

OECD Education Working Papers No. 154

Who Makes It Into PISA? Understanding the Impact of PISA Sample Eligibility Using Turkey as a Case Study (PISA 2003 - PISA 2012)

Nicholas Spaull

https://dx.doi.org/10.1787/41d175fc-en

OECD

Unclassified **EDU/WKP(2017)3**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

English - Or. English

DIRECTORATE FOR EDUCATION AND SKILLS

EDU/WKP(2017)3
Unclassified

WHO MAKES IT INTO PISA? UNDERSTANDING THE IMPACT OF PISA SAMPLE ELIGIBILITY USING TURKEY AS A CASE STUDY (PISA 2003 - PISA 2012)

OECD Education Working Paper No. 154

Nicholas Spaull, Stellenbosch University, South Africa

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

Nicholas Spaull, Senior Researcher
Stellenbosch University, South Africa
spaull@sun.ac.za / NicholasSpaull@gmail.com

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

OECD EDUCATION WORKING PAPERS SERIES

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcome, and may be sent to the Directorate for Education and Skills, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

---

www.oecd.org/edu/workingpapers

Copyright © OECD 2017

ACKNOWLEDGEMENTS

Nicholas Spaull acknowledges funding from the OECD Thomas J. Alexander Fellowship programme for carrying out this work.

ABSTRACT

Of the OECD countries that participate in the Programme for International Student Assessment (PISA), Turkey has one of the lowest levels of performance and the highest rates of improvement in PISA scores between 2003 and 2012. New evidence presented in this paper suggests that existing accounts have underestimated both progress and inequity in Turkey because they did not take into account the large proportion of 15-16 year-olds that are ineligible for the PISA sampling frame, either because they are no longer in school or because they are severely delayed. Using Turkey's Demographic and Health Survey (DHS) data for 2003, 2008 and 2013 we show that the proportion of 15-16 year-olds that were eligible for the PISA sample in Turkey has nearly doubled from 45% in PISA 2003, to 80% in PISA 2012. By combining DHS data on access/attainment and PISA survey data on learning outcomes we show that: (1) the improvement in the percentage of 15-16 year-olds reaching Level 2 in PISA (functional literacy and functional numeracy) is up to twice as large as that reflected in official PISA reports, (2) the gap in functional literacy rates between rich and poor youth in 2012 is 2.3 times as large as was previously thought, and (3) contrary to earlier research the gap between rich and poor has not declined between 2003 and 2012. The paper emphasises the importance of accounting for sample eligibility and representivity when making inter-country and inter-temporal comparisons using international assessment data, particularly for developing countries with expanding education systems.

RÉSUMÉ

Des pays de l'OCDE ayant participé au Programme international pour le suivi des acquis des élèves (PISA), la Turquie obtient un des niveau de performance les plus bas et un des taux d'amélioration le plus élevé dans les scores PISA entre 2003 et 2012.De nouvelles preuves dans cet article suggères que les récits existant sous-estiment à la fois le progrès et l'inégalité en Turquie car ils ne prennent pas en compte la proportion importante de jeunes de 15-16 ans qui ne sont pas éligibles dans l'échantillonnage PISA, soit parce qu'ils ne sont plus scolarisés soit parce qu'ils ont un retard important. En se servant de l'enquête démographique et de santé pour la Turquie en 2003, 2008 et 2013, nous démontrons que la proportion des 15-16 ans qui étaient en réalité éligible à l'échantillon PISA en Turquie a presque doublé allant de 45% au moment de l'enquête PISA 2003, à 80% dans PISA 2012. En combinant les données de l'enquête démographique sur l'accès/obtention et les données PISA sur les acquis d'apprentissage nous démontrons que : (1) l'amélioration dans le pourcentage des 15-16 ans ayant atteint le Niveau 2 dans PISA (l'alphabétisation fonctionnelle et la numératie fonctionnelle) est deux fois plus important que ce qui est

reflété dans les rapports de PISA, (2) l'écart dans les taux d'alphabétisation fonctionnelle entre les riches et les pauvres en 2012 est 2.3 fois plus important qu'il n'était pensé et (3) contrairement aux recherches antérieurs sur l'écart entre les riches et les pauvres n'a pas baissé entre 2003 et 2012. Cet article accentue l'importance de prendre en compte l'éligibilité dans l'échantillonnage et la représentation lorsqu'on effectue des comparaisons inter-pays et inter-temporaires en utilisant la base de données de PISA (ou de n'importe quel autre évaluation internationale), particulièrement lorsqu'il s'agit des pays en voie de développement ayant des systèmes d'éducation en expansion.

TABLE OF CONTENTS

OECD EDUCATION WORKING PAPERS SERIES....2
ACKNOWLEDGEMENTS....3
ABSTRACT....3
RÉSUMÉ....3
INTRODUCTION AND RESEARCH QUESTION....7
Structure of the paper....8
LITERATURE....9
PISA and sample eligibility....9
METHOD....14
Assumptions about the performance of PISA-ineligible students....17
Household survey data required for measures of access and attainment: Demographic and Health Survey (DHS) data....18
RESULTS....19
OVERARCHING FINDINGS....29
ROBUSTNESS CHECKS....36
POTENTIAL LIMITATIONS....38
CONCLUSION....39
REFERENCES....41

Tables

Table 1. The percentage of the total 15 year old population covered by the PISA sampling frame (Coverage Index 3) in selected countries....11
Table 2. Percentage of 15-16 year-olds in DHS enrolled in Gr7 or higher (PISA eligible), with Standard Errors (%)....25
Table 3. Percentage of 15-16 year-olds in DHS who never enrol in school (PISA ineligible), with Standard Errors (%)....25
Table 4. Percentage of 15-16 year-olds in DHS who have left school (due to drop out or matriculation) (PISA ineligible), with Standard Errors (%)....25
Table 5. Percentage of 15-16 year-olds in DHS who are still enrolled in Grades1-6 (PISA ineligible), with Standard Errors (%)....26
Table 6. Percentage of the PISA sample that is literate (Level 2+), with Standard Errors (%) - Uncorrected for those who are ineligible....26
Table 7. Percentage of the PISA sample that is numerate (Level 2+), with Standard Errors (%) - Uncorrected for those who are ineligible....27
Table 8. Percentage of 15-16 year-olds that have Access-to-Literacy (Level 2+), with Standard Errors (%)....27
Table 9. Percentage of 15-16 year-olds that have Access-to-Numeracy (Level 2+), with Standard Errors (%)....28

Figures

Figure 1. The educational status and PISA-eligibility of 15-16 year-olds in Turkey in DHS 2003, DHS 2008 and DHS 2013....8
Figure 2. Access to literacy (Level 2) in Turkey 2003 (PISA 2003 and DHS 2003) 20
Figure 3. Access to Literacy (Level 2) in Turkey 2009 (PISA 2009 and DHS 2008) 21
Figure 4. Access to Literacy (Level 2) in Turkey 2012 (PISA 2012 and DHS 2013 21
Figure 5. Figure 5: Access to Numeracy (Level 2) in Turkey 2003 (PISA 2003 and DHS 2003) 22
Figure 6. Access to Numeracy (Level 2) in Turkey 2009 (PISA 2009 and DHS 2008)....23
Figure 7. Access to Literacy (Level 2) in Turkey 2012 (PISA 2012 and DSH 2013) 23
Figure 8. Percentage of 15-16 year-olds reported to be literate (L2+) in OECD reports and corrected for sampling eligibility (Access to Literacy rates) in 2003 and 2012 30
Figure 9. Percentage of 15-16 year-olds reported to be Numerate (L2+) in OECD reports and corrected for sampling eligibility (Access to Numeracy rates) in 2003 and 2012....31
Figure 10. Percentage of 15-16 year-olds reported to be literate (L2+) in OECD reports and corrected for sampling eligibility (access to literacy rates) in 2003 and 2012....32
Figure 11. Percentage of 15-16 year-olds reported to be numerate (L2+) in OECD PISA reports and corrected for sampling eligibility (access to numeracy rates) in 2003 and 2012....33
Figure 12. PISA eligibility among Turkish 15-16 year-olds by region in the Turkish DHS 2003, 2008 and 2013 36
Figure 13. Kernel density of Turkey's PISA 2012 Reading Scores (Plausible Value 1) by grade of enrolment 37

INTRODUCTION AND RESEARCH QUESTION

Of the OECD countries that participate in the Programme for International Student Assessment (PISA), Turkey has one of the lowest levels of performance and the highest rates of improvement in PISA scores (2003-2012). Between 2003 and 2012 Turkey managed to improve learning outcomes by almost a full year of learning in Mathematics (25 points) while also decreasing the achievement gap between rich and poor (equity). While this is an impressive achievement in and of itself, it does not highlight the drastic expansions in access and attainment – and therefore PISA sample eligibility – that have taken place over the same period. PISA is supposed to "provide an assessment of the cumulative yield of education and learning at a point at which most young adults are still enrolled in initial education" (OECD and UNESCO-UIS, 2003, p. 249). While this enrolment criterion is true for most OECD countries, where at least 80% of 15-16 year-olds[1] are eligible for PISA, historically this has not been the case for Turkey or a number of other middle income countries such as Mexico, Brazil, Indonesia or Vietnam.

For example in PISA 2003 less than half (36-45%[2]) of Turkish 15-16 year-olds were actually eligible[3] for the PISA sampling frame. To put this another way, PISA in Turkey in 2003 was representative of less than half of the population of 15-16 year-olds in the country. This is because PISA does not include out-of-school 15-16-year-olds (those who never enrol, have dropped out or left school with a primary school diploma) or those 15-16-year-olds that are severely delayed (those who are still enrolled in Grades 1-6) (OECD, 2005, p. 46). Given that PISA-eligibility is not a random process – with under-performing and poorer students being more likely to be ineligible (Hanushek and Woessmann, 2011) – the research question emerges: How does Turkey's education system fare over the 2003-2012 period if one takes into account the changing proportions of PISA-eligible 15-16-year-olds?

In this paper we attempt to answer this question by incorporating information on the level and distribution of PISA-ineligible students from household-survey data at three points in time. Specifically we use the 2003, 2008 and 2013 rounds of the Turkish Demographic and Health Survey (DHS) data for information on the out-of-school and delayed 15-16-year-old population, and the 2003, 2009 and 2012 waves of PISA for the PISA-eligible 15-16-year-old population. Doing so allows us to validly compare the total population of 15-16 year-olds across these three time periods.

Traditional analyses of Turkey's PISA results ignore these large changes in PISA eligibility and thus provide inaccurate estimates of achievement, progress and inequality, as we will show. Figure 1 below provides just one indication of how large these changes have been. It reports on DHS data and shows that the proportion of 15-16 year-olds that were actually eligible for the PISA sample in Turkey (and thus the

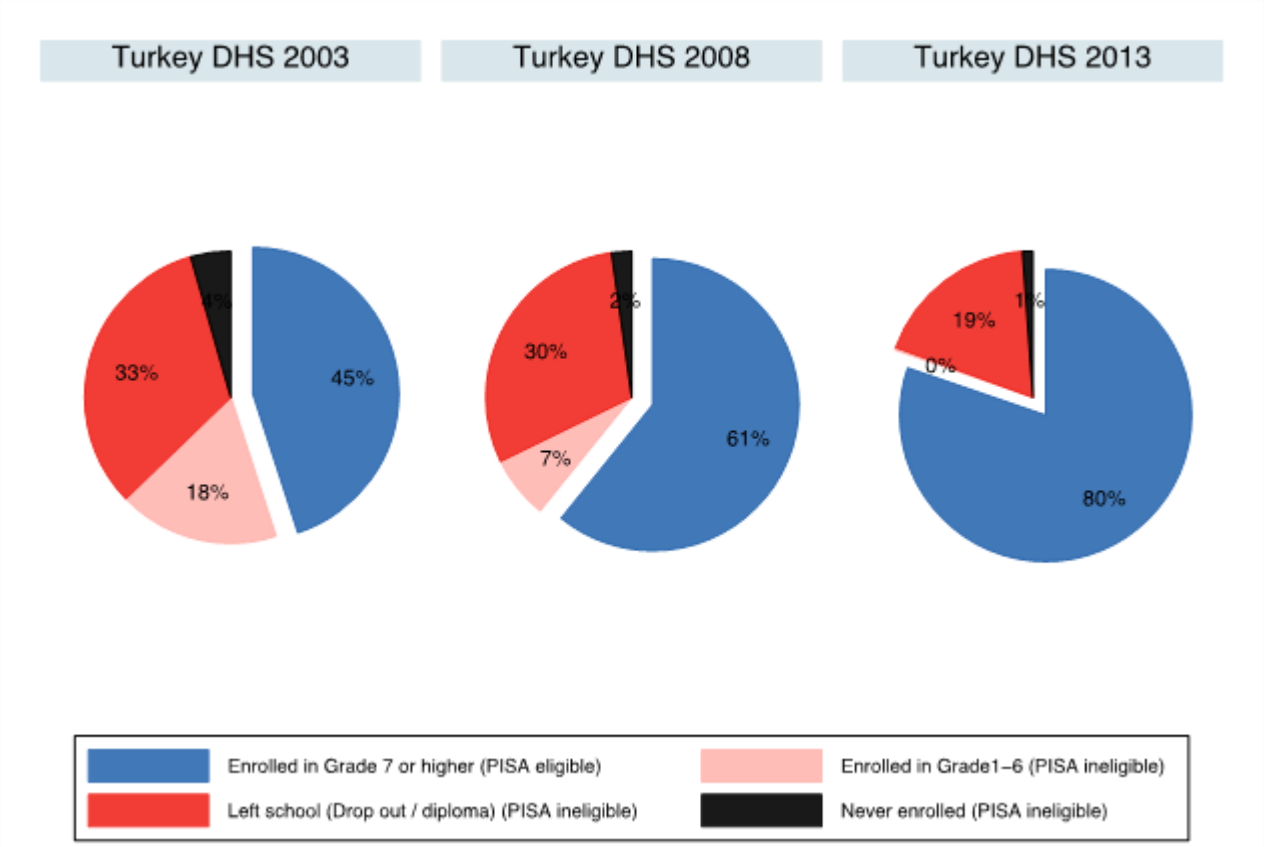
1 The sampling frame for the PISA assessment covers only those students that are aged 15 years and 3 (completed) months to 16 years and 2 (completed) months who are currently enrolled in Grade 7 or higher (OECD, 2005, p. 46). As is discussed later in the paper, using both 15 and 16 year-olds increases the sample size (and therefore decreases the standard errors) when using the household survey data.

2 The lower estimate reflects the OECD's own "Coverage Index 3" for Turkey 2003 (OECD, 2004, p. 321) while the upper estimate is taken from the Turkish Demographic and Health Survey (DHS) data for 2003. Both data sources are discussed further in the paper.

3 For ease of reading, throughout this paper I use 'sample eligibility' and 'sample coverage' interchangeably. In the technical terms used by the OECD these are marginally different because sample coverage is sample eligibility minus legitimate exclusions. These exclusions might be school-level exclusions (for inaccessibly remote schools), or for students that are intellectually disabled, for example. These exclusions are almost always less than 5% in developing countries (OECD, 2012b: 149). Thus wherever the paper refers to coverage it should technically be read as "coverage after exclusions among those eligible."

percentage of the population that PISA is representative of) nearly doubled from 45% in 2003, to 80% in 2013, with large declines in the proportion of out-of-school 15-16 year-olds and those who had severely delayed grade progression.

Figure 1. The educational status and PISA-eligibility of 15-16 year-olds in Turkey in DHS 2003, DHS 2008 and DHS 2013



Source: Author's own calculations from PISA and DHS datasets

Structure of the paper

The analysis begins by presenting a short review of some of the existing Turkish PISA literature, drawing specific attention to shortcomings around inter-temporal comparisons in Turkey. Thereafter the method used in the paper to combine DHS and PISA data – first employed by Spaull and Taylor (2015) - is discussed, and the two sources of data are reviewed. The main results are presented in Figures 2-7 and Tables 2-9, each providing a breakdown in access, attainment and quality for 2003, 2009 and 2012 for the country as a whole, as well as for three demographic sub-groups defined by: gender, socio-economic status and a gender-socio-economic-status interaction. By assuming that PISA-ineligible 15-16 year-olds would not have reached PISA proficiency Level 2 in either Reading or Mathematics – a relatively conservative assumption - it becomes possible to create what we call an "access-to-literacy" and "access-to-numeracy" rate. This is the proportion of the total population of 15-16 year-olds that acquire PISA Level 2 in Reading (functional literacy) and Mathematics (functional numeracy). After a discussion of the main findings we ask whether or not the above assumption is valid and present some empirical evidence and sensitivity analysis in support of it, as well as a penultimate section on the limitations of the present study. The concluding section highlights the contribution that this analysis makes both to our understanding of the

Turkish education system between 2003 and 2012, and also to the need for a more circumspect approach when interpreting PISA results in countries with large and changing percentages of ineligible 15-16 year-olds.

LITERATURE

There is a large body of uncontested research showing that Turkey has made significant progress in terms of learning outcomes, access and attainment. In a recent study, Köseleci (2015) provides a succinct overview of the Turkish education literature over the last 15 years, and shows that while there has been considerable progress in access and attainment (Aydagül, 2008a; Dülger, 2004), as well as in closing the gender gap (Tansel, 2002), these reforms have been less successful in the eastern and south eastern regions of Turkey (Kırdar, 2009; O'Dwyer, Aksit, and Sands, 2010).

In almost all cases, this rapid expansion was a direct response to deliberate policy initiatives such as 'Rapid Coverage for Compulsory Education Program of 1997' where over 1 million students enrolled in Grades 1-8 in the six year period 1997-2002 (Dülger, 2004), as well as progress in equity in response to the Ninth Development Plan (covering 2007-2013) which "prioritizes girls and students in rural areas and addresses dropouts as important policy issues" (Aydagül, 2008a, p. 402). This included bussing students, expanding boarding schools as well as free textbooks and school meals to low-income students (Dülger, 2004). Notwithstanding the above achievements, most authors also emphasise the high levels of inequality in the system. The major dimensions of inequality are typical of those found in the literature: household income, parental education and occupation, rural/urban status, ethnicity and school selectivity (Alacacı and Erbaş, 2010; Kırdar, 2009; Köseleci, 2015; Rankin and Aytac, 2006).

PISA and sample eligibility

Since its inception in 2000, PISA has acquired global significance with approximately 70 countries/economies participating in the latest round (2015). To that end, some have argued that PISA "has become the world's premier yardstick for evaluating the quality, equity and efficiency of school systems" (OECD, 2012b, p. 3). While this may be true for most OECD countries, where PISA is representative of more than 80% of 15-16 year-olds, PISA may be a less accurate yardstick in Turkey where, historically, a substantially lower proportion of 15-16 year-olds were actually eligible for PISA. The same can be said for other middle-income countries such as Mexico, Brazil, Indonesia and Vietnam.

Discussions around sample coverage in international assessments have been around for over two decades. For example, Rotberg (1995, p. 1446) argues that:

> "The basic problem is student selectivity: The fewer the students who take the test, the higher the average score. That score is not a valid measure of the overall quality of the education system. It simply reflects the fact that the students represented in the test comparisons have been much more highly selected in some countries than in others."

More recently, Ferreira and Gignoux explain that:

> "Selection would not be a problem if one were interested *exclusively* in the performance of 15 year-olds that are in school, and within a reasonable range of their expected grade of attendance.

> But this is likely to be an excessively narrow prism through which to assess a country's education system and – even more so – to make international comparisons." (2014: 16)

Two of the most prolific authors in the field of international assessment research, Eric Hanushek and Ludger Woessman (2011: 79), explain why this is the case: "Non-random differences in patterns of school enrolment, sample exclusions, and non-response are clearly able to influence rankings of countries on international league tables of average student achievement." Following their analysis of TIMSS 1995, 1999, and 2003, as well as PISA 2000 and 2003, they conclude as follows:

> "We find that countries having more schools and students excluded from the targeted sample, having schools and students who are less likely to participate in the test, and having higher overall school enrolment at the relevant age level tend to perform better on the international tests" (Hanushek and Woessmann, 2011:79).

While Hanushek and Woessman do flag the issue of sample coverage, they are also at pains to explain that accounting for these differences does not affect the conclusion that academic achievement is an important determinant in economic growth regressions. However, this is not the case in the education literature, at least not for Turkey. The findings presented in this paper show that the conclusions on both educational inequality and progress change substantially when accounting for PISA-ineligible students in Turkey.

The OECD is not unaware of the issues relating to sample representivity. In fact, in all PISA reports and technical reports the OECD includes a variety of coverage indices, including 'Coverage Index 3' which is "the proportion of the national population of 15 year-olds covered by the non-excluded portion of the student sample" (OECD, 2014b, p. 177). Or phrased differently, the proportion of the national 15 year old population (in and out of school) that is represented in the PISA sample. Although this statistic is readily available for all countries and all rounds of PISA, very few researchers have analysed it or discussed the implications for high, low or changing levels of this index. The only substantive treatment of this issue (prior to the PISA 2015 report which will be discussed later) can be found in a single paragraph found in all PISA reports from 2003 onwards, and in Volume II of the 2012 PISA Report in a section titled "Providing Access to Schooling to All 15-year-olds" (OECD, 2012b, p. 53). The paragraph in question is included verbatim below:

> "This high level of coverage contributes to the comparability of the assessment results. For example, even assuming that the excluded students would have systematically scored worse than those who participated, and that this relationship is moderately strong, an exclusion rate in the order of 5 per cent would likely lead to an overestimation of national mean scores of less than 5 score points … This assessment is based on the following calculations: If the correlation between the propensity of exclusions and student performance in 0.3, resulting mean scores would likely to overestimated by 1 score point if the exclusion rate is 1 per cent, by 3 score points if the exclusion rate is 5 per cent, and by 6 score points if the exclusion rate is 10 per cent. If the correlation between the propensity of exclusions and student performance is 0.5, resulting mean scores would be overestimated by 1 score point if the exclusion rate is 1 per cent, by 5 score points if the exclusion rate is 5 per cent, and by 10 score points if the exclusion rate is 10 per cent. For this calculation, a model was employed that assumes a bivariate normal distribution for the propensity to participate and performance" (OECD, 2004, p. 325)

Although this is taken from the PISA 2003 report, it can also be found in the PISA 2006, 2009 and 2012 (Vol. II) reports (on pages 24, 176 and 151 respectively). What this simulation seems to imply is that correcting for those omitted from the PISA sample would not make much of a difference for most countries. Indeed, this is explicitly stated in the PISA 2012 (Volume II) report: "Most countries would

show no or very small changes in average performance after taking into account the performance of those outside the school system" (OECD, 2012b, p. 53). It goes on to mention five countries/economies where changes in performance are largest: Hong Kong (China), Macao(China), Mexico, Shanghai-China, Turkey and Viet Nam, but stresses that in all cases except Viet Nam and Macao (China) this correction would not change the ranking by much. This conclusion is drawn from Table II.2.12 in the annex of the PISA 2012 (Volume II) report which compares coverage rates between 2003 and 2012 and simulates how the scores would change if those who were out of school achieved at Level 1 (OECD, 2012b, p. 210).

Unfortunately the Coverage Index 3 (CI-3) values for PISA 2003 that were used in the PISA 2012 calculations described above are the incorrect figures. There is an error in the calculation of the CI-3 figures presented in the PISA 2003 report (OECD, 2004, p. 322), and these differ from the (correct) figures presented in the PISA 2003 Technical Report (OECD, 2005, p. 166)[4]. To provide an indication of how large the differences are between the (incorrect) CI-3 figures of the PISA 2003 Report and the (correct) CI-3 figures of the PISA 2003 Technical Report, one can look at four countries. The correct figures are as follows (with incorrect figures in brackets): Indonesia 46% (73%), Brazil 54% (65%), Mexico 49% (58%), and Turkey 36% (54%). Clearly the incorrect figures are much higher and therefore all analyses based on these figures (like that in the PISA 2012 Report) will underestimate how much of a difference including the excluded students will make.

Using the OECD's assumption of a 0.5 correlation between the propensity of exclusions and student performance, Turkey's 2003 PISA scores would drop by 64 points, or the equivalent of approximately two years of schooling, if excluded 15-16 year-olds were accounted for.

To provide some examples of just how low the coverage rates are in some countries, and how quickly they have changed, Table 1 below provides the correct CI-3 rates for a sample of selected PISA countries and years with very low levels of sample coverage (Germany and Canada are included as reference countries). Note that Turkey has the lowest CI-3 rate of 36% in PISA 2003 and that this nearly doubled to 70% by PISA 2015. While a number of other countries have very low levels of sample coverage, the aim here is to focus on Turkey to show how incorporating data on PISA-eligibility and coverage can shift conclusions on education quite substantially. To calculate these percentages the OECD divides the weighted estimate of PISA-eligible students by the national population of 15-year-olds, the latter being provided by the National Project Manager in each country (OECD, 2014b, p. 177). This figure is calculated from the most recent census data, and may be somewhat dated depending on the year of the census (OECD, 2012b, p. 211).

Table 1. The percentage of the total 15 year old population covered by the PISA sampling frame (Coverage Index 3) in selected countries

	PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA 2015
Albania	40%			61%	55%	84%
Brazil	69%	53%	55%	63%	69%	71%
Canada	86%	83%	87%	84%	83%	84%
Chile	82%		78%	85%	83%	80%
Colombia			60%	58%	63%	75%
Costa Rica					50%	63%

4 I have confirmed this with OECD officials familiar with PISA sampling and the reports in question (Miyako Ikeda, personal communication, 22 June 2016).

Germany	89%	93%	95%	90%	95%	96%
Indonesia	39%	46%	53%	53%	63%	68%
Mexico	45%	49%	54%	61%	63%	62%
Peru	50%			73%	72%	74%
Thailand	54%	69%	72%	73%	72%	71%
Tunisia		92%	90%	89%	91%	93%
Turkey		36%	47%	57%	68%	70%
Uruguay		63%	69%	63%	73%	72%
Vietnam					56%	49%

.

Source: (Adams and Wu, 2002, p. 135 and own-calculations for non-OECD countries),(OECD, 2005, p. 169),(OECD, 2009, p. 180),(OECD, 2012a, p. 161),(OECD, 2014b, p. 179) (OECD, 2016a, p. 46)

While the CI-3 statistic is useful, and provides a broad indication of how representative the PISA sampling frame is, it tells us nothing about those youth that are omitted from the PISA sample. Looking at CI-3 and PISA data alone cannot tell us how those who make it into the PISA sampling frame differ from those that do not, or by how much. For example, they cannot tell us if girls or poorer students are more or less likely to be included in the PISA sampling frame or how this has changed over time.

Although traditionally the OECD has under-emphasised the importance of sample coverage when discussing trends and making policy recommendations, there are two notable exceptions:

1. **The new OECD initiative aimed at developing countries**: **PISA for Development (PISA-D)**. The aim of the PISA-D programme is to make PISA more relevant and accessible to a wider range of countries. It is also "pioneering new methods and approaches to include out-of-school youth in the assessment – a crucial first step towards addressing social exclusion" (OECD, 2016b, p. 2). This will take the form of a household-survey component which, together with the school-based survey, can provide a holistic picture of the achievement of 15 year-olds in the PISA-D countries. There are no immediate plans to include this component for 'traditional' PISA countries.

2. **The PISA 2015 Report**. The PISA 2015 Report (OECD, 2016a), launched in December 2016, spends a considerable amount of time reporting on issues relating to sample eligibility, representivity and coverage, which marks a break with all previous PISA reports. A simple discourse analysis of the last five PISA reports makes this point abundantly clear. Analysing the PISA 2003, 2006, 2009 and 2012 reports, one can find the word 'coverage' on 2 - 4% of the total pages[5]. This is in stark contrast to the PISA 2015 Report where the word appears on 32% of the report's pages and has a number of dedicated sections to coverage and related topics of equity. The report goes further to explain how this has the potential to impact representivity: "Overall, there are 20 countries in PISA 2015 where less than 80% of 15-year-olds are enrolled in school and eligible to participate in PISA. This implies that PISA results for these countries are not fully representative of their populations of 15-year-olds. It also signals that these school systems face serious challenges in becoming more inclusive and equitable" (OECD, 2016a, p. 210). While resident at the OECD headquarters in Paris in 2016 I presented this research twice to the OECD

5 The exact number of pages that the word 'coverage' can be found relative to the total number of pages in the report for each year are as follows: 2003 (8/471), 2006 (8/390), 2009 (10/276), 2012 (11/564), and 2015 (160/494).

> staff members tasked with writing the PISA 2015 Report, and argued for its inclusion in the (then) upcoming PISA report. It would seem that the present research, together with other reports (Bloem, 2015; Lockheed, Prokic-Breuer, and Shadrova, 2015) have shifted the conversation somewhat. While this is a welcome and important addition to the PISA reports, and one for which the OECD should be congratulated, it will take some time before there is widespread recognition that comparisons between countries with very different coverage rates are problematic. In a recent World Bank working paper titled "Unravelling a Secret: Vietnam's Outstanding Performance on the PISA Test", Parandekar and Sedmik (2016) put forward an impressive list of possible explanations for Viet Nam's high scores in PISA without once mentioning that more 15 year-olds had been excluded from the PISA sampling frame in Viet Nam than in any other country. A quite serious omission given that about half of all 15 year-olds are not represented by the PISA sample in the country. For the remainder of this paper we will focus on Turkey exclusively.

It is not only the OECD that has, up until very recently, given insufficient attention to the issues of sample eligibility in Turkey (or other countries). For example, the World Bank (2005) in their 'Education Sector Study' of Turkey made little mention of the PISA 2003 Turkey sampling issue. For example throughout the report the PISA results are generalised to all Turkish 15 year-olds when in fact only 36% of Turkish 15 year-olds were not covered in the PISA sample, according to the OECD. One excerpt from the World Bank's study: "Over half (55%) of Turkey's 15-year-old students were unable to perform above level 1" (2005, p. 17). This does not take into account that nearly one in five 15-16 year old students in Turkey were still in school but were ineligible for PISA in 2003 because they were severely delayed, as will be shown later in this paper using household survey data.

In the peer-reviewed academic literature there is very little discussion of the impact of sample eligibility on PISA results. One exception is a recent short article titled "PISA for low- and middle-income countries" where Bloem (2015, p. 485) explains that "High levels of out-of-school children mean PISA cannot capture a representative sample of the 15-year-old population. Measures, most importantly equity measures, would be less meaningful if most disadvantaged students were no longer in school at this age, and the PISA results would refer to an already relatively privileged student population" (see Bloem, 2013 for a more detailed discussion).

Looking at Turkey specifically, in a recent paper titled "A critique of the interpretation and utilisation of PISA results in Turkey" Gür, Çelik and Özoglu (2012) do not mention that less than half of Turkish 15 year-olds were actually eligible for PISA 2003 in the first place, which is perhaps the most self-evident critique. Similarly Alacaci and Erbas (2010) analyse PISA 2006 data and focus on inequality but do not mention that less than half of 15 year-olds (CI3=47%) were eligible for the PISA sample in Turkey in 2006. In fact they state that "the Turkish sample is representative of the 15-year-old students as the index of coverage of the national enrolled-in-schools population was .98 (out of 1)" (Alacacı and Erbaş, 2010, p. 184), citing Coverage Index 2. However Coverage Index 2 only records *eligible* students and thus excludes 15-16 year old students that are still enrolled but are in Grade 1-6 (OECD, 2004, p. 323). In 2008 about 18% of Turkish 15-16 year-olds were still in school but were ineligible for PISA due to delayed progression (according to the Turkish DHS 2008). The same can be said for a more recent analysis by Özdemir (2016) on "Equity in the Turkish Education System" which analyses PISA 2012 data, which again, is only representative of two of every three (CI3=68%) 15-16 year-olds, according to the OECD.

The only Turkish study that we are aware of that specifically mentions the lack of PISA coverage as a point of concern is that of Köseleci and Şaşmaz (2011, p. 128) who caution that "it should be kept in mind that PISA does not cover all 15-year-old children but students of this age." They use the 2009 Household Labor Survey and conclude that "only 55-60 percent of all 15-year-olds in Turkey attend school regularly." This figure is relatively close to the 57% Coverage Index 3 estimated by the OECD (2012a, p. 121) and the

67% of 15-16 year-olds that were still enrolled in school and in Grade 7 or higher estimates in this paper using the 2008 DHS. These two authors argue that "When children who are not enrolled in schools, who are most likely at a lower knowledge and skill level are also accounted for, Turkey's average performance and level distribution will be worse than what is reflected by PISA scores" (Köseleci and Sasmaz, 2011, p. 128). Unfortunately the authors do not attempt to estimate how much lower Turkey's performance would be if these out of school youth were accounted for. The present analysis goes some way to answering this question by combining both household survey and PISA data. The method for doing so is described in the next section.

METHOD

In PISA participating countries with low coverage rates, in order to gain an accurate picture of school system success and how this has changed over time, it becomes necessary to utilise (1) household survey data on enrolment and progression in conjunction with (2) PISA data on achievement for those in the sampling frame. Doing so allows one to disaggregate the total number of 15-16 year-old youth into four mutually-exclusive categories:

1. Those who never enrol in school,

2. Those who enrol in school initially but dropout or leave school with a primary school diploma before they are 15 years old,

3. Those who enrol and remain in school but do not reach Grade 7 by the time they are 15-16 years old,

4. Those who enrol remain in school and reach at least Grade 7 by the time they are 15-16 years old (i.e. those who make it into the PISA sampling frame).

In order to do this one needs two data-sources: (1) a measure of educational enrolment and attainment by age, and (2) a measure of learning outcomes at a particular age (15-16 years of age in the case of PISA). The reason for this is that cross-national studies of educational achievement such as PISA typically use schools as the sampling unit and thus cannot provide information on the out-of-school population. Information on topics such as non-enrolment, progression and dropout by gender, and particularly socio-economic status, are also not usually available in government administrative data (see the online appendix of Spaull and Taylor, 2015 for a full discussion). In a similar vein, household surveys such as the DHS have only rudimentary assessments of learning outcomes, or none at all. Combining data from household surveys and international assessments allows one to combine the benefits of each survey into a more meaningful metric of school-system success.

The present analysis builds on earlier work of mine where I have applied the same method to 11 African countries using Southern and Eastern African Consortium for Monitoring Education Quality (SACMEQ) data (Spaull and Taylor, 2015; Taylor and Spaull, 2015), as well as the conceptual insights of prominent scholars in the field (Filmer, Hasan, and Pritchett, 2006; Pritchett, 2013). There are four main stages in this process:

Step one - Access/attainment:

Calculate the proportion of 15-16 year-olds in DHS that have (a) never enrolled, (b) left school (dropped out or received a primary school diploma), (c) are delayed and currently in Grade 6 or below, and (d) are currently in Grade 7 or above. The following DHS variables are required to conduct this analysis (variable names in brackets): age (*HV105*), gender (*HV104*), highest education level (*HV106*), highest grade in highest level (*HV107*), current enrolment (*HV121*). As well as the following weighting variables: sample weight (*HV005*), primary sampling unit (*HV021*) and strata[6] (*HV023*). All estimates were calculated for the three sub-groups of gender (boys and girls), socio-economic status (Poorest 40%, Middle 40%, and richest 20%), and a gender-socio-economic-status interaction (i.e. Poorest 40% who are also girls, Poorest 40% who are also boys, Middle 40% who are also girls, etc.).

Step two - Learning outcomes:

Calculate the proportion of PISA students that attain some recognised benchmark in Reading and Mathematics. For the present estimates of learning outcomes we use PISA proficiency Level 2 in Reading and Mathematics. Level 2 is defined by the OECD as the baseline level of proficiency for all domains (OECD, 2014a, p. 69), what we refer to here as 'functional literacy' and functional numeracy' respectively. For example, in Reading the OECD (2010, p. 13) note that: "Level 2 can be considered a baseline level of proficiency at which students begin to demonstrate the Reading competencies that will enable them to participate effectively and productively in life. Students who do not reach Level 2 have difficulties locating basic information that meets several conditions, making comparisons or contrasts around a single feature, working out what a well-defined part of a text means when the information is not prominent, or making connections between the text and outside knowledge by drawing on personal experiences and attitudes." We calculate the proportion of students reaching Level 2 for the three sub-groups identified above. To do so we use the STATA programme *repest* (Avvisati and Keslair, 2014) with five plausible values for each year and each domain. The programme was designed specifically for use with OECD data, including PISA, and produces estimates that account for the complex survey design of PISA by using replicate weights.

Step three - Combining access and quality:

Once there are estimates of both learning outcomes for 15-16 year-olds (PISA) and an indicator of the proportion of students that were eligible for the PISA sampling frame (using DHS), one can calculate 'access-to-literacy' and 'access-to-numeracy' rates. These are the proportions of a cohort of all youth aged 15-16 years that achieve Level 2 in PISA Reading and Mathematics respectively. If one assumes that students that are ineligible for PISA (either due to non-enrolment, dropout, graduating with a primary qualification, or extremely delayed progression) and would not have achieved Level 2 in either Reading or Mathematics, then one can simply multiply the PISA Literacy Rate (Level 2 or above) with the proportion of 15-16 year-olds that were eligible for the PISA sampling frame (taken from DHS). For example in 2003, of the students that took the PISA Reading test, 63% achieved Level 2 or above. Yet according to the Turkish DHS 2003, only 45% of 15-16 year-olds were actually eligible for PISA in the first place. Thus the access-to-literacy rate would be 28% (63% x 45%). That is to say that when looking at the population of 15-16 year-olds as a whole (both in and out of school); only 28% would have reached PISA Level 2 or higher in 2003. In all cases standard errors are reported for both PISA estimates and DHS estimates (Tables 2-9). For the combined statistic of access-to-literacy and access-to-numeracy the two standard errors need to be combined in some way. Given that PISA and DHS use different samples in their surveys, the samples are independent and thus the composite standard errors are simply the square root of the sum

6 For DHS 2008 the 'HV023' variable is missing and the strata must be created using unique values of 'HV005' which creates the 39 strata identified in the official DHS 2008 report (Hacettepe-UIPS, 2009, p. 209).

of the squared standard errors of each sample, i.e. $(SE1_{DHS}^2 + SE2_{PISA}^2)^{1/2}$ where the first element is the standard error from DHS and the second element the standard error from PISA.

Step four – Correcting for differential access by socio-economic status:

Step three above illustrates how to calculate access-to-literacy at the country level, and can also be applied to males and females separately. However, existing literature has shown that there are large wealth-based differentials in both learning outcomes (OECD, 2012b) and attainment (Filmer and Pritchett, 1999). Therefore one cannot simply multiply the literacy rate for the poorest 40% of students in PISA with the access rate of the poorest 40% of 15-16 year-olds in the DHS because they do not represent the same underlying population. The poorest 40% of 15-16 year-olds in the DHS represent the poorest 40% of 15-16 year-olds *in the country*. Whereas the poorest 40% of the PISA sample represent the poorest 40% of 15-16 year-olds *that were still in school and in Grade 7 or higher* (i.e. eligible for PISA). These are two, sometimes very different, groups.

This is made clearer using a hypothetical example (see Spaull and Taylor, 2015 for a similar example). Assume that there are 1,000 15-16 year-olds and that the national percentage that is eligible for PISA is 85%, and that this proportion differs by socio-economic status. Among the poorest 400 15-16 year-olds only 300 are eligible for PISA (75% eligibility rate), among the middle 400 15-16 year-olds only 350 are eligible for PISA (87.5% eligibility rate), and among the wealthiest 200 15-16 year-olds all are eligible for PISA (100% eligibility rate). Under these conditions, the national percentage of 15-16 year-olds that are eligible for the PISA sampling frame would be 85% ([300 + 350 + 200]/1,000). If one were to ignore the difference in PISA eligibility and simply calculate the poorest 40%, middle 40% and richest 20% of 15-16 year-olds using the 850 students that actually make it into the PISA sampling frame, one would get categories of 340 students in the poorest 40%, 340 students in the middle 40%, and 170 students in the richest 20% category. However, only 300 of the 850 students are actually from the poorest 40% of households, not 340, and 200 students come from the richest 20% of households, not 170. Thus if one did not apply a correction to account for the difference in PISA eligibility by wealth the result would overestimate the access of the poorest 40% of 15-16 year-olds.

To overcome the above issue we apply a correction that amounts to a reweighting of the PISA sample based on the PISA-eligibility of each group found in DHS. This creates comparable groups in DHS and PISA. Analogously, this can be thought of as calculating new socio-economic-status-grouping cut points in the PISA sample to match those in the DHS sample. We first order the distribution of students in PISA by the 'family wealth' variable created by the OECD (2014b, p. 316)and then split this distribution according to the PISA-eligibility rates of each wealth group found in the DHS[7]. This can be shown mathematically by the following formula:

7 For Turkey DHS 2003 and 2013 we use the wealth quintiles created by DHS (HV270). For DHS 2008 the variable for wealth quintiles (HV270) significantly underperforms a standard asset index created using Multiple Correspondence Analysis (MCA) on 22 possession variables SH144B-SH144Y which includes a gas/electric oven, microwave, blender, dishwasher, LCD TV, Camera, Laptop, Internet, Cellular Telephone, Car, Air Conditioner etc. (Hacettepe-UIPS, 2009). This is most likely because the DHS asset index (HV271) on which the DHS wealth quintiles (HV270) are based, is discrete rather than continuous in 2008. We believe it is preferable to use the DHS' own wealth quintiles where possible which is why we used these in 2003 and 2013 but this is inadvisable for 2008. As a sensitivity check we also created the MCA asset index for 2003 and 2013 to see whether 15-16 year olds were classified differently in each wealth group (40:40:20) whether using the DHS' wealth index or the MCA index. In 2003 90% of students were classified in the same category irrespective of wealth-quintile method. In 2013 the figure is 92% of students.

$$Total\ PISA\ sample = \int_{0}^{\left(\frac{PER_{poor40}}{0.4*PER_{total}}\right)} CN_{ses} + \int_{\left(\frac{PER_{poor40}}{0.4*PER_{total}}\right)}^{\left(\frac{PER_{mid40}}{0.4*PER_{total}}\right)} CN_{ses} + \int_{\left(\frac{PER_{mid40}}{0.4*PER_{total}}\right)}^{\left(\frac{PER_{rich20}}{0.2*PER_{total}}\right)} CN_{ses}$$

Where PERpoor40 is the PISA Eligibility Rate for group 1 (poorest 40%), PERmid40 is the PISA Eligibility Rate for group 2 (middle 40%), PERrich20 is the PISA Eligibility Rate for group 3 (richest 20%), PERtotal is the national PISA Eligibility Rate. These eligibility rates are calculated from the DHS sample as the percentage of 15-16 year-olds that are still enrolled in Grade 7 or higher. The variable CNses is the cumulative distribution of weighted 15-16 year-olds in PISA sorted from poorest to wealthiest, using the PISA Student Population Weight (W_FSTUWT). The first integral represents the PISA students that correspond to the poorest 40% of students from DHS, the second integral represents the PISA students that correspond to the middle 40% of the DHS population and the last integral represents the PISA students that correspond to the richest 20% of the DHS population.

Assumptions about the performance of PISA-ineligible students

The above relies on a key assumption that PISA-ineligible students would not have achieved Level 2 or higher if they were actually assessed. We believe that both the quantitative and qualitative literature suggests that this is a reasonable assumption. Many students who dropout do so because they have failed previous grades or repeated grades multiple times (Lewin, 2007). Those who drop out due to income constraints or remoteness are also statistically less likely to be in the better performing part of the distribution before dropout (Lewin, 2007). Finally, given that many of the students who are eligible for PISA do not reach Level 2, it is unlikely that those who have dropped out would be a better performing group than their peers who remain in school. An important fact about Turkey over the 2003-2012 period is that there was only 8 years of compulsory schooling and that those who completed the 8 years were granted primary school degrees. In Turkey a number of youth leave school *after* receiving their primary school qualification, either to start work or, frequently in the case of girls, to get married (Celik and Lüküslü, 2010). Even in this situation there are selection effects in terms of who stays in schooling, who leaves to work, and who leave to get married.

The existing academic literature supports these arguments and shows that there are strong selection effects involved in determining who goes to school and who stays in school, with the 'strongest' (i.e. the wealthiest, most advantaged, and most able) remaining in the schooling system and progressing on time (Lambin, 1995; Lewin and Little, 2011; Lewin, 2007). The DHS data also confirm this. Among the richest 20% of 15-16 year-olds in Turkey 2003, only 8% had dropped out of school compared to 49% among the poorest 40% of 15-16 year-olds (see Table 4). We believe it is reasonable to assume that there is a monotonic relationship between performance and an ordinal measure of educational status. That is to say that, on average and in the large majority of cases, those who are delayed would perform worse on PISA (if they were assessed) than those who are not delayed; those who dropout would perform worse than those who are merely delayed; and those who never enrol in school are likely to perform worst of all.

Given that PISA uses age-based sampling criteria rather than grade-based criteria, it becomes possible to test this assumption empirically. Do students that are 1, 2 or 3 grades delayed perform worse on PISA than those that do not? Looking specifically at Turkey we find that delayed students that are in PISA perform significantly worse than those that are not delayed. In the PISA 2003 Reading assessment, 63% of non-delayed students (i.e. those with the modal grade) achieved Level 2 or higher compared to 13% of those who were delayed by two or three years relative to the modal grade. We return to this issue later in the paper when conducting robustness checks.

Household survey data required for measures of access and attainment: Demographic and Health Survey (DHS) data

For the estimates on access, attainment and progression we use the Turkish Demographic and Health Survey (DHS) data from 2003, 2008 and 2013. The 2003 data was available from the DHS Program website[8]. The 2008 and 2013 DHS data are only available on the Hacettepe University Institute of Population Studies website[9]. The DHS are an important source of data for public health and social science research and are widely used in both fields, including in education. In our previous work (Spaull and Taylor, 2015, p. 140), we have discussed some of the benefits and limitations of using the DHS data over other sources. Some of these benefits include:

1. Self-reported enrolment and grade completion rates are often more accurate than administrative records, the quality of which varies widely between countries (UNESCO Institute for Statistics, 2010). Unlike country-specific administrative data, the uniformity of the surveys means that DHS data are usually more comparable across countries and over time.

2. They can be linked with household characteristics like socio-economic status (Filmer and Pritchett, 2001) and not simply gender — which is one of the main limitations of administrative data.

3. Grade completion rates by age, which is one of the PISA sampling eligibility criteria, cannot be calculated reliably using administrative education data.

Some of the limitations include the introduction of sampling errors, household non-response, and exclusion of homeless children from the sampling frame, measurement error and problems with capturing school attendance. However, given that DHS data have been used in hundreds of peer-reviewed academic publications for a variety of purposes, including educational attainment (Filmer and Pritchett, 1999) and enrolment (Hanushek and Woessmann, 2008), we do not believe that any of these problems outweigh the serious limitations of the alternatives.

The DHS is stratified using a combination of first-level NUTS[10] regions, urban and rural variables, a sub-divided Istanbul (slum and non-slum), and separate strata for metropolitan cities. This creates 40 strata in DHS 2003 (Hacettepe-UIPS, 2003, p. 169), and 39 strata in both DHS 2008 (Hacettepe-UIPS, 2009, p. 209), and DHS 2013 (Hacettepe-UIPS, 2014, p. 197). All estimates presented in this paper have been weighted and stratified in accordance with the various DHS reports[11] using STATA's *svy* command for complex two-stage sampling.

The main aim of using a household survey data for the present analysis is to gain information on 15 year-olds that are ineligible for PISA, either due to delayed grade progression or because they are not currently enrolled in school. One problem with using DHS data for this purpose is the number of 15 year-olds included in the survey. The smaller the number, the less representative the data is and the larger the standard errors associated with each estimate. To overcome this problem we have included both 15 year-olds and 16 year-olds in the analysis. We do not believe that this is especially problematic given that PISA

8 Access was formally requested and granted under Project Number: 906922.

9 Formal access was granted to both data sets on the 17th of May 2016.

10 The 'Nomenclature of Territorial Units for Statistics' (NUTS) is a statistical classification of regions that is used by member countries of the European Union and was introduced in Turkey in 2002.

11 In DHS 2003 and 2013, strata are identified in variable 'HV023'. For DHS 2008 the 'HV023' variable in missing and the strata must be created using unique values of 'HV005'.

does actually sample some 16 year-olds in addition to 15 year-olds[12]. In Turkey the number of 16 year-olds in the PISA sample ranges from 27% (in 2009 and 2012) to 36% (in 2003). By including two age-years one can both increase representivity and reduce the standard errors. Furthermore, the estimates from a sensitivity analysis using only 15 year-olds does not differ substantially from those presented here. National point estimates, as well as those for boys and girls, differ by less than three percentage points when comparing 15 year-olds only and 15 and 16 year-olds together.

The total number of 15-16 year-olds and the number of unique households from which they came were as follows: In DHS 2003 there were *1,982* 15-16 year-olds from 1,800 unique households. In DHS 2008 there were *1,780* 15-16 year-olds from 1,629 unique households. In DHS 2013 there were *1,646* 15-16 year-olds from 1,535 households. Throughout this paper I use the *de jure* population of 15-16 year-olds found in the household survey, i.e. those who usually live in the selected household, rather than the *de facto* population, i.e. those who spent the previous night in the household (Hacettepe-UIPS, 2003, p. 17). This is one way to further increase the sample size since the *de jure* population is slightly larger than the *de facto* population.

It is worth highlighting one important feature of education statistics in Turkey[13], and one that helps explain the discrepancy between DHS estimates and OECD estimates of sample eligibility. Up until 2007 the Ministry of National Education (MoNE) used age-district level population projections based on Census 2000. From 2007 the Ministry of Interior introduced the Address Based Population Registry System (Turkish abbreviation: ADNKS). Once the MoNE started using ADNKS to compute enrolment rates it realised that they were drastically over-estimating the population of school-age children between 2000 and 2006. Given that the OECD uses the population and enrolment estimates provided by country staff this is one of the explanations for the difference between the DHS estimates and the Coverage Index 3 estimates in 2003. Consequently, the CI-3 statistics are not strictly comparable over time for Turkey, while the DHS estimates should be comparable. A further reason why the OECD estimates might be lower than the DHS estimates is that the Coverage Index 3 statistic reports the rate of coverage after sample exclusions have been taken into account. For example, if certain schools were ineligible for the sample because of conflict or because they were too remote, this would be reflected in the OECD estimates but not in the DHS estimates.

RESULTS

The results of the analysis are presented in Figures 2-7 and Tables 2-9 below. In each of these figures the percentage of 15-16 year-olds in each wave of the DHS are broken up into five mutually exclusive categories: (1) those who never enrol and are therefore ineligible for the PISA sampling frame, (2) those who have dropped out of school and are also ineligible, (3) those who are still in school but are enrolled in Grades 1-6 and therefore ineligible for PISA, (4) those who are still enrolled in school and in Grade 7 or

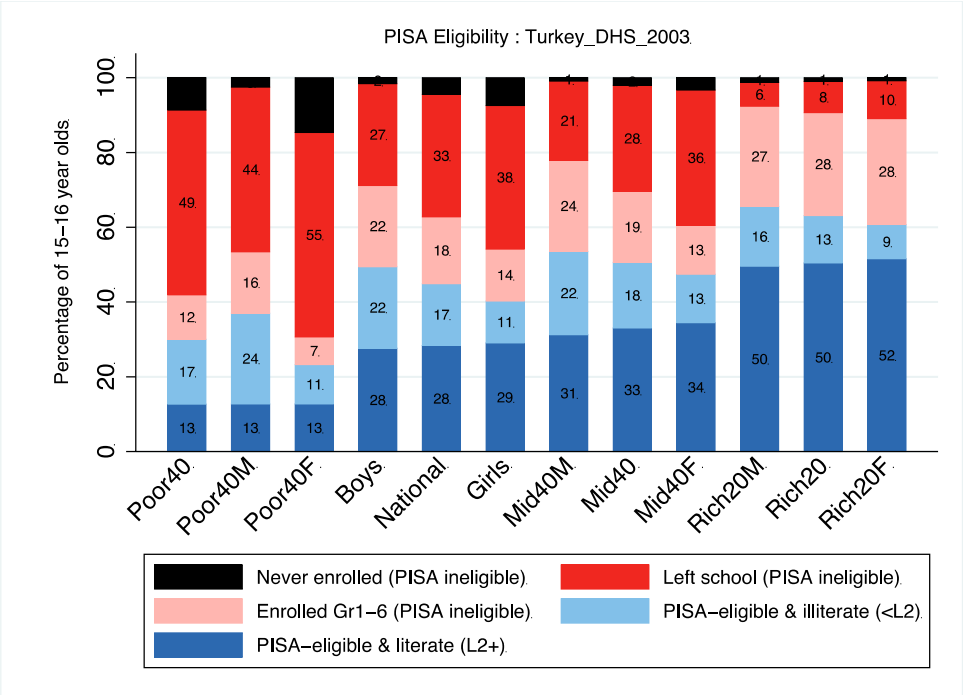
12 The sampling frame for the PISA assessment covers students that are aged 15 years and 3 (completed) months to 16 years and 2 (completed) months who are currently enrolled in Grade 7 or higher (OECD, 2005, p. 46).

13 Many thanks to Alper Dinçer for explaining this to me and for his helpful comments on the Turkish education system more generally.

higher (i.e. eligible for PISA) but who did not reach proficiency Level 2 (referred to as functionally illiterate and functionally innumerate for Reading and Mathematics), and (5) those who are eligible for PISA and *did* reach Level 2 (referred to as functionally literate and functionally numerate). This last group is said to have "access-to-literacy" or "access-to-numeracy" respectively. The tables provide the point estimates and standard errors for the percentage of 15-16 year-olds in DHS that were enrolled in Grade 7 or higher and were thus eligible for PISA (Table 2), those who never enrolled in school (Table 3), those who left school (due to drop out or because they had acquired their primary school qualification) (Table 4), and those who were still enrolled in Grades 1-6 (Table 5). These last three groups are all ineligible for PISA. Table 6 and Table 7 report the percentage of the PISA sample that acquired functional literacy (Level 2 Reading) and functional numeracy (Level 2 Mathematics). These PISA estimates are the 'uncorrected' estimates that match to those reported in the official OECD publications. Table 8 and Table 9 report the percentage of students that have 'access-to-literacy' and 'access-to-numeracy' respectively. These are the percentage of all 15-16 year-olds in the country that were still enrolled in Grade 7 or above and also attained PISA proficiency Level 2 in Reading and Mathematics respectively. As has been mentioned earlier, these estimates assume that those who are ineligible for PISA – because they are no longer in school or are extremely delayed – would not have reached Level 2 if they were assessed.

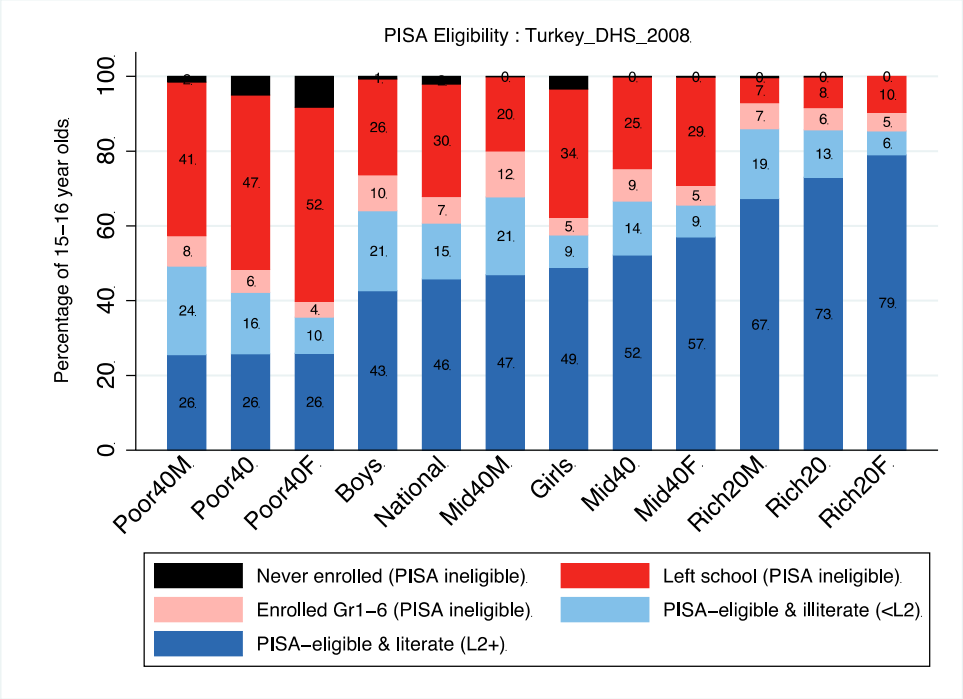
In each Table and graph estimates are provided for different sub-groups of the population as well as the country as a whole. The three groupings are gender (boys and girls), socio-economic status (poorest 40%, middle 40% and wealthiest 20%), and a socio-economic-status and gender interaction (poorest 40% that are also girls – Poor40F - the poorest 40% that are also boys – Poor40M – and so on). The standard errors for each category can be found in the tables that follow the graphs.

Figure 2. Access to literacy (Level 2) in Turkey 2003 (PISA 2003 and DHS 2003)



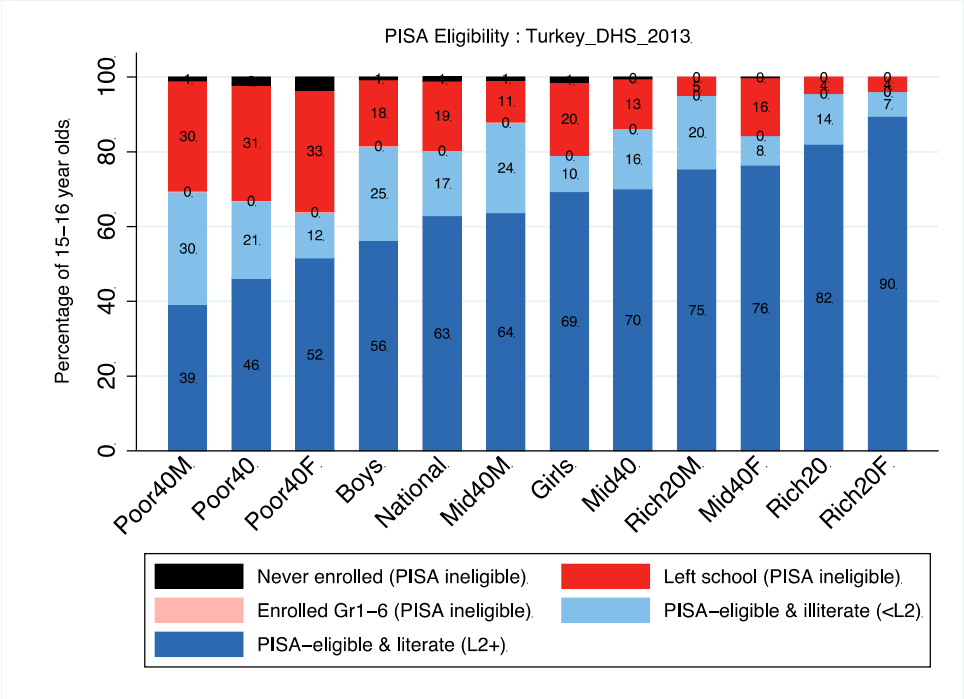
Source: Author's own calculations from PISA and DHS datasets

Figure 3. Access to Literacy (Level 2) in Turkey 2009 (PISA 2009 and DHS 2008)



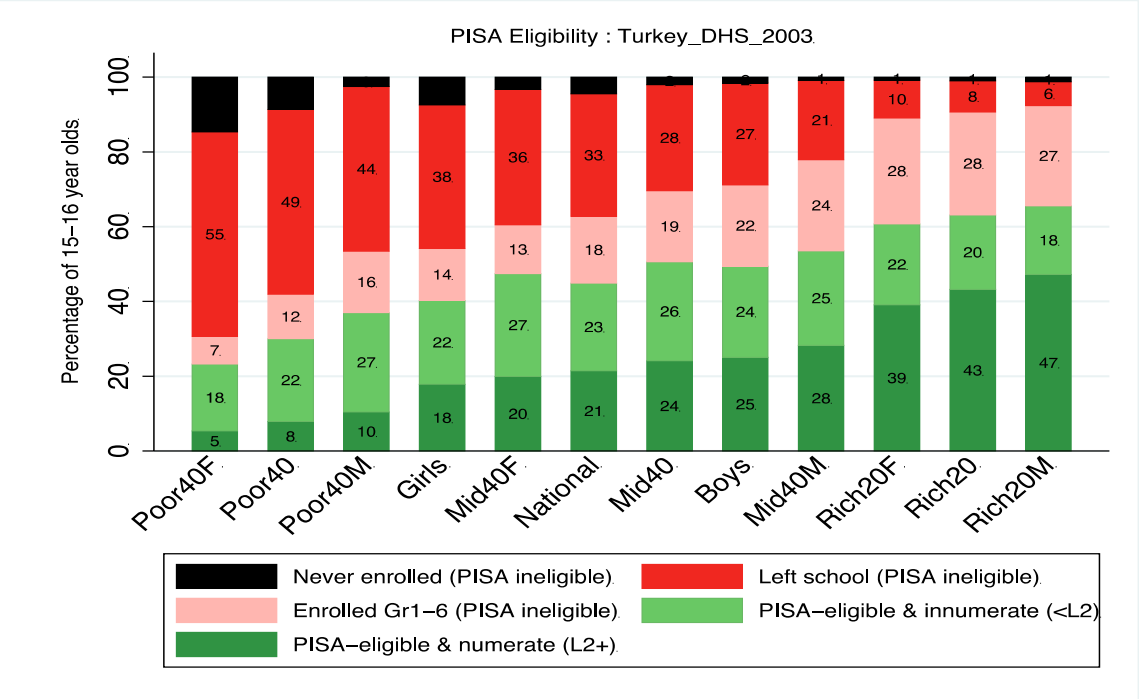
Source: Author's own calculations from PISA and DHS datasets

Figure 4. Access to Literacy (Level 2) in Turkey 2012 (PISA 2012 and DHS 2013



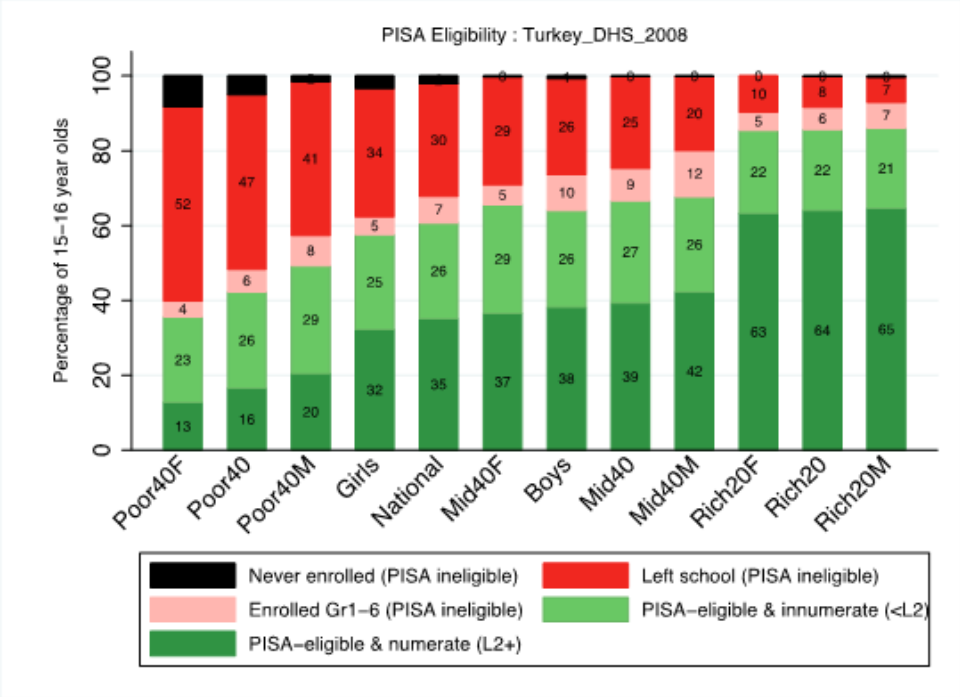
Source: Author's own calculations from PISA and DHS datasets

Figure 5. Access to Numeracy (Level 2) in Turkey 2003 (PISA 2003 and DHS 2003)



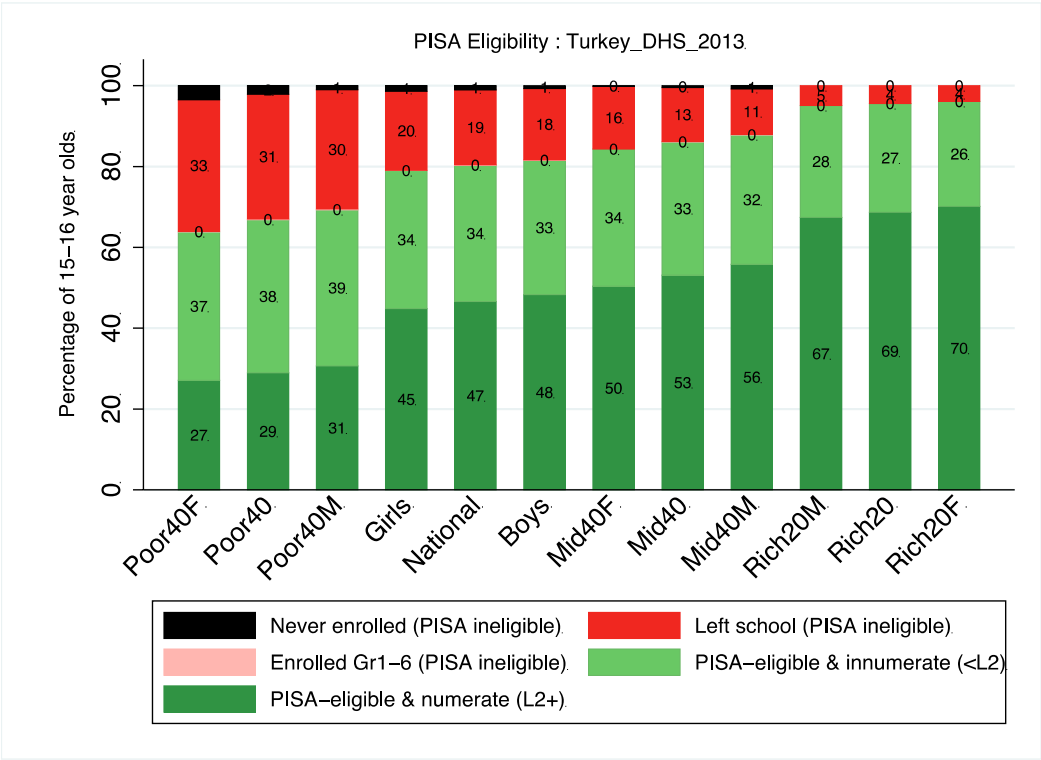
Source: Author's own calculations from PISA and DHS datasets

Figure 6. Access to Numeracy (Level 2) in Turkey 2009 (PISA 2009 and DHS 2008)



Source: Author's own calculations from PISA and DHS datasets

Figure 7. Access to Literacy (Level 2) in Turkey 2012 (PISA 2012 and DSH 2013)



Source: Author's own calculations from PISA and DHS datasets

Table 2: Percentage of 15-16 year-olds in DHS enrolled in Gr7 or higher (PISA eligible), with Standard Errors (%)

Country	National	*SE*	Males	*SE*	Females	*SE*	Poor40	*SE*	Mid40	*SE*	Rich20	*SE*
Turkey 2003	44,9	*1,5*	49,4	*1,9*	40,3	*2,0*	30,0	*1,9*	50,6	*2,4*	63,1	*3,0*
Turkey 2008	60,7	*1,7*	64,1	*2,0*	57,5	*2,5*	42,2	*2,3*	66,6	*2,7*	85,7	*2,2*
Turkey 2013	80,3	*1,4*	81,5	*1,6*	79,0	*1,9*	66,8	*2,3*	86,1	*1,9*	95,6	*1,7*

	Poor40M	*SE*	Poor40F	*SE*	Mid40M	*SE*	Mid40F	*SE*	Rich20M	*SE*	Rich20F	*SE*
Turkey 2003	37,0	*2,7*	23,3	*2,5*	53,5	*2,9*	47,4	*3,3*	65,5	*3,9*	60,7	*4,3*
Turkey 2008	49,2	*3,1*	35,6	*3,1*	67,7	*3,1*	65,6	*3,9*	86,0	*3,1*	85,4	*3,6*
Turkey 2013	69,3	*2,7*	63,8	*3,2*	87,8	*2,3*	84,3	*3,0*	95,1	*2,5*	96,1	*2,3*

Table 3: Percentage of 15-16 year-olds in DHS who never enrol in school (PISA ineligible), with Standard Errors (%)

Country	National	*SE*	Males	*SE*	Females	*SE*	Poor40	*SE*	Mid40	*SE*	Rich20	*SE*
Turkey 2003	4,5	*0,6*	1,6	*0,4*	7,4	*1,1*	8,7	*1,3*	2,0	*0,6*	1,0	*0,6*
Turkey 2008	2,1	*0,3*	0,7	*0,2*	3,4	*0,6*	5,0	*0,8*	0,2	*0,1*	0,2	*0,2*
Turkey 2013	1,0	*0,3*	0,7	*0,3*	1,4	*0,4*	2,1	*0,6*	0,5	*0,4*	0,0	*0,0*

	Poor40M	*SE*	Poor40F	*SE*	Mid40M	*SE*	Mid40F	*SE*	Rich20M	*SE*	Rich20F	*SE*
Turkey 2003	2,5	*0,7*	14,7	*2,3*	0,9	*0,6*	3,3	*1,0*	1,2	*1,0*	0,8	*0,8*
Turkey 2008	1,6	*0,6*	8,3	*1,5*	0,1	*0,1*	0,2	*0,2*	0,4	*0,4*	0,0	*0,0*
Turkey 2013	1,0	*0,5*	3,5	*1,1*	0,8	*0,7*	0,2	*0,2*	0,0	*0,0*	0,0	*0,0*

Note: Poor40 refers to the poorest 40% of individuals in the country, Mid40 the middle 40% and Rich20 the richest 20%. 'SE' is the standard error. An 'M' or 'F' after the wealth bracket refers to 'Males' or 'Females', respectively. Values shown are percentages

Table 4: Percentage of 15-16 year-olds in DHS who have left school (due to drop out or matriculation) (PISA ineligible), with Standard Errors (%)

Country	National	*SE*	Males	*SE*	Females	*SE*	Poor40	*SE*	Mid40	*SE*	Rich20	*SE*
Turkey 2003	32,8	*1,5*	27,2	*1,8*	38,4	*2,1*	49,4	*2,2*	28,4	*2,3*	8,3	*1,7*

Turkey 2008	30,1	*1,6*	25,6	*1,8*	34,3	*2,3*	46,6	*2,1*	24,5	*2,4*	8,2	*1,9*
Turkey 2013	18,5	*1,3*	17,6	*1,6*	19,5	*1,8*	30,9	*2,2*	13,3	*1,9*	4,4	*1,7*
	Poor40M	***SE***	**Poor40F**	***SE***	**Mid40M**	***SE***	**Mid40F**	***SE***	**Rich20M**	***SE***	**Rich20F**	***SE***
Turkey 2003	44,0	*3,0*	54,7	*3,1*	21,2	*2,6*	36,2	*3,3*	6,4	*2,1*	10,1	*2,6*
Turkey 2008	41,1	*2,8*	51,9	*3,1*	19,8	*2,8*	28,9	*3,6*	6,7	*2,5*	9,6	*3,0*
Turkey 2013	29,5	*2,7*	32,5	*2,9*	11,2	*2,2*	15,5	*2,9*	4,9	*2,5*	3,9	*2,3*

Table 5: Percentage of 15-16 year-olds in DHS who are still enrolled in Grades1-6 (PISA ineligible), with Standard Errors (%)

Country	**National**	***SE***	**Males**	***SE***	**Females**	***SE***	**Poor40**	***SE***	**Mid40**	***SE***	**Rich20**	***SE***
Turkey 2003	17,9	*1,0*	21,8	*1,5*	13,9	*1,3*	11,9	*1,5*	19,0	*1,8*	27,6	*2,6*
Turkey 2008	7,1	*0,8*	9,6	*1,2*	4,8	*0,9*	6,1	*1,0*	8,7	*1,4*	6,0	*1,5*
Turkey 2013	0,1	*0,1*	0,2	*0,1*	0,1	*0,1*	0,2	*0,1*	0,1	*0,1*	0,0	*0,0*
	Poor40M	***SE***	**Poor40F**	***SE***	**Mid40M**	***SE***	**Mid40F**	***SE***	**Rich20M**	***SE***	**Rich20F**	***SE***
Turkey 2003	16,4	*2,1*	7,4	*1,6*	24,4	*2,6*	13,1	*1,9*	26,8	*3,6*	28,3	*3,6*
Turkey 2008	8,1	*1,5*	4,2	*1,2*	12,3	*2,1*	5,3	*1,6*	7,0	*2,2*	5,0	*2,3*
Turkey 2013	0,2	*0,2*	0,2	*0,2*	0,2	*0,2*	0,0	*0,0*	0,0	*0,0*	0,0	*0,0*

Note: Poor40 refers to the poorest 40% of individuals in the country, Mid40 the middle 40%, and Rich20 the richest 20%. 'SE' is the standard error. An 'M' or 'F' after the wealth bracket refers to 'Males' or 'Females', respectively. Values shown are percentages.

Table 6: Percentage of the PISA sample that is literate (Level 2+), with Standard Errors (%) - Uncorrected for those who are ineligible

Country	**National**	***SE***	**Males**	***SE***	**Females**	***SE***	**Poor40**	***SE***	**Mid40**	***SE***	**Rich20**	***SE***
Turkey 2003	63,2	*2,4*	55,9	*3,0*	72,2	*2,6*	47,3	*2,5*	69,3	*2,4*	83,0	*2,5*
Turkey 2009	75,5	*1,4*	66,6	*1,9*	85,0	*1,4*	65,0	*2,1*	80,6	*1,4*	86,2	*1,9*
Turkey 2012	78,4	*1,4*	69,1	*2,1*	87,8	*1,0*	71,1	*1,8*	82,1	*1,4*	85,4	*2,0*
	Poor40M	***SE***	**Poor40F**	***SE***	**Mid40M**	***SE***	**Mid40F**	***SE***	**Rich20M**	***SE***	**Rich20F**	***SE***

Turkey 2003	38,3	*3,2*	59,3	*3,4*	63,5	*3,0*	75,8	*2,8*	78,2	*3,4*	88,8	*2,7*
Turkey 2009	54,9	*2,6*	77,0	*2,4*	72,7	*2,0*	88,5	*1,5*	79,6	*2,7*	92,8	*1,8*
Turkey 2012	59,0	*2,7*	82,3	*1,5*	73,6	*2,3*	91,2	*1,2*	78,5	*2,6*	93,2	*1,7*

Table 7: Percentage of the PISA sample that is numerate (Level 2+), with Standard Errors (%) - Uncorrected for those who are ineligible

Country	**National**	***SE***	**Males**	***SE***	**Females**	***SE***	**Poor40**	***SE***	**Mid40**	***SE***	**Rich20**	***SE***
Turkey 2003	47,8	*2,6*	50,7	*2,9*	44,3	*3,1*	31,3	*2,0*	50,7	*2,8*	74,9	*3,2*
Turkey 2009	57,9	*1,8*	59,6	*1,8*	56,0	*2,7*	42,7	*2,1*	63,6	*2,1*	76,6	*2,5*
Turkey 2012	58,0	*1,9*	59,2	*2,2*	56,8	*2,4*	46,1	*2,1*	63,0	*2,3*	71,9	*2,4*
	Poor40M	***SE***	**Poor40F**	***SE***	**Mid40M**	***SE***	**Mid40F**	***SE***	**Rich20M**	***SE***	**Rich20F**	***SE***
Turkey 2003	21,1	*2,5*	12,7	*1,8*	35,3	*2,7*	24,5	*3,4*	54,2	*3,6*	43,3	*3,9*
Turkey 2009	45,0	*2,4*	40,0	*3,1*	66,9	*2,1*	60,4	*3,1*	77,0	*2,8*	76,1	*3,2*
Turkey 2012	47,6	*2,7*	44,7	*2,5*	64,3	*2,5*	61,6	*3,1*	70,4	*2,7*	73,6	*3,4*

Note**:** Poor40 refers to the poorest 40% of 15-16 year olds in the country, Mid40 the middle 40%, and Rich20 the richest 20%. 'SE' is the standard error. An 'M' or 'F' after the wealth bracket refers to 'Males' or 'Females', respectively. Values shown are percentages.

Table 8. Percentage of 15-16 year-olds that have Access-to-Literacy (Level 2+), with Standard Errors (%)

Country	**National**	***SE***	**Males**	***SE***	**Females**	***SE***	**Poor40**	***SE***	**Mid40**	***SE***	**Rich20**	***SE***
Turkey 2003	28,4	*2,8*	27,6	*3,5*	29,1	*3,3*	12,7	*3,2*	33,0	*3,3*	50,4	*4,0*
Turkey 2009	45,8	*2,2*	42,6	*2,7*	48,9	*2,8*	25,8	*3,2*	52,2	*3,1*	73,0	*2,8*
Turkey 2012	62,9	*2,0*	56,3	*2,7*	69,4	*2,2*	46,1	*3,0*	70,1	*2,4*	82,0	*2,5*
	Poor40M	***SE***	**Poor40F**	***SE***	**Mid40M**	***SE***	**Mid40F**	***SE***	**Rich20M**	***SE***	**Rich20F**	***SE***
Turkey 2003	12,7	*4,4*	12,8	*4,5*	31,3	*4,1*	34,5	*4,2*	49,6	*5,2*	51,5	*5,2*
Turkey 2009	25,6	*4,2*	25,9	*4,3*	46,9	*3,8*	57,0	*4,3*	67,3	*4,0*	79,0	*4,0*
Turkey 2012	39,1	*4,1*	51,6	*3,7*	63,7	*3,4*	76,4	*3,2*	75,4	*3,6*	89,5	*2,7*

Table 9. Percentage of 15-16 year-olds that have Access-to-Numeracy (Level 2+), with Standard Errors (%)

Country	National	*SE*	Males	*SE*	Females	*SE*	Poor40	*SE*	Mid40	*SE*	Rich20	*SE*
Turkey 2003	21,5	*3,0*	25,1	*3,5*	17,8	*3,7*	7,9	*3,0*	24,1	*3,3*	43,2	*4,7*
Turkey 2009	35,1	*2,5*	38,2	*2,7*	32,2	*3,7*	16,5	*3,2*	39,3	*3,4*	63,9	*3,2*
Turkey 2012	46,6	*2,4*	48,3	*2,7*	44,8	*3,1*	29,0	*3,1*	53,1	*2,9*	68,7	*3,0*
	Poor40M	*SE*	**Poor40F**	*SE*	**Mid40M**	*SE*	**Mid40F**	*SE*	**Rich20M**	*SE*	**Rich20F**	*SE*
Turkey 2003	10,5	*4,0*	5,4	*3,9*	28,3	*3,9*	20,0	*4,6*	47,2	*5,4*	39,1	*6,2*
Turkey 2009	20,5	*4,0*	12,7	*4,4*	42,2	*3,9*	36,6	*5,0*	64,6	*3,9*	63,3	*4,8*
Turkey 2012	30,7	*4,0*	27,1	*4,1*	55,8	*3,3*	50,4	*4,2*	67,4	*3,8*	70,1	*3,9*

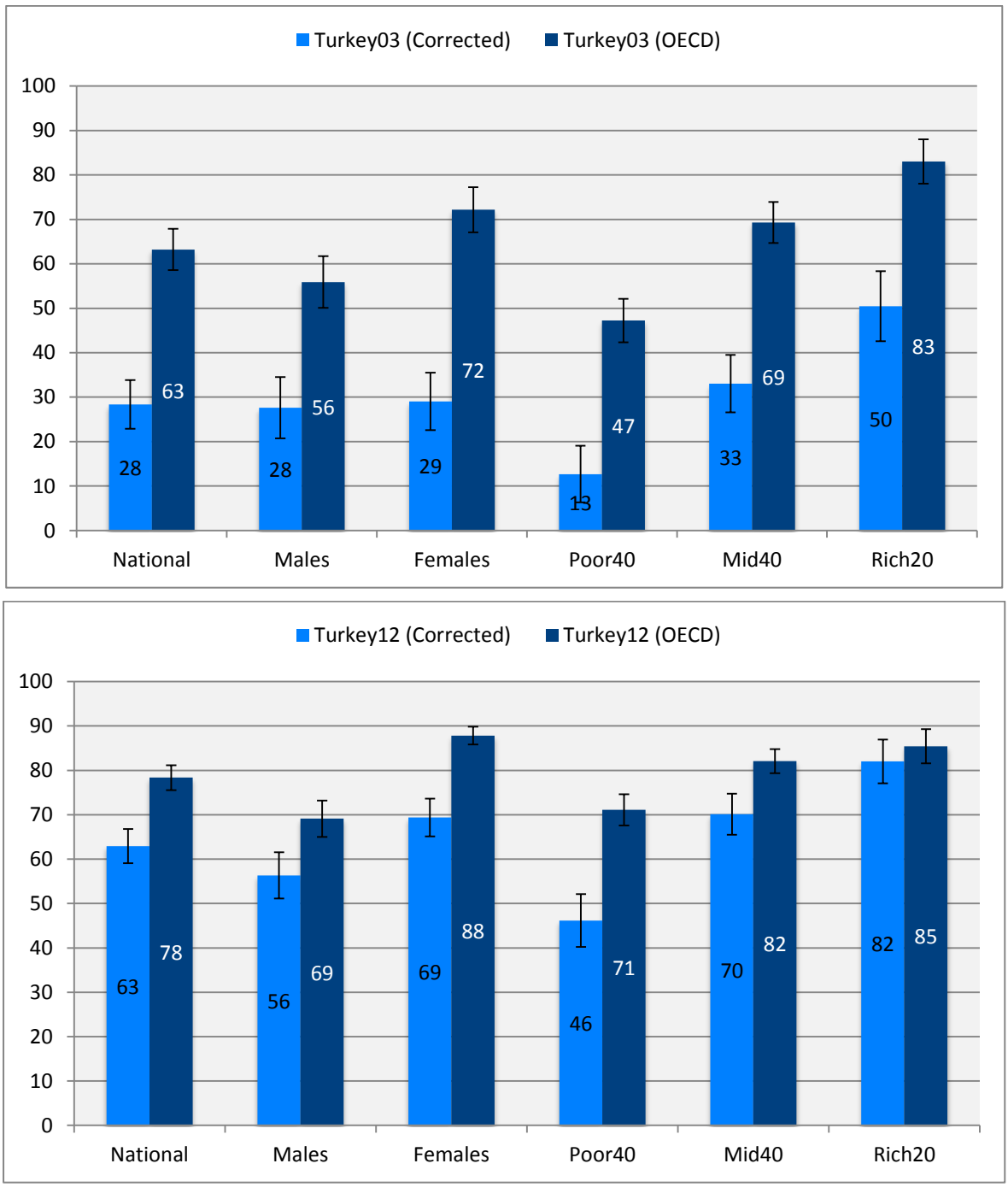
Note: Poor40 refers to the poorest 40% of individuals in the country, Mid40 the middle 40%, and Rich20 the richest 20%. 'SE' is the standard error. An 'M' or 'F' after the wealth bracket refers to 'Males' or 'Females', respectively. Values shown are percentages.

OVERARCHING FINDINGS

If one applies the assumptions outlined above to the OECD PISA data, one can compare the percentage of 15-16 year-olds that are reported to be functionally literate (PISA Level 2 or higher) and functionally numerate (PISA Level 2 or higher) in OECD reports and when one corrects for sample eligibility. By assuming that those who are ineligible for the PISA sampling frame are functionally illiterate and functionally innumerate, one can 'correct' the reported PISA estimates. This is done by incorporating information from household surveys on which students (and how many of them) are ineligible in selected demographic groups. The figures below compare the percentage of 15-16 year-olds that are reported to be literate (Figure 8) and numerate (Figure 9) in 2003 and 2012 in the OECD reports and once corrected for sample ineligibility as described above (these figures are from Table 6/7 and Table 8/9). For example, the OECD reports that 48% of the Turkish PISA sample in 2003 achieved at least Level 2 in Mathematics (OECD, 2004: 354), but once this is corrected for sample eligibility the figure drops to 21% (see 'National' in Figure 9).

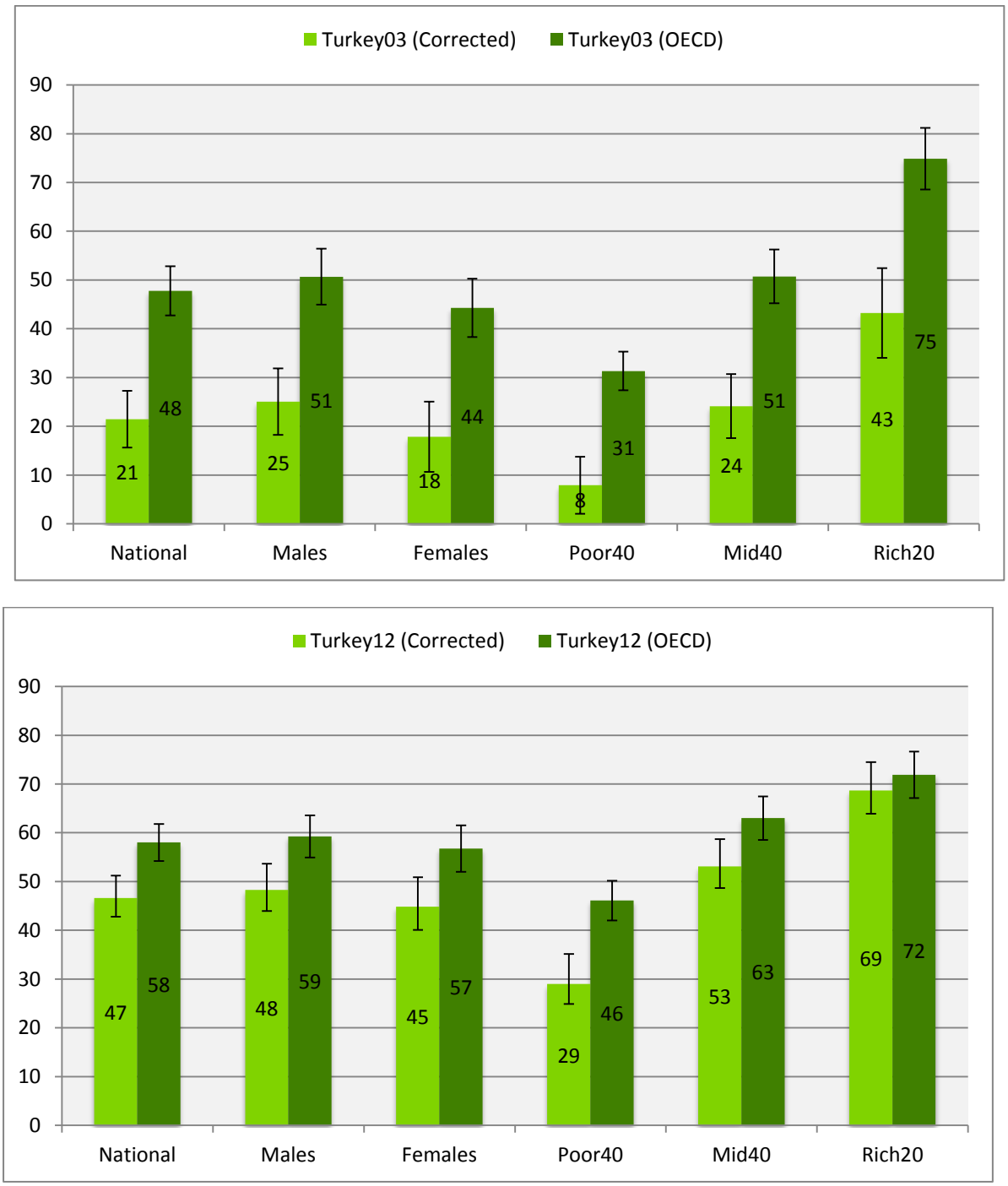
What is most evident in Figure 8 and 9 are the large discrepancies between the OECD estimates (which ignore sample ineligibility) and the corrected figures, particularly in 2003. It is also worth noting that the gap between the traditional estimates and the corrected estimates is largest for the poorest 40% of the population. For example, while the OECD reports that 47% of the poorest 40% of students reached Level 2, once this is corrected for sample eligibility the figure drops to 13%. It could be argued that it is erroneous to compare the 'poorest 40%' estimates of the OECD with the 'poorest 40%' once corrected for sampling eligibility. This is because the 'poorest 40%' in the OECD reports are technically 'the poorest 40% of those that were in the PISA sample.' The reason why the comparison is important and legitimate is that the reports refer to the 'poorest 40%' as though they were the poorest 40% of the population, not of the sample.

Figure 8. Percentage of 15-16 year-olds reported to be literate (L2+) in OECD reports and corrected for sampling eligibility (Access to Literacy rates) in 2003 and 2012



Source: Author's own calculations from PISA and DHS datasets

Figure 9. Percentage of 15-16 year-olds reported to be Numerate (L2+) in OECD reports and corrected for sampling eligibility (Access to Numeracy rates) in 2003 and 2012

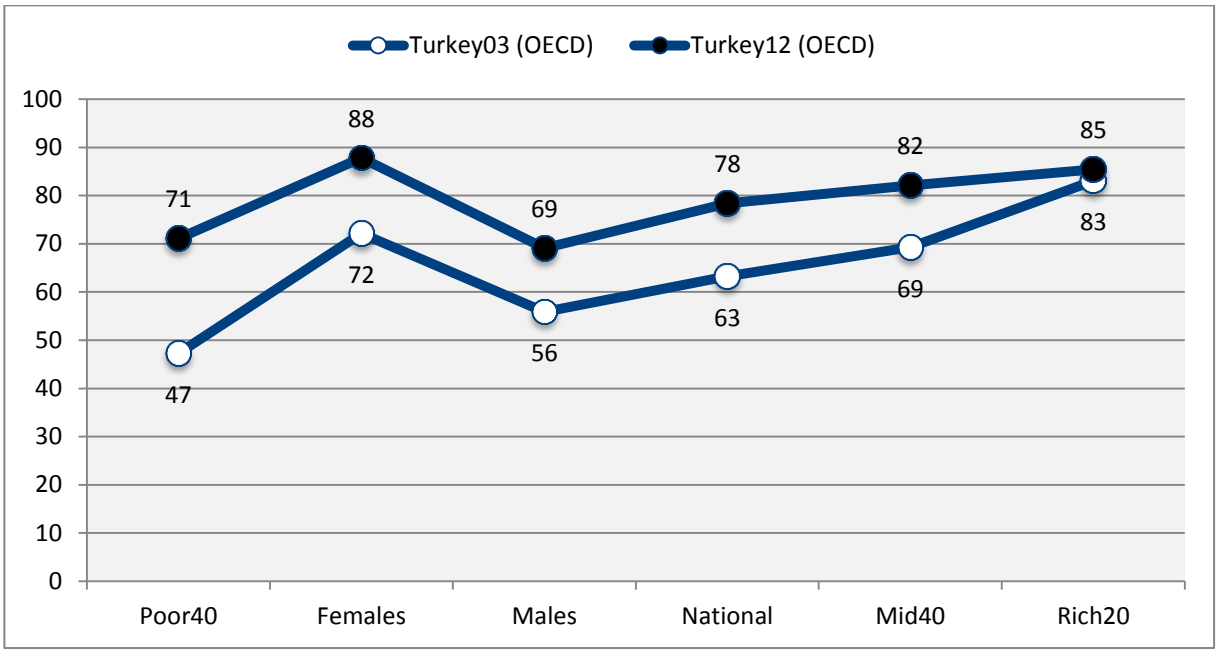


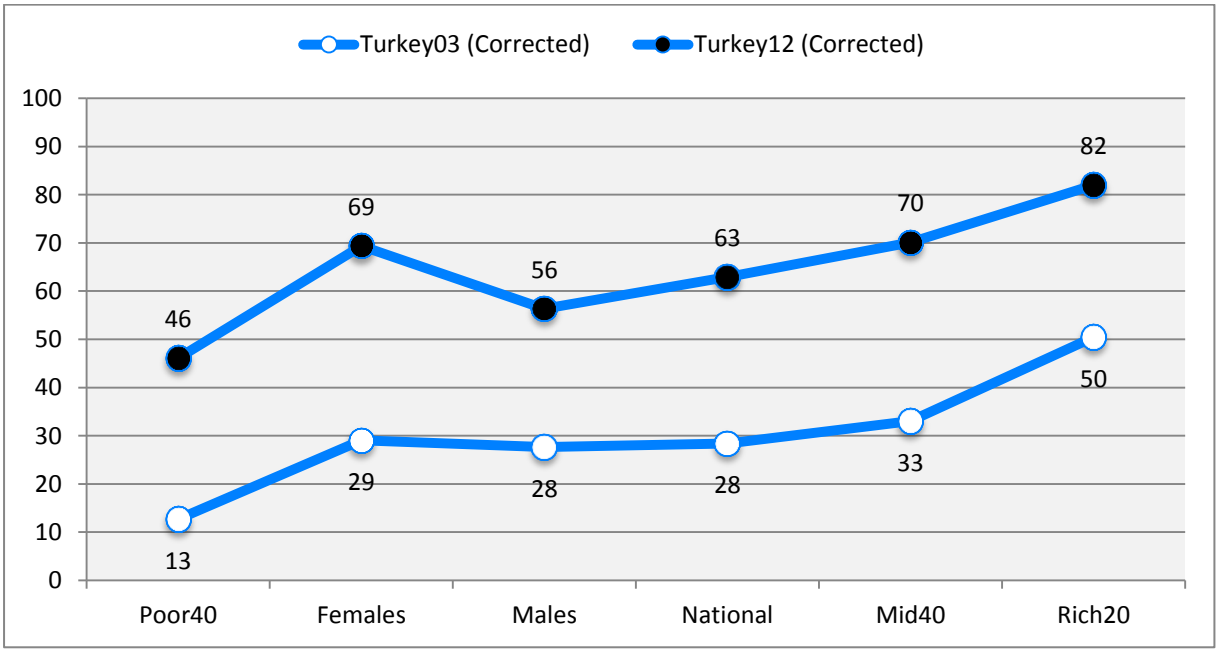
Source: Author's own calculations from PISA and DHS datasets

Figure 10 and 11 below show the changes in percentage of 15-16 year-olds that are functionally literate (Figure 10) and functionally numerate (Figure 11) as reported by the OECD and when corected for sample eleigibility. These graphs clearly show that the increase in functional literacy and numeracy rates have been much greater than was previously reported by the OECD. Between 2003 and 2012 OECD reports indicate that the percentage of 15-16 year-olds reaching Level 2 in literacy increased from 47% to

71% while in reality the percentage increase was five times as large, increasing from 13% in 2003 to 46% in 2012.

Figure 10. Percentage of 15-16 year-olds reported to be literate (L2+) in OECD reports and corrected for sampling eligibility (access to literacy rates) in 2003 and 2012

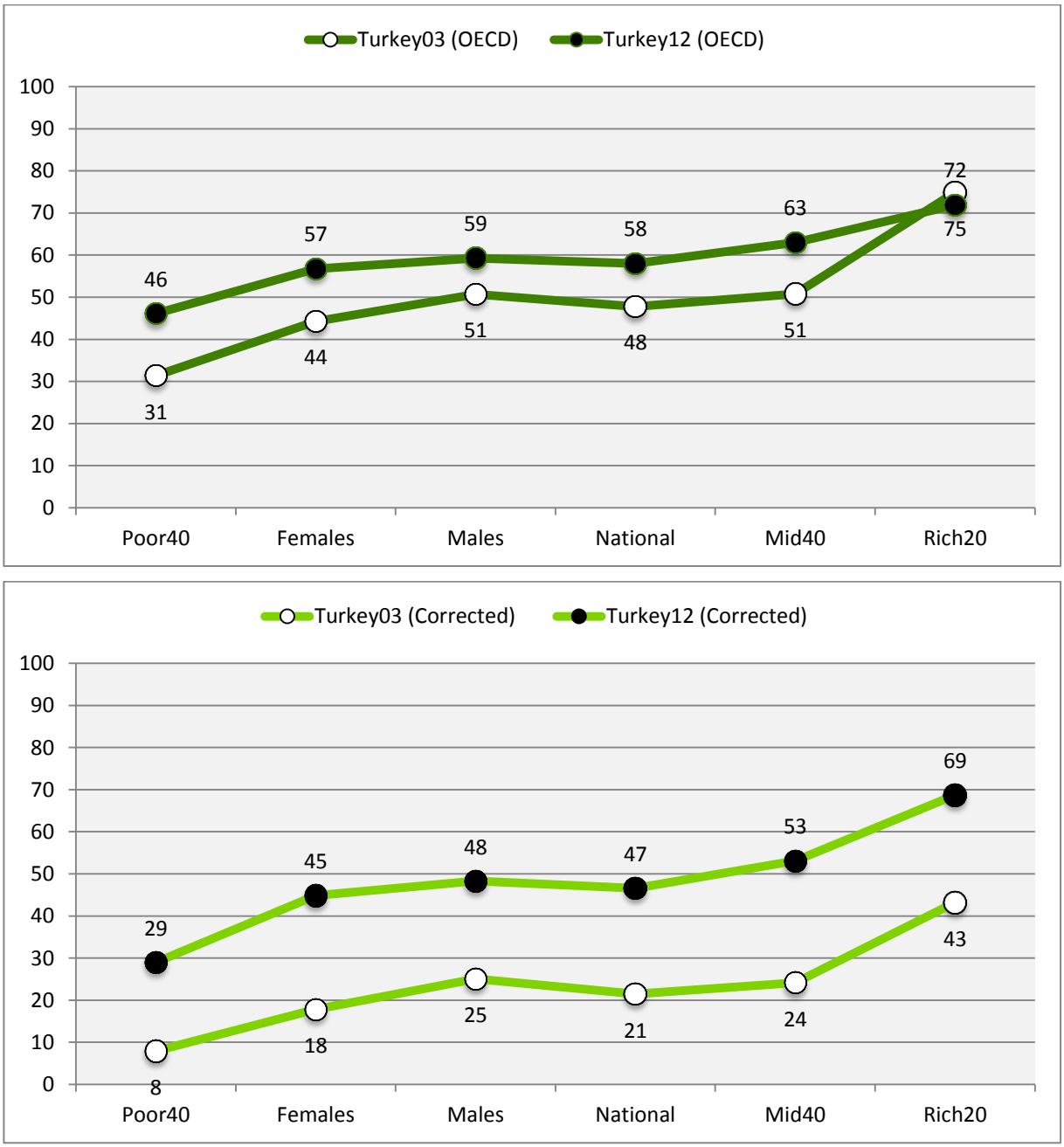




The same trends can be seen in Mathematics, with the exception that girls perform slightly worse than boys (although this is not statistically significant). Again, the increases in functional numeracy rates are orders of magnitude larger than was previously thought or reported in OECD publications. The largest proportional increases in functional literacy and numeracy rates can be found among the poorest 40% of 15-16 year-olds and amongst females. Traditional (uncorrected) estimates of functional literacy and

numeracy also significantly underestimate how much the richest 20% of 15-16 year-olds benefited over the period.

Figure 11. Percentage of 15-16 year-olds reported to be numerate (L2+) in OECD PISA reports and corrected for sampling eligibility (access to numeracy rates) in 2003 and 2012



Source: Author's own calculations from PISA and DHS datasets

Close scrutiny of the estimates provided in the tables and figures above point to nine main findings:

1. **Significant increases in PISA sample eligibility over time**: Across all groups there have been large and statistically significant increases in PISA sample eligibility in Turkey over the 2003-

2012 period. While only 45% of 15-16 year-olds were eligible for PISA in 2003, by 2013 this had nearly doubled to 80% (Figure 1). The largest growth in PISA eligibility was among the poorest 40% of 15-16 year-olds where PISA eligibility increased from 30% to 67% over the period (Table 2). This agrees with the existing literature that highlights the pro-poor policies focussed on access (Aydagül, 2008b; Dülger, 2004).

2. **The improvement in the proportion of 15-16 year-olds reaching Level 2 in PISA is between twice as large (for numeracy) and five times larger (for literacy) than previously thought**: The OECD reports that between 2003 and 2012 the proportion of Turkish students acquiring at least Level 2 in PISA increased from 63% to 78% in literacy and from 48% to 58% in numeracy (OECD, 2014a). However this ignores the large expansion in PISA eligibility in Turkey over the same period. By taking this expansion into account, one can see that access to literacy among all 15-16 year-olds (not only those that are eligible for PISA) actually doubled, increasing from 28% to 63% in literacy and from 22% to 47% in numeracy (Table 8 and 9, Figure 10 and 11).

3. **There are considerably more functionally illiterate and functionally innumerate Turkish 15-16 year-olds than PISA reports indicate**: Because the PISA sampling frame only considers those 15-16 year-olds that are still enrolled in school and in Grade 7 or higher, it ignores the students that are either severely delayed, have dropped out of school, or matriculated and left school already. These students typically perform much worse than their 'on-track' peers (Lewin and Little, 2011). If one assumes that PISA ineligible students would not have reached Level 2 in Reading or Mathematics if they were assessed, one can calculate the 'true' proportion of functionally illiterate and functionally innumerate 15-16 year-olds in the Turkish population (not only in school). This is simply 100 minus the access-to-literacy or access-to-numeracy rate. While the PISA reports indicate that the proportion of Turkish 15-16 year-olds that were functionally illiterate was 37% in 2003 and 22% in 2013 (Table 6), the corrected estimates presented here show that these figures are drastic underestimates. By including the ineligible 15-16 year-olds these figures increase to 72% in 2003 and 37% in 2012. Thus, although there have been large decreases in the proportion of functionally illiterate Turkish 15-16 year-olds, the figures are still much higher than those reported by the OECD (OECD, 2014a).

4. **Severe delays in progression have been virtually eliminated in Turkey**: In 2003 nearly 1 in 5 (18%) 15-16 year-olds in Turkey had severely delayed grade-progression and were still enrolled in Grades 1-6. These students were not eligible for PISA. By 2008 this had dropped to 7% and by 2013 there were essentially no 15-16 year-olds (0.1%) that were severely delayed (Table 5).

5. **Dropout / early school-leaving among 15-16 year-olds declined by 42% over the period**: In 2003 a third (33%) of Turkish 15-16 year-olds had left school, with the highest concentration among the poorest 40% of girls where more than half (55%) had left school before age 15 (Table 4). By 2012 the national figure had dropped from 33% to 19% and among the poorest 40% of girls it had declined from 55% to 33% over the same period. This again would seem to point to the success of government policies focussing on inclusion and gender (Köseleci, 2015; Tansel, 2002).

6. **The rich are much more likely to be sampled in PISA**: In all three waves of PISA the poorest 40% of 15-16 year-olds are statistically significantly less likely to be in the PISA sample than the richest 20%. In 2003, 63% of the richest group were eligible for PISA compared to 30% of the poorest group (Table 2). Thus the rich were twice as likely to be sampled as the poor in 2003. This differential – the ratio between the rich and poor who are eligible for PISA - has been declining from 2.1 (in both 2003 and 2008) to 1.4 (2013), as can be calculated from Table 2. Notwithstanding the above, the rich are still considerably less likely to leave school early than the

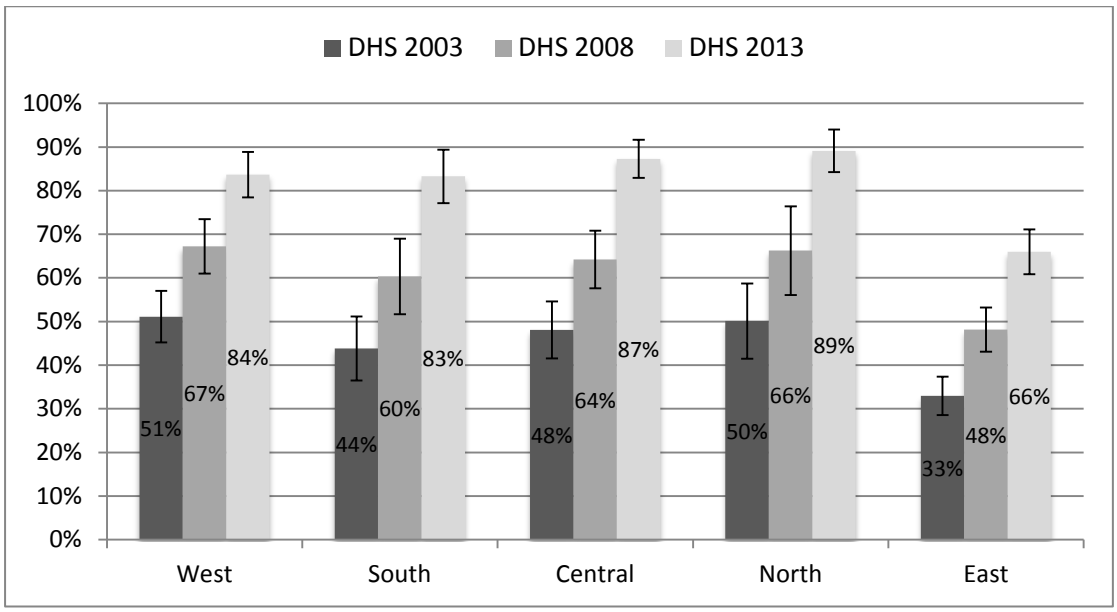
poor are, with only 4% of the richer 15-16 year-olds leaving school early in 2013, compared to 31% among the poor.

7. **The gap in access to literacy and access to numeracy between rich and poor has not changed and is larger than previously thought**: In 2003 wealthy students were four times as likely to have access-to-literacy (50%) compared to poorer students (13%), and five times as likely to have access to numeracy (43%) compared to poorer students (8%) (Table 9). This gap between rich and poor declined only marginally between 2003 and 2012, from 37 percentage points to 35 percentage points in access-to-literacy. In access-to-numeracy the gap actually increased over the period. In 2003 the gap was 35 percentage points (43% vs 8%), which grew to 40 percentage points in 2012 (29% compared to 69%). Thus while the poor have experienced the largest proportional growth in access to literacy and numeracy, the gap between wealthy and poor remains unaffected. Note that these gaps are about 2.3 times larger than would be found if one only looked at the PISA reports (Table 6 and 7), i.e. if one ignored PISA-ineligible students.

8. **Although both boys and girls benefited from increased access to literacy and numeracy, girls benefited more:** Between 2003 and 2012 the access to literacy and numeracy rates doubled for boys and more than doubled for girls (2.4 times). A big part of the explanation is that drop out / school-leaving rates among girls halved from 38% in 2003 to 20% in 2013. Although girls and girls have almost equal probabilities of being in PISA in 2012 – 82% for boys and 79% for girls (Table 2), girls do much better in literacy than boys and about the same in Mathematics (Table 6 and 7). Because of these two factors, girls are statistically significantly more likely to have access-to-literacy than boys, with 69% of girls acquiring functional literacy skills in 2012 compared to 56% of boys (Table 8 and 9). The slight pro-boy advantage in access-to-numeracy is not statistically significant (Table 9).

9. **15-16 year-olds in the East region are less likely to make it into the PISA sample:** Although region was not one of the comparison criteria in the main results, given the considerable attention in the Turkish academic literature to the underperformance of the East region (Kırdar, 2009; O'Dwyer et al., 2010), a brief analysis on PISA sample eligibility by region is included below (Figure 12). Across all three DHS surveys, 15-16 year-olds in the East region were statistically significantly less likely to be eligible for PISA than those in the West, Central or North regions, and in 2013 this is true relative to all four of the other regions. In 2013, 15-16 year-olds in the East region were 23 percentage points less likely to be eligible for PISA than 15-16 year-olds in the North region.

Figure 12. PISA eligibility among Turkish 15-16 year-olds by region in the Turkish DHS 2003, 2008 and 2013



Source: Author's own calculations from PISA and DHS datasets

ROBUSTNESS CHECKS

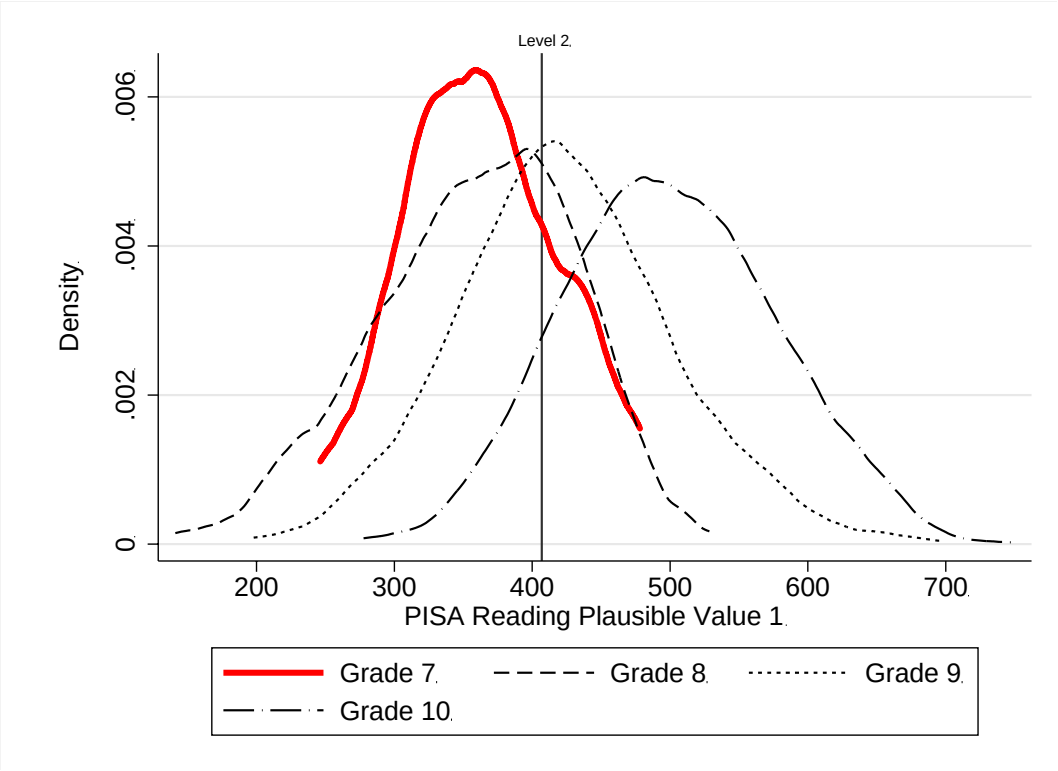
The method employed in this paper relies quite strongly on the assumption that PISA ineligible 15-16 year-olds would not reach Level 2 in either Reading or Mathematics if they were actually assessed. This allows for the multiplication of the literacy or numeracy rate with the eligibility/access rate to produce an access-to-literacy and access-to-numeracy rate. But how plausible is this assumption? There is a considerable amount of literature testifying to the correlates of student, out and severe delay. A review of the academic and grey literatures on student achievement shows considerable theoretical[14] and empirical evidence in support of the argument that students who leave school early or do not progress steadily over the grades are considerably more likely to come from poorer backgrounds, have lower levels of parental education and household wealth (Filmer and Pritchett, 1999), live in rural areas (Lewin, 2007), be female (Lewis and Lockheed, 2006) and are more likely to be under-achieving students prior to dropout (Lambin, 1995; UNESCO, 2005). Köseleci's (2015) recent review of the Turkish literature on educational participation shows the same list of correlates in Turkey, with the addition of the East region as a particularly problematic area.

14 For a theoretical framework addressing access, progression and transition see the CREATE research initiative (Lewin, 2007, 2009) and associated publications, as well as Volume 31 (2011) of the International Journal of Educational Development (see also the editorial by Lewin and Little, 2011).

Although the existing literature can tell us that those who dropout, leave school after only 8 years, or are severely delayed, underperform those who are 'on track', it cannot tell us how much worse they perform or document the skills they have acquired (or not) before dropping out. Given that PISA assesses 15-16 year-olds in Turkey from Grades 7-11, it is possible to compare those 15-16 year-olds in PISA that are delayed (behind the modal grade), to those that are not delayed (have the modal grade). And specifically to determine how many delayed students in PISA do reach Level 2 in Reading and Mathematics. This will provide some evidence of how much worse delayed students perform relative to 'on track' students.

The modal grade in Turkey in PISA 2012 was Grade 10, with 66% of the sample being in this grade (OECD, 2012b, p. 155). Figure 13 below shows the distribution of Reading achievement by grade of enrolment in Turkey in PISA 2012. Clearly those 15-16 year-olds that are delayed perform significantly worse than those that are not delayed. The mean Reading score for the 120 students who were in Grade 7 or Grade 8 was 362 (SE: 9.6), compared to those who were in Grade 9 who scored 425 (SE:5.4) and those who were in Grade 10 who scored 500 (SE: 4,1). Given that Turkey's standard deviation in Reading is approximately 86 points, those students that were at least two years delayed scored 1.6 standard deviations below those in Grade 10.

Figure 13. Kernel density of Turkey's PISA 2012 Reading Scores (Plausible Value 1) by grade of enrolment



Source: Author's own calculations from PISA and DHS datasets

Of those 15-16 year-olds that were eligible for the PISA exam in 2003 (i.e. still in school and enrolled in Grade 7 or higher), 87% of students that were delayed by 2 or 3 years were functionally illiterate compared to 37% of students among those in the modal grade. In 2009, 86% of those that were delayed by 2 or 3 years were functionally illiterate, compared to 11% of those in the modal grade (Grade 10). In 2012, 70% of those that were delayed by 2 or 3 years were functionally illiterate, compared to 11% of those in the modal grade. Clearly there is strong empirical evidence within PISA that delayed students perform

significantly worse than their 'on-track' peers, and these are among those that are actually still in school. There is good reason to believe that those who have dropped out (or are so severely delayed as to be ineligible for PISA), would perform even worse than this under-performing group.

Another way of phrasing the above assumption is to assume that students that have dropped out, left school early or are severely delayed would achieve at PISA Level 1a or 1b – an assumption that the OECD itself has put forward (OECD, 2012b, p. 53). Given that this is still below PISA Level 2, all of the above estimates would remain unchanged. If one chose to impute scores for those that are severely delayed or who have , left school early or dropped out, the most logical counter-factual would be the group of students that are in that round of PISA but who are also delayed by 2-3 years. For example one could use the percentage of students that are delayed by 2-3 years that *do* reach Level 2 as an upper-bound proxy for those that are delayed, left school early, and those that have dropped out. In which case the access-to-literacy rate would increase from 28% to 35% in 2003, from 46% to 51% in 2009, and from 63% to 69% in 2012, but never more than an additional 5-7 percentage points. Note that in all cases the upper-bound imputed estimate is still substantially lower than the percentages reaching Level 2 in the official PISA reports. For example in 2003 the imputed estimate above is 35% but in the PISA report it is 63%. In 2009 the imputed estimate is 51%, but in the PISA report it is 76%. And in 2012 the imputed estimate is 69%, but in the PISA report it is 78%. The convergence over time is due to the decline in delay, early school leaving, and dropout.

One could perform a similar imputation for Mathematics and find again that the increases in access-to-numeracy are marginal. Increasing from 21% to 26% in 2003, from 35% to 36% in 2009, and from 47% to 51% in 2012 if one assumes that those who are severely delayed, left school early or drop out would have the same numeracy rate as those that were in PISA but delayed 2-3 years. All of these are considerably lower than the numeracy rates provided in the PISA reports.

While more sophisticated methods of imputation do exist (Ferreira and Gignoux, 2014, p. 18), they all assume PISA sample selection is fully explained by observable variables such as gender, parental education and father's occupation. These methods do not fully account for the fact that those who drop out or leave school early are systematically different from those that do not, and – importantly - in unobservable ways such as pre-dropout academic performance. The aim of the present paper is not to try and improve on the imputation techniques used in earlier studies, but rather to highlight the extent of sample non-representivity in one country as well as how this changes across gender and socio-economic status groupings, and over time.

POTENTIAL LIMITATIONS

There are a few limitations to the data and method employed in this paper that are worth highlighting. The first is the use of two age-groups (15-16 years) in the DHS data. It is possible that for whatever reason these two age groups are not nationally representative given that the aim of the DHS is for the entire survey to be nationally representative, rather than a sub-set of the data. That being said we cannot think of why the DHS sample of 15-16 year-olds would be systematically non-representative of the 15-16 year old population. Secondly, while PISA uses the age range of 15 years and 3 (completed) months to 16 years and 2 months (OECD, 2005, p. 46), we have opted to use all 15 and 16 year-olds in the DHS group. We believe that the technical gains from doing so (increasing the sample size and decreasing the standard errors) are worth having a non-exact comparison group. Similarly, one could argue that it is not ideal that

there is a mismatch in the DHS years corresponding to the PISA 2009 and PISA 2012 cycles (where we use the 2008 and 2013 DHS surveys). However the alternative here is to exclude both of these rounds of PISA and only look at PISA 2003 in Turkey. We would argue that the DHS of 2008 and 2013 are good enough proxies of the access/attainment situation in Turkey in 2009 and 2012 respectively to warrant their inclusion.

One potentially confounding factor, that is prominent specifically in Turkey, is the extent of distance education, or what is referred to as 'open' schooling in Turkey. These students do not attend a physical school and do not have teachers in any traditional sense but rely on information communication technologies (ICT) to receive and submit information. Demiray and Adiyaman (2002) provide an overview of the literature on Open High Schools in Turkey, although it should be emphasised that this is a field with scant quantitative research (for one exception see the dissertation by Köksal, 2014). Of the little literature that exists on large-scale international assessments in Turkey, almost none of it mentions the size or potential impact of this hidden sector, with the exception of Özdemir (2016) who states that about 20% of secondary education students are in distance/open education and also that these students are not included in international assessments. In fact statistics released by the Turkish Ministry of National Education show that of the 5,807,643 students in secondary school in 2015/16, 26% (1,536,135) were in 'Open secondary education' (MoNE, 2016, p. 123). However, the majority of these students are adults working full time and trying to get a degree. Unfortunately it was not possible to obtain the Turkish Household Labour Force Survey to provide exact estimates of the size of this group. However, one economist[15] familiar with the education sector in Turkey estimated that at a maximum this would be 10-15% of 15 year-olds. Unfortunately the DHS does not have any questions relating to the type of school attended, and thus this cannot be explored empirically. The reader should bear this in mind when interpreting the figures above. Notwithstanding the above, similar caveats and logic apply to these students as well. A host of international and Turkish literature attest to the fact that those who drop out of traditional schools (for whatever reason) are, on average, weaker students that come from poorer areas with less educated parents, even if they do on to undertake distance education.

CONCLUSION

The principal aim of the present analysis has been to draw attention to the issue of PISA sample eligibility and sample representivity in one middle-income country: Turkey. Whether one chooses to use the Coverage Index 3 figures reported by the OECD itself, or those calculated from various DHS surveys, the conclusion is the same: there are large and changing proportions of students that do not make it into the PISA sampling frame and this has a substantial effect on the validity of inter-country and inter-temporal comparisons.

Perhaps unsurprisingly, the analysis showed that when PISA-ineligible 15-16 year-olds are accounted for the gaps between rich and poor are bigger than was previously thought, and the improvements over time are larger than traditionally reported by the OECD. In the discourse on educational reform in large middle-income countries, Turkey is a useful case study of efficacious education policies. Over the last two decades the educational reforms implemented by the Ministry of National Education have had a considerable impact, and one that is evident in both household surveys and in PISA. The DHS surveys

15 (Personal Communication, Alper Dinçer, 2 September 2016)

analysed here show large increases in grade attainment and substantial declines in drop out and delay, all of which are statistically significant at the 95% confidence level.

If one looks at the full population of 15-16 year-olds in Turkey, this study has shown that the percentage of those who have 'access-to-literacy' has more than doubled from 28% in 2003 to 63% in 2012, and similarly those with 'access-to-numeracy' has increased from 21% to 47% over the same period. Perhaps most encouragingly, the proportional gains were most pronounced for those in the poorest 40% of households, and particularly for girls in those households, where access to literacy rates increased four-fold in under a decade. Yet inequalities remain unabated. The gaps between rich and poor in access-to-literacy are persistent and in access-to-numeracy they have grown over the period. In 2012/13 the rich were 43% more likely to make it into the PISA sampling frame, and were twice as likely to acquire access-to-literacy as the poor.

The evidence presented here has not sought to discredit the PISA results in Turkey or secondary analysis on that data, but rather to emphasise that additional caution is needed when interpreting the PISA results in countries with low and/or changing proportions of PISA eligible students, such as Turkey. It is true that there is little that the OECD can do to increase the representivity of PISA in developing countries, but there is much that they can do to highlight which countries are least representative, or at the very least, which countries fall below some threshold, such as 70-80% coverage of the overall population of 15 year-olds. In the most recent PISA 2015 Report, this is in fact the approach that the OECD has taken (OECD, 2016a). Furthermore, the OECD can advise caution when using PISA data for comparisons that include countries that have especially low coverage rates. As one possible suggestion, the OECD may wish to 'asterisk' countries with less than 70% coverage rates in all comparisons, tables, graphs etc. and include a footnote highlighting that in these countries sample coverage is problematically low and comparisons should be made with caution. The OECD's latest PISA for-Development initiative is a notable exception to the above critique since this specifically takes account of out-of-school children in the survey design.

Finally, by drawing attention to the large changes in PISA sample ineligibility in Turkey, it is our hope that this evidence will prompt other researchers in the OECD and in the academic community to undertake similar analysis in other countries with equally low and changing coverage rates, such as Brazil, Costa Rica, Mexico, Indonesia, Peru and Vietnam. And at the very least will help combat the apparent widespread ignorance relating to PISA sampling, coverage and representivity in middle-income countries.

REFERENCES

Alacacı, C., and A. K. Erbaş, (2010), "Unpacking the inequality among Turkish schools: Findings from PISA 2006", *International Journal of Educational Development*, Vol 30/2, pp. 182–192, http://doi.org/10.1016/j.ijedudev.2009.03.006.

Avvisati, F., and F. Keslair, (2014), "REPEST: Stata module to run estimations with weighted replicate samples and plausible values", Boston: *Statistical Software Components, Boston College Department of Economics*, https://ideas.repec.org/c/boc/bocode/s457918.html.

Aydagül, B. (2008a), "No shared vision for achieving Education for All: Turkey at risk", *PROSPECTS*, Vol. 38/3, pp. 401–407, http://doi.org/10.1007/s11125-009-9085-2.

Aydagül, B. (2008b), "No shared vision for achieving Education for All: Turkey at risk", *PROSPECTS*, Vol. 30/3, pp. 401–407, http://doi.org/10.1007/s11125-009-9085-2.

Bloem, S. (2013), "PISA for low- and middle-income countries", *OECD Education Working Papers*, Vol. 93, pp. 40.

Bloem, S. (2015), "PISA for low- and middle-income countries", *Compare: A Journal of Comparative and International Education*, Vol. 45/3, pp. 481–486, http://doi.org/10.1080/03057925.2015.1027513.

Celik, K., and D. Lüküslü, (2010), "Spotlighting a Silent Category of Young Females: The Life Experiences of 'House Girls' in Turkey", *Youth and Society*, Vol. 44/1, pp. 28-48.

Demiray, U., and Z. Adiyaman, (2002), "A Review of the Literature on the Open High School in Turkey Between the Years on its 10th Anniversary (1992-2002)", Istanbul, Turkish republic: *Ministry of National Education.*

Dülger, I. (2004), "Case Study on Turkey: Rapid Coverage for Compulsory Education Program" , *Scaling Up Poverty Reduction*, pp. 1–46, Shanghai, China.

Ferreira, F. H. G., and J. Gignoux, (2014), "The Measurement of Educational Inequality: Achievement and Opportunity", *The World Bank Economic Review*, Vol. 28/2, pp. 210–246, http://doi.org/10.1093/wber/lht004.

Filmer, D., A. Hasan, and L. Pritchett, (2006), "A Millennium Learning Goal: Measuring Real Progress in Education"; *SSRN Electronic Journal*, http://doi.org/10.2139/ssrn.982968.

Filmer, D., and L. Pritchett, (1999), "The effect of household wealth on educational attainment: evidence from 35 countries", *Population and Development Review*, Vol. 25, pp. 85–120, http://doi.org/10.1111/j.1728-4457.1999.00085.

Filmer, D., and L. Pritchett (2001), "Estimating wealth effects without expenditure data—Or tears: An application to educational enrollments in states of India*", *Demography*, http://link.springer.com/article/10.1353/dem.2001.0003.

Gur, B., Z. Celik, and M. Ozoglu, (2012), "Policy options for Turkey: a critique of the interpretation and utilization of PISA results in Turkey"; *Journal of Education Policy*, Vol. 27/1, pp. 1–21.

Hacettepe-UIPS. (2003), *Turkey Demographic and Health Survey 2003*, Ankara, Turkey: Hacettepe University Institute of Population Studies.

Hacettepe-UIPS. (2009), *Turkey Demographic and Health Survey, 2008*, Ankara, Turkey: Hacettepe University Institute of Population Studies.

Hacettepe-UIPS. (2014). *2013 Turkey Demographic and Health Survey*. Ankara, Turkey: Hacettepe University Institute of Population Studies.

Hanushek, E. A., and L. Woessmann, (2011), "Sample selectivity and the validity of international student achievement tests in economic research", *Economics Letters*, Vol. 110/2, pp. 79–82, http://doi.org/10.1016/j.econlet.2010.10.022.

Hanushek, E., and L. Woessmann, (2008), "The role of cognitive skills in economic development", *Journal of Economic Literature*, Vol. 46/3, pp. 607–668., http://www.jstor.org/stable/27647039.

Kırdar, M. G. (2009), "Explaining Ethnic Disparities in School Enrollment in Turkey", *Economic Development and Cultural Change*, Vol. 57/2, pp. 297–333, http://doi.org/10.1086/592841.

Köksal, S. (2014), *Open and Distance Education in Turkey*, Galatawaray University.

Köseleci, N. (2015), "Progress towards good-quality education for all in Turkey: a qualified success?" *Comparative Education*, Vol. 51/4, pp. 555–574, http://doi.org/10.1080/03050068.2015.1087211.

Köseleci, N., and A. Sasmaz, (2011), "PISA 2009: Where Does Turkey Stand?" *Turkish Policy Quarterly*, Vol. 10/2, pp. 125–134.

Lambin, R. (1995), "What can Planners Expect from International Quantitative Studies ?", *Reflections on Educational Achievement: Papers in Honor of T. Neville Postlethwaite*, pp. 169–182.

Lewin, K. M. (2007), "Transitions and Equity. Improving Access , Equity and Transitions in Education : Creating a Research Agenda (No. 1)", *CREATE*.

Lewin, K. M. (2009), "Access to education in sub-Saharan Africa: patterns, problems and possibilities", *Comparative Education*, Vol. 45/2, pp. 151–174, http://doi.org/10.1080/03050060902920518.

Lewin, K. M., and A.W. Little, (2011), "Access to education revisited: Equity, drop out and transitions to secondary school in South Asia and Sub-Saharan Africa", *International Journal of Educational Development*, Vol. 31/4, pp. 333–337.

Lewis, M., and M. Lockheed, (2006), "Inexcusable absence: Why 60 million girls still aren't in school and what to do about it", *Center for Global Development*, http://scholar.google.com/scholar?hl=enandbtnG=Searchandq=intitle:Inexcusable+Absence:Why+60+Million+Girls+Still+Aren't+in+School+and+What+to+Do+About+It#0.

Lockheed, M., T. Prokic-Breuer, and A. Shadrova, (2015), *The Experience of Middle-Income Countries Participating in PISA 2000-2015*, Paris: Organisation for Economic Cooperation and Development.

MoNE. (2016), *National Education Statistics: Formal Education 2015/16*, Istanbul: Turkish Republic Ministry of National Education.

O'Dwyer, J., N. Aksit, and M. Sands, (2010), "Expanding educational access in Eastern Turkey: A new initiative", *International Journal of Educational Development*, Vol. 30/2, pp. 193–203, http://doi.org/10.1016/j.ijedudev.2009.03.005.

OECD (2004), *PISA Report 2003: Learning for Tomorrow's World*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264010543-en

OECD (2005), *PISA 2003 Technical Report*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264010543-en

OECD (2010), *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Volume I)*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264091450-en

OECD (2012a), *PISA 2009 Technical Report*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264167872-en

OECD (2012b), *PISA 2012 Results: Excellence through Equity: Giving Every Student the Chance to Succeed (Volume II)*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264201132-en

OECD. (2014a), *PISA 2012 Results: What Students Know and Can Do. Student Performance in Mathematics Reading and Science (Volume I)*, OECD Publishing, Paris. http://dx.doi.org/ 10.1787/9789264208780-en

OECD (2014b), *PISA 2012 Technical Report*, OECD Publishing, Paris.

OECD (2016a), *PISA 2015 results (Volume 1): Excellence and Equity in Education*, OECD Publishing, Paris. http://dx.doi.org/ 10.1787/9789264266490-en

OECD (2016b), PISA for Development: Benefits for participating countries. PISA for Development Brief, Vol. 2, OECD Publishing, Paris. http://www.oecd.org/pisa/aboutpisa/PISA-for-Development-Benefits-for-participating-countries-PISA-D-Brief2.pdf

OECD and UNESCO-UIS. (2003), *PISA 2000 Report*, OECD Publishing, Paris. http://doi.dx.org/ 10.1787/9789264102873-en

Özdemir, C. (2016), "Equity in the Turkish education system: A multilevel analysis of social background influences on the mathematics performance of 15-year-old students", *European Educational Research Journal*, Vol 15/2, 193–217, http://doi.org/10.1177/1474904115627159.

Parandekar, S., and E. Sedmik, (2016), *Unraveling a Secret: Vietnam's outstanding performance on the PISA test*, Washington, DC.

Pritchett, L. (2013), *The Rebirth of Education: Schooling Ain't Learning*. *2013*. Washington, DC.: Center for Global Development.

Rankin, B. H., and I.A. Aytac, (2006), "Gender Inequality in Schooling: The Case of Turkey", *Sociology of Education*, Vol. 79/1, pp. 25–43. http://doi.org/10.1177/003804070607900102.

Rotberg, I. (1995), "Myths About Test Score Comparisons", *Science*, Vol. 270/5241, pp. 1446–1448, http://doi.org/10.1126/science.270.5241.1446.

Spaull, N., and S. Taylor, (2015), "Access to What? Creating a Composite Measure of Educational Quantity and Educational Quality for 11 African Countries", *Comparative Education Review*, Vol. 59/1, pp. 133–165.

Tansel, A. (2002), "Determinants of school attainment of boys and girls in Turkey: individual, household and community factors", *Economics of Education Review*, Vol 21/5, pp. 455–470, http://doi.org/10.1016/S0272-7757(01)00028-0.

Taylor, S., and N. Spaull, (2015), "Measuring access to learning over a period of increased access to schooling: The case of Southern and Eastern Africa since 2000", *International Journal of Educational Development*, Vol. 41, pp. 47–59.

UNESCO, (2005), *Education For All Global Monitoring Report 2005: The Quality Imperative*. Paris.

UNESCO Institute for Statistics, (2010), "Measuring educational participation: Analysis of data quality and methodology based on ten studies" (Technical Paper 4).

World Bank, (2005), *Turkey - Education Sector Study*, Washington, DC.