that the average PISA scores of “non-participants” (those 15-year-olds not in school, those who were excluded from the study, those whose school chose not participate and student who chose not to participate) would be lower than those of the students who actually sat the test. For instance, students having intellectual disabilities was the main reason for student exclusion in Canada – a group clearly with low levels of academic achievement. Similarly, previous research has shown how pupil absence is more common amongst low academic achievers (e.g. Gottfried 2009), including in the context of PISA tests conducted in Canada in previous cycles (Knighton, Brochu and Gluszynski 2010). It has also been shown that weaker schools may be more likely to opt out of PISA (Micklewright and Schnepf 2006). We hence believe that our assumption of non-participating students being weaker academically (on average) than participating students is likely to hold.

However, one does not know how much lower non-participating students in Canada would have scored on the PISA assessment. Consequently, our sensitivity analysis essentially investigates how Canada’s PISA scores would change under different assumptions made about the achievement of not enrolled/excluded/non-participating students. We are particularly interested in comparisons with four other high-performing OECD nations (Estonia, Finland, Japan and South Korea) where student exclusions and school/student non-response is much lower (see section 2 for further details). This approach is similar in spirit to investigations of the number needed to disturb a finding (Gorard and Gorard 2016): what would the average score of non-participants need to be in order for Canada and (for example) South to be equally ranked on the PISA test?

This approach is implemented as follows. First, we take the total number of 15-year-olds in Canada from the PISA 2015 technical report (396,966) and divide this into two groups: the number of participants weighted by the final student weight (210,476) and the weighted number of non-participants (186,490)<sup>13</sup>. For the participants, we simply use their PISA scores<sup>14</sup> as recorded within the international database, but deflate the final student weight so that it totals 210,476. Then, for unobserved excluded/non-participating students, we randomly draw

---

<sup>13</sup> From this point forward, we use the terms “participants” and “non-participants” for brevity. We note that “participants” are technically those who actually took the test. But, throughout this section, we use the term “participants” to mean the weighted total of those who actually took the PISA test.

<sup>14</sup> We use ‘scores’ in this context to refer to plausible values. To simplify the process, our analysis focuses upon the first plausible value only.

186,490 scores from a normal distribution, assuming different values for the mean (detailed below), with the standard deviation taking the same value as for participants (e.g. 93 points in the case of reading). The values we use for the mean of this normal distribution correspond to different percentiles of the observed PISA score distribution for Canada. Specifically, we report results when assuming the mean score of excluded/non-participating students is equal to:

- The observed 45<sup>th</sup> percentile (assumed mean of non-participants = 519 in reading).
- The observed 40<sup>th</sup> percentile (assumed mean of non-participants = 507 in reading).
- The observed 35<sup>th</sup> percentile (assumed mean of non-participants = 494 in reading).
- The observed 30<sup>th</sup> percentile (assumed mean of non-participants = 480 in reading).
- The observed 25<sup>th</sup> percentile (assumed mean of non-participants = 465 in reading).
- The observed 20<sup>th</sup> percentile (assumed mean of non-participants = 449 in reading).
- The observed 15<sup>th</sup> percentile (assumed mean of non-participants = 429 in reading).
- The observed 10<sup>th</sup> percentile (assumed mean of non-participants = 402 in reading).

For each of these different scenarios, the randomly drawn scores of the 186,490 unobserved excluded/non-participating students are appended to the database with the observed data for the participants. Key results for Canada – most notably mean scores and inequality as measured by the gap between the 90<sup>th</sup> and 10<sup>th</sup> percentiles – are re-estimated, incorporating the simulated effect of exclusion/non-participation. This, in turn, allows us to consider how Canada’s performance in PISA would change, particularly in comparison to other high-performing countries with much lower exclusion/non-response rates, under a set of different plausible scenarios. We do not argue that any of our alternative scenarios are ‘correct’, but that some of them are at least as plausible as the results used to construct the PISA rankings, while resulting in quite different conclusions. For additional simulations under different assumptions, see Appendix E.

As Canada was the highest-performing OECD country in reading in 2015, we focus upon the robustness of scores within this domain when reporting our results. Appendix C provides analogous results for science and mathematics. Key findings are presented in Table 4. Columns (1) and (2) provide information about the simulated average PISA reading scores of non-participants, while columns (3) to (6) illustrate revised estimates of the mean, 10<sup>th</sup> and 90<sup>th</sup> percentile of PISA reading scores in Canada following the simulated inclusion of the not-enrolled/excluded/non-participating students.

<< **Table 4** >>

Even under the most moderate of our assumed performance distributions of excluded/non-participating students, reading scores in Canada decline dramatically with their simulated inclusion. For instance, if we assume non-participants have only slightly lower levels of achievement than participants (i.e. they would achieve the same score as those at the 40<sup>th</sup> percentile of participants) then the mean score in Canada falls to 517. This is below the average for Finland (526) and now level with South Korea (517) and Japan (516). Hence, the scores of non-participants in Canada do not need to be particularly low (only 507, which is still substantially above the OECD average of 493) to eliminate any difference between Canada and these other high-achieving nations. If we alter the assumption so that non-participants score at (on average) the 30<sup>th</sup> percentile of participants (480 points) the average score for Canada would decline to 505, which is similar to Germany (509), Poland (506) and Slovenia (505). Indeed, under the scenario that non-participants would have achieved an average score of 465 points (equivalent to the 25<sup>th</sup> percentile among participants) the mean score for Canada (497) would be similar to the OECD average (493).

A similar finding emerges with respect to inequality in reading achievement, as measured by the gap between the 90<sup>th</sup> and 10<sup>th</sup> percentiles. Using the data from participants only, inequality in reading achievement in Canada (238 points) is around 11 points lower than in the average OECD country (249 points). Yet, using plausible assumptions about the likely scores of non-participants, there is potentially no difference between Canada and the OECD average at all. For instance, were non-participants in Canada to achieve reading scores that were (on average) around the 30<sup>th</sup> percentile (480 points) then inequality in reading scores in Canada (247) and across the OECD (249) would effectively be the same.<sup>15</sup>

What do these sensitivity analyses imply for how one should interpret the Canadian PISA 2015 results? Our interpretation is that, although it remains perfectly plausible that average reading scores in Canada are above the OECD average (and inequality in achievement below the OECD average), there is not the strength of evidence to classify this country as an ‘education super-power’. Moreover, under reasonable assumptions, average PISA scores fall below those of four other genuinely high-performing OECD countries (Finland, Canada, Japan and South Korea) in all three core PISA domains (see Appendix C for further details regarding science and

---

<sup>15</sup> Note that results referring to inequality are also sensitive to the standard deviation used for non-participants used in the simulations. For reading, we have kept this at 93 points throughout (i.e. the same standard deviation as for participants) as it is not clear whether the reading score of non-participants would be more or less equal.



mathematics). Likewise, it is plausible that inequality in educational achievement in Canada is quite similar to the average across OECD countries.

#### **4. Conclusions**

PISA is an influential large-scale study of 15-year-olds achievement in reading, science and mathematics which is now conducted in more than 70 countries and economies across the world. Results from PISA are widely reported by international media and have had a significant influence upon policymakers (and policymaking). High-performing PISA countries have received much attention, with Finland and a group of high-performing East Asian nations (e.g. Japan, South Korea, Singapore, Hong Kong) being prominent examples. Canada has also performed extremely well on the PISA tests, being lauded for its high average scores and low levels of inequality in achievement. This is striking because – given its similar language, culture, economy, political system and population size – it is a more natural comparator for many Western education systems than some of the high-performing East Asian nations. Canada has hence been held up as an example of a high-quality, equitable education system which leads the Western world (Coughlan 2017).

Yet are Canada's PISA results really as strong as they first seem? This issue has been explored in this paper, considering critical elements of the quality of the Canadian PISA data. We have highlighted how Canada only just meets the minimum threshold the OECD sets for several criteria, with the PISA data for this country suffering from a comparatively high student exclusion rate, low levels of school participation and high rates of student absence. The combination of these factors leads us to believe that there are serious problems with comparing the PISA 2015 data for Canada to other countries. It is hence suggested that additional sensitivity analyses should be applied to the Canadian PISA results, particularly if it is going to be compared to other high-performing OECD countries where exclusion rates are much lower and participation rates are much higher (most notably Estonia, Finland, Japan and South Korea). Our analysis shows that, under plausible scenarios, average PISA scores in Canada drop below those of these other world-leading systems, while inequality in achievement draws close to the OECD average. We hence conclude that, although it remains plausible that educational achievement in Canada is higher than in the average OECD country, there is not the strength of evidence to put in the same class as the world's genuine top performers.

This case study of Canada has wider implications for how the PISA data are reported by the OECD. Three particular issues arise. First, the criteria the OECD sets for a country's inclusion

in the results need to be tightened and how they apply these rules needs to be more transparent. In our opinion, the minimum student response rate should be raised from 80% to 90% and the 5% criterion for student exclusions much more strictly applied. Likewise, given that a school response rate below 65% is labelled ‘unacceptable’, countries with school participation below this level (such as the Netherlands in 2015) should be excluded. We also believe that the OECD should introduce a new criterion for the overall coverage rate to be above some minimum level, in order to avoid the situation that has emerged for Canada (where we believe the cumulative impact of being on the border line for several of the OECD’s rules has led to problems).

Second, Non-Response Bias Analyses (NRBA) need to become much more thorough, open and transparent. We wholeheartedly believe that comparisons of respondents to non-respondents in terms of observable characteristics is a sensible and insightful approach and that such analyses should be undertaken by *all* countries as a matter of course (i.e. not just for countries that fall into the ‘intermediate’ or ‘unacceptable’ response zones). This should be done at both the school and student levels wherever possible, given that non-response amongst either could generate bias in the results. We also advocate an increased focus on the magnitude of differences between participants and non-participants, rather than on statistical significance. However, most importantly, full details and results from the NRBA must be routinely published by the OECD as part of their technical report. The current brief, nebulous summaries provided within the technical report and individual country reports are not fit for purpose. The only way the OECD (and individual countries) will inspire greater confidence in their data is by becoming more transparent about such issues. In an effort to push the OECD in this direction, we have used freedom of information laws to gain access to selected school-level NRBA that have been produced by England (for PISA 2009), New Zealand<sup>16</sup>, the Netherlands and Canada for PISA 2015. We provide a selection of these in online Appendix F-H for readers to inspect. This is (to our knowledge) the first time such evidence has been made available in the public domain, and hence will help readers reach their own conclusion about potential bias in the PISA sample due to school non-response.

Finally, it is unreasonable to expect non-specialist audiences to go through the same level of detail as this paper, or to have the necessary technical understanding (and time) to decipher details that can only be found in the depths of the PISA technical report. Therefore, for each country, the OECD should provide a ‘security rating’ for the quality of the data that is presented

---

<sup>16</sup> We are unable to publish the NRBA for New Zealand that was received, due to a confidentiality agreement that had to be signed in order to obtain the relevant documentation.

in the same table as the headline PISA results. These could be based upon existing information collected (e.g. exclusion rates, school response rates, student response rates, population coverage) and be presented on a simple 1\* to 5\* scale. A similar system is already being used by some organisations devoted to research use amongst the education community, such as the Education Endowment Foundation (EEF) in England, and have generally been well-received.<sup>17</sup> Given the importance and wide interest and influence of PISA, we believe that the introduction of such a system would significantly improve understanding about the uncertainties surrounding the results.

---

<sup>17</sup> The EEF use a 1\* to 5\* security rating system to help non-specialist audiences understand the ‘quality’ of the randomised controlled trials (RCTs) that they conduct.

## References

- Coughlan, S. (2017). How Canada became an education superpower. BBC News Website. Accessed 08/04/2019 from <https://www.bbc.co.uk/news/business-40708421>
- Durrant, G. and Schnepf, S. (2018). Which schools and pupils respond to educational achievement surveys? A focus on the English Programme for International Student Assessment sample. *Journal of the Royal Statistical Society Series A* 181(4): 1,057–1,075.
- Education Datalab. (2017). Why does Vietnam do so well in PISA? An example of why naive interpretation of international rankings is such a bad idea Accessed 08/04/2019 from <https://ffteducationdatalab.org.uk/2017/07/why-does-vietnam-do-so-well-in-pisa-an-example-of-why-naive-interpretation-of-international-rankings-is-such-a-bad-idea/>
- Feniger, Y. and Lefstein, A. (2014) How *not* to reason with PISA data: an ironic investigation, *Journal of Education Policy*, 29:6, 845-855.
- Gorard, S. and Gorard, J. (2016). What to do instead of significance testing? Calculating the ‘number of counterfactual cases needed to disturb a finding’. *International Journal of Social Research Methodology* 19(4): 481-490.
- Gottfried, M. (2009). Excused versus Unexcused: How Student Absences in Elementary School Affect Academic Achievement. *Educational Evaluation and Policy Analysis* 31(4): 392-415.
- Hendrickson, K. A. (2012). Learning from Finland: Formative assessment. *The Mathematics Teacher*, 105(7), 488–489.
- Imbens, G. M. & D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York, NY, Cambridge University Press.
- Jerrim, J. (2015) Why do East Asian children perform so well in PISA? An investigation of Western-born children of East Asian descent, *Oxford Review of Education*, 41:3, 310-333, DOI: [10.1080/03054985.2015.1028525](https://doi.org/10.1080/03054985.2015.1028525)
- Jerrim, J. (2013). The Reliability of Trends over Time in International Education Test Scores: Is the Performance of England's Secondary School Pupils Really in Relative Decline? *Journal of Social Policy* 42(2): 259-279.
- Kankaraš, M., & Moors, G. (2014). Analysis of Cross-Cultural Comparability of PISA 2009 Scores. *Journal of Cross-Cultural Psychology*, 45(3), 381–399.
- Kastberg, D.; Lemanski, N.; Murray, G.; Niemi, E. and Ferraro, S. Technical Report and User Guide for the 2015 Program for International Student Assessment (PISA). Data Files and Database with U.S.-Specific Variables. *National Center for Education Statistics report 095*. Accessed 08/04/2019 from <https://nces.ed.gov/pubs2017/2017095.pdf>
- Knighton, T.; Brochu, P. and Gluszynski, T. (2010). Measuring up. Canadian results of the OECD PISA study. Accessed 08/04/2019 from <https://www.cmec.ca/Publications/Lists/Publications/Attachments/254/PISA2009-can-report.pdf>
- Masri, Y.; Baird, J. and Graesser, A. (2016) Language effects in international testing: the case of PISA 2006 science items, *Assessment in Education: Principles, Policy & Practice* 23(4): 427-455

- Micklewright, J. and Schnepf, S. (2006). Response bias in England in PISA 2000 and 2003. Department for Education and Skills research report 771. Accessed 08/04/2019 from <https://webarchive.nationalarchives.gov.uk/20130323024553/https://www.education.gov.uk/publications/eOrderingDownload/RR771.pdf>
- Micklewright, J., Schnepf, S. V., and Skinner, C. J. (2012) Non-response biases in surveys of school children: the case of the English PISA samples, *Journal of the Royal Statistical Society. Series A (General)*, 915-938.
- OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264266490-en>
- OECD (2017). PISA 2015 technical report. OECD:Paris. Accessed 08/04/2019 from <http://www.oecd.org/pisa/data/2015-technical-report/>
- O'Grady, K.; Deussing, M.; Scerbina, T.; Fung, K. and Muhe, N. (2016). Measuring up: Canadian Results of the OECD PISA Study. 2015 First Results for Canadians Aged 15. Accessed 08/04/2019 from <https://www.cmec.ca/publications/lists/publications/attachments/365/pisa2015-cdnreport-en.pdf>
- Pereira, M. 2011. An analysis of Portuguese students' performance in the OECD programme for international student assessment (PISA). Accessed 08/04/2019 from [https://www.bportugal.pt/sites/default/files/anexos/papers/ab201111\\_e.pdf](https://www.bportugal.pt/sites/default/files/anexos/papers/ab201111_e.pdf)
- Raffe, D. (2011). Policy borrowing or policy learning? How (not) to improve education systems. CES Briefing, 57. Accessed 24/04/2019 from <http://www.ces.ed.ac.uk/PDF%20Files/Brief057.pdf>.
- Rutkowski, L., & Rutkowski, D. (2016). A Call for a More Measured Approach to Reporting and Interpreting PISA Results. *Educational Researcher*, 45(4), 252–257.
- Schnepf, S. (2018). "Insights into survey errors of large scale educational achievement surveys" Working Papers 2018-05, Joint Research Centre, European Commission (Ispra site).
- Spaull, N. (2018) Who makes it into PISA? *Understanding the impact of PISA sample eligibility using Turkey as a case study (PISA 2003–PISA 2012)*, *Assessment in Education: Principles, Policy & Practice*, DOI: [10.1080/0969594X.2018.1504742](https://doi.org/10.1080/0969594X.2018.1504742)
- Takayama, K., Waldow, F., & Sung, Y.-K. (2013). Finland Has it All? Examining the Media Accentuation of 'Finnish Education' in Australia, Germany and South Korea. *Research in Comparative and International Education*, 8(3), 307–325.

**Table 1. PISA 2015 reading scores compared across OECD countries**

<b>Country</b>	<b>Confidence</b>				
	<b>Mean</b>	<b>interval</b>	<b>P10</b>	<b>P90</b>	<b>P90-P10</b>
Canada	527	522-531	404	642	238
Finland	526	521-531	401	640	239
Ireland	521	516-526	406	629	223
Estonia	519	515-523	404	630	226
South Korea	517	511-524	386	637	251
Japan	516	510-522	391	629	238
Norway	513	508-518	381	636	255
New Zealand	509	505-514	368	643	275
Germany	509	503-515	375	634	259
Poland	506	501-511	386	617	231
Slovenia	505	502-508	382	621	239
Netherlands	503	498-508	368	630	262
Australia	503	500-506	365	631	266
Sweden	500	493-507	364	625	261
Denmark	500	495-505	383	608	225
France	499	494-504	344	637	293
Belgium	499	494-503	360	623	263
Portugal	498	493-503	374	614	240
UK	498	493-503	372	621	249
USA	497	490-504	364	624	260
Spain	496	491-500	379	603	224
Switzerland	492	486-498	360	614	254
Latvia	488	484-491	374	595	221
Czech Republic	487	482-492	352	614	262
Austria	485	479-490	347	611	264
Italy	485	480-490	359	602	243
Iceland	482	478-485	350	607	257
Luxemburg	481	479-484	336	616	280
Israel	479	472-486	326	621	295
Hungary	470	464-475	338	593	255
Greece	467	459-476	334	590	256
Chile	459	454-464	342	572	230
Slovak Republic	453	447-458	312	583	271
Turkey	428	421-436	322	535	213
Mexico	423	418-428	321	523	202

**Table 2. Exclusions, school, pupil and overall participation rates in PISA 2015 amongst OECD countries**

	% of 15-year-olds enrolled in school	% Excluded	School response % before replacement	School response % after replacement	Student participation rate (%)
South Korea	100	0.9	99 (100)	99 (100)	99
Mexico	62	0.9	95 (95)	97 (98)	95
Turkey	83	1.1	90 (97)	96 (99)	95
Portugal	91	1.3	84 (86)	94 (95)	82
Belgium	99	1.7	81 (83)	95 (95)	91
Chile	96	1.8	89 (92)	97 (99)	94
Greece	100	1.9	90 (92)	99 (98)	94
Austria	94	2.1	99 (100)	99 (100)	71
Germany	100	2.1	96 (96)	99 (99)	93
Czech Republic	100	2.4	99 (98)	99 (98)	89
Japan	98	2.4	95 (94)	99 (99)	97
Poland	95	2.4	89 (88)	99 (99)	87
Finland	100	2.8	99 (100)	100 (100)	93
Ireland	98	3.1	99 (99)	99 (99)	89
Slovenia	98	3.1	95 (98)	95 (98)	91
Spain	94	3.2	99 (99)	100 (100)	89
Hungary	95	3.3	92 (93)	97 (99)	92
USA	95	3.3	67 (67)	83 (83)	90
Israel	95	3.4	89 (91)	91 (93)	91
Iceland	99	3.6	95 (99)	95 (99)	86
Netherlands	100	3.7	62 (63)	92 (93)	85
Italy	92	3.8	78 (74)	87 (88)	89
France	96	4.2	91 (91)	95 (94)	88
Slovak Republic	99	4.3	92 (93)	98 (99)	91
Switzerland	98	4.4	91 (93)	97 (98)	93
Denmark	99	5.0	88 (90)	89 (92)	87
Latvia	98	5.1	86 (86)	92 (93)	90
Australia	100	5.3	91 (94)	92 (95)	81
Estonia	98	5.5	100 (100)	100 (100)	93
Sweden	99	5.7	99 (100)	99 (100)	91
New Zealand	95	6.5	69 (71)	84 (85)	80
Norway	100	6.8	95 (95)	95 (95)	91
Canada	96	7.5	70 (75)	72 (79)	81
Luxemburg	96	8.2	100 (100)	100 (100)	96
UK	100	8.2	85 (84)	91 (93)	88
<b>OECD average</b>	<b>96</b>	<b>3.7</b>	<b>90 (91)</b>	<b>95 (96)</b>	<b>89</b>
<b>OECD median</b>	<b>98</b>	<b>3.3</b>	<b>91 (93)</b>	<b>97 (98)</b>	<b>91</b>

Notes: Both weighted and unweighted school response rates are provided (the former appear in brackets).

**Table 3. Estimated size of the 15-year-old population and the (weighted) number of students assessed across OECD countries**

	<b>15-year-olds in population</b>	<b>15-year-olds assessed (weighted)</b>	<b>Weighted students assessed / population</b>
Japan	1,201,615	1,096,193	91%
Finland	58,526	53,198	91%
South Korea	620,687	559,121	90%
Germany	774,149	685,972	89%
Switzerland	85,495	74,465	87%
Estonia	11,676	10,088	86%
Greece	105,530	89,588	85%
Ireland	61,234	51,947	85%
Sweden	97,749	82,582	84%
Luxemburg	6,327	5,299	84%
Slovenia	18,078	15,072	83%
Hungary	94,515	77,212	82%
Slovak Republic	55,674	45,357	81%
Czech Republic	90,391	73,386	81%
Spain	440,084	356,509	81%
Belgium	123,630	99,760	81%
Iceland	4,250	3,365	79%
Poland	380,366	300,617	79%
Israel	124,852	98,572	79%
Norway	63,642	50,163	79%
France	807,867	611,563	76%
Netherlands	201,670	152,346	76%
Latvia	17,255	12,799	74%
Chile	255,440	189,206	74%
Denmark	68,174	49,732	73%
Australia	282,888	204,763	72%
Austria	88,013	63,660	72%
UK	747,593	517,426	69%
Portugal	110,939	75,391	68%
Turkey	1,324,089	874,609	66%
USA	4,220,325	2,629,707	62%
New Zealand	60,162	36,860	61%
Italy	616,761	377,011	61%
Mexico	2,257,399	1,290,435	57%
Canada	396,966	210,476	53%
<b>OECD average</b>	<b>453,543</b>	<b>317,841</b>	<b>77%</b>
<b>OECD median</b>	<b>110,939</b>	<b>89,588</b>	<b>79%</b>

Notes: Source of data is Table 11.1 ‘all 15-year-olds’ and Table 11.7 ‘number of students assessed’ from the PISA 2015 technical report. Canada highlighted along with four other high-performing OECD countries. See Appendix I for further discussion of the derivation of the percentages presented in the final column.

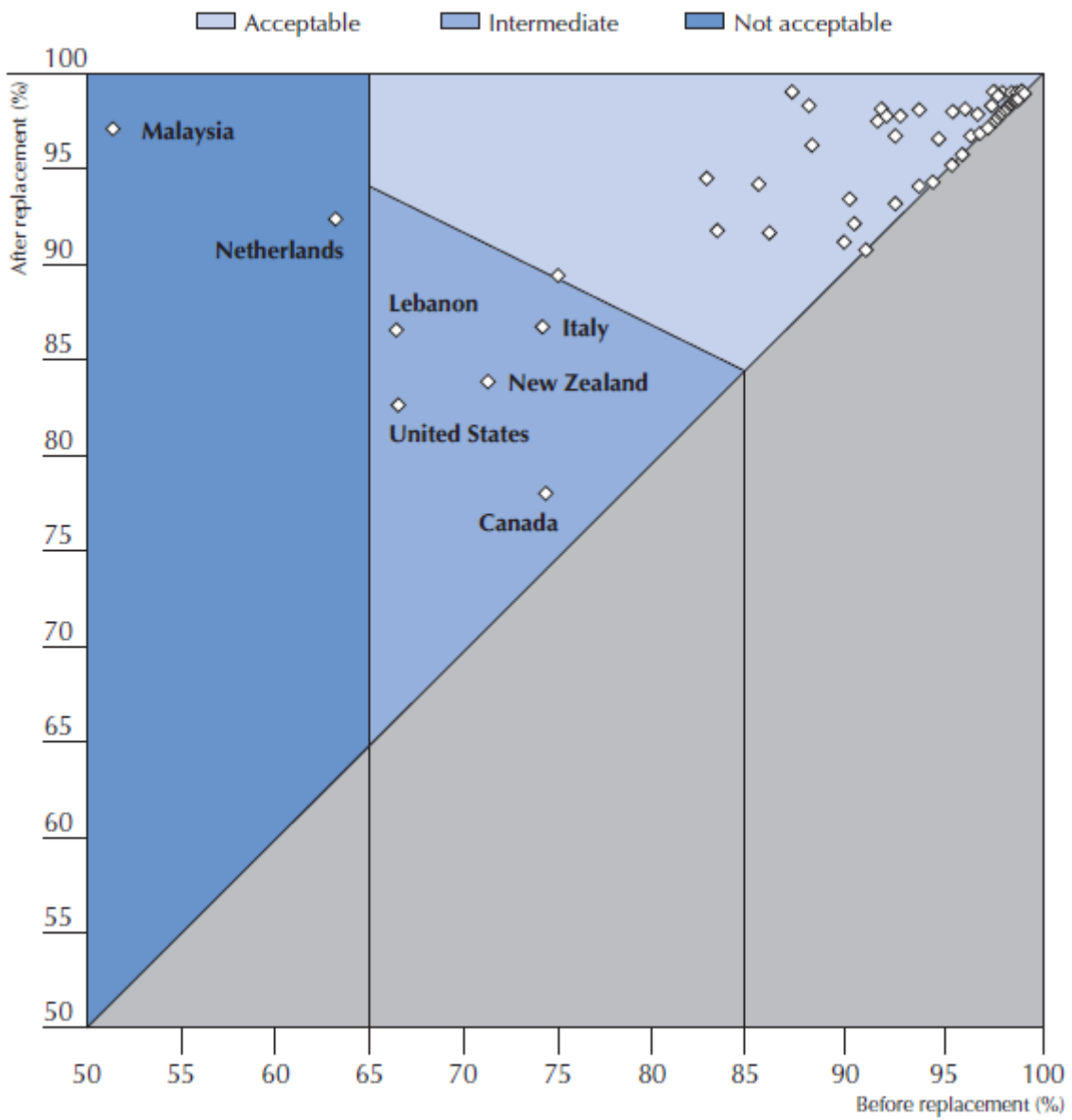


**Table 4. Simulated PISA reading scores under differing assumptions about the likely average scores of non-participants**

1. Non-participants achievement as a percentile of observed Canadian distribution	2. Assumed average score of non-participants	Revised PISA scores			
		3. Mean	4. P10	5. P90	6. P90-P10
Original	527	527	404	642	238
45	519	523	402	640	238
40	507	517	395	635	240
35	494	512	388	631	243
30	480	505	380	626	247
25	465	497	370	622	252
20	449	490	358	619	260
15	429	481	343	615	272
10	402	468	321	612	291

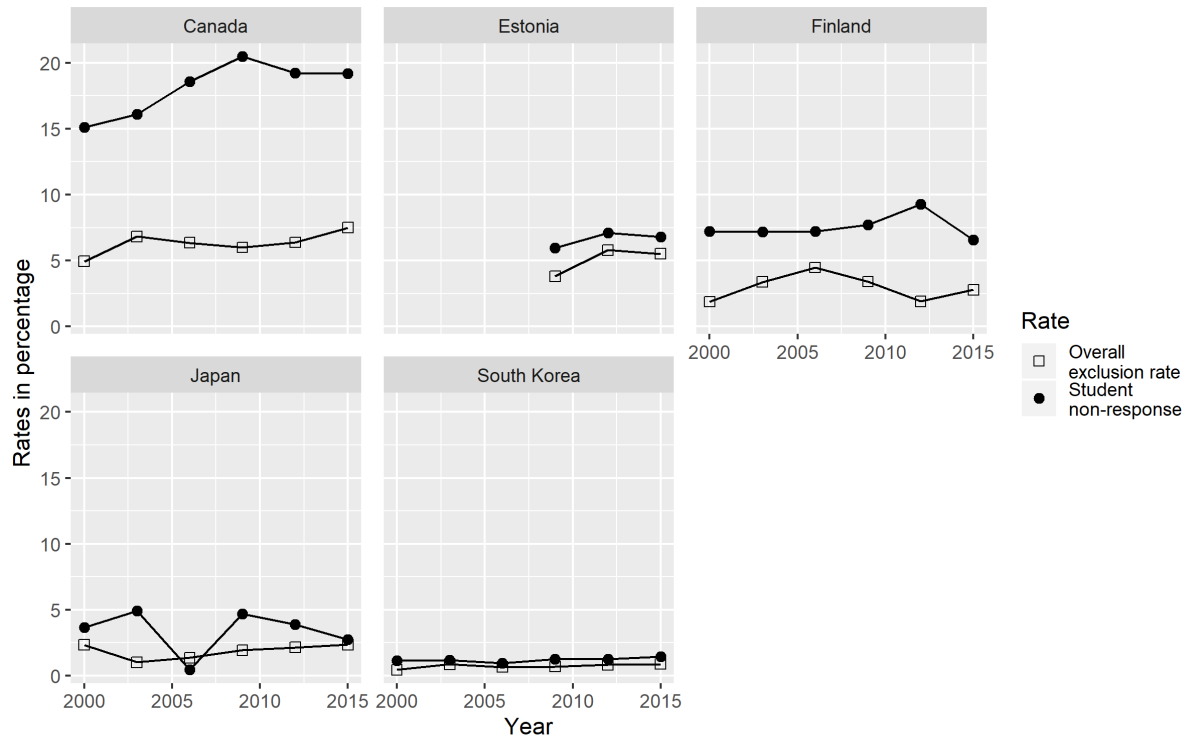
Notes: Column 1 refers to the percentile of the Canadian PISA reading score distribution that the average non-participant would have achieved had they sat the test (column 2 illustrate the actual PISA score this corresponds to). Columns 3 to 6 then illustrate how PISA reading scores for Canada would change under the different scenarios.

**Figure 1. School-response rates in PISA 2015.**



Source: PISA (2015) technical report. Figure 14.1.

**Figure 2. Student exclusion and non-response rates over time in selected high-performing countries**



**Appendix A. Number of eligible and (weighted) number of participating students in PISA by Canadian province**

<b>Province</b>	<b>Eligible students</b>	<b>Participating students</b>	<b>%</b>
Newfoundland and Labrador	5,579	3,959	71%
Prince Edward Island	1,625	1,164	72%
Nova Scotia	9,594	6,882	72%
New Brunswick	8,068	5,488	68%
Quebec	72,433	28,941	40%
Ontario	152,406	92,974	61%
Manitoba	13,554	9,191	68%
Saskatchewan	12,851	8,637	67%
Alberta	42,814	23,559	55%
British Columbia	47,475	29,678	63%
<b>Canada total</b>	<b>366,399</b>	<b>210,473</b>	<b>57%</b>

Notes: Source is Table A1.a and A2 of PISA 2015 national Canadian report (<https://www.cmec.ca/publications/lists/publications/attachments/365/pisa2015-cdnreport-en.pdf>). Figures for Canada different from Table 3 as we are now considering eligible 15-year-old students (rather than all students) and there being some discrepancies between the figures reported within the Canadian national report and OECD international technical report.

**Appendix B. Countries having to do a non-response bias analysis in PISA since 2000**

<b>Country/Year</b>	<b>% school response</b>	<b>Included in report?</b>
<b>2000</b>		
Netherlands	27%	Excluded
USA	56%	Included
UK	61%	Included
Belgium	69%	Included
New Zealand	77%	Included
Poland	79%	Included
<b>2003</b>		
UK	64%	Excluded
USA	65%	Included
Canada	80%	Included
<b>2006</b>		
United States	69%	Included
Scotland	64%	Included
United Kingdom	76%	Included
<b>2009</b>		
Panama	84%	Included
United Kingdom	70%	Included
United States	67%	Included
<b>2012</b>		
Netherlands	74%	Included
United States	67%	Included
<b>2015</b>		
Malaysia	51%	Excluded
Netherlands	63%	Included
Lebanon	67%	Included
United States	67%	Included
Canada	74%	Included
New Zealand	71%	Included
Italy	74%	Included

Notes: School response rate reported before replacement.

## Appendix C. Sensitivity analysis results for science and mathematics

### Science

Non-participants achievement as a percentile of observed Canadian distribution	Assumed average PISA score of non-participants	Revised PISA scores			
		Mean	P10	P90	P90-P10
Original	-	528	404	644	240
45	518	523	401	640	239
40	506	518	394	635	240
35	493	512	388	631	243
30	480	505	379	627	248
25	464	498	370	623	253
20	447	490	358	620	261
15	426	480	343	616	273
10	399	467	320	614	293

Note: Mean science score in Finland = 531; Estonia = 534; Japan = 538; South Korea = 516; OECD average = 493.

### Mathematics

Non-participants achievement as a percentile of observed Canadian distribution	Assumed average PISA score of non-participants	Revised PISA scores			
		Mean	P10	P90	P90-P10
Original	-	516	400	627	227
45	505	511	397	623	225
40	494	506	391	618	227
35	482	500	385	614	230
30	469	494	377	611	234
25	455	487	368	606	238
20	440	480	357	603	246
15	422	472	344	599	255
10	400	461	325	597	272

Note: Mean mathematics score in Finland = 511; Estonia = 520; Japan = 532; South Korea = 524; OECD average = 490.

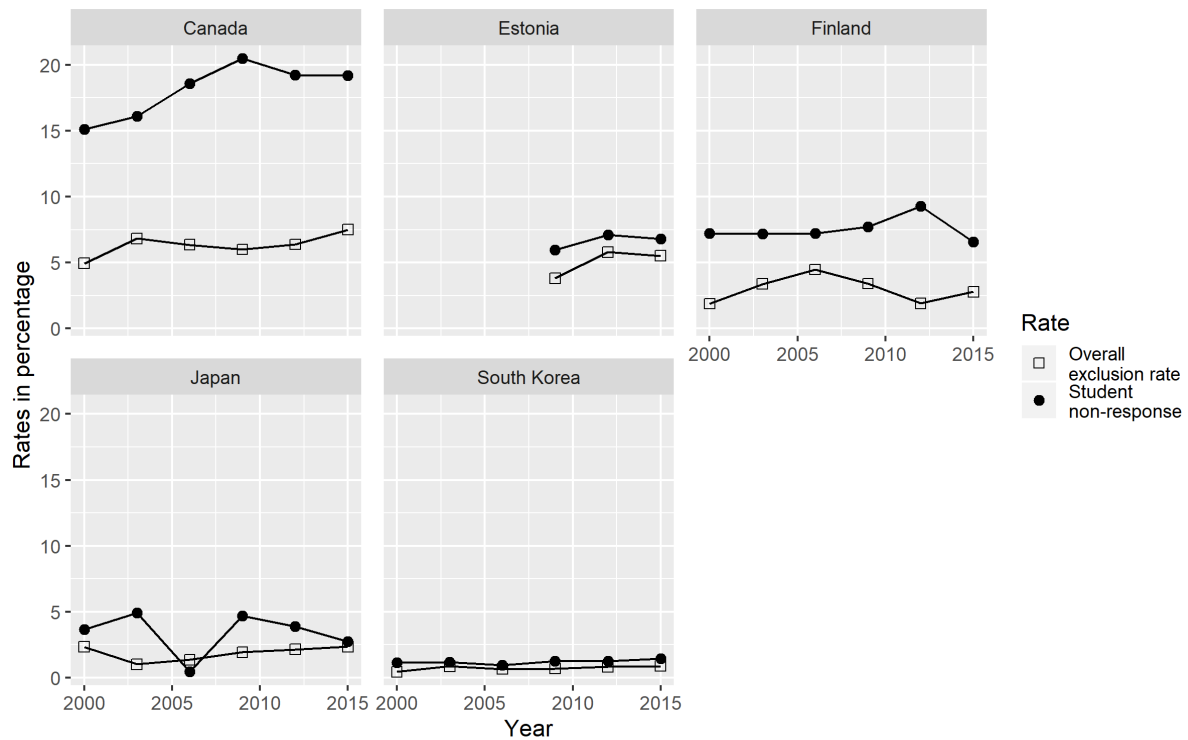
## **Appendix D. Participation and exclusion rates over time**

As explained in Section 2, the participation and exclusion rates differ substantially between countries in PISA 2015 (for more details, see also Table 2). This appendix details key changes in the participation rates of Canada and four other high performing countries (Estonia, Finland, Japan and South Korea) over time. The following numbers and graphs are based on the official figures reported in the PISA technical reports (see OECD 2017).

Student non-participation can occur for two reasons. First, exclusion before the test. Even if students are eligible and form part of the target population, they can be excluded from taking the PISA test by their school (e.g. their school deems them to have insufficient language skills or are intellectually disabled). Countries are also allowed to exclude a small number of schools if they are, for instance, inaccessible or in a very remote location. In Figure D1 we illustrate how student exclusion rates compare across five high-performing countries over the 2000-2015 PISA cycles. The rates are consistently higher in Canada than some other high-performing countries (e.g. Japan and South Korea). As argued elsewhere, this could introduce bias into comparisons of educational achievement across these countries if less able students (e.g. those with learning disabilities) were more likely to be excluded.

The second reason for non-participation in PISA is pupil absence on the day of the test. Figure D1 also illustrates how in Canada up to a fifth of the students were absent from the PISA assessment, which is a much larger proportion than in other high-performing countries. For instance, in South Korea the non-attendance rate was less than 2% across all six PISA cycles. As absence from school is likely to be correlated with student performance (e.g. if low-performing students are more likely to play truant), variations in absence-related non-participation will also lead to bias in comparisons between countries.

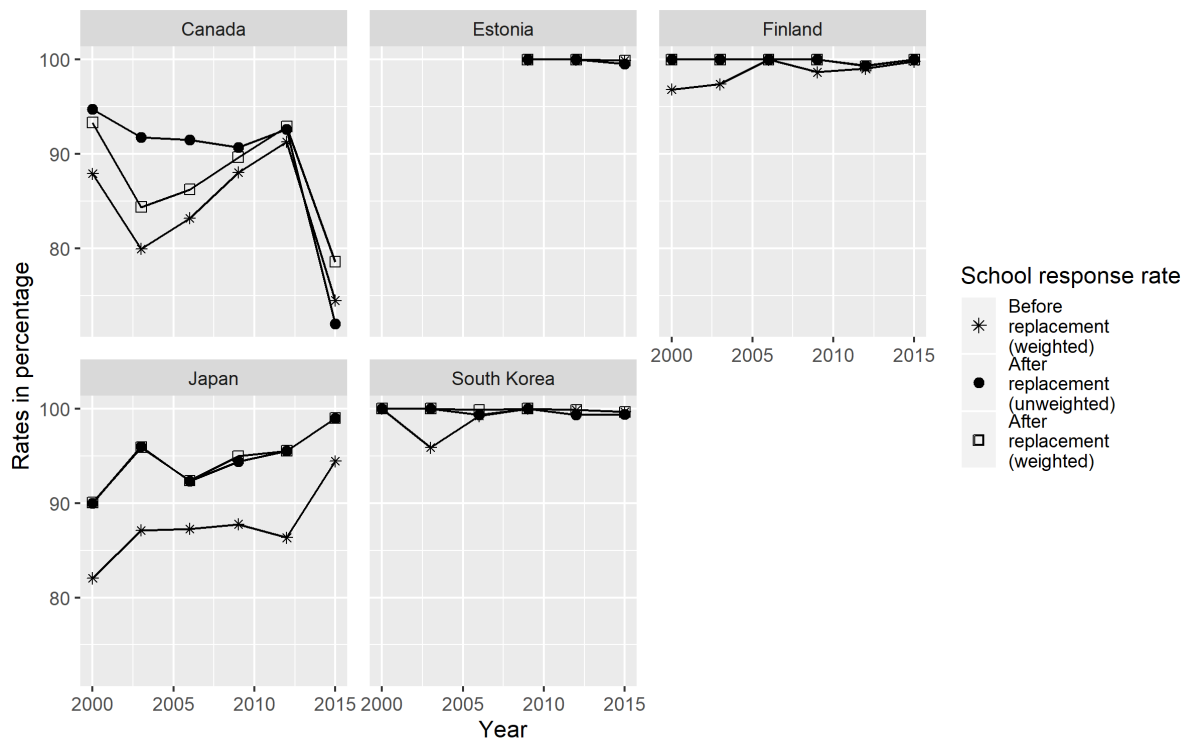
**Figure D1. Student exclusion and non-response rates over time in selected high-performing countries**



In PISA, schools are first sampled and then asked to participate. They may, of course, decline to participate. Indeed, school participation may depend upon the staff’s attitude towards PISA and that of their national government. Figure D2 illustrates how Estonia, Finland and South Korea generally have high school response rates (after allowing replacement schools, they often reach close to 100% response rates). In Japan, the response rate before replacement has usually been between 80% and 90%, with this increasing to between 90% and 100% once replacement is allowed; the after-replacement school response rate has also increased over time in Japan, from 90% in 2000 to 99% in 2015. The situation is different for Canada. The unweighted school response rate fell dramatically, from around 90% in PISA cycles between 2000-2012 to a low of 72% in 2015.



**Figure D2. School response rates over time in selected high-performing countries**



The findings presented in Figure D2 warrant further scrutiny of possible selection issues in the PISA 2015 school sample for Canada. In particular, it is important to consider whether there has been a systematic bias in school-level characteristics due to the high school non-response rates in PISA 2015.

Unfortunately, PISA does not provide any information about the characteristics of non-responding schools. The only exception to this is non-response bias analyses, which are not routinely disclosed. As a proxy, we examine whether the proportion of schools with select background characteristics changed between 2009 and 2015. We focus on three variables, each measured at the school level: the school average of the PISA Economic, Social and Cultural Status (ESCS) index, the proportion of immigrants in the school, and the percentage of students who have repeated at least one grade.

With respect to school-average of the ESCS index, Table D1 suggests that it is higher for Canada in PISA 2015 than 2012, but similar to 2009. Results with respect to this variable are hence inconclusive.

**Table D1. Distribution of the school-average of the ESCS index**

	Median	Mean	SD
2009	0.439	0.435	0.428
2012	0.369	0.374	0.425
2015	0.439	0.436	0.406

Notes: ESCS index average is calculated at the school-level for Canada's participating schools. Number of schools in 2009 is 978 (including three without valid ESCS index), in 2012 it is 885, and in 2015 it is 759 (including three without valid ESCS index).

Table D2 presents a similar analysis with respect to the percentage of first- and second-generation immigrants in each school. There has been an increase in the proportion of immigrant students in the Canadian PISA sample over this six-year period. For instance, the average school had around seven percent of first-generation immigrant students in 2009, increasing to around 10 percent in PISA 2015.

**Table D2. School-average percentage of immigrant students in the Canadian PISA sample**

	PISA 2009	PISA 2012	PISA 2015
Second generation	8.3%	8.9%	9.0%
First generation	6.9%	8.3%	10.2%

Notes: Percentage of immigrant students is calculated at the school-level for Canada's participating schools. Immigrant students are defined as in the PISA 2015 international report: 'First generation' is defined as foreign-born students with foreign-born parents and 'second generation' is defined as students born in the country with both parents foreign-born (OECD, 2016: p.243). We use the pre-computed variable 'IMMIG' from the PISA data sets. Number of schools in 2009 is 978, in 2012 it is 885, and in 2015 it is 759.

Another way to assess the school's situation is the proportion of students who have repeated a grade. Table D3 illustrates that the average school's percentage of repeater students decreased substantially from PISA 2009 and 2012 to 2015, from around nine percent to around six percent. As this variable is likely to be strongly related to academic achievement (repeater students are, by definition, low academic achievers) this suggests more bias could be introduced into the Canadian PISA 2015 data than in previous cycles.

**Table D3. School-average percentage of students who have repeated a grade in the Canadian PISA sample**

	PISA 2009	PISA 2012	PISA 2015
Repeated at least one grade	9.4%	9.0%	6.1%

Notes: 'Repeated at least one grade' is captured by the pre-computed variable 'REPEAT' in 2012 and 2015 and by the variables 'ST07Q01' (repeated a grade in ISCED 1), 'ST07Q02' (repeated a grade in ISCED 2) and 'ST07Q03' (repeated a grade in ISCED 3) in 2009, which are used to compute the equivalent of 'REPEAT'. It measures if at least one grade was repeated at any time during ISCED 1, 2 and 3. Number of schools in 2009 is 978 (three without valid answers), in 2012 it is 885 (one without valid answers), and in 2015 it is 759 (three without valid answers).

## Appendix E. Additional Simulations

### *E1. Introduction*

In this appendix we show how Canada's 2015 PISA scores would change by conducting a series of simulations. In these simulations, we make some assumptions about how the pupils who did not take the test, for the reasons previously discussed, would have performed if they had. The goal is to show how Canada's final PISA ranking changes if these students had participated in the test.

Recall that bias in the PISA scores could potentially arise from three main sources:

1. Exclusion rate
2. School response rate (before and after replacement)
3. Student participation rate

The PISA Technical Report shows how some variables are used to adjust for imbalance using the student weights provided with the PISA data (OECD 2017). However, this re-weighting is performed only on a few *ex ante* observable variables, i.e. gender and grade when accounting for student non-response. Certain subgroups of students might be more likely to participate in the PISA study than others. For instance, high performing students might be more likely to participate than low performing students. This is not taken into account when the final student weights are computed. If a country has high non-response rates, the resulting bias might alter the final PISA scores significantly.

For the purposes of the simulations in this appendix, it is helpful to think of the Canadian student population as being in two groups: students who participated in the PISA study and those who did not. For those who took part in the 2015 PISA study, the final reading score is observed. For those students who did not take part, we do not observe a final reading score, so we impute a mean score for this group based on a series of assumptions outlined below. An overall, simulated Canadian reading score, combining the observed and the imputed scores, is then obtained using the adjusted student weights.

The simulations described in this section address two key questions:

1. What happens to the Canadian PISA scores, if we assume certain groups to be more likely to take the test than others?

2. How strong would those imbalances need to be in order to observe a specific drop in the Canadian PISA score?

All simulations assume 53% of Canadian 15-year-olds are covered within the PISA 2015 study (as explained in the main body of the paper). This means that 47% of the data are simulated according to assumptions made on potential selection bias.

The following sections introduce simulations based on selection on ability (section E2) and selection on background variables such as socio-economic status and class repetition (section E3). For those simulations, the original PISA student weights are adjusted according to some assumptions on the selection bias. Details about the simulation method can be found in section E4.

### *E2. Selection on Ability*

Table E1 shows how different combinations of an assumed ability level below which students are more likely not to participate in the PISA study. Depending on the assumptions made, the Canadian reading score could drop between four and 22 points. The most important determinant of the magnitude of the drop appears to be the strength of selection bias rather than the ability level below which it is assumed to be observed.

**Table E1. Changes in mean scores under different assumptions about selection on ability into PISA sample**

		Ability level below which selection bias occurs						
		400	420	440	460	480	500	527
Strength of selection bias	Original	527	527	527	527	527	527	527
	1.25	523	522	521	521	520	520	521
	1.50	519	517	516	515	515	515	516
	1.75	513	511	510	510	510	511	513
	2.00	506	505	505	505	506	507	510

Notes: Data from Canada's PISA 2015. In each column a different assumption is made on below which ability level selection bias occurs, from below ability of 400 (around 10% of Canadian students) up to below the average value of the reading score in Canada of 527. Each row assumes a different strength of selection bias from none to twice as likely not to be covered by the PISA study.

Introducing selection bias has an effect on the simulated quantile distance P90-P10 (see Table E2). The inequality in reading scores remains at the original level of 238 or increases up to values around 271 for different sets of assumptions. Simulations of P90-P10 are very sensitive to both changes in the assumption of ability level below which selection bias occurs and the strength of the selection bias. Under the assumption that students of low ability level are subject to selection bias, the simulation of the inequality in scores shows the largest change. The inequality in scores does not increase if selection bias is assumed for all students performing below average, regardless of the strength of selection bias.

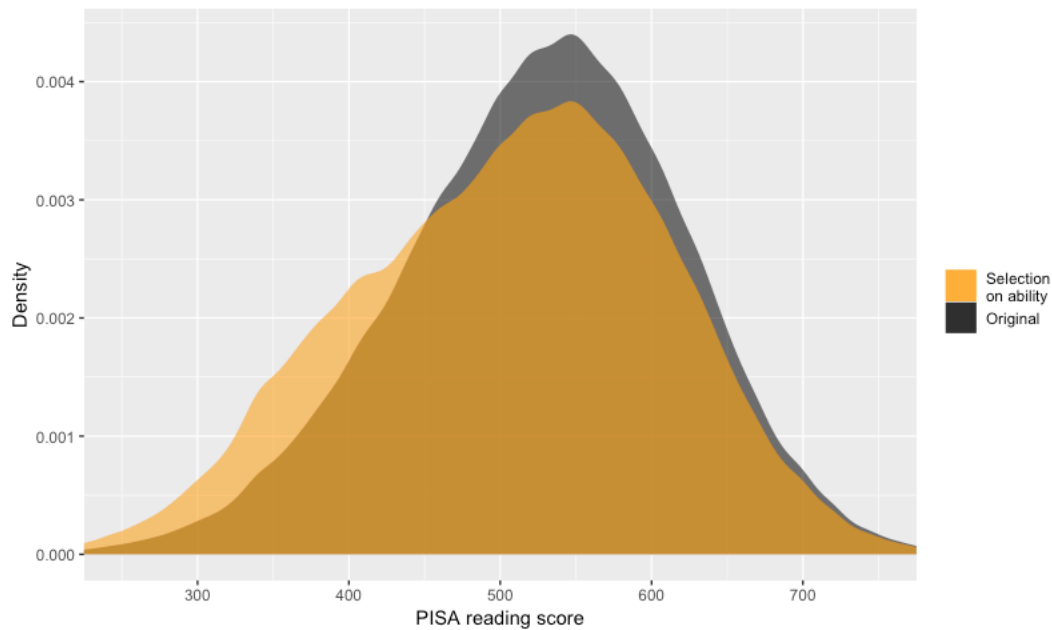
**Table E2. Changes in quantile distance P90 - P10 under different assumptions about selection on ability into PISA sample**

		Ability level below which selection bias occurs						
		400	420	440	460	480	500	527
Strength of selection bias	Original	238	238	238	238	238	238	238
	1.25	245	245	245	244	243	242	239
	1.50	254	254	252	250	246	243	239
	1.75	263	260	257	253	248	244	238
	2.00	271	266	260	255	249	244	238

Notes: Data from Canada's PISA 2015. Columns represent different ability levels below which a selection bias is assumed and the rows show different assumptions on the strengths of the selection bias. See Table E1 for further details.

Taking an extreme case, if we assume that students with low ability (PISA score of all three subjects lower than 400, i.e. lowest 10% of the distribution) to be twice as likely to not participate in the PISA study than all other students, the distribution of scores changes as presented in Figure E1. Under these assumptions, the Canadian PISA reading score is reduced from the originally reported 527 to 506, moving Canada from the top position to the same score as Poland, and just behind Germany. Furthermore, the quantile difference P90-P10 increases to 271, one of the largest among high performing countries. Even if we assume that there are no issues arising from a high exclusion rate and low school response rate, and only focus on assumed bias arising from the 20% of Canadian pupils that do not show up on the day of the test, the Canadian reading score still drops to 523. If we make a stronger assumption, that lowest ability students are four times as likely to stay at home than all other students, we see a decline in PISA score of 12 points, causing Canada to fall behind Japan.

**Figure E1. How selection on low ability students could affect the distribution of Canada's 2015 PISA scores**



Notes: Kernel density plot of the distribution of the 2015 Canadian PISA reading scores. Based on simulation in which strength of selection bias assumed to be twp and occurs for students with an ability level below 400 (corresponding with bottom left cell in Tables E1 and E2).

### *E3. Selection on Background Variables*

#### **Socio-economic status**

The following tables show the effects of selection on socio-economic status on the average reading score and the quantile difference P90-P10. As Table E3 shows, under all assumptions made in our simulations the Canadian reading score drops by up to eight points. This is the case if we assume a strong selection bias (twice as likely not to participate) on the lowest 20 to 30% in terms of SES. Other combinations result in a smaller shift in Canada's average PISA reading score.

Similarly, the quantile distance only increases by up to four points if a strong selection bias is assumed and remains constant for most cases covered by the simulations (Table E4).

**Table E3. Simulated mean PISA reading score under different assumptions about selection on SES into PISA sample**

		Assumed quantile of ESCS variable, below which selection bias occurs								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
Strength of selection	Original	527	527	527	527	527	527	527	527	527
	1.25	526	525	525	525	525	525	525	526	526
	1.50	524	523	523	523	523	523	524	525	526
	1.75	522	521	521	521	522	523	523	525	526
	2.00	520	519	519	520	521	522	523	524	526

Notes: Data from Canada’s PISA 2015. Cells contain simulated mean PISA reading scores under the reported assumptions. Each column reports these simulated means with a different quantile of socio-economic status (ESCS variable) cut off below which selection bias is generated in simulation; each row reports these simulated means with differing strengths of selection bias below the column’s ESCS cut off.

**Table E4. Simulated P90-P10 quantile distance in PISA reading score under different assumptions about selection on SES into PISA sample**

		Assumed quantile of ESCS variable, below which selection bias occurs								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
Strength of selection	Original	238	238	238	238	238	238	238	238	238
	1.25	239	239	239	239	239	238	238	238	238
	1.50	240	240	240	239	239	238	238	238	238
	1.75	241	241	240	240	239	238	238	238	238
	2.00	242	241	240	239	239	238	238	238	238

Notes: Data from Canada’s PISA 2015. Cells contain simulated P90-P10 quantile distance in PISA reading scores under the reported assumptions. Each column reports these quantile distances with a different quantile of socio-economic status (ESCS variable) cut off below which selection bias is generated in simulation; each row reports these simulated quantile differences with differing strengths of selection bias below the column’s ESCS cut off.

### Class repetition

Table E5 shows how different assumptions on selection bias on class repetition affect the Canadian PISA reading scores. Around 5% of Canadian students in the PISA data report having repeated a class, with these students being (by definition) lower-achievers. If those students are more likely to not be covered by the PISA study, this affects Canadian reading scores as well as the quantile distance P90-P10 as shown in Table E4. For instance, if we assume students who have repeated a class to be twice as likely not to be covered by PISA, the Canadian reading score drops by 10 points.



**Table E5. Simulated means and quantile distances under different assumptions about selection on repeating a class**

		Mean Score	P90 – P10
Strength of selection bias	<b>Original</b>	527	238
	<b>1.25</b>	525	240
	<b>1.50</b>	523	242
	<b>1.75</b>	521	246
	<b>2.00</b>	517	250

Notes: Data from Canada’s PISA 2015. The cells report simulated mean scores (first column) and the quantile distance P90-P10 (second column) under the assumption reported in the row i.e. how much more likely students who previously repeated a class are not to be covered by the PISA study.

#### *E4. Technical details of Simulations*

Starting from the observed sample and the student participation rate  $\pi$ , it is assumed that there is a certain degree of selection bias in participation rates. Students with certain background characteristics or ability levels might be more or less likely to stay at home at the day of test than others. Similarly, students with low ability levels might be more likely to fall under certain exclusion criteria.

The method used to simulate potential corrections to this selection bias is based on changing the original student weights in the PISA data: students’ weights with characteristics that we assume to be underrepresented in the PISA data are inflated compared to their peers that are overrepresented. Those new weights can then be used to compute the mean reading score and certain distribution characteristics such as the difference between the tenth and ninetieth percentiles (“P90 - P10”).

For the ease of communicating the simulation results, the aim is to produce statements such as “if students with ability level lower than 400 are 50% more likely to not be covered by the PISA study, the average PISA score for Canada changes to...”. Hence, the reweighting needs to be based on (among others) an assumed difference in the likelihood of participation.

In the following, we derive the formulas used for the reweighting of students. The notation is as follows:

- $\pi$ : Student participation rate

- $G$ : Binary variable of interest, taking values 0 and 1
- $w_i$ : student weight as observed in the dataset
- $v_i$ : student weight in unobserved missing data
- $u_i$ : new student weight taking into account observed and unobserved students
- $f$ : factor, with which original weights  $w_i$  are multiplied if  $G = 1$

To highlight the notation, Table E6 provides a fictional hypothetical dataset.

**Table E6. Hypothetical dataset**

Student ID	W: PISA student weight	G: belongs to underrepresented group	V: weights within underrepresented group	U: final corrected weights
1	$w_1$	1	$v_1 = f \cdot w_1$	$u_1$
2	$w_2$	0	$v_2 = w_2$	$u_2$
3	$w_3$	1	$f \cdot w_3$	$u_3$
4	$w_4$	0	$w_4$	$u_4$
5	$w_5$	0	$w_5$	$u_5$
6	$w_6$	0	$w_6$	$u_6$
7	$w_7$	1	$f \cdot w_7$	$u_7$
8	$w_8$	0	$w_8$	$u_8$
9	$w_9$	1	$f \cdot w_9$	$u_9$
10	$w_{10}$	0	$w_{10}$	$u_{10}$

Notes: Example table highlighting the notation as well as the basic structure of the PISA data. Each student with a student ID is assigned a weight  $w$  and may or may not belong to a group  $G$  which is subject to selection bias (first three columns). Weights of students in underrepresented group are then multiplied by a factor  $f$  and from this, by the steps described in this section, the final weights  $u$  are computed.

Table E6 shows how the original student weights  $W$ , the indicator variable  $G$  and the weights for the unobserved group  $V$  are related. Note that only those original weights belonging to student in the group of interest are manipulated by a factor  $f$ ; for students with  $G = 0$  they remain unchanged. Hence, the total sum of weights denoted as  $\bar{w}$  and  $\bar{v}$  is different for  $W$  and  $V$

Next, the weights are summed for both  $G = 1$  and  $G = 0$ :

$$W_1 = \sum_{G=1} w_i \quad 1$$

$$V_1 = \sum_{G=1} v_i = f \cdot W_1$$

and

$$W_0 = \sum_{G=0} w_i$$

$$V_0 = \sum_{G=0} v_i = W_0.$$

Note that the sum of weights for  $G = 0$  is identical for both the observed students with weights  $W$  and the unobserved students with weights  $V$ .

Next, consider the total sum of weights from observed and unobserved students:

$$\bar{w} = \sum w_i = W_0 + W_1 \text{ and } \bar{v} = \sum v_i = W_0 + f \cdot W_1.$$

Now the proportion of students in the group of interest (with  $G = 1$ ) is computed. As before, we distinguish between those students within the PISA sample ( $W$ ) and out-of-sample ( $V$ ). Additionally, the combination of those students covered by PISA and those excluded or absent is shown ( $U$ ):

$$R_W[G = 1] = \frac{W_1}{W_0 + W_1} = \frac{W_1}{\bar{w}}$$

$$R_V[G = 1] = \frac{f \cdot W_1}{W_0 + f \cdot W_1} = \frac{f \cdot W_1}{\bar{v}}$$

$$R_U[G = 1] = \pi \cdot R_W[G = 1] + (1 - \pi) \cdot R_V[G = 1].$$

Recall that the aim is to produce simulations that allow for a statement such as “if students with ability level lower than 400 are 50% more likely to not be covered by the PISA study, the average PISA score for Canada changes to...”. For this, it is of interest how the likelihood of participation differs for  $G = 1$  and  $G = 0$ . The probability of one student with  $G = 1$  not to participate is:

$$P[\text{non-participation}|G = 1] = \frac{(1 - \pi)R_V[G = 1]}{R_U[G = 1]} = \frac{(1 - \pi)f\bar{w}}{\pi\bar{v} + (1 - \pi)f\bar{w}}$$

The probability of student with  $G = 0$  not to participate is:

$$P[\text{non-participation}|G = 0] = \frac{(1 - \pi)R_V[G = 0]}{R_U[G = 0]} = \frac{(1 - \pi)\bar{w}}{\pi\bar{v} + (1 - \pi)\bar{w}}$$

This allows for the computation of the ratio between the two. If, for example, 50% of male students and 25% of female students do not participate, then we could divide the probability of non-participation by males with the probability of non-participation of females and arrive at a ratio of two. This would mean males are twice as likely (100% more likely) not to participate than females. Formalised, this ratio looks as follow:

$$X = \frac{P[np|G = 1]}{P[np|G = 0]} = f \cdot \frac{\pi\bar{v} + (1 - \pi)\bar{w}}{\pi\bar{v} + (1 - \pi)\bar{w}f}$$

In the course of the simulations we choose  $X$  and compute the factor  $f$  used for reweighting. This can be done as follows:

$$f = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q},$$

where

$$p = \frac{1}{\pi W_1} [W_0 + (1 - \pi)W_1 - X((1 - \pi)W_0 + W_1)]$$

and

$$q = -\frac{XW_0}{W_1}.$$

Note that only positive factors are chosen and the negative ones from the analytical solution discarded.

After the factor  $f$  is applied in the computation of  $V$ , the adjusted weights  $u_i$  can now be computed using the participation ratio  $\pi$ :

$$u_i = \pi \cdot w_i + (1 - \pi) \frac{\bar{w}}{\bar{v}} v_i$$

The above procedure allows us to assume a difference in likelihood, which can be easily communicated. In the case of Canada, we  $\pi = .53$ . If we are interested in what happens to the PISA score if students whose parents are in the lower education categories are twice (three times, four times, ...) as likely not to participate in the study than those whose parents are in

the highest category, we just have to identify the two groups ( $g_1$  and  $g_0$ ) and to set  $X = 2$  ( $X = 3, X = 4, \dots$ ).

### **Combination of variables**

Variables cannot be easily combined using this approach. Communication of the results is very easily understandable thanks to the link to “likelihood” of non-participation, but this comes at a price: as students can be part of more than one group (for instance repeated a class AND low SES), one of the groups or none of the groups, simply multiplying the factors will distort the ease of communication. A work around to this issue is to use nested intervals to adjust the factors, such that for all variables involved in the combination, the initial claim still holds true (e.g. “X times as likely”).

## **Appendix I. Calculating coverage rates in PISA 2015**

The PISA technical report (Chapter 11) presents the following two figures:

- An estimate of the total number of 15-year-olds in the country (Table 11.1 – column “all 15-year-olds”).
- A weighted estimate of the number of students who took the PISA test (Table 11.7 – column “Number of students assessed (Weighted) (NUMSTW3)”).

In Table 3 of the paper, we take the latter as the numerator and the former as the denominator to create the percentages in the final column (the weighted number tested divided by the weighted number of all 15-year-olds).

What the PISA 2015 technical report does not describe clearly, however, is how one moves from the former (total number of 15-year-olds in the country) to the latter (the weighted number of 15-year-olds who took the test). In this appendix, we provide our best effort to reproduce the OECD figures.

Table I1 reproduces Table 2 from the report. We have calculated how one moves from the total number of 15-year-olds in a country to the percentage covered by the survey as the product of the following four components:

- % of 15-year-olds enrolled in school
- % exclusion
- % school response rate (after replacement)
- % student participation rate

In other words, it is the combination of four factors: 1. 15-year-olds in the country not being part of the target population (i.e. enrolled in school); 2. Schools or students being excluded from the sample (e.g. due to having Special Educational Needs); 3. School non-response; 4. Student non-participation.

For example, the calculation for Canada following the above would be<sup>18</sup>:

$$0.96 * 0.925 * 0.72 * 0.81 = 0.51$$

Table I2 provides a comparison of the figures depending upon which method is used. Specifically, the left-hand column uses the method described above, while the right-hand column uses figures directly from the PISA 2015 technical report (dividing the weighted number of students assessed in PISA by the total 15-year-old population).

From this table, there are two key points to note. First, the figure for Canada is very similar, regardless of the method used (51% versus 53%). Indeed, we have experimented making various changes to our calculation method, and always found the figure to be between 50% to 60%. Second, the final figure in most cases is similar whichever method is used (correlation = 0.93) with the figures for most countries differing by just a few percentage points. Although there are some exceptions where the difference is bigger (e.g. Italy, South Korea, Chile,

---

<sup>18</sup> This uses the unweighted school response percentage after replacement (0.72). If one were to use the weighted figure instead, the final figure for Canada would increase to 0.56.

Turkey, Austria) we believe that the calculation this appendix has presented broadly reflects how the final coverage rate of the population has been derived.

**Appendix Table I1. Exclusions, school, pupil and overall participation rates in PISA 2015 amongst OECD countries**

	1. % of 15-year-olds enrolled in school	2.% Excluded	3.School response % before replacement	4.School response % after replacement	5.Student participation rate (%)
South Korea	100	0.9	99 (100)	99 (100)	99
Mexico	62	0.9	95 (95)	97 (98)	95
Turkey	83	1.1	90 (97)	96 (99)	95
Portugal	91	1.3	84 (86)	94 (95)	82
Belgium	99	1.7	81 (83)	95 (95)	91
Chile	96	1.8	89 (92)	97 (99)	94
Greece	100	1.9	90 (92)	99 (98)	94
Austria	94	2.1	99 (100)	99 (100)	71
Germany	100	2.1	96 (96)	99 (99)	93
Czech Republic	100	2.4	99 (98)	99 (98)	89
Japan	98	2.4	95 (94)	99 (99)	97
Poland	95	2.4	89 (88)	99 (99)	87
Finland	100	2.8	99 (100)	100 (100)	93
Ireland	98	3.1	99 (99)	99 (99)	89
Slovenia	98	3.1	95 (98)	95 (98)	91
Spain	94	3.2	99 (99)	100 (100)	89
Hungary	95	3.3	92 (93)	97 (99)	92
USA	95	3.3	67 (67)	83 (83)	90
Israel	95	3.4	89 (91)	91 (93)	91
Iceland	99	3.6	95 (99)	95 (99)	86
Netherlands	100	3.7	62 (63)	92 (93)	85
Italy	92	3.8	78 (74)	87 (88)	89
France	96	4.2	91 (91)	95 (94)	88
Slovak Republic	99	4.3	92 (93)	98 (99)	91
Switzerland	98	4.4	91 (93)	97 (98)	93
Denmark	99	5.0	88 (90)	89 (92)	87
Latvia	98	5.1	86 (86)	92 (93)	90
Australia	100	5.3	91 (94)	92 (95)	81
Estonia	98	5.5	100 (100)	100 (100)	93
Sweden	99	5.7	99 (100)	99 (100)	91
New Zealand	95	6.5	69 (71)	84 (85)	80
Norway	100	6.8	95 (95)	95 (95)	91
Canada	96	7.5	70 (75)	72 (79)	81
Luxemburg	96	8.2	100 (100)	100 (100)	96
UK	100	8.2	85 (84)	91 (93)	88
<b>OECD average</b>	<b>96</b>	<b>3.7</b>	<b>90 (91)</b>	<b>95 (96)</b>	<b>89</b>
<b>OECD median</b>	<b>98</b>	<b>3.3</b>	<b>91 (93)</b>	<b>97 (98)</b>	<b>91</b>

Notes: Both weighted and unweighted school response rates are provided (the former appear in brackets).

**Appendix Table I2. Comparison of how overall “population coverage” are derived.**

	Derived figures using method above	Figures taken directly from OECD report	Difference
South Korea	97%	90%	7%
Mexico	57%	57%	0%
Turkey	75%	66%	9%
Portugal	69%	68%	1%
Belgium	84%	81%	3%
Chile	86%	74%	12%
Greece	91%	85%	6%
Austria	65%	72%	-7%
Germany	90%	89%	1%
Czech Republic	86%	81%	5%
Japan	92%	91%	1%
Poland	80%	79%	1%
Finland	91%	91%	0%
Ireland	83%	85%	-2%
Slovenia	82%	83%	-1%
Spain	81%	81%	0%
Hungary	82%	82%	0%
USA	68%	62%	6%
Israel	76%	79%	-3%
Iceland	78%	79%	-1%
Netherlands	75%	76%	-1%
Italy	69%	61%	8%
France	77%	76%	1%
Slovak Republic	85%	81%	4%
Switzerland	84%	87%	-3%
Denmark	73%	73%	0%
Latvia	77%	74%	3%
Australia	70%	72%	-2%
Estonia	86%	86%	0%
Sweden	84%	84%	0%
New Zealand	60%	61%	-1%
Norway	80%	79%	1%
Canada	51%	53%	-2%
Luxemburg	84%	84%	0%
UK	73%	69%	4%
<b>Mean</b>	78%	77%	2%
<b>Median</b>	80%	79%	0%