The Effects of System Type and System Characteristics on Skills Acquisition in Upper Secondary Education and Training

Andy Green, Neil Kaye, Nicola Pensiero and Hao Phan

LLAKES Research Paper 69

Centre for Learning and Life Chances (LLAKES)
LLAKES is an ESRC-funded Research Centre - grant reference ES/T001526/1.

To request printed copies of this paper or other LLAKES Research Papers, please contact the LLAKES Centre - llakescentre@ucl.ac.uk.

To view or download this and other LLAKES Research Papers online, please visit www.llakes.ac.uk

LLAKES Research Paper 69

The Effects of System Type and System Characteristics on Skills Acquisition in Upper Secondary Education and Training

Andy Green, Neil Kaye, Nicola Pensiero and Hao Phan

The Centre for Learning and Life Chances (LLAKES).
LLAKES is an ESRC-funded Research Centre - grant reference ES/T001526/1.

The research for this paper is supported by funding from the Institute for Adult Learning, Singapore University of Social Sciences.

Contents

Tables

Figures

Executive Summary

This report examines the effects of upper secondary system types and characteristics on literacy and numeracy skills acquisition during the upper secondary phase of education and training. Whereas there is a substantial literature on system effects on skills during the primary and lower secondary phases of education, much less has been written about these effects in relation to the upper secondary phase. However, with the arrival of the OECD's Survey of Adult Skills (SAS), which has now tested adults in over 40 countries and regions, it is now possible to explore how far education system characteristics explain the substantial variation across countries in changes in skills levels and inequalities during upper secondary phase.

In this report we seek to build on earlier work and provide more robust evidence on system effects during the upper secondary phase in three ways. Firstly, we use data from the larger sample of countries in both waves 1 and 2 of SAS. Secondly, we test the effects of a considerably wider range of system indicators. Thirdly, we use a variety of statistical methods to analyse the relationships across countries between upper secondary system types and characteristics and changes in levels and distributions of skills between age 15 (in PISA) and the end of the upper secondary phase. Whereas our previous work analysed changes using quasi-cohort analysis of published data on skills from PISA (at age 15) and SAS (at age 25-29), thus allowing compounding effects from tertiary education and employment, here we use customised data from OECD on skills scores at age 18-20 to capture more precisely the skills at the beginning and end of upper secondary education and training.

Following a review of the literature on system effects, we identify a range of factors deemed to influence skills acquisition in the upper secondary phase and six upper secondary system types based on common and distinctive characteristics. The subsequent sections provide descriptive statistics on system characteristics by country/region and by system type and a statistical analysis, using both OLS regressions and Difference-in-Difference methods to estimate the effects of different types and characteristics on relative changes in skills levels and inequalities during the upper secondary phase.

Main Results

Our analysis identifies broadly three upper secondary characteristics, which vary quite systematically across system types, which have significant effects on our outcomes, and thus

go some way to explaining the variation across countries in changes in levels and distributions of skills during the upper secondary phase. These are:

- 'Vocational prevalence' – a measure of the proportion of upper secondary students undertaking vocational learning;
- 'Standardised curriculum' – measuring the degree to which systems mandate the learning of Maths and the national language across different tracks; and
- 'Teacher workloads – an indicator capturing teacher time in the classroom and the number of students taught.

Various indicators for typical duration of studies and graduation ages, and for the parity of funding across vocational and general programmes also seem relevant, although their impacts are less consistent in statistical terms.

As predicted by our typology, and supported by dominant theories, systems perform better in raising average skills levels and mitigating skills inequalities where they achieve relative 'parity of esteem' across vocational and general programmes. This can be achieved through high levels of participation in high quality vocational learning, particularly in apprenticeships, or through high levels of system integration, including through the use of common core curricula and the standardisation across tracks of resourcing and the length of programmes.

The strongest finding from our analysis is that the Type 2a (North American comprehensive) systems, including those in the US and Canada, and the Type 4 (mixed) systems (including those in Australia, Chile, Ireland, Israel, New Zealand, Spain, and UK countries) perform significantly worse than other system types in both literacy and numeracy skills acquisition between age 15 and age 18-20. They also do less, on average, than systems in other system types to mitigate skills inequalities during the upper secondary phase.

Compared with the systems which perform relatively well in raising skills levels and reducing skills inequalities, Type 2a and Type 4 systems tend to have lower rates of participation in vocational learning, particularly in high quality apprenticeships, and (in the case of the mixed systems) relatively low resourcing of vocational tracks relative to general tracks. Teacher workloads tend to be higher in both North American and mixed systems. The mixed systems also tend to have significantly lower levels of curriculum standardisation.

In our conclusions we also note the importance of other factors relating to pedagogy and culture which we were unable to measure.

Introduction

This report examines the effects of upper secondary system types and characteristics on skills acquisition during the upper secondary phase of education and training.

There is an extensive cross-country literature on education system effects on skills acquisition during the primary and lower secondary phases of education. This draws on data on tested knowledge and skills, and on education system characteristics, from three international surveys: The Progress in International Reading Literacy Study (PIRLS) - which tests 4th graders at five-yearly intervals; Trends in International Mathematics Study (TIMSS) - which tests 4th and 8th graders at four-yearly intervals and the Programme for International Student Assessment (PISA) which conducts tests of proficiency in Reading, Mathematics and Science amongst students aged 15. The wide range of international data available on skills and education systems covering two decades (and more for TIMSS) has allowed researchers to make comparative estimates of country-level changes in the levels and distributions of core skills during primary and lower secondary phases of education (e.g., Hanushek & Woessmann, 2006). It has also made it possible to explore how far education system characteristics explain the substantial variation across countries in these changes during the primary and lower secondary phases of education.

Much less comparative research, however, has been conducted on skills acquisition during upper secondary education and training, partly, as we explain later, due to the relative paucity of international data for this phase. However, this is now beginning to change with OECD's Survey of Adult Skills (SAS - also known as PIAAC) which was first fielded in 24 countries and regions in 2011/12 and now tests proficiency in literacy, numeracy and problem solving in over 40 countries and regions. A number of reports were published using data from the first round of SAS (including OECD, 2013; Borgonovi, 2017; Green et al., 2014, 2015, 2016; Pensiero & Green, 2018) drawing tentative conclusions about the role of system characteristics in explaining the variation across countries in changes in skills levels and distributions during the upper secondary phase.

In this report we seek to build on earlier work and provide more robust evidence on system effects during the upper secondary phase in three ways. Firstly, we use data from the larger sample of countries in both waves 1 and 2 of SAS. Secondly, we test the effects of a considerably wider range of system indicators. Thirdly, we use a variety of statistical methods to analyse the relationships across countries between upper secondary system types and characteristics and changes in levels and distributions of skills between age 15 (in PISA) and the end of the upper secondary phase. Whereas our previous work analysed changes using quasi-cohort analysis of published data on skills from PISA (at age 15) and SAS (at age 25-29), thus allowing compounding effects from tertiary education and employment, here we use customised data from OECD on skills scores at age 18-20 to capture more precisely the skills at the beginning and end of upper secondary education and training.

The report is structured as follows. Section 2 provides a review of the literature on education system effects on skills during lower and upper secondary phases of education and training and a typology of upper secondary education and training systems for the period under scrutiny (roughly 2008-2012). The latter builds on previous typologies, modified according the most recent research system characteristics, as reviewed in Section 6. This is followed by sections on Hypotheses and Data and Methods. Sections 5 and 6 present descriptive statistics from PISA and SAS for skills outcomes at ages 15 and 18-20 respectively, focusing on average skills levels and measures of inequalities of skills outcomes for each system. We also provide descriptive statistics on the characteristics of different systems and how systems cluster. Section 7 presents our statistical analysis of the effects of system types and characteristics. Section 8 provides a discussion of our conclusions and their policy implications.

1. Literature Review

There is now a substantial literature on the effects of education system types and characteristics on skills acquisition during the primary and lower secondary phases of education. Much less research has been conducted on the factors associated with changes in aggregate skills levels and distributions during the upper secondary phase of education and training, in part due to the relative paucity of comparable longitudinal or time series data across countries on skills after age 15. However, since the introduction of the OECD's Survey of Adult Skills (SAS/PIAAC) in 2011/2, it has become possible to make comparative estimates of skills acquisition in the upper secondary phase across countries using measurements at the country level of skills at age 15 from relevant PISA surveys and at different points in the adult life course (after age 16) from SAS and the OECD's International Adult Literacy Survey (IALS), the precursor survey to SAS conducted in the mid-1990s. Unlike PISA and TIMSS, these adult skills surveys, based on household samples, do not collect information on education system characteristics, but some international data are available from other sources (UNESCO, OECD Education at a Glance, Eurydice) which relate specifically to the upper secondary phase. In addition, some of the comparative data on lower secondary education systems can be deemed relevant to upper secondary education systems. As a result of these advances in data availability, it is now possible to conduct comparative analysis of the effects of upper secondary systems on skills outcomes, and there is now a small but burgeoning research literature which seeks to do this, either by focussing on the effects specific system characteristics or by developing typologies of education systems. A multitude of system characteristics have been identified which may explain the variation across countries in changes in the levels and distributions of skills of students (and those not in education) during the upper secondary phase. These can be grouped under four broad headings:

- *Institutional Structures and Organisation*
- *Governance, Regulation and Funding*
- *Curriculum, Pedagogy and Assessment*
- *Teacher Quality*

Our review focuses predominantly on the research literature which provides direct evidence of the impact of system factors on Reading, Maths and Science scores at age 15 from PISA and on literacy, numeracy and problem-solving skills after the end of upper secondary education from SAS. We also consider the main theories which seek to explain variations across countries in changes during this phase. We divide the review into two: the first part focussing on evidence on factors affecting changes in aggregate skills levels and the second part focusing on evidence on factors affecting changes in skills inequalities.

1.1 Factors Affecting Changes in Mean Skills Levels during the Upper Secondary Phase

Institutional Structures

The organisation of educational institutions was a major pre-occupation of education policy makers during the half century after the end of WW2. With the gradual and uneven introduction of non-selective, comprehensive secondary schools in much of the developed world, policy debates and research were dominated by discussions of the relative merits of selective and non-selective forms of school organisation. A long-standing contention about the effects of school organisation on student performance was that tracked systems raise overall standards of achievement by encouraging more homogeneous classrooms where learning is more effective (Lynn, 1988; Maaz et al., 2008). However, more recent studies cast doubt on this. Some intra-regional comparative studies - such as Annermueller et al. (2005) on Central and Eastern Europe and Rivas on Latin America - find that early tracking and the proliferation of private schools is associated with higher average test scores. However, other studies using performance data for a wider range of countries find no clear relationship between tracking and mean levels of achievement and educational attainment (Micklewright and Schnept, 2007; Schutz et al, 2008; Woessmann, 2008). Some studies even find a negative effect from early selection. Hanushek and Woessmann (2006), for instance, using a Difference-in-Difference (DID) approach with data from PIRLS, TIMSS and PISA, find that whilst early tracking tends to increase family background effects on achievement, there is no trade-off with increased efficiency. Van de Werfhorst and Mifs (2010), reviewing the literature based on the evidence from multiple international surveys, conclude that early tracking leads to lower, rather than higher, average levels of achievement across a range of subjects/domains.

These, and other findings on the effects of ability grouping in schools, remain somewhat inconclusive, although recent research does tend to the conclusion that early tracking into selective schools depresses overall performance (OECD, 2013). However, the research is largely based on the lower secondary phase of education and may have limited applicability to upper secondary education and training which is tracked in all countries to some degree - either through selection by ability to different types of school, or through selection, or self-selection, of students into different tracks within comprehensive or multilateral schools (Van Houtte & Demanet, 2012). Upper secondary education is typically more differentiated because it is a phase of transition - either to the labour market, or to different types of further and higher education - which requires greater occupational or disciplinary specialisation (OECD, 2004). The prevalence of tracking may, therefore, have a different significance at the upper secondary level than it does during earlier phases (Busemeyer, 2014; Brunello and Checchi, 2007; Green and Pensiero, 2016).

Research on the effects of tracking in upper secondary education and training is rather limited to date and tends to point in different directions, arguing that both tracked systems and more integrated systems may perform relatively well in raising mean levels of skills provided that they are characterised by a relative 'parity of esteem' between general ‘academic’ and vocational provision. Raffe et al. (2001) first developed the theory of 'parity of esteem' with specific reference to upper secondary education and training. They argued that whilst during the compulsory phase schools and educational programmes are organised hierarchically, according to a monotonic scale of academic prestige, in the upper secondary phase the tracks are valued according to more differentiated criteria, whether it be their success in getting graduates into skilled jobs or into higher levels of education and training. Some more integrated systems, as with the Nordic comprehensive high schools, could be effective in raising average skills levels through greater standardisation of curricula in core areas and through generating higher normative expectations, through the provision of flexible progression routes and a common qualification framework. Some more tracked systems, as in countries combining academic high schools with the Dual Systems of apprenticeship, could also raise mean skills levels skills through establishing high normative standards in their vocational tracks by virtue of the prestige associated with high quality apprenticeships and the peer effects generated by the wider ability mix amongst students recruited to these apprenticeships.

Comparative political economists who study skills formation after lower secondary school (Busemeyer and Iversen, 2011; Busemeyer and Trampusch, 2012; OECD, 2001) also argue that Dual Systems of apprenticeship increase overall standards through the positive feedback effects from the labour market - apprentices are motivated to increase their skills because they know their qualifications will be valued in the labour market and will lead to well-remunerated skilled jobs which utilise the higher skills they can attain through apprenticeship training. The utilisation of literacy skills at work has been shown to be associated with the maintenance or improvement of literacy skills at the individual level (OECD and Statistics Canada, 2000; Rubenson, 1987; Desjardins and Ederer, 2015), suggesting that apprentice training in work contexts may also boost core skills learning.

Pensiero and Green (2018) sought to test these theories in their cross-country analysis of changes in levels and distributions of literacy and numeracy skills using quasi-cohort data on skills changes after age 15, and a variety of indicators for upper secondary system characteristics. They found that the system types which performed best in improving their relative position in country skills rankings for literacy and numeracy scores between PISA and

SAS tests were the Nordic comprehensive systems and the tracked systems which included academic high schools and Dual Systems of apprenticeship (i.e. Austria and Germany).

Governance, Regulation and Funding

There are a number of comparative studies investigating how the governance and regulation of secondary schools affects student performance. Most prominent in recent years has been the research on the effects of the suite of policies - including those for increasing school diversity, competition and choice, and school autonomy and accountability - associated with so-called neo-liberal education policy, or what Pasi Sahlberg (2015) calls the 'Great Education Reform Movement'. However, there is little consistency in the findings on the effects of these policies.

One plank of the school diversification agenda is the encouragement of a wide range of private schools. The research does not provide any strong evidence that the prevalence of private schools at primary and secondary levels has a significant impact on overall student achievement. Studies by Woessmann (2005) and Hanushek and Woessmann (2009), using data from PISA, conclude that countries which combine private management of schools with public funding tend to perform better in international tests in some contexts. However, the OECD (2013) analysis of the PISA 2012 data shows no relationship between the percentage of students enrolled in private schools and overall system performance. Nor do their analyses show positive effects on performance from increasing between-school competition. Using a simple general measure of the degree of competition between schools in PISA (based on asking school principals if their school competes with other schools for enrolments), the OECD finds no relationship between levels of between-school competition and average system performance across countries (but a negative effect of educational equality) (OECD, 2013).

More evidence has been forthcoming on the positive effects of school autonomy. Bol and Van de Werfhorst (2013) combine a range of measures from PISA relating to the level of autonomy exercised by schools in decisions about courses offered, course content, and textbooks. They find a negative effect on performance across countries from the 'standardisation of inputs,' concluding that systems which give schools more autonomy tend to have higher overall performance. The OECD analysis of PISA 2012 data also suggests that systems giving schools more autonomy in relation to course content and textbook choices tend to perform better; but they also note that school autonomy over budgets has no significant effect on overall performance in Maths (OECD, 2013). The effects of different school accountability measures introduced in tandem school autonomy policies are also variable, according to the research. A number of studies find a positive effect on performance from externally-controlled student assessment, and particularly from ‘centralised exit examinations’ (Bol & Van de Werfhorst, 2013, Woessmann, 2005b). Woessmann’s analysis shows that the effect is strongest in systems which also have high levels of school autonomy in pedagogic and budgetary matters and that the impact on performance is greatest towards the end of lower secondary education. However, not all accountability measures are associated with better performance. The OECD (2013) find that the prevalence of external administrative monitoring of school performance (through league tables publicising school performance in examinations etc.) is negatively associated with overall performance in Maths, although this may arise from the fact the principals of under-performing schools are more likely to report such monitoring.

As with research on school organisation and tracking, research on the effects of school governance and regulation policies has focused almost exclusively on lower secondary schools. This is partly because we lack comparable quantitative cross-country data on these characteristics for upper secondary systems, except for centralised exit examinations. Pensiero and Green's (2018) difference-in-difference analysis of the effects of system characteristics on changes in literacy and numeracy score ranking in PISA 2000 and SAS (2011/12) found that centralised exit exams and private school enrolments had significant *negative* effects on skills levels. The use of centralised exit exams was associated with a reduction in relative gains in literacy and numeracy skills, while private school enrolments had the same effect on literacy (with significant levels at $P<0.2$ and $P<0.3$ respectively).

Curriculum, Pedagogy and Assessment

Cross-country comparative research on the effects of different forms of curriculum, pedagogy and assessment is predominantly qualitative in methods, particularly as regards pedagogy, since these phenomena can be hard to define and measure. The research also tends to focus on broader educational outcomes, rather than core skills acquisition, and utilises evidence mostly from the primary and lower secondary phases of education which are typically less differentiated than the upper secondary phase and thus easier to compare across countries. The literature provides rather little direct evidence of impacts on skills acquisition during the upper secondary phase, but this does not mean that such factors are irrelevant to explaining cross-country variations in skills acquisition at this level.

Recent policy-oriented debates on the curriculum have often revolved around the relative merits of 'curriculum control' versus 'school autonomy' or 'professional autonomy' in curriculum matters. While many countries continue to exercise tight curriculum control - through the central prescription of detailed curricula and standards, and the use of centrally authorised textbooks (and even centrally issued teaching materials), the dominant international discourse on school reform (Sahlberg, 2015) favours increasing the degree of school and professional autonomy over the implementation of the curriculum, albeit within the framework of national curricula and prescribed standards or through centralised 'steering by goals'. OECD research based on PISA (OECD, 2013) finds that the greater the number of schools having some autonomy in curriculum 'elaboration' the better the overall performance of the school system (when controlling for GDP at least), but also notes that the benefits of this approach are highly dependent on the professional competence of teachers and on system support for continuing teacher professional development. An alternative view is that too much autonomy in curriculum matters may undermine 'curriculum entitlement' and the normative expectations on students' achievement (Sinnema, 2016) or deprive some students of access to 'powerful knowledge' (Young and Muller, 2010; Young et al, 2014). Another dimension of the debate relates to the importance of 'curriculum coherence', achieved, amongst other things, through a close alignment between the curriculum, textbooks, and teacher coverage of the curriculum in the classroom. Schimdt and and Prawat (2006) conclude from research on the 37 countries participating in TIMSS that curriculum coherence is strongly associated with high performing education systems. They do not claim that this is necessarily achieved through central curriculum control, but Oates (2010) argues that a certain amount of central curriculum control is important for high performing education systems.

Curriculum control and coherence are characteristics often associated with a number of East Asian systems (including those in Japan, South Korea and Singapore) which typically occupy the top positions in the country rankings of student skills in Reading, Maths and Science based on PISA surveys. Research on the most effective forms of pedagogy is also influenced by the association of high performing systems in East Asia (and also in Finland) with traditional didactic modes of instruction. These are said to utilise teacher-led, whole-class teaching methods, combined with an emphasis on drilling and memorisation, to cultivate student mastery of key concepts and skills, and particularly those necessary for excelling in examinations and international skills tests. Teaching in Finland (until recent reforms) has traditionally been characterised as teacher-led and didactic (Norris et. al., 1996). Teaching in Singapore's primary and secondary schools is noted for strong 'curriculum control' and 'curriculum coherence' and for the use of 'mastery methods' of instruction, which Hogan's (2011) research suggests all contribute towards Singapore's high performance in PISA tests. In a rare example of experimental research on the effects of pedagogy, Jerrim and Vignoles conducted a Randomised Control Trial of an intervention to introduce Singapore-style mastery methods of teaching Maths in a selection of 90 primary schools and 50 secondary schools in England in 2012/13. Comparing the learning gain amongst the intervention group with that of the control group, they found a small positive impact from the intervention, although this did not reach statistical significance at traditional levels. Arguably, even a small impact from the introduction of such methods in the English context is notable, given the absence of the cultural contexts which some researchers argue underpin their effectiveness in East Asia (Zongyi Deng & Gopinathan, 2016).

There has been only limited research to test the applicability of the research findings above in the context of upper secondary education and training. The research by Pensiero and Green (2018), drawing on earlier work by Hodgen et al. (2010), found a statistically significant ($P<0.01$) positive impact on aggregate skills gain after age 15 in both literacy and numeracy from the mandatory inclusion of maths and the national language in the curriculum across upper secondary tracks. Replicating the findings of Bol and Van der Werfhorst (2013), they also found significant ($P<0.01$) positive effects on relative skills gain from the participation of students in work-based learning. This points to the relevance for upper secondary skills acquisition of 'situated learning' and what Lave and Wenger (1991) call 'communities of practise'. Mindful of the importance of these learning contexts and relationships for skills acquisition many countries now seek to enhance the elements of work-based learning in upper secondary programmes through offering extended work placements and what we refer to later as 'hybrid apprenticeships' (Verdier, 2013; Méhaut, 2013).

Teacher Quality

International policy debates about the characteristics of 'high performing education systems' have become increasingly preoccupied with the question of teacher quality. A series of highprofile reports from Mckinsey and Company explore the role of teachers in boosting student performance and propose systematic approaches to developing a high-quality teaching force through case studies of 'high performing systems' in Finland, Singapore and South Korea. Their 2007 report (Barber & Mourshed, 2007) argues that countries scoring highest in PISA recruit teachers from the top third of each graduating cohort and suggest three fundamental components of a successful strategy: getting the right people to become teachers; developing them into effective instructors and ensuring that systems give the best possible instruction to every child. A number of quantitative research studies provide support for these contentions (Woessmann, 2003; Dolton and Macanaro-Gutierrez, 2012; and Hanushek et. al., 2019).

Dolton and Marcenaro-Gutierrez (2011) draw on country-level data from international tests between 1995 and 2006 to explore how teacher salaries, in absolute and relative terms (compared to national wages), impact on student achievement. Their OLS regression results suggest a linear relationship between teacher pay and student performance across countries: a 10 percent increase in real pay increases student performance by the same fraction and a 5 percent increase in the relative position of teachers in the national salary distribution likewise. Hanushek et. al. (2019) go further, extracting country level measures of teacher quality (literacy and numeracy skills) from SAS and using student micro level data from PISA to test the relationship between teacher cognitive skills and student performance at 15 in Reading and Maths across 31 countries. They confirm that higher relative teacher pay is systematically related across countries to higher teacher cognitive skills. These are in turn correlated with higher student performance across countries (even after controlling for a wide range of system characteristics and student demographic characteristics).

These studies are not able to match individual teachers with particular classes or students and cannot take account of how the sorting of teachers to different classes and schools affects overall student performance or what differential effects from teacher cognitive skills may apply in different subjects, or with students in different school grades or with different levels of attainment. Nor do the studies take account of the different recruitment procedures for teachers in upper secondary education and training, where many of those in vocational institutions are necessarily recruited on the basis of their technical skills and experience. Nevertheless, they are consistent in general terms with the findings of many qualitative studies which suggest the importance for student performance of recruiting well-qualified teachers and supporting these through continuing professional development (see for instance Law (2015) on teacher quality in Singaporean vocational institutions).

1.2 Factors Affecting Changes in Skills Distributions During the Upper Secondary Phase

Institutional Structures

Much of the comparative work which seeks to explain variations across countries in inequalities in education builds on the seminal work of Raymond Boudon in the 1970s. According to Boudon's (1974) influential 'positional' theory, social stratification has both primary and secondary effects within the education system. Primary effects occur as a result of the transmission of cultural capital within the family, so that children who acquire high levels of cultural capital at home achieve better in schools that value the same forms of cultural capital. Secondary effects occur as a result of children from different backgrounds making different choices within the education system, whereby children from higher status families, for instance, are more likely to choose pathways that lead to higher status qualifications than children of similar ability from lower status families. The first process tends to occur, arguably, in a similar way in all societies and education systems (Jackson, 2013). However, the second process may be more conditional on the nature of the particular education system. As Boudon cogently argued, in societies structured by social class and other inequalities, the greater the variety of different routes through the education system - i.e. the more 'branching-off' points - the greater the likelihood that socially differentiated aspirations and expectations, engendered from outside the education system, will structure student choices, even in a situation of ostensibly meritocratic access, so that educational opportunities and outcomes will be structured along lines of social class, ethnicity and gender (Green et. al., 2015).

In more recently elaborated theories of 'persistent inequalities' in education, elite social groups maintain their educational advantages as education systems expand in two ways. According to the theory of Maximally Maintained Inequality (MMI) (Raftery and Hout, 1993), as a phase of the education systems expands, higher social groups can maintain their advantage so long as their participation in that phase of education grows as fast as, or faster, than that of lower groups. However, when participation by elite students reaches saturation levels then positional competition tends to shift to a higher level of education.

At the same time, according to the Effectively Maintained Inequality (EMI) theory (Lucas, 2001), mass provision at the lower level develops more differentiated pathways, increasingly organised into a status hierarchy, with elite students tending to colonise the most prestigious tracks with the best progression routes to higher education.

When applied to upper secondary education and training, these theories would suggest that differentiated and diverse systems – with more branching points – will be likely to increase inequality both in skills opportunities and skills outcomes, whereas more standardised systems will have the opposite effect. Social origins effects will be enhanced in systems with more differentially valued pathways as family background differences encourage socially stratified patterns of pathway choice. Educational (and skills) outcomes will become more unequal as social segregation in tracks encourages greater differentiation in curricula and aspirational norms across tracks. A Special Edition of Oxford Review of Education (Heath & Sullivan, 2011) analyses what has happened to social inequalities in upper secondary education as this phase of education has been 'democratised' over the past 20 years. In a number of the cases examined, including France and Japan, social inequality has persisted at similar levels, despite massification, through the processes described in EMI theory.

A second hypothesis derives from the comparative literature specific to upper secondary education and training and suggests that tracking may not necessarily have the expected Boudonian effect in the upper secondary phase. As discussed above, this literature focuses on the so-called 'parity of esteem' between the academic and vocational tracks (Lasonen & Young, 1998; Raffe et al., 1998; and Raffe et al., 2001) and argues that where this is relatively greater there will be more mitigation of inequality during this phase. Where vocational tracks are of high quality and attract students from across the ability spectrum this is likely to lead to a mitigation of skills inequality. Peer effects are likely to help raise the standards of the lower achievers in the vocational track. Whereas the previous hypotheses are based on the Boudonian assumption that greater standardisation of curricula across pathways will reduce inequality, the parity of esteem argument qualifies the standardisation argument in as much as it allows for certain forms of tracking in upper secondary education which reduce inequality by virtue of the fact that the pathways are not aligned in a single status hierarchy, as tends to be the case in lower secondary education.

Governance and Regulation

The dominant position within current cross-country research on school systems and skills inequality is that more unequal outcomes are likely to occur when there is early selection to differentiated tracks and types of school; a higher proportion of entirely privately funded schools; a lack of standardisation in curricula and assessment; and a federal system where funding is devolved to the regional level (e.g. Hanushek & Woessmann 2006, 2010; Salverda et al., 2014; Schütz et al., 2008; OECD, 2010; Woessmann, 2005). According to this research, early tracking increases inequality as combined peer effects and school effects raise aspirations amongst students in high status tracks and schools and depress aspirations amongst students in lower status tracks and schools (Horn, 2009; Schütz et al., 2008; Werhorst and Mifs, 2010). Private schools promote inequality as families with high incomes are able to buy higher-quality education for their children in schools with smaller class sizes, better resources and more highly paid teachers. Lack of standardisation in curricula and assessment systems promotes inequality because school practises become more differentiated according to the social and ability composition of their intakes, thus exacerbating variation in school and peer effects across schools (Van de Werhorst and Mifs, 2010; Woessmann, 2005). Regionalised funding in state school sectors increases inequality in school quality as richer areas can spend more on education than poorer areas (for the U.S. see: OECD, 2010; Winkler, 1993).

From a Boudonian perspective the mechanisms above create greater differentiation and more branching points within systems and are therefore likely to raise inequality by allowing inequalities external to schools permeate the school system. Market-oriented policies for school diversification, competition and autonomy are considered in this light in much of the research on lower secondary schools and some of this research does indeed find that countries which have been most inclined to adopt these policies tend to produce more unequal skills outcomes (Green, 2008; Green, Green and Pensiero, 2015). Relatively little research has been carried out on these relationships in the context of the upper secondary phase of education and training but what there is suggests that some different logics may apply here.

Green and Pensiero's (2015) quasi-cohort analysis of the effects of system types and characteristics on relative changes across countries in skills inequality after age 15 (using PISA and SAS), suggested that various measures of system standardisation reduce skills inequality. In their difference-in-difference analysis, a high rate of full ISCED 3 completion (representing the near universalisation of long-cycle programmes) was negatively associated with both inequality of skills outcomes (at $P<0.05$ for literacy and $P<0.01$ for numeracy) and inequality of skills opportunities (at $P<0.02$ for literacy and $P<0.05$ for numeracy). Another measure of system standardisation was the mandatory inclusion of Maths and the national language across upper secondary programmes. The analyses found that this also had a significant negative effect on both inequality of skills outcomes (at $P<0.05$ for literacy and $P<0.01$ for numeracy) and inequality of skills opportunities (at $P< 0.05$ for literacy and $P<0.01$ for numeracy). However, paradoxically, the research also showed that it was the countries with Dual Systems of apprenticeship (which have high levels of school diversity and tracking) which were the most effective in reducing skills inequality, having negative effects on both inequality of skills outcomes ($P<0.05$ for literacy and $P<0.02$ for numeracy) and inequality of skills opportunity ($P<0.3$ for literacy). The system type representing countries (e.g. Norway and Sweden) with the most institutionally standardised (comprehensive) high school systems had no significant effect in reducing inequalities of skills outcomes and was associated with increases in inequality of skills opportunities. In this account some aspects of school standardisation reduce skills inequality at the upper secondary level while other aspects increase it; the same applies, conversely, for school diversification.

Pedagogy and Teacher Quality

Although there is little direct evidence for the upper secondary phase of the effects on skills inequalities of teacher quality and different pedagogic practices, some inferences can be drawn from the literature on prior phases of education and from the more theoretical work on 'situated learning'. A substantial body of qualitative research on situated learning - stretching from the US (Lave and Wenger, 1990) to Europe (Evans and Niemeyer, 2006) and Asia (Koike and Inoki, 1990) - uses ethnographic approaches to explore the processes by which technical skills are acquired in work environments, stressing the importance of social contexts and relationships ('communities of practise') in the transfer of 'tacit' knowledge and skills. This may apply equally to the development of core literacy and numeracy skills, since, as noted above, these are enhanced by utilisation in the workplace (Desjardins and Ederer, 2015). Work-based learning is accessed mostly by students from vocational tracks which recruit more lower-achieving students and may therefore be contributing to narrowing the skills distribution by raising the skills of those at the bottom end. A similar argument has been made with reference to traditional teacher-led instructional methods which may be particularly beneficial to lower attaining students with less skill in autonomous learning.

By contrast, all students may be assumed to benefit from access to high quality teachers. The relevant question for upper secondary education and training, then, is how these teachers are distributed across different tracks and institutions. Unfortunately, only limited cross-country data are available on teacher qualifications and skills by type of institution and programme. However, if we assume that teacher pay is a reasonable proxy for teacher quality, then we can expect that more equal resource distribution between different types of school in some systems (measured for instance by spending per student) will contribute towards reducing skills inequality. A number of East Asian systems (e.g. Japan and South Korea) have been notable for resource equalisation policies which require the periodic rotation of teachers between schools by the regional authorities (National Centre on Education and Economy). Schools in Singapore have somewhat more discretion in the hiring of teachers but, as Schleicher notes (2011), vocational schools benefit from resource parity with academic schools and also from the maintenance of teacher skills through the extended CPD undertaken by teachers in schools of all sorts (see also Chong, 2014; Law, 2015).

1.3 A Typology of Upper Secondary Education and Training Systems

Our typology is developed on the basis of the theories and typologies commonly found in the comparative education and training literature which classifies systems according to institutional structures, forms of curriculum, pedagogy and assessment, and modes of governance and regulation (Busemeyer and Iversen, 2011; CEDEFOP, 2008; Dumas et al., 2013; Green, 2003; Greinert, 2004; Lasonen and Young, 1998; Maurice et al., 1986; McLean, 1999; OECD, 1985; Raffe et al., 2001; Verdier, 2013). However, we modify the typology in accordance with the latest empirical data on relevant indicators for the period under investigation.

We can identify six broad types of upper secondary education and training systems in OECD countries relating to systems as they were during the period when the 18-20 years olds tested in SAS (in 2011-2014) were going through upper secondary education (i.e., roughly 2008 - 2014). The country groups emerging from this bear a close resemblance to the classifications of economies and welfare systems in the comparative political economy literature, with liberal mostly Anglophone countries representing the core of one type, the social market and social democratic countries representing two further types, and southern Europe and East Asia, in some of the literature at least, being accredited with separate, though less distinctive, political economy models (see Esping-Andersen, 1990; Green et al., 2010; Hall, and Soskice, 2001; Pontussen, 2008). The connections between the two forms of classification analysed in the more recent literature should not be surprising since education systems form an integral part of welfare systems in general and since their functioning is substantially affected by the way the external contexts, such as labour market and welfare institutions, interact with them (see Green and Janmaat, 2011; Busemeyer, 2014).[1]

Type 1. These are predominantly school-based systems with general academic and vocational provision in different types of dedicated upper secondary institution and with apprenticeships representing separate but residual systems. This is the modal type in southern European countries and other western countries influenced historically by the French education system and also, through more complex genealogies, in central and eastern European (CEE) and East Asian countries (Green, 2013). Programmes in upper secondary institutions normally last for two or - more usually - three years from the age of 15, as in the original model of the modern French lycée,[2] and end with a qualification which gives access to general university higher education (ISCED 5A) in the case of general education students, and vocational tertiary education (at ISCED Level 4 or 5B) for vocational students. The curricula in different general and vocational programmes today generally share certain common core elements but programmes are typically organised around a cluster of subjects specific to the disciplinary or vocational orientation of programme. Diplomas are normally based on externally administered ‘grouped awards’ which require passes in a range of subjects, including core areas of language, Maths (and sometimes Civics). The majority of continental European and east Asian OECD countries have systems of this type (e.g., in our sample - Cyprus, Czech Republic, Denmark,

[1] The main difference between the two classification systems is that France tends to be classified with the southern European model in the education typologies, whereas it is normally grouped with north-western European countries in the comparative political economy literature. The advantage of considering the classification systems from the comparative education literature in addition to those from the comparative political economy literature is that it provides more fine-grained distinctions between the different types of school-based system, and the effects of their different pedagogic approaches, whereas the latter tends to focus primarily on a single contrast between apprenticeship systems and school-based systems and the effects of the contrasting forms of linkage between skills formation and labour market systems in the two cases.

[2] In some countries courses can last for 4 or 5 years (Italy) and in countries with grade repeating, such as France, some student may not graduate until age 19 or 20.

Estonia, Flanders (Belgium,) Finland, France, Greece, Italy, Japan, Korea, Lithuania, Netherlands, Poland, Russian Federation, Slovak Republic, Singapore and Slovenia.) [3]

Type 1b. A subset of these Type 1 countries (Denmark, Netherlands, Finland and Singapore) is also distinguished by having a substantial element of work-based training as part of their vocational school provision and represent a growing trend towards hybridisation of upper system types (Verdier, 2013). This may be organised in different ways. Denmark includes the option of hybrid school-based apprenticeships in all four types of academic High School (STX; HF; HHX; HTX) and also provides options within vocational colleges for work placements or apprenticeships organised on a 2 plus 2 model (two years school followed by two years apprenticeship (Méhaut, 2013). Students at Finland's vocational institutes generally receive school-based provision but a minority (around 10 percent) opt to take the same programme by an apprenticeship which may be organised by the Institute. Singapore's Institute for Technical Education also provides apprenticeship routes for some qualifications, and a minority of students undertake such apprenticeships, in firms but organised by the ITE, in conjunction with the school-based studies. In the Netherlands students following MBO vocational courses can opt to spend most of their time in apprentice training with their employer with day- or block - release for study in the vocational school.

Type 2. These are predominantly comprehensive, school-based systems with academic and vocational provision within the same institution and with, again, apprenticeships representing a largely residual alternative form. Provision is organised either as a standardised, core plus options programme, as in most North American high schools, or in differentiated programmes with distinctive subject specialisms but overlapping cores of general education, as in Norway or Sweden. Study durations tend to be standardised at two or three years across all tracks in the US and Canada and at three years in the Nordic countries. These systems share most of the characteristics of Type 1 systems but tend to have a higher degree of integration of curricula and assessment across the range of provision (Raffe et al., 2001). They can be regarded generally as relatively standardised on one level – since there is only one main type of upper secondary institution and all programmes tend to have long cycles.

[3] Denmark could be considered a borderline case because up to 30% of young people take a form of apprenticeship, but many of the apprentices are not on traditional apprentice contracts as in Dual System apprenticeships.

However, governance and regulation varies considerably between the US and Scandinavian contexts, with school choice and diversity policies in the federal US system leading to much greater institutional variation than would be found in the more standardised and unitary Nordic systems. US high school configurations vary by state, but across states include private high schools, public comprehensive high schools, Charter schools, Magnet schools and, in a few cases, vocational high schools. There is also considerable variation in levels of funding across districts and schools. Norway and Sweden, by contrast, have one type of public high school (*gymnasieskola*) and there are relatively few private schools (although this has increased in Sweden during the past decade). The two Nordic countries have a relatively high proportion in vocational programmes in high school, whereas the US classify the vast majority of their students as being in general programmes even where they are taking vocational options.

A number of countries have some comprehensive upper secondary institutions (*eniaia lykeio* in Greece; *lycées polyvalents* in France and tertiary colleges in England) but only four OECD countries have this type of institution as the dominant institutional form (Canada; Norway; Sweden and the USA). Because of differences in governance and regulation Type 2 systems are best divided between Type 2a for the North American variant and Type 2b for the Nordic variant.

Type 3. These are systems with participation distributed relatively equally between school-based general education and employment-based Dual Systems of apprenticeship and are found exclusively in social market political economies. In this kind of system, the provision at upper secondary level may be of similar duration across the different tracks (as with the normatively three-year apprenticeships and final stage *Abitur* courses in Germany), and the vocational track contains significant mandatory components of general education in all Dual System apprenticeships (Solga et al., 2014). However, the general and vocational tracks remain very distinctive, with sharp differences in forms of regulation, curricula and assessment, and with clearly differentiated final qualifications and subsequent progression possibilities in education, training and work (e.g. in Germany, university for *Abitur* graduates from the *Gymnasium* and *Realschule* and skilled jobs or higher technical courses for apprentice graduates). In respect of their Dual Systems, Type 3 systems have distinctive forms of regulation based on social partner organisations. This means that apprenticeship systems are closely integrated with labour market institutions and the world of work, and this has important effects on the labour market value of the qualifications they offer and the consequent incentives this provides for apprentices (Busemeyer and Iversen, 2011).

Dual System apprenticeships are generally considered to be of high quality and the programmes attract a large number of students, coming from across the ability range, including a substantial proportion graduating from the academic *Gymnasium* or even university (Schneider and Tieben, 2011). These higher achieving entrants add to the prestige of the vocational system, the quality of its outputs, and the value of its qualifications on the labour market, the latter being boosted in addition by strong labour market demand for intermediate skills (Hall and Soskice, 2001). With such apprenticeship systems, it is argued (Raffe et al., 2001), there will be greater parity of esteem between the academic and vocational tracks, and consequently expectation and achievement in the vocational tracks will be higher.

A number of countries have traditional Dual Systems of apprenticeship, where provision is regulated by the social partners, and with apprentices recruited by firms and placed on employment contracts (including Austria, Belgium, Denmark, Germany, Ireland, Luxembourg, Netherlands, and Switzerland). But it is only in Austria, Germany, and Switzerland where a third or more of young people participate in them (OECD, 2008, p. 331, Table C1.1; Steedman, 2001, 2010). In the varieties of capitalism literature these three countries present the core cases of the social market model of political economy.

Type 4. These are 'Mixed Systems' which include many different school- and employment-based programmes of variable length and quality but with dominant academic tracks. Systems of this type tend to have pronounced status gaps between academic and vocational programmes with the most qualified students entering academic programmes and the least qualified confined to vocational programmes which are often shorter in duration and do not necessarily offer progression routes into higher level programmes or high-quality jobs. Wage returns to vocational qualifications on the labour market tend to be relatively low (see for the UK: Greenwood et al., 2007). Where, as in the UK, there remains considerable labour market demand for low skills recruits, participants on these programmes may have little incentive to raise their skills levels (Keep & Mayhew, 2014), except in the limited range of high-quality apprentice programmes that offer long cycle training up to ISCED level 3.

Mixed systems are notable for their lack of curriculum standardisation across programmes since they lack a common core curriculum and do not mandate the study of Maths and the national language across all tracks. The general education component of vocational courses tends to be quite limited and vocational courses, particularly in Anglophone countries, are often competence-based. Students are assessed on the basis of their ability to demonstrate competences rather than on their knowledge of a syllabus, and programmes often do not have a prescribed duration. Assessment in general subjects can be through elective single subject awards (as with the A-levels in England) or, in a few cases, by grouped awards which specify a given combination of subjects to be assessed, as in *Bachillerato* in Spain. Regulation and governance in mixed systems is generally more liberal and market-oriented than in other systems, with much diversity in programmes and types of providers, including private training organisations and, in the case of the UK, private awarding bodies. Systems in this group tend to have lower participation rates amongst 17- and 18-year-olds [4] and relatively high rates of early school leaving (defined by the European Commission as those who leave education without qualifications above the ISCED 3C (short) level.) [5]

Systems broadly conforming to this type can be found in Australia, Chile, England, Northern Ireland, Ireland, Israel, Spain, and New Zealand. With the exception of Israel and Spain, these countries all belong to the liberal model of political economy identified in the varieties of capitalism literature.

2. Hypotheses

Our review of the research literature and our typology of education systems points to some tentative hypotheses, as follows:

Skills Levels

System characteristics affecting changes in levels of skills during the upper secondary phase:

[4] Spours, Hodgson and Rogers (for Edge Foundation) note short duration for England and Spain particularly.

[5] See Eurostat data at:
http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/School_enrolment_and_early_leavers_from_education_and_training

- The prevalence of private schools, and other forms of school choice and diversity, have little effect on relative changes in average levels of skills;
- Some aspects of curriculum standardisation (e.g., mandatory learning of Maths and the national language; prevalence of long-cycle level 3 programmes) will enhance average performance in core skills, while others (state control over textbooks) may not;
- Traditional mastery methods of instruction may enhance performance with certain groups of students (and thus overall standards) in particular cultural contexts;
- The prevalence of worked-based learning and relative parity of esteem between academic and vocational tracks will tend to enhance the performance of vocational students (thus raising average performance levels overall).

System type effects on changes in skills levels:

- Type 2b and Type 3 systems will perform well in enhancing average skills levels, because they combine a degree of standardisation in key areas (e.g., core curricula, long-cycle programmes) with an emphasis vocational learning and high teacher quality;
- Type 2a and Type 4 systems will perform less well in raising mean skills levels because of the relative absence of vocational learning and the relative lack of parity of esteem between academic and vocational tracks.
- Type 4 systems will be additionally disadvantaged by the lack of curriculum standardisation in core areas.

Skills Inequalities

Factors affecting changes in skills inequality during the upper secondary phase:

- Greater school diversification, through the proliferation of private schools and selective public schools may increase skills inequality, except in cases where academic and vocational school retain relative 'parity of esteem';
- Certain elements of school standardisation (such as in core curricula, resourcing and teacher quality) may reduce skills inequality.
- Extended vocational learning, particularly through apprenticeships, may raise skills levels amongst vocational students and thus reduce overall skills inequality.

System type effects on skills inequality:

- Type 2b systems and Type 3 systems will reduce skills inequality because of curriculum standardisation in key areas (mandatory literacy and language provision; length of courses) and due to relative parity of esteem between academic and vocational tracks (particularly in Type 3 where work-based learning raises skills of vocational students).
- Type 4 systems will be less effective in reducing skills inequality because of the institutional fragmentation and lack of curriculum standardisation in core areas, both of which undermine normative standards. Comparatively low recruitment to vocational programmes seen to be of low quality, and the large status gaps between these and academic programmes, will also tend to sustain skills inequalities in the upper secondary phase.
- Type 2a systems will be less effective in reducing skills inequality because of the relative absence of vocational learning.

3. Data and Methods

The focus of our empirical analysis lies in a comparison of average skills acquisition outcomes in literacy and numeracy for young people undertaking upper secondary education across different system-types. Second, we also examine the specific effect of individual system characteristics on acquisition of skills at upper-secondary level.

A further strand of investigation looks at the distribution of skills outcomes to assess differences in skills inequalities within system-types and the effect of system characteristics on changes in inequalities across upper-secondary education.

Countries included in the dataset are those for which data are available in the international PISA and PIAAC assessments. Both of these datasets examine cognitive skills in numeracy and literacy. PISA assesses students at age 15 – when they are at the beginning of their upper secondary education or, in some countries, at the end of their lower secondary studies. Whilst PIAAC samples the entire working-age population (ages 16-65), our analysis is restricted to participants within the 18-to-20 years age range – when the vast majority of young people will have completed their upper-secondary education. A quasi-cohort study approach is taken by matching cohorts taking part in the PISA survey at age 15 in 2006 and 2009 with the corresponding round of PIAAC in 2011/12 and 2014/15, when the same birth-cohort will be at least 20 years old.

Table 1: Countries included in the sample

Country	Country code	PISA wave	PIAAC round	Country	Country code	PISA wave	PIAAC round
Australia	AU	2006	2011/12	Italy	IT	2006	2011/12
Austria	AT	2006	2011/12	Japan	JP	2006	2011/12
Canada	CA	2006		Korea	KR	2006	2011/12
Canada (English)	CA-en		2011/12	Lithuania	LT	2009	2014/15
Canada (French)	CA-fr		2011/12	Netherlands	NL	2006	2011/12
Chile	CL	2009	2014/15	New Zealand	NZ	2009	2014/15
Czech Republic	CZ	2006	2011/12	Northern Ireland (UK)	UK-ni	2006	2011/12
Denmark	DK	2006	2011/12	Norway	NO	2006	2011/12
England (UK)	UK-en	2006	2011/12	Poland	PL	2006	2011/12
Estonia	EE	2006	2011/12	Russian Federation	RU	2006	2011/12
Finland	FI	2006	2011/12	Singapore	SG	2009	2014/15
Flanders (Belgium)	BE-fl	2006	2011/12	Slovak Republic	SK	2006	2011/12
France	FR	2006	2011/12	Slovenia	SI	2009	2014/15
Germany	DE	2006	2011/12	Spain	ES	2006	2011/12
Greece	GR	2009	2014/15	Sweden	SE	2006	2011/12
Ireland	IE	2006	2011/12	United States	US	2006	2011/12
Israel	IL	2009	2014/15				

In total, 32 systems are included. Note that this includes the sub-national systems of Flanders in Belgium, and England and Northern Ireland in the UK. Scores for PIAAC are available separately for English- and French-speaking Canada, whilst corresponding PISA scores are available for Canada as a single national unit. Where possible, sub-national indicators and controls are included for these regions of Canada, whilst where data are not available, national-level measures are included as proxies.

The degree to which PISA and PIAAC scores can be directly compared has been discussed in the literature (Hanushek & Woessmann, 2012; Gal & Tout, 2014; Cathles et al., 2018) and it is clear that the relationship is not strictly like-for-like. Nonetheless, studies have concluded that it is possible to compare results from the two tests with matched quasi-cohorts, whilst acknowledging the limitations inherent in this approach (Cathles et al., 2018, p. 8). Further, recent work by Pokropek and Borgonovi (2020) has sought to link 'pseudo-equivalent' groups using statistical matching procedures to arrive at robust concordances between PISA and PIAAC scores on numeracy and literacy scales.

Our analytical approach uses standardised scores for PISA and PIAAC to compare outcomes between students at the start and end of upper-secondary education, whilst advising caution when interpreting the strict comparability of the data.

As discussed in section 2.3 above, the 32 countries/territories in our sample can be categorised into system types according to the main characteristics of their education system at upper secondary level.

Table 2: Typology of upper secondary systems

System type	System type	No of territories	Country code
Type 1	Differentiated, dedicated upper secondary system	13	CZ, BE-fl, EE, FR, GR, IT, JP, KR, LT, PL, RU, SK, SI
Type 1b	Differentiated, dedicated upper secondary system, with school-based apprenticeships	4	DK, NL, FI, SG
Type 2a	Comprehensive system (N. American model)	2 (3)	CA (CA-en, CA-fr), US
Type 2b	Comprehensive system (Nordic model)	2	NO, SE
Type 3	Dual system	2	AT, DE
Type 4	Mixed system	8	AU, CL, UK-en, UK-ni, IE, IL, NZ, ES
All systems		**31 (32)**	

In addition to these indicators of skills levels and inequalities derived from the PISA and PIAAC datasets, our analysis includes key information regarding the main characteristics of the upper secondary education system, which differ according to systems typology, as well as in relation to various policy, cultural and historical factors. Grounded in the empirical and theoretical literature, these characteristics have been posited as relevant factors in explaining variation in the effectiveness of education systems for skills acquisition. These variables can be broadly divided into four groupings – timing and duration of upper secondary education; financial and human resources levied at upper secondary education; provision of education at upper-secondary level; and, finally, the vocational prevalence and completion rate of students.

Data on these indicators are derived, as far as possible, from publicly available datasets on educational statistics (OECD, 2021; UNESCO, 2021), and refer to time-points close to the start of our cohorts' upper-secondary education. Where suitable indicators cannot be found in the existing international literature to match relevant factors identified in the literature review, we have developed new indicators drawing on statistics from sources such as Eurydice or from examination of information in county-specific sources, whether from official Government statistics or from published research studies. Descriptive statistics for these variables can be found in the appendices and are presented in more detail in Section 7, below.

3.1 Analytic Approach

To assess the cross-system variation in skills outcomes and distributions, we adopt two parallel empirical strategies. First, we employ ordinary least-squares (OLS) regression to model the relationship between outcomes in numeracy and literacy and system type, following the specification:

$$Y = \alpha + \beta_1\, system\ type + \beta_2\, PISA\ score + \varepsilon$$

The dependent variable, *Y*, is the standardised score for literacy or numeracy as derived from the PIAAC dataset for participants aged 18-to-20. The parameter of interest relates to the regression coefficient (β_1) for system type, which is presented for each of the distinct systems types (Type 1b, 2a, 2b, 3 and 4), in comparison to the reference group (Type 1 systems). We control for the appropriate standardised PISA score (β_2) for matched cohorts at age 15 (reading scores for models regressing on literacy; Maths scores for those on numeracy) in order to account for the overall prior skills level within the cohort.

An identical approach is taken to assessing skills distributions by systems type but with the dependent variable, *Y*, taking the value of the Gini coefficient for the distribution of scores in literacy and numeracy within the PIAAC cohort. PISA distributions are controlled for by including the corresponding Gini coefficients as independent variables in these models and estimates are presented for each of the distinct system types in comparison to the Type 1 reference group.

Second, we seek to account for the effect of a wide range of cultural, historical and policy differences between countries included in our sample by adopting a difference-in-differences (DID) approach. As noted elsewhere (Green & Pensiero, 2016), OLS estimators are likely to be biased as they do not take into account systematic differences between countries at the start of upper secondary education. A DID approach controls for time-invariant factors that are likely to affect skills outcomes and inequalities over and above the estimated effect of the independent variable alone.

The parameter of interest in our DID models relates to the difference in average score (inequality) in one group before and after 'treatment' (defined in this case as the upper secondary phase of education) minus the difference in average score (inequality) in a reference group before and after 'treatment':

$$\gamma = \left(\bar{\iota}_{group_1 time_2} - \bar{\iota}_{group_1 time_1}\right) - \left(\bar{\iota}_{group_{ref} time_2} - \bar{\iota}_{group_{ref} time_1}\right)$$

This is implemented using a regression framework, with the following specification:

$$Y = \alpha + \gamma_0 \, system\ type + \beta_1 \, T + \gamma_1 \, T.system\ type + \varepsilon$$

As above, the dependent variable, *Y*, refers to the standardised score for literacy or numeracy (or Gini coefficient relating to the distribution of these scores). The variable, *T*, is a dummy that takes the value of 1 for observations relating to 18-20-year-olds (i.e. PIAAC) and 0 for age 15 observations (i.e. PISA). The parameter of interest, then, is given by the interaction term [γ_1 *T, system type]*, which refers to the change in the coefficient of the independent variable over the time period examined.

Whilst we expect results from the DID analysis to be more robust, given that they account for the time invariant factors that may affect outcomes and distributions of skills that are omitted in ordinary least-squares models, the OLS regression estimates can provide at least indicative evidence of an effect even if traditional thresholds of statistical significance are not attained in the DID analysis.

The next sections (5 and 6) present descriptive statistics for our outcome measures and key independent variables (system characteristic indicators). The analysis then moves on to present findings of our statistical models on the effect of system types and characteristics on skills outcomes and inequalities in skill distributions (Section 7).

4. Skills Outcomes

The following descriptive analysis looks at the average skills outcomes in literacy and numeracy for 18-to-20-year-olds, as measured in the PIAAC survey. We examine, firstly, the average scores in both domains across the different countries/regions in our sample. Then, we examine the underlying trends according to the system type in which these countries and regions are grouped.

In addition, we highlight the correlation between skills outcomes at the end of upper secondary education and the scores in reading and mathematics for matched cohorts in the PISA survey of 15-year-olds.

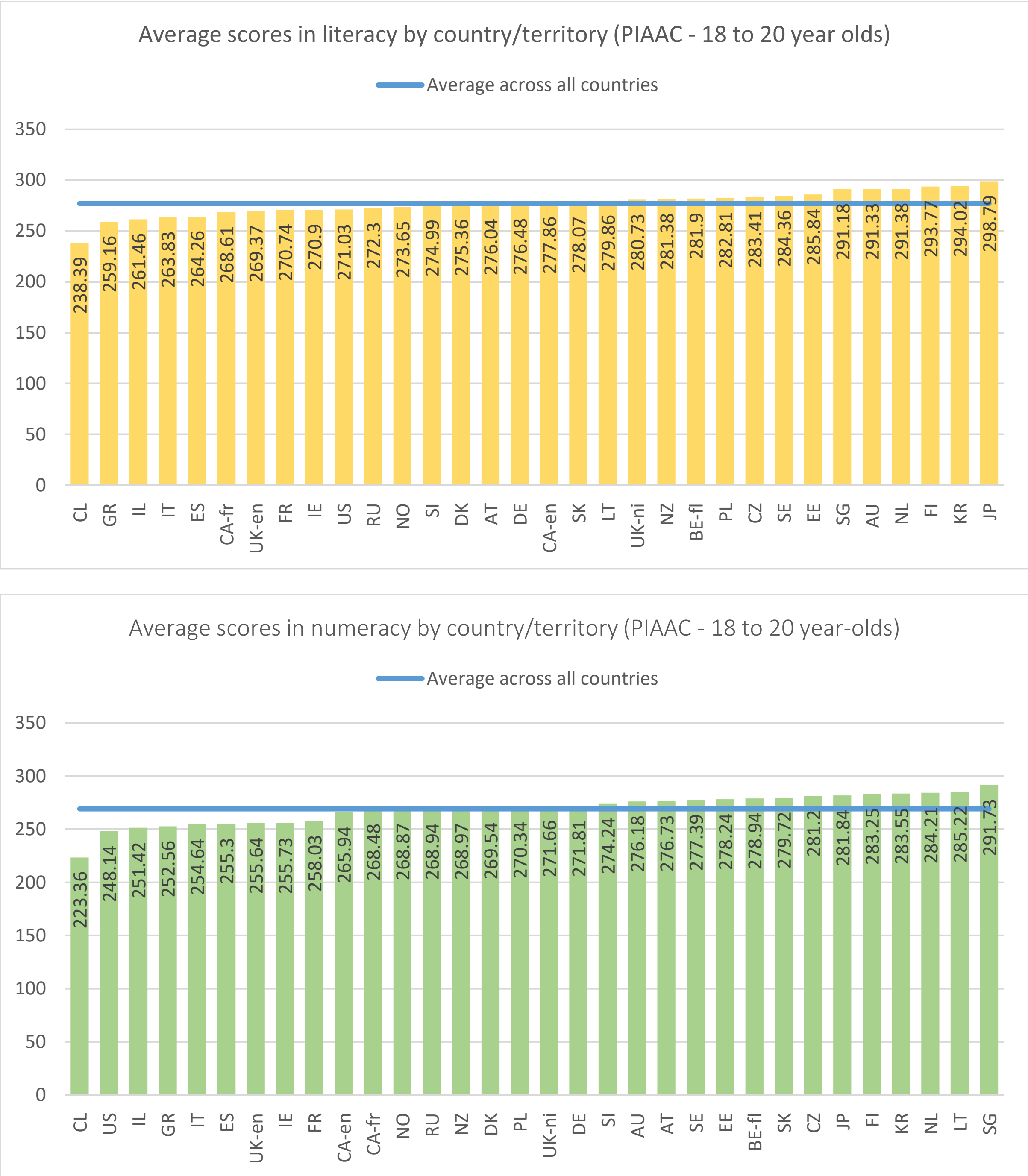
Following this, we quantify the distribution of skills outcomes within a country (and within system type) by using Gini coefficients of skills inequality. Our descriptive analysis looks at whether inequality in literacy/reading and numeracy/mathematics has increased between the time of the PISA and PIAAC surveys and the relative performance of different system types.

4.1 Average Scores in Literacy and Numeracy

The following graphs (figure 1) show average literacy and numeracy scores at age 18-20 by country/region. Possible scores for PIAAC outcomes range from 0 to 500. Across all countries in our sample the mean score for literacy was 276.98 (SD = 12.18), and for numeracy was 269.12 (SD = 14.19).

Countries with the highest average scores for literacy include Japan, South Korea, Finland, Netherlands and Austria, whilst those with the lowest scores include Chile, the US, Greece, Israel and Spain. For numeracy, highest average scores are seen in Singapore, Lithuania, the Netherlands, Korea and Finland, whilst Chile, the United States, Israel, Greece and Spain have the lowest scores. Finland and the Netherlands are notable for appearing in the top five in both literacy and numeracy.

Figure 1: Average scores in literacy and numeracy by country/territory (PIAAC – 18-20-year olds)



Grouping countries by system type allows us to identify some underlying trends in relation to average scores in literacy and numeracy.

Type 4 systems score lower, on average, than those in other system types, and below the average for all systems by 7 points in literacy and 12 points in numeracy. It should be noted that there is a large standard deviation around this average score for Type 4 systems. These mixed systems form a particularly disparate group and relatively low scores in literacy in Chile (238.4), have the effect of decreasing the overall average, compared to relatively-high-scoring countries, such as Australia (291.3) and New Zealand (281.4). Nevertheless, even removing the outlier Chile, the average for the group remains below the average for all systems by 2 points in literacy and 7 points in numeracy.

Type 2a systems (the US and the two regions in Canada) score, on average, below the country average in both literacy and numeracy, and below the average for other types, except Type 4. Anglophone Canada performs better than the other systems in the group in literacy, but all systems in the group perform below the system average in numeracy.

Type 1b systems perform, on average, higher than other system types in both domains, with average scores respectively 11 and 13 points higher in literacy and numeracy. This group includes Singapore, the Netherlands, Denmark and Finland.

Type 1 systems – the largest group of countries/territories – perform at around the average seen across the sample in both literacy and numeracy.

Table 3: Average scores in literacy and numeracy by system type (PIAAC – 18-20-year olds)

		Literacy		**Numeracy**	
	N	**Mean**	**SD**	**Mean**	**SD**
Type 1: Differentiated systems	*13*	278.90	11.03	272.88	11.25
Type 1b: Differentiated systems, with school-based apprenticeships	*4*	287.92	8.46	282.18	9.24
Type 2a: Comprehensive systems – North American	*3*	272.50	4.80	260.85	11.08
Type 2b: Comprehensive systems – Nordic	*2*	279.01	7.57	273.13	6.03
Type 3: Dual systems	*2*	276.26	0.31	274.27	3.48
Type 4: Mixed systems	*8*	269.73	16.06	257.28	16.45
TOTAL	***32***	**276.98**	**12.18**	**269.12**	**14.19**

Type 3 systems (Austria and Germany) perform similarly in both domains. Their average is close to the system average in literacy scores, but 5 points above the system average in numeracy.

4.2 Correlation between PISA Scores and PIAAC Scores

It is to be expected that average scores of students at age 15 in reading and Maths will correlate highly with scores, respectively, in literacy and numeracy amongst 18- to 20-year-olds. In order to account for the different scales on which the PISA and PIAAC surveys are scored, average country-level scores are standardised to have a mean of zero and a standard deviation of one.

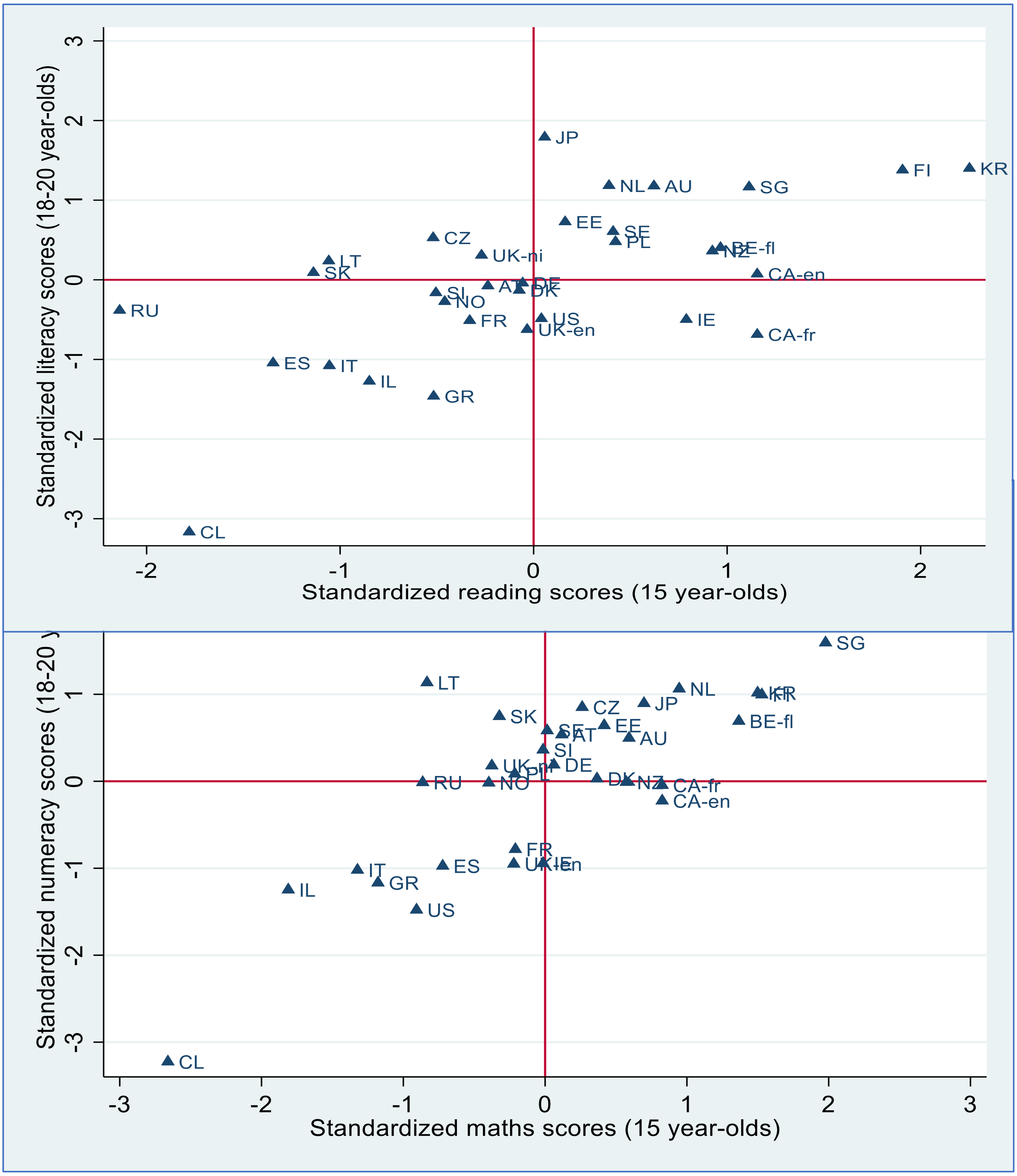
The graphs below (figure 2) show a strong correlation between PISA and PIAAC scores for literacy ($r = .616$, $p < .001$) and numeracy ($r = .780$, $p < .001$). Countries such as South Korea, Finland and Singapore are relatively high-performing in literacy in PISA and at the end of upper-secondary education, whilst in Chile, Israel, Italy and Spain scores show a consistent lower-than-average performance. In numeracy, Singapore, the Netherlands and Korea are consistently high performing in both tests, whilst Israel, Italy, Greece and the US are consistently low performing.

Interestingly, there are a small number of countries – notably Lithuania and (in literacy) Slovakia – that appear to rapidly 'catch-up' during upper-secondary education, with relatively poor performance in PISA being transformed to much higher skills outcomes by age 18-20. Austria performs at close to the average in numeracy at age 15 but considerably above the average by age 18-20. By contrast, some countries, e.g., Canada and (in literacy) Ireland, do relatively well in PISA at age 15 whilst this does not translate into similar performance in post-upper secondary skills outcomes. England performs at the average in literacy and numeracy at age 15 but some way below the average by age 18-20.

Interestingly, there are a small number of countries – notably Lithuania and (in literacy) Slovakia – that appear to rapidly 'catch-up' during upper-secondary education, with relatively poor performance in PISA being transformed to much higher skills outcomes by age 18-20. Austria performs at close to the average in numeracy at age 15 but considerably above the average by age 18-20. By contrast, some countries, e.g., Canada and (in literacy) Ireland, do

relatively well in PISA at age 15 whilst this does not translate into similar performance in post-upper secondary skills outcomes. England performs at the average in literacy and numeracy at age 15 but some way below the average by age 18-20.

Figure 2: Scatterplot of standardised scores in reading/literacy and maths/numeracy by country/territory (PISA – 15-year olds plotted versus PIAAC – 18-20-year olds)



4.3 Score Distributions in Literacy and Numeracy

Average scores tend to belie the level of within-country inequality in performance by summarising score distributions with a single value. Gini coefficients provide an indication of the inequality of skills distributions with each country or territory. This can allow us to examine the extent to which there is variation around the mean scores presented above and see where countries (and system types) appear to be more or less equal across the cohort. Gini coefficients closer to one indicate a greater level of inequality in the distribution of scores; those closer to zero indicate the opposite.

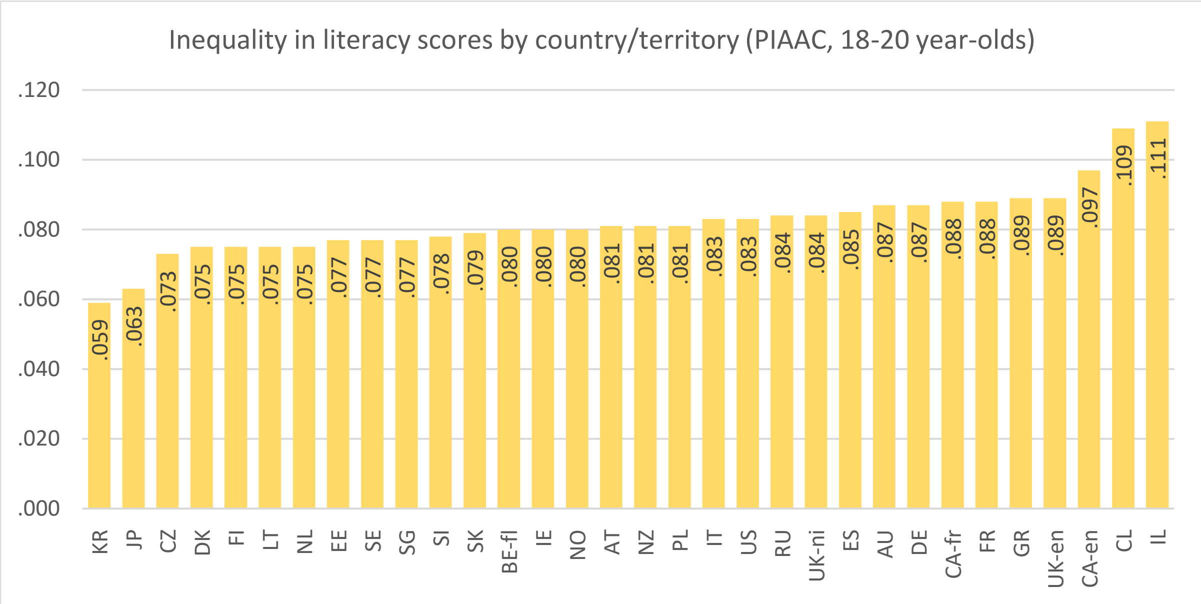
Overall, across our sample, the countries have a Gini coefficient of .082 in literacy and .095 in numeracy, indicating a slightly greater level of inequality with regards to numeracy scores compared to scores for literacy.

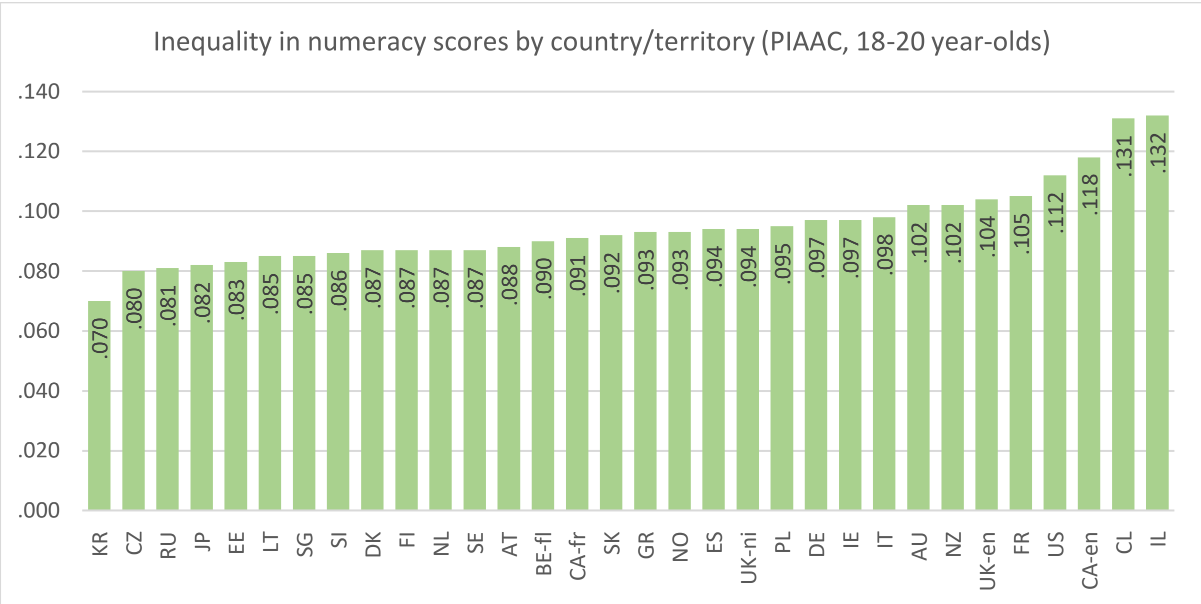
Countries with the highest levels of inequality for literacy scores include Israel, Chile, English-speaking Canada, England and Greece. The least unequal score distributions can be seen in South Korea, Japan, the Czech Republic, Germany and Finland.

For numeracy scores, the greatest levels of inequality can be seen, again, in Israel, Chile, English-speaking Canada and also in the US, France and England. Similarly, the least unequal score distributions can be seen in Japan, South Korea, and Czech Republic (as seen for literacy scores), as well as in Russia [6] and Estonia.

[6] It should be noted that results for Russia may not be reliable due to the exclusion of the Moscow region from the sample (see: OECD, 2016)

Figure 3: Gini coefficient for inequality in literacy and numeracy by country/territory (PIAAC – 18-20-year olds)

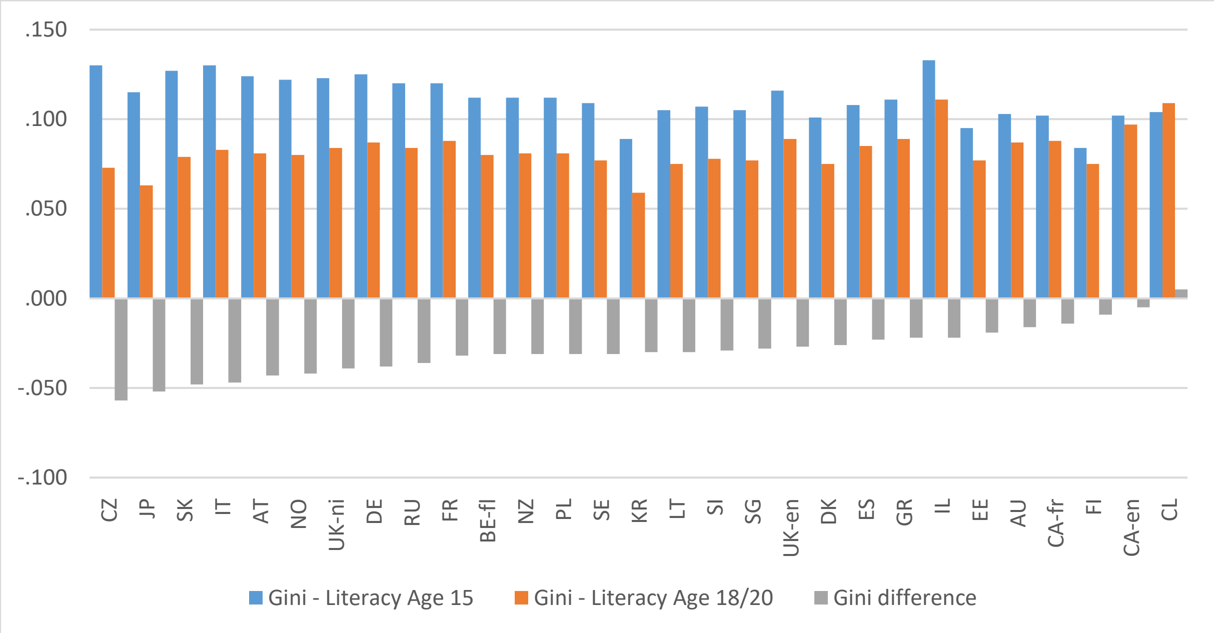


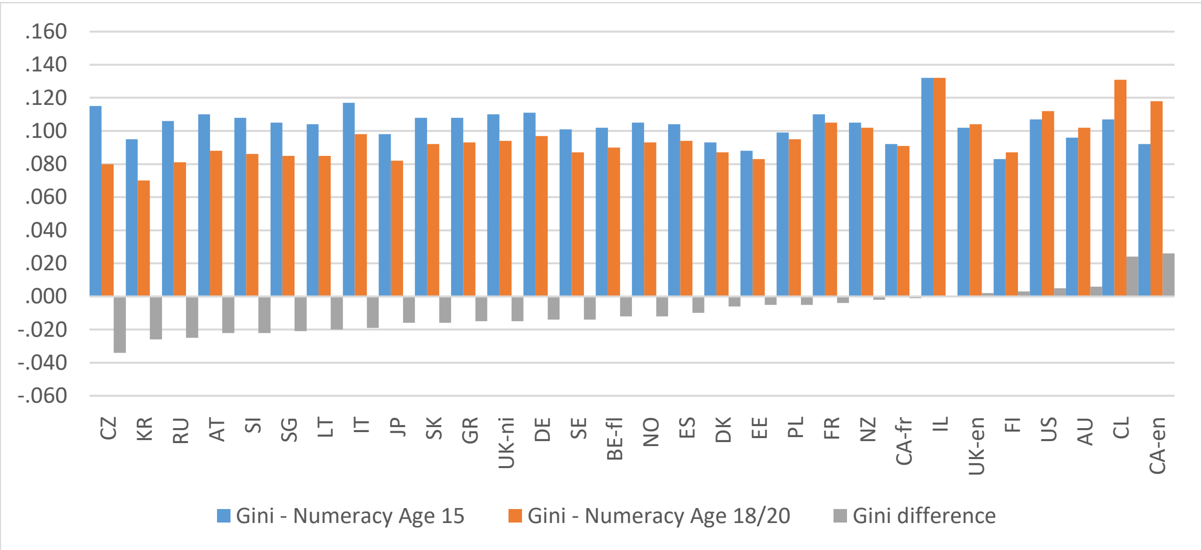


4.4 Correlation between Inequality in PISA Scores and PIAAC Scores

Following on from the previous section, we now look at how measures of skills inequalities change between age 15 and following the end of upper secondary education. The figures below (figures 4) show the difference between Gini coefficients in literacy and numeracy for young people aged 18 to 20 included in the PIAAC survey, compared to 15-year-olds in the matched PISA cohort.

Figure 4: Gini coefficients of inequality in relation to scores in reading/literacy and maths/numeracy by country/territory (PISA – 15-year olds and PIAAC – 18-20-year olds)





In literacy, we can see that all countries with the exception of Chile managed to reduce inequality in scores during the upper-secondary stage, although there is considerable variability in terms of the magnitude of this reduction. For numeracy, six countries see an increase in inequality, whilst for countries that see a reduction in inequality, these are typically more modest than those seen in literacy scores.

The countries seeing the greatest reduction in equalities for literacy (Czech Republic, Japan, Slovakia and Italy) all have Type 1 systems, whilst Type 3 systems (Austria and Germany) also improve their relative inequality of scores. At the other end of the spectrum, the pattern is less clear-cut with the smallest reductions (or even increases) in inequality being seen in Chile and Australia (Type 4), Canada (Type 2a) and Finland (Type 1b).

Similarly, for numeracy, countries with Type 1 systems (again the Czech Republic, but also this time Russia, Korea and Slovenia) see the greatest reduction in inequality, along with Austria and Singapore. The same countries that perform relatively badly in literacy are also those seeing increases in inequality for numeracy – Canada (English-speaking), Chile, Australia and Finland. In addition, the United States and England also experience small increases in inequality for numeracy scores in upper-secondary education.

Comparing levels of inequality between PISA scores at age 15 and PIAAC scores amongst 18-to-20-year olds by system type (table 4), we can see that overall, for both reading/literacy and maths/numeracy, the distribution of scores has become less unequal. This reduction in inequality is notable in Type 1 and Type 2b systems and, particularly, in Type 3 systems, who see the most unequal scores at age 15, yet have substantially less unequal distribution of scores at the end of upper-secondary education. By contrast, Type 2a North American systems are the only group of systems where levels of inequality get more pronounced throughout upper-secondary education. Type 4 mixed systems see a reduction in literacy inequality but flatline in numeracy inequality.

Table 4: Gini coefficients of inequality in relation to scores in reading/literacy and maths/numeracy by system type (PISA – 15-year olds and PIAAC – 18-20 year olds)

	Age 15 (PISA)	**Age 18-20 (PIAAC)**	**Age 15 (PISA)**	**Age 18-20 (PIAAC)**
	Reading	Literacy	Maths	Numeracy
Type 1: Differentiated systems	.113	.077	.104	.088
Type 1b: Differentiated systems, with school-based apprenticeships	.097	.076	.094	.086
Type 2a: Comprehensive systems – North American	.068	.090	.097	.107
Type 2b: Comprehensive systems – Nordic	.115	.079	.103	.090
Type 3: Dual systems	.125	.084	.111	.093
Type 4: Mixed systems	.114	.091	.108	.107
TOTAL	**.108**	**.082**	**.104**	**.095**

5. System - characteristic Indicators

Our analysis focuses on the effect of ten characteristics of upper secondary education systems, which differ according to systems typology, as well as in relation to various policy, cultural and historical factors. Grounded in the empirical and theoretical literature, these characteristics have been posited as relevant factors in explaining variation in the effectiveness of education systems for skills acquisition.

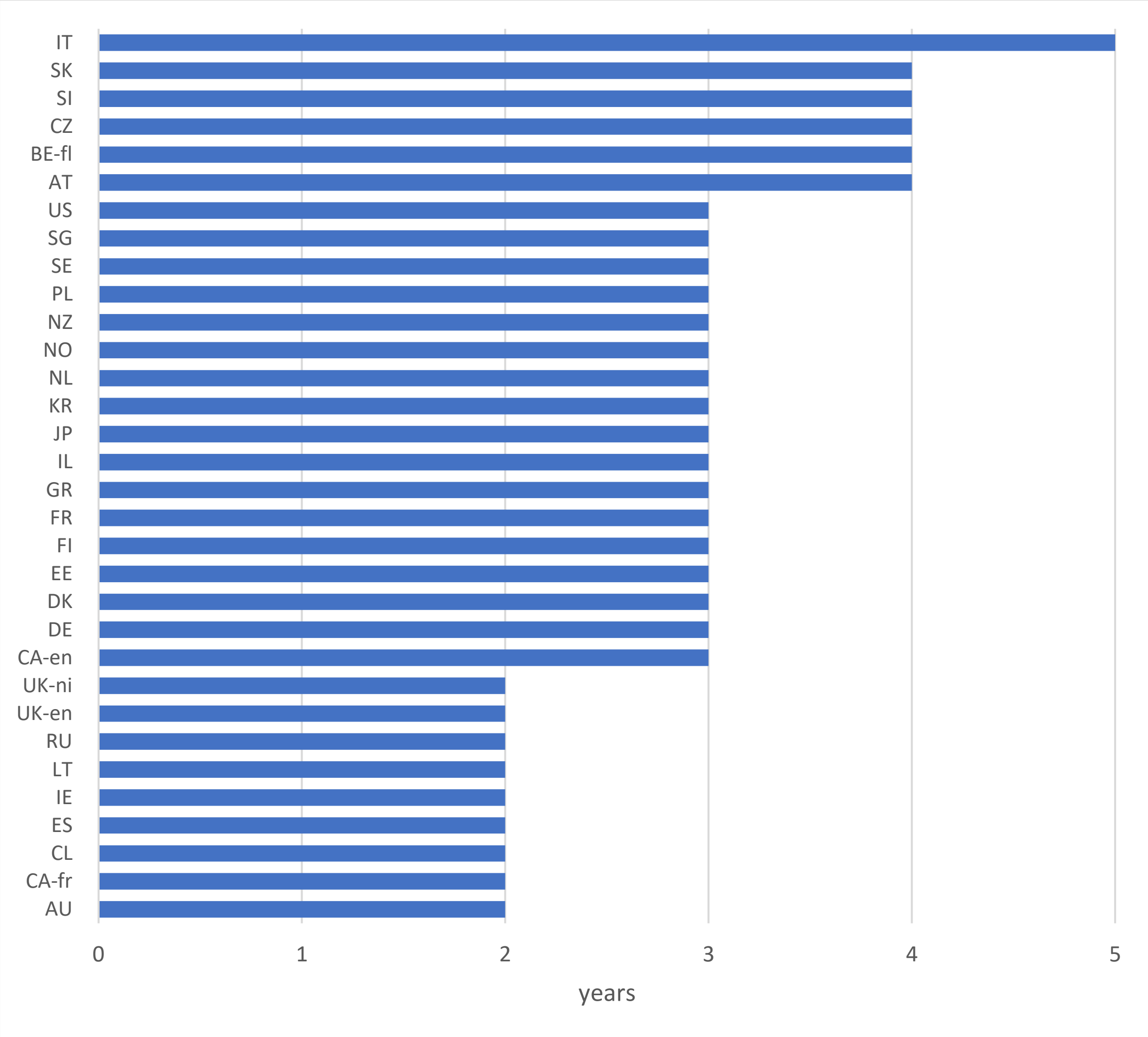
These variables can broadly be divided into four groupings: the first relates to the timing of education at upper secondary level – that is, the duration of courses, the stage at which selection by track is first applied, and the age at which students are expected to complete their education. Second, we examine the financial and human resources devoted to upper secondary education, including the average starting salary for teachers at this level, the proportion of per capita GDP spent on upper secondary education (and its distribution between different types) and the average overall workload of teachers. The third grouping relates to educational provision at upper secondary level; specifically, it takes into account the number of school types in which this stage of education occurs and whether students are obligated to study courses in Maths and/or their national language (a proxy for the level of standardisation within the curriculum). Finally, we consider the overall completion rate at upper secondary education and the prevalence of vocational education and training in the system as measures of the proportion of the upper secondary cohort who achieve a qualification at this level and the proportion who take programmes of vocational training.

5.1 Duration and Timing within Upper Secondary Education

Data are included in our analysis on three variables relating to the duration and timing of key transition points within the upper secondary system.

Theoretical duration of upper secondary education refers to the number of years students are expected to spend in upper secondary education. This figure is the theoretical duration based on when students in the main tacks may be expected to start and progress through this stage of education assuming no grade repetition or skipping of years (UNESCO).

Figure 5: Duration of upper secondary education by country/territory



The duration of upper secondary education in regions included in our sample ranges from 2 years (e.g., Australia, the UK, French-speaking Canada, Spain) to 5 years (Italy). The modal duration is 3 years, which is the case in 17 out of the 32 territories (53.1%), including for Singapore.

Looking at the duration of upper secondary education by system type (below), we can see that Differentiated (Type 1), Dual systems (Type 3) and Nordic comprehensive systems (Type 2b) have longer-than-average upper secondary programmes – 3.3 years, 3.5 years and 3 years respectively - with programmes in Comprehensive systems in North America (Type 2a; 2.67 years) and Mixed systems (Type 4; 2.25 years) lasting less time, on average. This is consistent with the system typology, as detailed in section 2.3.

Table 5: Duration of upper secondary education by system type

	N	Mean	SD
Type 1: Differentiated systems	13	3.31	.86
Type 1b: Differentiated systems, with school-based apprenticeships	4	3.00	.00
Type 2a: Comprehensive systems – North American	3	2.67	.58
Type 2b: Comprehensive systems – Nordic	2	3.00	.00
Type 3: Dual systems	2	3.50	.71
Type 4: Mixed systems	8	2.25	.46
TOTAL	32	2.94	.76

Stage at first selection provides an indicator for the degree of specialisation present in the education system overall. Students may be required to select (or be selected for) different tracks based on specialised curricula (e.g., vocational or technical programmes). Typically, School-based systems with differentiated tracks select students on entry to lower- or upper-secondary education. Comprehensive systems may have no selection or students select different programmes within a single school type during their upper-secondary education.

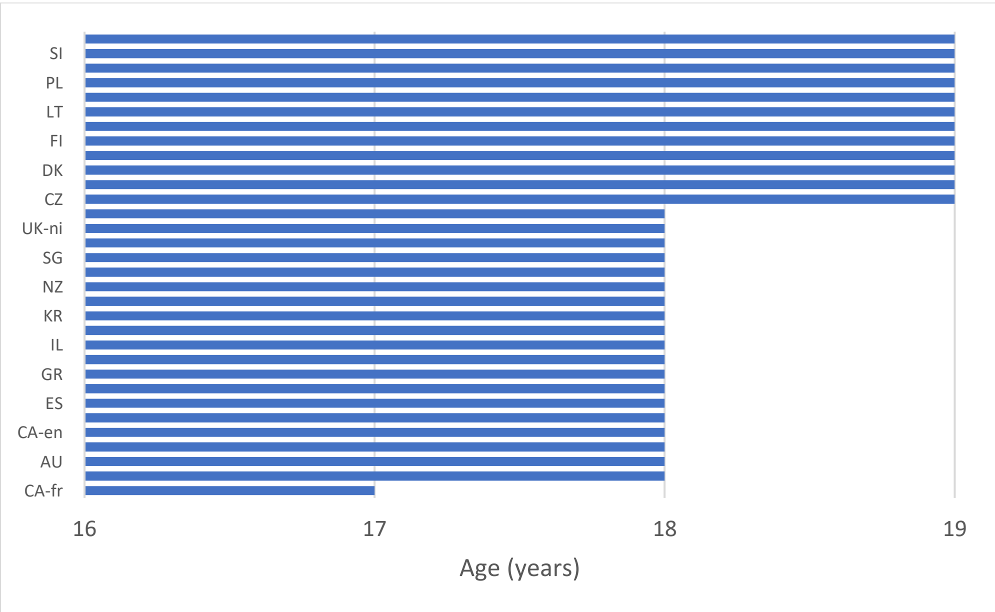
Table 6: Stage of first selection by system type

	On entry to lower secondary		On entry to upper secondary		During upper secondary	
	N	**row %**	**N**	**row %**	**N**	**row %**
Type 1: Differentiated systems	4	30.8	9	69.2	0	0.0
Type 1b: Differentiated systems, with school-based apprenticeships	2	50.0	2	50.0	0	0.0
Type 2a: Comprehensive systems – North American	0	0.0	1	33.3	2	66.7
Type 2b: Comprehensive systems – Nordic	0	0.0	0	0.0	2	100.0
Type 3: Dual systems	2	100.0	0	0.0	0	0.0
Type 4: Mixed systems	1	12.5	6	75.0	1	12.5
TOTAL	**9**	**28.1**	**18**	**56.3**	**5**	**15.6**

The majority of systems first track students on entry to upper secondary education (56.3%). This is predominantly the case in Type 1 (Differentiated) and Type 4 (Mixed) systems. Type 3 systems all select on entry to lower secondary education. A significant minority of Type 1 systems (Czech Republic, Slovakia, Flanders and Lithuania) also have selection on entry to lower secondary education, as do the Netherlands and Singapore (from Type 1b). By contrast, with the exception of French-speaking Canada, the Type 2 (Comprehensive) systems have selection during stages of upper secondary education.

Theoretical age at completion of upper secondary education is a variable derived from two UNESCO Institute for Statistics indicators: official entrance age to upper secondary education; and theoretical duration of this educational stage (UNESCO, 2020). It is calculated as the simple summation of these two numbers, such that a student entering upper secondary education at age 15 and undertaking a three-year course, for example, may be theoretically expected to complete their programme at age 18 (as above, assuming no grade repetition of skipping of years). All of the territories included in our analytical sample, with the exception of French-speaking Canada (17 years), have a theoretical completion age of upper secondary of either 18 or 19. By far the most common age at completion is 18 years old (59.4%), with 37.5% of systems having upper secondary completion at age 18. It is important to recognise that this variable does not take into account the statutory age at which students may legally leave education, which may be prior to upper secondary education, on the attainment of upper secondary education or once a young person reaches a specific age, irrespective of their educational progress. It also takes no account of those who never enter upper secondary education or who drop out early.

Figure 6: Theoretical age at completion of upper secondary by country/territory



The table below shows that students in mixed systems typically graduate at 18, those in differentiated and dual systems complete upper secondary education at relatively older ages (on average 18.5 and 19, respectively). There is a degree of variability for comprehensive systems – with Nordic systems having students complete at age 19 and North American systems typically finishes when students are 17 or 18.

Table 7: Theoretical age at completion of upper secondary education by system type

	Age 17		Age 18		Age 19	
	N	**row %**	**N**	**row %**	**N**	**row %**
Type 1: Differentiated systems	0	0.0	6	46.2	7	53.9
Type 1b: Differentiated systems, with school-based apprenticeships	0	0.0	2	50.0	2	50.0
Type 2a: Comprehensive systems – North American	1	33.3	2	66.7	0	0.0
Type 2b: Comprehensive systems – Nordic	0	0.0	0	0.0	2	100.0
Type 3: Dual systems	0	0.0	1	50.0	1	50.0
Type 4: Mixed systems	0	0.0	8	100.0	0	0.0
TOTAL	**1**	**3.1**	**19**	**59.4**	**12**	**37.5**

5.2 Financial and Human Resources in Upper Secondary Education

Starting salary for teachers refers to the annual statutory teachers' starting salaries in public institutions at upper secondary level. Salaries are given in US dollars, converted using PPP for private consumption (OECD, 2012). Data are available for 29 of the 32 regions in our sample, with an overall mean starting salary across these countries of $29,500 (SD = $10,200).

The graph below shows that Germany, Denmark and Spain have the highest levels of salaries for teachers at upper secondary level, with the average for new starters exceeding $40,000. By contrast, new teachers in Israel, Chile and the Eastern European countries (Czech Republic, Poland, Estonia and Slovakia) receive amongst the lowest teacher salaries – on average, less than $20,000.

The highest starting salaries for teachers at upper secondary level are found within Type 3 systems (US$ 43,300) (driven by Germany) and in Type 1b systems. On average, teachers in

Type 1 (US$ 22,900) and Type 4 (US$ 28,800) systems receive starting salaries that are below the average for all systems. The Comprehensive Type 2 systems have the least variability in starting salaries, with these systems also having higher-than-average salaries compared to all system.

Figure 7: Average teacher starting salary (US$, PPP) in upper secondary by country/territory

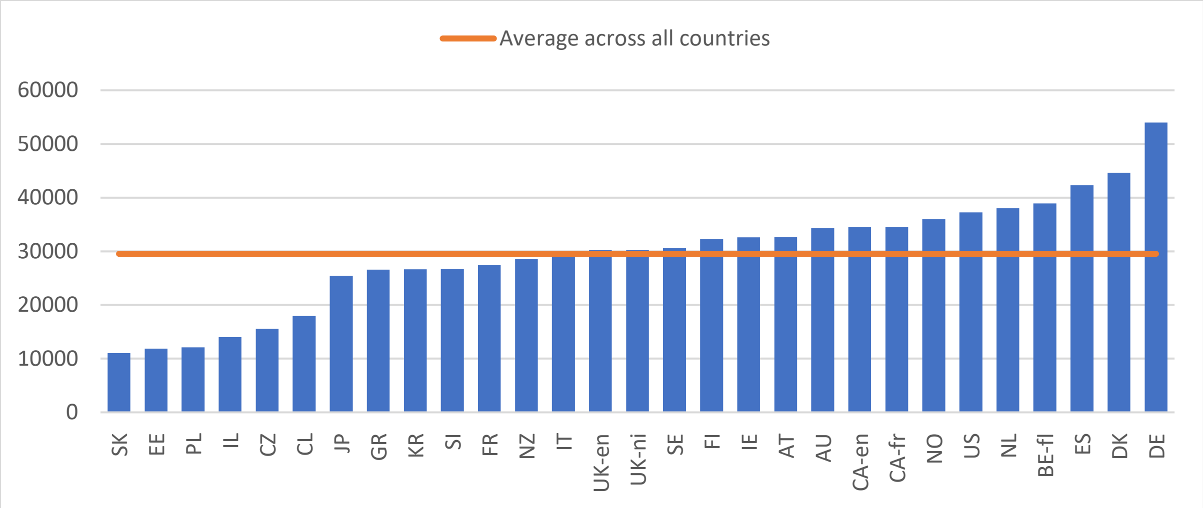


Table 8: Average teacher starting salary (US$ PPP) in upper secondary by system type

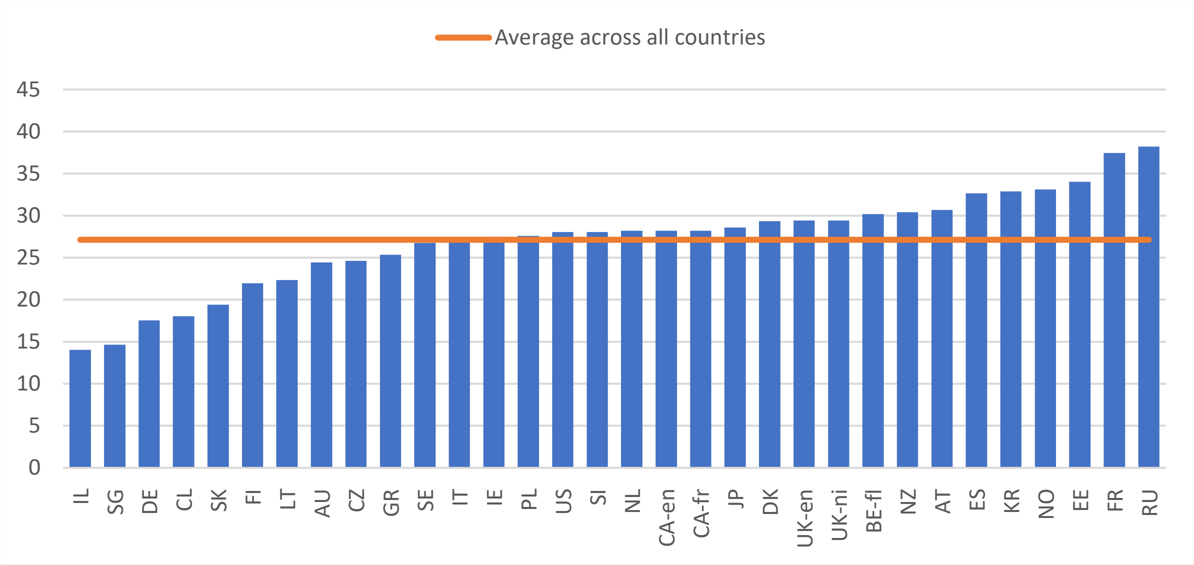
	N	Mean	SD
Type 1 (Differentiated systems)	11	22,857.61	8,933.19
Type 1b (Differentiated systems, with school-based apprenticeships)	3	38,305.87	6,187.60
Type 2a (Comprehensive systems – North American)	3	35,481.15	1,546.45
Type 2b (Comprehensive systems – Nordic)	2	33,320.42	3,776.64
Type 3 (Dual systems)	2	43,321.36	15,049.09
Type 4 (Mixed systems)	8	28,765.82	9,015.55
TOTAL	**29**	**29,524.31**	**10,229.35**

Expenditure on education at upper secondary level measures public expenditure on educational institutions per full-time equivalent student as a proportion of per capita GDP (OECD, 2013). The figures refer to upper secondary level education.[7] On average, across our sample, countries spend around 27% of per capita GDP on upper secondary education.

[7] The data refer to spending per student over the course of their upper secondary programmer. OECD estimate this by multiplying the annual spending by the theoretical duration of courses. (See OECD EAG (2008a), p. 45.

As a proportion of per capita GDP, Russia (38.2%) and France (37.4%) spend the most on educational institutions (per FTE student). Israel (14.1%) and Singapore (14.6%) spend relatively much less on upper secondary education.

Figure 8: Expenditure per FTE student at upper secondary level (% per capita GDP) by country/territory



Overall, Type 1 and Type 2a/2b systems appear to spend relatively more as a proportion of per capita GDP on upper secondary education than Type 1b, Type 3 and Type 4 systems, although there is considerable variability within each system-type grouping. It should be noted that Type 3 systems with substantial apprentice training also benefit from the net spending of employers on training, not included in these figures.

Table 9: Expenditure per FTE student at upper secondary level (% per capita GDP) by system type. Data for 2016.

	N	**Mean**	**SD**
Type 1 (Differentiated systems)	13	28.89	5.58
Type 1b (Differentiated systems, with school-based apprenticeships)	4	23.54	6.76
Type 2a (Comprehensive systems – North American)	3	28.14	.10
Type 2b (Comprehensive systems – Nordic)	2	29.93	4.51
Type 3 (Dual systems)	2	24.11	9.27
Type 4 (Mixed systems)	8	25.73	6.52
TOTAL	**32**	**27.13**	**5.84**

Spending per student in vocational tracks as proportion of spending per student in general tracks. This is a derived variable based on OECD data (EAG, 2018) on spending per student by track in 2015 (OECD, 2018, p. 245). It can be a taken an indicator for Government priorities regarding general and vocational education and therefore as one proxy for 'parity of esteem'.

Figure 9: Spending per student in vocational tracks as proportion of spending per student in general tracks

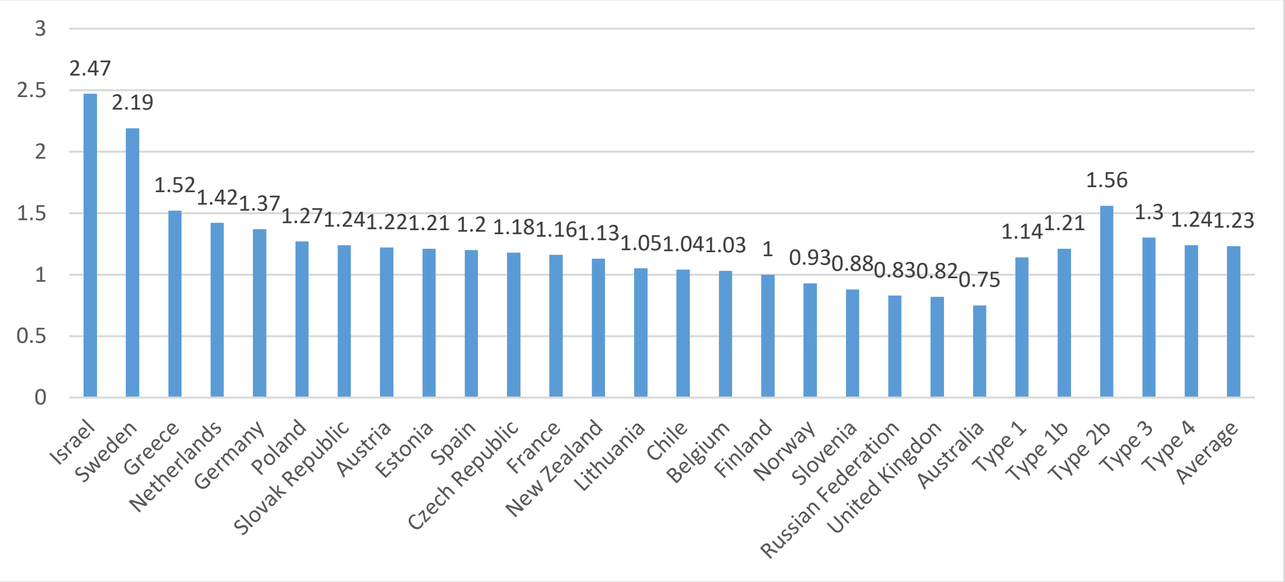


Figure 9 shows that most countries spend more per student in vocational upper secondary education than in general education. On the average across sampled OECD countries, public spending per student is 1.23 times higher for students in vocational upper secondary education than for those in general upper secondary education. Countries spending relatively highly on their vocational students include Israel, Sweden, Greece, Netherlands and Germany. Those spending relatively less include Australia, the UK, the Russian Federation and Slovenia. Type 2b countries and Type 3 countries have the highest relative spending on vocational students, consistent with the literature which marks out these countries as achieving greater 'parity of esteem' than others. Type 1 systems spend least on vocational students relative to general students. Type 4 countries are somewhere in the middle, with an outlier country, Israel, driving up the average. Excluding Israel brings the average below that of other groups (at 0.98), reflecting the relatively low spend on vocational students by Australia and the UK in particular.

Teacher workload is a derived variable calculated from two measures: average student-teacher ratio and average number of teaching hours per year (OECD, 2012). Both measures refer to upper secondary level only. These two values are multiplied together to obtain a measure of the average number of teaching hours undertaken, accounting for the average number of students for which teachers are responsible. This provides a reasonable proxy for the measurement of teacher workload across different systems. Complete data is available for 29 of the 32 regions in our sample.

Figure 10 shows that teacher workload in Chile is far greater than seen in the other territories included in our sample. The Netherlands and the US also have relatively high teacher workloads. By contrast, teachers in Greece, Lithuania, Norway and Denmark have relatively-less-demanding workloads.

Figure 10: Teacher workload (student-hours per year) by country/territory

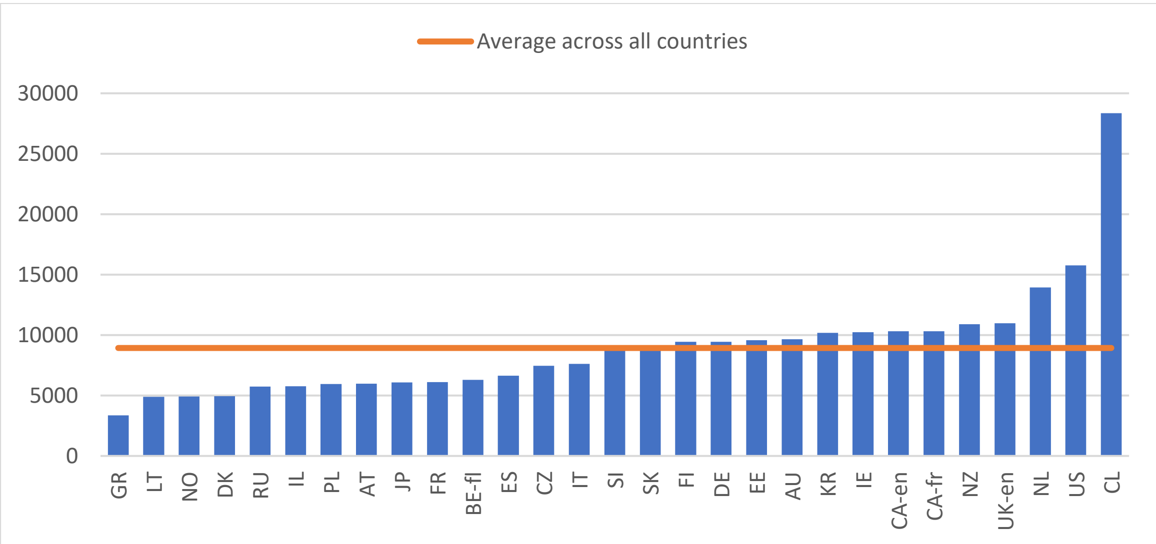


Table 10: Teacher workload (student-hours per year) by system type

	N	**Mean**	**SD**
Type 1 (Differentiated systems)	13	7,032.08	2,010.16
Type 1b (Differentiated systems, with school-based apprenticeships)	3	9,440.06	4,508.29
Type 2a (Comprehensive systems – North American)	3	12,121.43	3,143.17
Type 2b (Comprehensive systems – Nordic)	1	4,928.69	.
Type 3 (Dual systems)	2	7,706.36	2,458.39
Type 4 (Mixed systems)	7	11,788.61	7,592.53
TOTAL	29	8,929.77	4,687.33

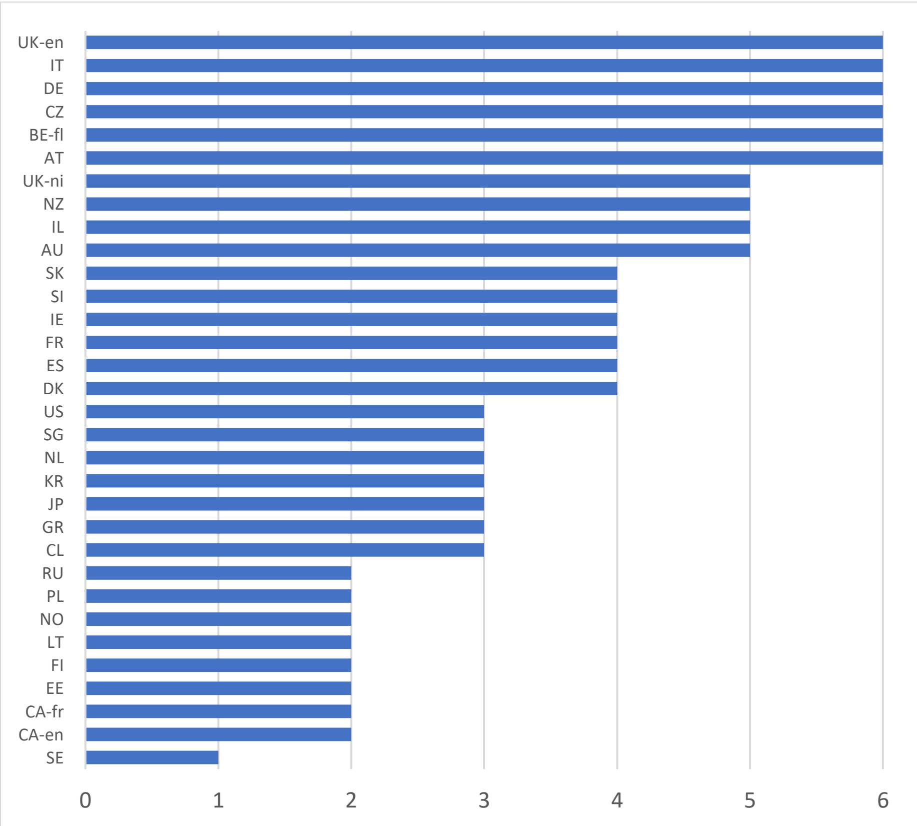
Looking at teacher workload by system types, we can see that teachers in Type 2a systems have the highest workloads, with those in Type 4 also experiencing higher-than-average workloads (although this is skewed considerably by data on Chile, reflected in the high standard deviation). By contrast Norway (the only included Type 2b system) and the Type 3 and Type 1 systems have lower-than-average workloads overall.

5.3 Provision of Upper Secondary Education

Number of school types provides an indication of how differentiated an education system is at upper secondary level. The data come from a range of sources, including Eurydice. The values are derived by counting the main distinctive types of public and private schools in the sector.

The number of school types ranges from one (in Sweden) to six (e.g., England (UK), Italy and Germany). The modal number of types of school is two, which is the case in 8 out of the 32 territories (25%).

Figure 11: Number of school types by country/territory



It is clear to see that comprehensive systems have the fewest number of school types – typically only 1 or 2 – whilst countries with Dual systems have the largest (6 different types of school). Differentiated systems also typically have 3 or more types of schools, whilst amongst the Mixed systems, there is also a relatively high number of school types (on average, 4.6), indicative a more decentralised and fragmented system of educational provision.

Table 11: Number of school types by system type

	N	**Mean**	**SD**
Type 1 (Differentiated systems)	13	3.62	1.56
Type 1b (Differentiated systems, with school-based apprenticeships)	4	3.00	.82
Type 2a (Comprehensive systems – North American)	3	2.33	.58
Type 2b (Comprehensive systems – Nordic)	2	1.50	.71
Type 3 (Dual systems)	2	6.00	.00
Type 4 (Mixed systems)	8	4.63	.92
TOTAL	**32**	**3.69**	**1.53**

Standardised curriculum score is a quantitative score ranging from zero to four, derived in relation to whether students are obliged to take courses in Maths and their national language at upper secondary level. A country's system is scored '1' if most of its programmes require compulsory mathematics/national language and '2' if all of its programmes have such a requirement. Scores for mathematics and national language are added together for a total score, which provides a proxy measure of curriculum standardisation.

Data are obtained via Hogden et al. (2010) and Eurydice (2019) and relate to the situation prevailing around 2010 (when our PIACC sample enters upper secondary education). Values are available for 30 out of 32 countries in our sample. Values range from zero (e.g., Australia, England, New Zealand and Ireland) to four (e.g., United States, Sweden and Poland). Type 4 systems appear to have the least standardised curriculum in relation to compulsory Maths and national language (mean score of .57), whilst the Type 2b Nordic systems both have the maximum score of 4 and type 1 systems close with 3.5, as consistent with our typology.

Table 12: Standardised curriculum score by system type

	N	Mean	SD
Type 1: Differentiated systems	12	3.50	.80
Type 1b: Differentiated systems, with school-based apprenticeships	4	2.50	1.29
Type 2a: Comprehensive systems – North American	3	2.00	1.73
Type 2b: Comprehensive systems – Nordic	2	4.00	.00
Type 3: Dual systems	2	2.00	.00
Type 4: Mixed systems	7	0.57	.98
TOTAL	**30**	**2.47**	**1.53**

5.4 Upper Secondary Completion rates and Vocational Prevalence

These two variables give an indication as to the extent to which students complete their studies in each country and what proportion of them go through upper secondary on vocational programmes, as compared to general academic courses.

Upper secondary completion rate is derived from the PIAAC dataset and relates to the proportion of young people aged 20 to 24 who have completed an upper secondary level qualification over two or more years (at ISCED 3 or higher). Whilst these data come from the PIAAC survey, and so are subject to potential sampling bias, to ensure cross-national comparability, we have selected the age group of 20-to-24-year olds, which should include the vast majority of leavers from upper secondary education.

The highest rate of ISCED 3 completion within the 20-to-24-year olds are seen in Singapore (98%), South Korea (97%), and Russia (96%); The lowest rates are found in Spain (62%), Italy (66%), Northern Ireland (73%) and Denmark (73%).

Overall, an average of 85% of students complete upper secondary education. Completion rates are slightly higher than average in Type 1 and Type 2a systems, at close to 90%; they are slightly lower than average in Type 2b systems (82%) and Type 4 systems (81%).

Figure 12: Upper secondary completion rate by country/territory

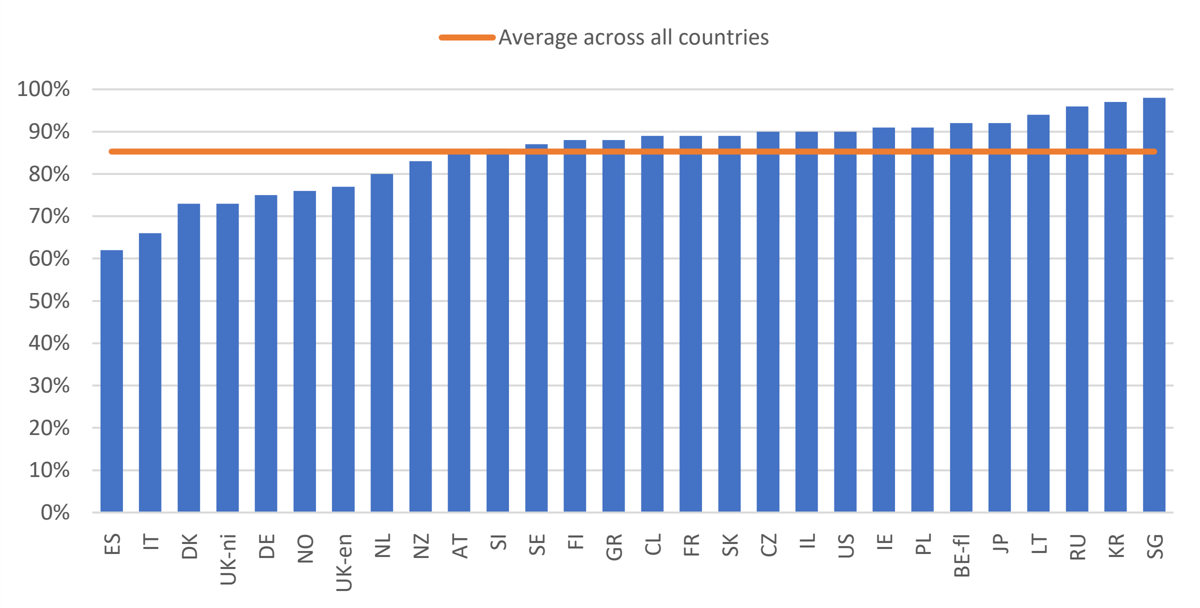


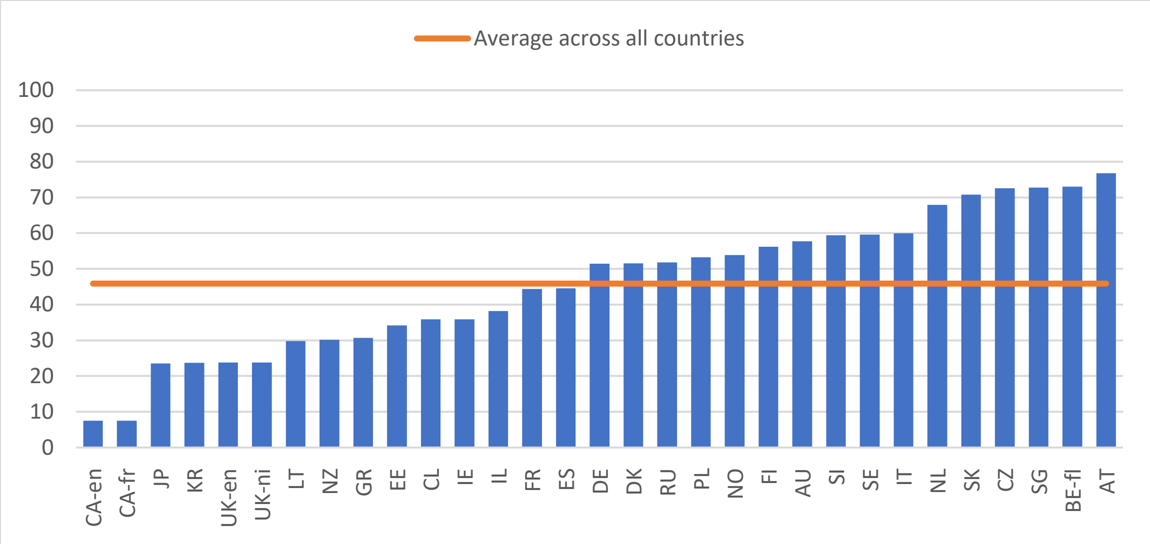
Table 13: Upper secondary completion rate by system type

	N	Mean	SD
Type 1: Differentiated systems	12	0.89	.08
Type 1b: Differentiated systems, with school-based apprenticeships	4	0.85	.11
Type 2a: Comprehensive systems – North American	1	0.90	.
Type 2b: Comprehensive systems – Nordic	2	0.82	.08
Type 3: Dual systems	2	0.80	.07
Type 4: Mixed systems	7	0.81	.11
TOTAL	**28**	**0.85**	**.09**

Vocational prevalence refers to the proportion of students enrolled on upper-secondary level programmes who are in vocational education (UNESCO, 2020). Data are available for all countries except for the United States.[8] Overall, across all countries in our sample, approximately 46% of students are enrolled in vocational courses of study at upper-secondary level. However, there is a great degree of variability between countries and systems. In Canada, for example, only around 7% of students are in a vocational programme, whilst in Austria, the figure is more than 75%.

[8] In the US all high school completers are categorised as having achieved upper secondary education.

Figure 13: Vocational prevalence by country/territory



As shown in the table below, Type 3 (Dual system) and Type 1b (Differentiated with school-based apprenticeships) have the highest proportion of vocational students – 64% and 62% respectively. By contrast, Type 4 (Mixed) and Type 2a (Comprehensive North American – which includes only Canada here) have significantly lower vocational prevalence. Interestingly, the Comprehensive Nordic systems (Type 2b) have more than half of the upper secondary cohort in vocational education (although the majority of these are taking vocational programmes within comprehensive high schools). This exceeds the proportion of those on vocational programmes in Type 1 (Differentiated) systems, where students on general programmes are typically still the majority.

Table 14: Vocational prevalence by system type

	N	**Mean**	**SD**
Type 1: Differentiated systems	13	48.24	18.46
Type 1b: Differentiated systems, with school-based apprenticeships	4	62.10	9.90
Type 2a: Comprehensive systems – North American	2	7.49	.00
Type 2b: Comprehensive systems – Nordic	2	56.71	4.04
Type 3: Dual systems	2	64.13	17.88
Type 4: Mixed systems	8	36.26	11.24
TOTAL	**31**	**45.88**	**19.42**

7. Modelling the Effect of System Characteristics on Skills Acquisition

As mentioned in Section 4, we adopt two parallel modelling approaches to analyse the effect of system characteristics on skills acquisition and inequalities in skills outcomes. First, we undertake ordinary least-squares (OLS) regression to assess the extent to which different system types (and individual system-characteristic indicators) are associated with better or worse performance in literacy and numeracy at the end of upper secondary education. We include system-level PISA scores in (respectively) reading and Maths as control variables in order to account for the strong correlation between performance at age 15 and outcomes following completion of upper secondary schooling.

Additionally, in order further to control for omitted systemic time-invariant factors that are likely to affect outcome scores, we adopt a difference-in-difference (DID) approach. In this analysis, we compare change over time (from age 15 to post-upper secondary) in average scores across different systems (or according to the level of different system-characteristic indicators). The DID estimator is equivalent to the interaction term of the outcome score with a binary indicator that is '0' at $time_1$ (age 15) and '1' at $time_2$ (age 18-20).

7.1 The Effect of System Type on Skills Outcomes

The following table (15) presents the results of OLS regression analysis to model scores for literacy and numeracy at age 18-20 in PIAAC, controlling for the equivalent score at age 15 in PISA, and comparing different system types to the reference group (Type 1: Differentiated systems). All scores are standardised to have a standard deviation of 1 and a mean of 0.

There is, as expected, a strong correlation between country-level average scores for reading and Maths at age 15 with scores, respectively, in literacy and numeracy at age 18-20. In comparison to the reference group, average scores for literacy are significantly lower in Comprehensive systems (North American) and in Mixed systems. Lower average scores are also seen in Comprehensive systems (Nordic) and Type 3 systems, although these are not found to be statistically significant. A similar relationship can be seen in numeracy scores, with Comprehensive systems (North America) and Mixed systems having significantly lower average scores compared with Differentiated systems.

Table 15: OLS regression analysis of the effects of system type on literacy and numeracy scores

	Literacy – Age 18-20 (Standardised values)		**Numeracy – Age 18-20 (Standardised values)**	
	OLS estimate	**S.E.**	**OLS estimate**	**S.E.**
System type (ref: Type 1 – Differentiated systems)				
Type 1b	.024	(.449)	-.282	(.343)
Type 2a	**-1.210****	**(.494)**	**-1.074*****	**(.350)**
Type 2b	-.147	(.558)	.119	(.413)
Type 3	-.292	(.557)	-.010	(.413)
Type 4	**-.765****	**(.330)**	**-.709*****	**(.252)**
Literacy – Age 15 (standardised values)	.654***	(.147)		
Numeracy – Age 15 (standardised values)			.745***	(.115)
Constant	.329	(.207)	.306*	(.151)
Observations	32		32	
R-squared	.566		.762	

Statistically significant: * p < .05; ** p < .01; *** p < .001

Undertaking the same analysis but using a DID approach yields similar results. The DID models (table 16) show that, compared to Differentiated Type 1 systems, countries with Type 2a Comprehensive systems (North American) perform statistically-significantly worse in both literacy and numeracy. Whilst not attaining standard levels of statistical significance (as seen in the OLS regression above), Type 4 Mixed systems also have a relatively large negative effect size on literacy and numeracy, in comparison to Type 1 systems (respectively .772 and .575 standard deviations).

Of note is the relatively large (-.604 standard deviations) negative effect on numeracy outcomes for students in countries with Type 1b systems, which was found to be much smaller under the OLS approach.

Table 16: DID estimates of the effects of system type on literacy and numeracy scores

	Literacy (Standardised values)		**Numeracy (Standardised values)**	
	DID estimate (γ1Y.age1820)	**SE**	**DID estimate (γ1Y.age1820)**	**SE**
System type (ref: Type 1 – Differentiated systems)				
Type 1b	-.355	(.786)	-.604	(.712)
Type 2a	**-1.571***	**(.880)**	**-1.151**$^{+}$	**(.798)**
Type 2b	-.229	(1.044)	.154	(.947)
Type 3	-.332	(1.044)	-.047	(.947)
Type 4	-.772	(.618)	-.575	(.548)
Constant	-.262	(.270)	-.055	(.245)
Observations	32		32	
R-squared	.208		.325	

Statistically significant: $^{+}$ $p < .2$; * $p < .1$; ** $p < .05$

The graphs below illustrate the effect of system type on literacy and numeracy according to the difference-in-difference analysis, with effect sizes shown by the difference in standardised scores between age 15 and age 18-20 (unbroken lines), relative to the difference that would have been expected to be seen were scores to have changed in line with those for Type 1 systems (dotted lines).

The graphs provide a clear illustration of the relative performance amongst countries with Type 2a system and – to a lesser extent – in those with Type 4 and Type 1b systems, compared to the most common Type 1 systems.

Figure 14: Difference-in-difference effects of system type on literacy and numeracy

7.2 The Effects of System Characteristics on Skills Outcomes

What system-specific characteristics, then, can account for these relative differences in performance? The subsequent analysis seeks to model the effect of individual characteristics of educational systems at upper secondary level to allow for a more fine-grained examination of specific indicators that may influence some systems to improve their performance more than others.

Table 17 reports the OLS and DID estimates for twelve models that examine the independent effect of system-level characteristics summarised in section 7. Models 1 to 3 look at the effect of duration and timing of key transition points within the upper secondary system. Whilst the theoretical duration of upper secondary programmes (model 1) does not appear to have a significant effect on either literacy or numeracy scores, the educational stage at which students first select from among a range of programmes (model 2) does. The OLS estimates show that, compared to systems in which selection occurs on entry to lower secondary education, there is a negative effect on literacy and numeracy outcomes of selecting students later – on entry to upper secondary (-0.67 for literacy, significant to the $p < .05$ level; -0.72 for numeracy, significant to the $p < .01$ level), or during upper secondary (-0.69 for literacy, not-significant; -0.77, significant to the $p < .1$ level). The DID estimates show similar effect sizes, whilst they do not attain statistical significance.

Model 3 examines the effect of the theoretical age at which students are expected to complete their upper secondary education. The estimates are positive, with effect sizes of between 0.4 and 0.6 in literacy and 0.6 and 0.7 in numeracy, which indicates an improvement in skills outcomes for students expected to continue in upper secondary to age 19. However, this association is only statistically significant for the OLS model for numeracy scores (0.66; $p < .01$).

Models 4 to 6 assess the effect of financial and human resources in upper secondary education. The effect of teachers' salaries (model 4) appears to be negligible and non-significant in all cases. Similarly, there is no significant effect on the variability seen in our sample of countries according to levels of educational expenditure at upper secondary (relative to per capita GDP; model 5). By contrast, the derived measure of teacher workload does have a small-yet-significant negative effect of students' skills acquisition (-0.07 for both literacy and numeracy), at least in the OLS models. Whilst the effect size remains comparable within the DID estimates, these are no longer statistically significant. This does imply, however, a correlation between higher average workloads and poorer student performance.

Number of school types, whilst a fair indicator of the level of differentiation within a system, has a negative effect on literacy and numeracy in both OLS and DID analyses but none of these effects are significant (model 7). The effect of having a more standardised curriculum (model 8) is modest (effect sizes between 0.12 and 0.18) yet positive, indicating that students who are required to study Maths and/or their national language outperform those within systems with less-standardised curricula. This appears to be more strongly the case in relation to numeracy, where the effect size is between 0.17 and 0.18 and, in the case of the OLS estimate, statistically significant to the $p < .05$ level.

The proportion completing upper secondary education (model 9) does not appear to have any significant effect on skills levels. This remains the case, even when a composite indicator of cohort years (combining theoretical duration of upper secondary education with the actual proportion of students who complete this) is included in the analysis (model 11).

Table 17: OLS and DID estimates of the effect of system characteristics on skills outcomes

		Literacy (Standardised values)				**Numeracy (Standardised values)**			
		OLS estimate (S.E.)	**R-sq.**	**DID estimate γ1Y.age1820 (S.E.)**	**R-sq.**	**OLS estimate (S.E.)**	**R-sq.**	**DID estimate γ1Y.age1820 (S.E.)**	**R-sq.**
Model 1 (N: 32)	Theoretical duration of upper secondary education	.167 (.190)	.396	.156 (.338)	.010	.178 (.151)	.626	.137 (.334)	.035
Model 2 (N: 32)	Educational stage of first selection (ref: on entry to lower secondary)		.472		.057		.714		.151
	On entry to upper secondary	-.661** (.312)		-.668 (.580)		-.721*** (.236)		-.562 (.550)	
	During upper secondary	-.686 (.432)		-.870 (.792)		-.771* (.316)		-.672+ (.752)	
Model 3 (N:32)	Age at completion of upper secondary (ref: 17 or 18 years)		.435		.020		.713		.046
	19 years	.483 (.286)		.576 (.520)		.659*** (.202)		.672 (.513)	
Model 4 (N: 29)	Teachers' starting salary (*US$ thousands*)	-.009 (.015)	.434	-.015 (.026)	.028	-.019 (.010)	.734	-.022 (.025)	.044
Model 5 (N: 32)	Expenditure per FTE student at upper secondary (*% per capita GDP*)	.018 (.025)	.390	.017 (.044)	.006	-.012 (.020)	.612	-.017 (.044)	.009
Model 6 (N: 29)	Teaching workload (*thousand student-hours per year*)	-.074** (.031)	.477	-.074 (.058)	.059	-.070*** (.024)	.687	-.055 (.053)	.148
Model 7 (N: 32)	Number of school types	-.035 (.096)	.383	-.0003 (.167)	.019	-.039 (.076)	.611	-.028 (.168)	.010
Model 8 (N: 30)	Standardised curriculum score (*0–4*)	.123 (.097)	.419	.148 (.174)	.013	.165** (.072)	.679	.177 (.171)	.020

Model 9 (N: 28)	Upper secondary completion rate (%)	.012 (.017)	.460	.002 (.030)	.071	.010 (.039)	.652	.005 (.031)	.059
Model 10 (N: 31)	Vocational prevalence (%)	.015** (.007)	.468	.018+ (.014)	.031	.012** (.006)	.657	.011 (.013)	.062
Model 11 (N: 28)	Cohort years of upper secondary (Duration of upper secondary*completion rate)	.002 (.002)	.473	.001 (.004)	.076	.002 (.002)	.662	.001 (.004)	.130
Model 12 (N: 31)	Voc. prevalence*Edu. stage at first selection (ref: entry lower secondary)		.542		.122		.790		.228
	On entry to upper secondary	.026+ (.019)		.036 (.035)		.028** (.013)		.036 (.032)	
	During upper secondary	.025 (.024)		.035 (.044)		.042** (.016)		.053+ (.040)	

Statistically significant: + $p < .2$; * $p < .1$; ** $p < .05$; *** $p < .01$

What does appear to matter is the proportion of students in a system who are enrolled on vocational programmes. The effect sizes for vocational prevalence (model 10) are small – between 0.01 and 0.02 yet significant in most cases (with the exception of the DID estimates for numeracy scores). Small positive model estimates indicate that those systems with relatively more students in vocational education at upper secondary see modest improvements in literacy and numeracy outcomes. Moreover, when combined with stage of first selection (model 12), these effect sizes increase, notably in the models for numeracy, where for countries with higher prevalence of vocational students, later selection, on entry to upper secondary education or during this stage, is associated with better outcomes.

7.3 The Effect of System Types and Characteristics on Inequality of Skills Outcomes

The following analysis looks at the effect of system types and characteristics on changes in inequality of skills outcomes. The analysis uses the same principles as above, adopting two parallel modelling approaches: ordinary-least-squares regression and difference-in-difference. In this case, our dependent variable is the Gini coefficient of literacy and numeracy scores as calculated from scores in the PIAAC survey of 18-to-20-year olds.

Whilst the OLS models include Gini inequality coefficients in the PISA survey of 15 years olds as a control, the DID analysis uses an approach that assumes change over time in inequality (from 15-year olds to post-upper secondary) can be attributed to the independent variable(s), whilst holding other time-invariant factors constant.

Table 18 reports the findings of thirteen models, which examine the effect of system type (model 1), timing and duration of upper secondary education (models 2 to 4), financial and human resource factors (models 5 to 7), the provision/organisation of education (models 8 and 9), completion rates (model 10) and vocational prevalence (model 11) on inequality of skills outcomes in literacy and numeracy. The final two models examine interaction effects of theoretical duration of upper secondary and completion rates (as a measure of the overall cohort years of study at upper secondary level); and of vocational prevalence and educational stage of first selection (as a measure of the extent to which early or later selection affects systems with a greater or lesser proportion of vocational students).

Similar to the analysis of the outcomes themselves, we can see that systems type does have a significant effect on changes in inequality in the distribution of these outcomes. In particular, compared to the most-common Type 1 systems, Type 2a Comprehensive (North American) systems and Type 4 Mixed systems become more unequal between age 15 and post-upper secondary in both literacy and numeracy outcomes. In addition, for literacy outcomes, Type 1b systems are significantly more unequal at the end of upper secondary education, compared to the reference Type 1 group.

Model 2 shows that theoretical duration of upper secondary has a significant effect on inequality in literacy (between -0.005 and -0.009, $p < .05$) which is negative in the OLS estimation and positive in the DID estimation. For numeracy scores, whilst both models show negative effects, returning similar effect sizes, the estimates attain statistical significance only in the DID model (-0.0067, $p < .1$). This implies that systems with longer programmes in mainstream upper secondary education return less-unequal skills outcomes.

The effect of educational stage at selection differs slightly for literacy and numeracy outcomes. For literacy, the DID model reveals a significant ($p < .05$) and positive effect on inequality for systems selecting on entry to upper secondary, compared to those selecting on entry to lower secondary. The OLS models for numeracy however only reveal a significant (and positive) effect on inequality for systems selecting during upper secondary. The DID models provide evidence for an increasing (and statistically-significant) positive effect on inequality as selection happens later – with the largest effect size being seen for systems selecting during upper secondary, as compared to those selecting on entry to lower secondary (0.02, $p < .05$).

Table 18: OLS and DID estimates of the effect of system characteristics on inequality of skills outcomes

		Inequality in literacy scores (Gini coefficient)				Inequality in numeracy scores (Gini coefficient)			
		OLS estimate (S.E.)	**R-sq.**	**DID estimate γ1Y.age1820 (S.E.)**	**R-sq.**	**OLS estimate (S.E.)**	**R-sq.**	**DID estimate γ1Y.age1820 (S.E.)**	**R-sq.**
Model 1 (N: 32)	System type (ref: Type 1: Differentiated)		.490		.760		.595		.477
	Differentiated, w/sch-based app's	.00309 (.00613)		.0147* (.00869)		.00482 (.00690)		.00911 (.00868)	
	Comprehensive (N. American)	.0183** (.00680)		.0232** (.0101)		.0237*** (.00688)		.0267*** (.00915)	
	Comprehensive (Nordic)	.000514 (00657)		-.00101 (.0109)		.00326 (.00767)		.00375 (.0109)	
	Dual	.00354 (.00681)		-.00458 (.0109)		.00112 (.00780)		-.00134 (.0109)	
	Mixed	.0144*** (.00405)		.0122* (.00657)		.0189**** (.00480)		.0161** (.00656)	
Model 2 (N: 32)	Theoretical duration of upper secondary education	-.00543** (.00255)	.213	00856** 00372)	.673	-.00573 (.00322)	.240	-.00673* (.00399)	.175
Model 3 (N: 32)	Educational stage of first selection (ref: on entry to lower secondary)		.155		.673		.334		.222
	On entry to upper secondary	.00721 (.00480)		.0142** (.0066)		.0109 (.00542)		.0128* (.00689)	
	During upper secondary	.00687 (.00653)		.0124 (.00924)		.0182** (.00710)		.0200** (.00929)	
Model 4 (N:32)	Age at completion of upper secondary (ref: 17 or 18 years)		.186		.654		.242		.198
	19 years	-.00711 (.00379)		-.00662 (.0059)		-.00862 (.00477)		-.00713 (.00606)	

Model 5 (N: 29)	Teachers' starting salary (*US$ thousands*)	-.0000282 (.000216)	.066	.000068 (.00031)	.627	-.0000061 (.000258)	.176	.000041 (.00032)	.113
Model 6 (N: 32)	Expenditure per FTE student at upper secondary (*% per capita GDP*)	-.000553 (.000313)	.174	-.000402 (.00049)	.661	-.000587 (.000422)	.208	-.00033 (.00049)	.233
Model 7 (N: 29)	Teaching workload (*thousand student-hours per year*)	.00123*** (.000437)	.309	.00151** (.00067)	.654	.00167*** (.000470)	.448	.00159** (.00067)	.244
Model 8 (N: 32)	Number of school types	-.000128 (.00152)	.076	-.00329* (.00173)	.706	-.000544 (.00185)	.153	-.00163 (.00190)	.222
Model 9 (N: 30)	Standardised curriculum (*0–4*)	-.00312** (.00112)	.262	-.00191 (.00181)	.698	-.00333** (.00153)	.179	-.00298* (.00178)	.258
Model 10 (N: 28)	Upper secondary completion rate (*%*)	-.000105 (.000234)	.144	.000128 (.000338)	.691	-.000157 (.000266)	.289	-.000111 (.00035)	.217
Model 11 (N: 31)	Vocational prevalence (*%*)	-.000169 (.000103)	.162	-.000303** (.00015)	.663	-.000274** (.000123)	.283	-.000321** (.00016)	.208
Model 12 (N: 28)	Cohort years of upper secondary (Duration of upper secondary*completion rate)	-.00006* (.0000293)	.275	-.0000736⁺ (.0000463)	.693	-.0000697* (.0000345)	.387	-.0000693⁺ (.0000475)	.245
Model 13 (N: 31)	Voc. prevalence*Edu. stage at first selection (ref: entry lower secondary)		.262		.688		.464		.286
	On entry to upper secondary	.000066 (.0002498)		.000079 (.0003931)		.0000346 (.0002783)		.0000206 (.0004025)	
	During upper secondary	-.0003453 (.0003131)		-.0004231 (.000492)		-.0006226* (.0003493)		-.0006478 (.0005039)	

Statistically significant: ⁺ $p < .2$; * $p < .1$; ** $p < .05$; *** $p < .01$; **** $p < .001$

Model 4, examining the effect of age at completion of upper secondary education does not yield any statistically significant differences between systems where students are expected to complete at age 19 and those where they finish at age 17 or 18.

Neither of the models that look at teachers' salaries (model 5) or public expenditure on upper secondary education (model 6) return statistically significant results, although there is non-significant negative association of public expenditure the skills inequality. By contrast, whilst the effect sizes are small, there is a statistically significant ($p < .05$) positive effect of teacher workload on inequality in both literacy and numeracy scores (model 7) – teachers under greater workload pressure being associated with worsening inequalities in score distributions.

Model 8, examining the effect of the number of types of school, indicates that systems with fewer school types are associated with greater skills inequalities. Nonetheless, the estimates only attain statistical significance in the DID model for literacy scores ($p < .1$).

The effect of standardised curriculum scores (model 9) indicates that systems with less-standardised curricula have greater inequality of skills scores for both literacy and numeracy; results with are significant except in the case of the DID estimation for literacy. The magnitude of this effect appears to be greater for numeracy than for literacy, however, as students required to take courses in numeracy in particular suggests a more uniform level of skills acquisition across the cohort.

Whilst there appears to be no statistically significant effect of the overall upper-secondary completion rate within a country (model 10), when examined in conjunction with programme duration to assess overall cohort year of upper secondary (model 12), there is a small-yet-(marginally) significant effect. This indicates that fewer cohort-years of upper secondary education are associated with greater inequality in score distributions for literacy and numeracy.

For systems with greater prevalence of vocational students in upper secondary education (model 11), there appears to be a negative effect on inequality, suggesting that skills acquisition is more equally distributed in systems with higher levels of vocational provision. Whilst these results are statistically significant for literacy in the DID model only, the effect can be seen for numeracy in both the OLS and DID models.

For systems with greater proportion of vocational students (model 13), there is a slight negative effect of selecting students during upper secondary education, compared to on entry to lower secondary on inequality of outcomes in numeracy. These systems see slightly more equal numeracy scores as vocational prevalence increases.

8. Discussion and Conclusions

8.1 Effects of System Types and Characteristics on Changes in Mean Scores between Ages 15 and 18-20.

The most notable finding from our analysis is that the Type 2a (North American comprehensive) systems and the Type 4 (mixed) systems perform consistently poorly relative to other system types in both literacy and numeracy skills acquisition between age 15 and age 18-20.

Both system types show large negative effects, relative to the reference group, on mean scores in both of the two domains and on each method of estimation. Type 2a performs least well of all the system types with significant negative effects in both domains on both types of calculation. Type 4 performs slightly better than Type 2a but worse than other system types, with significant negative effects for both domains in the OLS calculations.

These results for the Type 2a and Type 4 systems broadly confirm the findings in Pensiero and Green's (2018) study, using different methods, that these two system types perform worse than other types in improving skills after age 15. However, in this case Type 4 systems perform better than type 2a systems, rather than vice-versa.

This difference may be partly explained by the fact that the earlier study used PISA 2000 scores as the baseline, whereas this study uses PISA 2006 and 2009. PISA 2000 scores were widely believed to flatter UK countries because tests were conducted during the winter months when there was a disproportionately high rate of school absenteeism amongst lower attaining students which was likely to have biased the sample of those taking the test (Prais, 2003). The higher base line scores for the UK in the earlier study will no doubt have contributed to the exceptionally low skills gain after 16 found in that study and this will have negatively affected the average for the Type 4 category in which UK countries were prominent. The results for the Type 4 countries in this report may therefore be considered more reliable.

The findings on the effects of the other system types are less consistent across the different types of calculation and rarely reach levels of statistical significance. However, the effect sizes are quite sometimes quite substantial. The strongest results for each domain are as follows:

For literacy skills acquisition:

- system types 1 and 1b perform well relative to other groups (each in 1st or 2nd place in the 'ranking' on each of the calculation methods);
- system types 2b and 3 perform in the middle of the country type ranking on each method.

For numeracy skills acquisition:

- Type 2b performs consistently well relative to other groups (1st in the rankings on both methods);
- Type 3 also performs well (2nd in ranking on both methods);
- Type 1b performs less well than other groups except Type 2a and Type 4 (with similar results on both calculation methods).

These findings are consistent with those in Pensiero and Green (2018) in respect of numeracy. Type 2b and Type 3 systems performs relatively well in both studies. However, they diverge in respect of Type 2b performance in literacy, where it performs better in the earlier study than the current study.

In our study we seek to explore the mechanisms behind the differential effects of system types on skills acquisition by examining the effects of specific system characteristics across countries. Our analysis identifies seven system characteristics which impact on skills acquisition in the same direction for both numeracy and literacy (on one or both calculation methods). There are three which do not reach significance levels in any of the calculations (number of school types, upper secondary completion rate and cohort years). Our interpretation therefore focuses on the remaining five variables which are:

- theoretical age of completion of upper secondary education (significant only for numeracy in OLS regression);
- stage of first selection (significant only for numeracy in OLS);
- teacher workload (significant for literacy and numeracy in OLS);
- standardised curriculum (significant for numeracy in OLS);
- vocational prevalence (significant for literacy and numeracy in OLS).

These factors, in combination, go some way to explaining our main finding - that Type 2a and Type 4 system performs less well than other systems type in skills acquisition after age 15.

According to the theoretical framework which informs our typology and hypotheses, systems will perform better when the participation in vocational learning is relatively high and where greater parity between the vocational and general tracks is established through the provision of consistently high-quality apprenticeships and/or through system integration measures which promote standardisation in key areas of the curriculum.

The higher-performing system types in our study (Types 1, 1b, 2b and 3) each have most, if not all, of these characteristics.

Type 1 and 1b systems have high participation in vocational learning as well as a substantial degree of curriculum standardisation. Learning of Maths and the national language is normally mandated in all tracks which are also typically long in duration with later graduation ages. Type 1b systems (including Denmark, Finland, the Netherlands and Singapore) have the highest participation in vocational learning and might be expected to perform best overall on account of this, but they do so only in literacy. The fact that mandatory core skills learning is less prevalent in these systems than in Type 1 systems may account for this since the absence of mandatory core skills learning will have more consequence in numeracy than in literacy, because students will continue to study and be assessed in the national language even where it is not mandated as a subject. Type 1b systems also face a high bar to achieve high relative skills gain in the upper secondary phase since they already perform best of the system types at age 15 in PISA.

Type 2 (Nordic comprehensive) systems (Norway and Sweden) have high participation rates in vocational learning, with relatively well-funded vocational tracks, and combine this with higher levels of curriculum standardisation than other types of system. Graduation at 19 years is typical across all programmes and Maths and the national language learning are mandated across programmes. On these criteria Type 2b systems might be predicted to perform better than all other system types, but in fact, whilst skills gain is higher in numeracy than in other system types, this is not the case in literacy, where the performance is only average in our analysis here. This finding diverges somewhat from previous research by Pensiero and Green (2018) which suggested that relative skills gain for Type 2b systems is highest in both domains. Their more mixed performance across the two skills domains in this study be because it capture better the fact that while Type 2 systems have high rates of vocational learning, only a portion of this is in apprenticeship programmes which are particularly effective in raising numeracy skills. Apprenticeships are more widespread in Norway than Sweden which may account for the fact that Norway achieves higher relative skills gain in the upper secondary phase than Sweden (see Figure 2).

Type 3 systems (Austria and Germany) also have high participation rates in well-funded worked-based learning and a degree of curriculum standardisation, with normatively long cycle programmes in which the learning of Maths and the national language is usually mandated. However, compared with Type 2b systems, they have more variability in the length of programmes (between 2 and 3.5 years in the Dual system) which do not all require the learning of core skills in the Berufsschule (see Table 12). Like the Type 2b systems they perform well in numeracy but at the average for different system types in literacy. In this case, Type 3 systems may be benefitting mostly from the quality and prevalence of their apprentice training.

The lower-performing Type 2a (North American) and Type 4 (mixed) systems are distinctive both for the low rates of participation in vocational learning and for their weak or partial curriculum standardisation measures. The Type 2a systems have the advantage in the latter in that their high school programmes include compulsory Maths and English, and are typically of the same duration, which is not the case in Type 4 systems. However, graduation is typically at age 18, as in Type 4 systems, rather than at age 19, which our analysis suggests is associated with higher performance. In Type 4 systems participation in worked-based learning is also relatively low, though substantially higher than in Type 2a systems, and (except in Israel) is relatively under-funded. The relatively poor performance of Type 4 systems may derive from the lack of curriculum standardisation and limited vocational learning whereas in Type 2a systems it can be mostly attributed to the lack of vocational learning.

The literature on mixed systems (and particularly in the UK) also points to other aspects of system fragmentation - including the great diversity of institutional types and qualifications - which are said to reduce system transparency and undermine normative expectations of students. Our analysis of one indicator for this (number of school types) does show the expected negative effects in both estimation methods but they are not statistically significance. A full examination of this would require a wider range of indicators for system fragmentation but unfortunately the available international data does not allow this.

8.2 Effects of system types and characteristics on changes in skills inequality mean between Ages 15 and 18-20.

The clearest finding from our analysis of system type effects on skills inequalities is, as with the discussion of the effects on skills means above, that system types 2a and 4 stand out from the other system types. They not only perform worse than other system types in raising skills levels; they also do less to reduce skills inequalities during the upper secondary phase than other system types.

Type 2a (north American comprehensive) systems preform worse than all other system types in this regard. Our analysis shows that, relative to the reference type, Type 2a systems have a significant positive effect on skills inequality in both literacy and numeracy on both methods of calculation. Type 4 (mixed systems) perform similarly (with significant positive effects in all calculations), but with slightly lower effect sizes.

The results for the other groups are less clear cut.

Type 1 systems (the reference group) would appear to be most effective in reducing skills inequality since other systems do worse by comparison on most calculations. Type 1b systems appear to do slightly less to mitigate skills inequality, since they increase skills inequality, relative to the reference type, in both domains on both calculations. But the effect size is small, and the relationship is significant in only one of the calculations.

These findings are consistent with those in Green and Pensiero (2016), in as much as they confirm that mixed systems tend to increase skills inequality. However, they differ in relation to the effects of system Types 2a and 3, with Type 2a systems not performing significantly differently in the earlier analysis from Type 1 systems, and Type 3 systems performing better in reducing inequality than other types of system. In the current analysis Type 3 systems cannot be said to differ significantly from Type 1 systems, showing only small non-significant relative reductions in skills inequality in both domains in the DID calculations but the opposite on the OLS calculations. Given that we find no significant difference now between Type 3 systems and the best-performing Type 1 systems, we can still infer that Type 3 system perform relatively well on reducing skills inequality, but the differences are not as clear cut as in the earlier analysis.

Our results with regard to Type 2b comprehensive systems, as with the Type 3 systems, is that they do not perform significantly differently from the best performing Type 1 reference system. For literacy there is a small, non-significant effect which is positive in the OLS regression but negative in the DID calculation. For numeracy the effects are again small and non-significant but positive on both calculations.

We can dig deeper into the analysis of the differential performance of system types by looking again at the effects across systems of particular system characteristics. Those characteristics with the most robust effects include:

- *Theoretical duration of upper secondary programmes* (which significantly reduces inequality in numeracy (on both calculations) and in literacy (on the OLS calculation)
- *Education stage of first selection*: later selection (on entry to or during upper secondary education) significantly increases inequality in numeracy (significantly so in the DID calculation) but not significantly in literacy. We can infer from this that the selection on entry to lower secondary education (the reference) does not necessarily increase inequality in the upper secondary phase (confirming the argument of Raffe et al. (2001);
- *Teacher workload*, which significantly increases skills inequality in both domains on both calculation methods;
- *Curriculum standardisation* which reduces skills inequality in both domains in both calculation methods (significantly so, except in the DID calculation for literacy);
- Vocational prevalence which reduces skills inequality in both domains on both estimation methods (significantly so except in OLS regression for literacy);
- *Cohort years of upper secondary education* which reduces inequality in both domains on both estimation methods (but more significantly so on the OLS regressions estimates).

The factors found to reduce skills inequality are broadly in line the predictions from the literature and typology.

As argued by Raffe et al (2001), early selection to different tracks in lower secondary education is not necessarily a barrier to skills inequality reduction in the upper secondary phase. What matters more is that upper secondary systems achieve relative 'parity of esteem' between the academic and vocational tracks. This can be achieved in comprehensive systems through greater curriculum standardisation across programmes in key areas relating to core skills learning and the duration of programmes. It can also be achieved through higher participation rates in vocational tracks, particularly in high quality apprenticeships.

The inequality reducing effects of certain standardisation measures in the first part of the hypothesis are confirmed in our analysis of the effects of three indicators which proxy system standardisation. *Theoretical duration* provides a measure of the normal duration of the main upper secondary tracks. *Cohort years* takes account of those not entering upper secondary education, or dropping out early, and thus provides a fuller measure of the degree of universalisation of completion of long cycle upper secondary programmes (defined by OECD as two or more years). *Curriculum standardisation* refers to the normalisation of learning Maths and the national language across all programmes areas. All are associated with inequality reduction across upper secondary systems. *Vocational prevalence* refers the ubiquity of vocational learning amongst upper duration secondary participants. This is also shown to reduce inequality, particularly where it takes the form of high-quality apprenticeship provision.

How these characteristics are combined in upper secondary education systems goes some way to explaining why certain types of system are more effective in closing skills gaps than others. The systems which do this best in our analysis belong to Types 1, 2b and 3.

Type 1 and 2b systems reduce inequality in much the same way. They have relatively integrated upper secondary systems with a substantial degree of standardisation around the length of programmes and the common aspects of their curricula. Courses generally last two or three years and a high proportion of young people complete these long cycle courses and acquire full ISCED Level 3 qualifications. They also have quite high rates of participation in vocational programmes with work-based learning (in the form of placements, work-experience etc.) which will also contribute to reducing skills inequality. Type 1b systems also have a significant proportion of students undertaking school-based or 'hybrid' apprenticeships which we predicted would make a further contribution to reducing skills inequality. According to our analysis it does not do so, but this may be because in two of the systems in the group (Finland and Singapore) the numbers of students participating is rather small, at less than 10 percent). Denmark has up to 30 percent of students taking some form of hybrid apprenticeship and does relatively better in skills inequality mitigation than Finland or Singapore (see: figure 4).

Type 3 systems also do relatively well in reducing skills inequalities because they have high participation in high-quality and well-funded work-based apprenticeships. The prestige of the vocational route is enhanced by the fact that it attracts a relatively balanced mix of students from different backgrounds and with different levels of achievement and offers the prospect of well-paid skilled jobs. Although tracked, Type 3 systems form part of a relatively integrated upper secondary system in as much as they share common core curriculum with other tracks and courses in each have a similar duration. Both set of factors generate relative parity of esteem and serve to reduce skills gaps to some degree.

The systems which perform worst in mitigating skills inequalities (in Type 2a and Type 4), by contrast, tend to lack the characteristics which our analysis shows tend to reduce skills inequality.

Systems in Type 2a have some characteristics of system integration, for instance, in ensuring the programmes are of a similar length and all contain learning in core skills. However, their programmes are generally shorter than those in Type 1 and 2b systems. With the exception of the system in francophone Canada, these systems also have very low rates of participation in vocational learning which may be reducing their capacity to narrow skills gaps. Type 4 mixed systems, as we have seen, have relatively low rates of participation in vocational learning and score on average very low on the measures of system integration which our analysis suggests reducing skills inequality. They are also more institutionally fragmented than other systems which may undermine normative expectations of achievement in upper secondary education and this increase inequality of skills outcomes, although our analysis lacks the indicators to test this adequately.

8.3 Other Factors Affecting Skills Acquisition Requiring Further Research

We have been unable to analyse statistically a number of other factors which are discussed in the literature review and which research suggests may affect changes in levels and inequalities of skills.

Certain pedagogic methods, such as the 'mastery learning' methods discussed in the literature review, may be effective for the acquisition of core skills like literacy and numeracy, particularly for lower achieving students, and particularly when they are embedded in certain cultural contexts. If so, they would play a part in raising levels of skills overall and reducing skills gaps since they may benefit the least skilled the most. There is no cross-country statistical analysis to support this at present because we do not have quantitative measures of the prevalence of the use of this pedagogy for a wide range of countries. However, there is a strong logical argument that these methods are effective in the east Asian education systems where they are widely practised. All of the more developed east Asian countries and territories included in PISA and PIAAC (Japan, Hong Kong, Korea, Singapore) score exceptionally highly in the tests of literacy and numeracy. They share many system characteristics with our Type 1 systems which do not, on average, score so highly (excepting those in the Type 1b subset of which Singapore is a member), suggesting that a regional factor may be in play which enhances performance. This may have to do with their pedagogic methods (including mastery learning), or certain cultural dispositions common throughout the region, such as the respect for teachers and high educational aspirations frequently invoked in the literature. It may also be due to a combination of both, suggesting that pedagogic approaches, like mastery learning, work particularly well in cultural contexts characterised by high educational aspirations, respect for teachers, and certain forms of social and collective discipline (including as practised in the classroom). While we have a body of qualitative research which supports this thesis, more quantitative comparative research needs to be done to substantiate these claims.

Another regional characteristic which has so far eluded quantitative comparative analysis is the high rate of participation in supplementary private tutorial schools (or so-called 'shadow education') which is particularly characteristic of east Asian societies (Bray, 1999). Our analysis suggests that literacy and numeracy levels are enhanced by their inclusion throughout longer upper secondary programmes, implying that time spent learning matters. This proposition is supported by the comparative research from the IEA which finds that 'hours spent learning' explains cross-country variation in performance in educational tests for lower secondary school students more consistently than any other factor analysed through decades of research (IEA, 1988). The hours east Asian students spend learning in private tutorial classes add to the already relatively high annual hours spent learning in school and through homework (by comparison with western countries) and this may also contribute to raising average performance. We currently lack comprehensive international data on learning time amongst upper secondary students. Further research will be needed to establish how far cross-country variation in learning time amongst upper secondary students contributes to the explanation of differences in performance across countries.

8.4 Teachers' Workload and Salaries

Our analysis is also unable to provide a satisfactory explanation of the effects of teacher salaries on student performance in upper secondary education. However, we do produce a significant new finding in relation to the negative effects of teacher workload which was not anticipated in our review of the literature.

Whilst our analysis reveals a significant effect of teacher workload on students' skills outcomes and inequalities, this does not appear to be a prevalent theme in the cross-national literature. A notable exception relates to Liang and Akiba's (2017) recent study of teacher employment contracts, workload, and national achievement levels in maths, science and reading, which nonetheless fails to garner statistically significant effects of workload on achievement. Other studies have focused instead on teachers' salaries, (Nir & Naphcha, 2007; Akiba et al., 2012), or other aspects of teachers' working conditions, such as instructional time (Baker et al., 2004), teacher 'quality' (Akiba et al., 2007), school resources and class size (Hanushek & Woessmann, 2017), in so far as that they impact upon student achievement.

The small number of within-country studies specifically focusing on the effect of teacher workload (Attah & Adebayo, 2018; Ost & Schiman, 2017; Rose & Sika, 2019; Toropova et al., 2020) have, however, found significant negative effects of high teacher workload on student achievement, in both developed- (e.g. Sweden and the US) and developing-country (e.g. Kenya and Nigeria) settings.

The findings of our analysis contribute to this evidence with a cross-national emphasis: we find that teacher workload has a significant negative effect on both literacy and numeracy outcomes for students across the upper-secondary phase of education (OLS estimates show this to be the case, whilst the DID analysis also indicates the same relationship, whilst the findings do not attain statistical significance due to larger standard errors).

Our findings also provide evidence for a significant independent effect of teacher workload on skills distributions both in literacy and numeracy. These effects are of a similar magnitude in both domains and remains statistically significant when a DID approach counteracts the effect of time-invariant factors (e.g., external cultural and policy effects).

The implication of these findings is somewhat intuitive – students of teachers experiencing greater workloads perform worse in absolute terms; moreover, widening inequalities of skills distributions suggests that students who require greater teacher assistance (towards the bottom of the distribution) are more negatively impacted by teacher workload levels, compared to students higher up the distribution, who are less reliant on teacher support.

At the same time, there appears to be no significant effect of teacher salaries on either skills outcomes or inequalities. This would suggest that teachers' working conditions are somewhat more important that levels of pay in terms of its effect on students' performance (this is somewhat congruent with Hanushek and Rivkin's (2007) findings which emphasise a lack of correlation between teachers' pay and quality).

Two clear caveats must be highlighted, however: first, our measure of teacher salaries refers to average starting salary within a country or territory and, as such, does not take into account the variability of salaries according to experience or quality of teaching. Second, we use data from the OECD, which presents salaries in terms of US dollars, converted using purchasing-power parity. Whilst this enables some degree of cross-national comparability, it does not allow for a more nuanced analysis of the effect of teacher salaries relative to a country's overall wage distribution (cf. Dolton & Marcenaro-Gutierrez, 2011).

Given the relative scarcity of empirical evidence relating to the independent effect of teacher workload on students' performance at a cross-national level, this is an important area of future research to ascertain the extent to which improving teachers' working conditions can positively influence students' skills acquisition.

9. Some Policy Implications

This report has focussed on how comparative analysis of system characteristics and system types may be used to explain cross-country variation in changes in skills acquisition during upper secondary education and training. Unlike most previous research in this under-researched area, we use data on skills from surveys which measure competence levels at the beginning and end of this phase (15-year-olds in PISA and 18-20-year olds in SAS) so as to capture more accurately the changes during this phase.

One our most consistent findings is that countries with Type 2a comprehensive systems and those with Type 4 mixed systems tend, on average, to be less successful than countries with other types of system in raising average skills levels and mitigating skills inequalities. Our analysis suggests that this has to do with their relatively low rates of participation in vocational learning and lack of system integration (in the case of Type 4), among other possible factors. Comparisons across the better performing groups would suggest that most of the English-speaking countries/regions in Type 2a and Type 4 (excluding Australia perhaps) would benefit from increasing participation in well-funded and high quality worked-based learning. The anglophone Type 4 countries might further benefit from enhancing the learning of core skills across all tracks and, in the English case, from standardising the duration of programmes across tracks.

Comparative analysis across system types can be seen as a somewhat blunt instrument since it obscures the variation within types which, in some cases, is considerable. We therefore give equal weight in our analysis to the variation of system characteristics across countries. However, from a policy perspective intra-type comparative analysis can be quite useful. It allows us to compare the performance of systems which are generally quite similar but vary on one factor which may then be used to explain the different outcomes (see Landman & Carvalho 2017). If the object in policy terms is to borrow policies which have been shown to work in other countries, then it also makes sense to look first at the more successful countries which share similar types of system, since borrowing and adapting policies is generally most effective between countries sharing a range of similar contexts. With this in mind, we conclude by making a few observations about where some specific countries might look to improve their performance.

Type 1 systems perform relatively well on average in raising mean scores and reducing skills inequalities. But it is the east Asian Countries (along with Finland) which perform best in raising skills levels. Given that the systems in these countries are quite similar in most respects to those in other Type 1 countries and are mostly obviously differentiated by their use of more traditional 'mastery' type learning pedagogies, other countries in the group might benefit from seeing what can be learnt from their pedagogic approaches. This would be contrary to the direction recently taken by Finland, the European 'champion' in PISA, which has been seeking to enhance individualisation of the curriculum; but this has apparently so far not enhanced average performance in international tests.

Estonia is a Type 1 country that performs relatively well in raising skills levels, but less well in reducing skills inequality, especially compared with the Czech Republic and Slovakia. One explanation for this may be Estonia's low rate of participation in vocational learning by comparison with other countries in Type 1, and especially the Czech Republic and Slovakia. Singapore in one of the best performing countries in average skills levels in both PISA and SAS and will find it challenging to improve its relative position for skills gains during the upper secondary phase because of its high baseline position. However, it may do more to close skills gaps during this phase by increasing participation in its hybrid-type apprenticeships, along the lines of Denmark, where high rates of such participation appear to be successful in reducing skills inequality in numeracy (see figure 4).

Amongst the Type 4 systems, it is Australia that perform highest in mean skills levels of students aged 18-20 and in terms of skills gain during the upper secondary phase. England, and some other UK countries do considerably less well. However, Australia, like most other countries in the group (except Northern Ireland), still do relatively poorly in reducing skills inequality during the upper secondary phase and retain relatively large skills gaps amongst students aged 18-20. Given that the system in Australia shares many features in common with the anglophone UK systems, it is not clear which characteristics of its system may explain the relatively high levels of measured skills. The one possible explanatory feature which stands out from our analysis is that there is a high rate of participation in vocational learning in Australia (57% of students) compared with the other countries, albeit equally underfunded by comparison with general education. However, somewhat paradoxically, this does not seem to be leading to significant reduction in skills gaps. Although the evidence here is somewhat mixed, it would certainly warrant further investigation from UK policy-makers seeking policies to improve performance in the UK upper secondary systems.

References

Akiba, M., Chiu, Y.-L., Shimizu, K., & Liang, G. (2012) ‘Teacher salary and national achievement: A cross-national analysis of 30 countries’, International *Journal of Educational Research,* 53, 171–181.

Akiba, M., Le Tendre, G. K., & Scribner, J. P. (2007) ‘Teacher quality, opportunity gap, and national achievement in 46 countries’, Educational *Researcher*, 36 (7), 369–387.

Ammermüller, A., Heijke, H. & Woessmann, L. (2005) ‘Schooling quality in Eastern Europe: Educational production during transition’, *Economics of Education Review*, 24 (5), 579-599.

Attah, B. G., & Adebayo, O. J. (2018) 'Effects of training and workload of teachers on 'achievement of students’ in mathematics in senior schools in Jos Metropolis, Plateau state, Nigeria', *KIU Journal of Humanities*, 3 (1), 209-218.

Baker, D. P., Fabrega, R., Galindo, C., & Mishook, J. (2004) ‘Instructional time and national achievement: Cross-national evidence’, *Prospects*, 34 (3), 311-334.

Barber, M & Mourshed, M. (2007) *How the World’s Best-Performing School Systems Come out on Top,* McKinsey & Company.

Bol, T. & Van de Werfhorst, H. G. (2013) ‘Educational systems and the trade-off between labor market allocation and equality of educational opportunity’, *Comparative Education Review*, 57 (2), 285-308.

Boudon, R. (1974) *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society,* Wiley, New York.

Borgonovi, F., Pokropek, A., Keslair, F., Gauly, B. & Paccagnella, M. (2017) *Youth in Transition: How Do Some of The Cohorts Participating in PISA Fare in PIAAC?* OECD Education Working Papers, 155, OECD, Paris.

Bray, M. (1999) *The Shadow Education System: Private Tutoring and its Implications for Planners.* Fundamentals of Educational Planning 61. UNESCO International Institute for Educational Planning (IIEP), Paris.

Brunello, G., & Checchi, D. (2007) ‘Does school tracking affect equality of opportunity? New international evidence', *Economic Policy,* 22 (52), 781-861.

Busemeyer, M. (2014) *Skills and Inequality: Partisan Politics and the Political Economy of Education Reforms in Western Welfare States*, Cambridge University Press, Cambridge.

Busemeyer, M. & Iversen, T. (2011) ‘Collective skills systems, wage bargaining and labour market stratification’ In: M. Busemeyer & C. Trampusch (eds) *The Political Economy of Collective Skill Formation*, Oxford University Press, Oxford.

Busemeyer, M. & Trampusch, C. (eds.) (2012) *The Political Economy of Collective Skill Formation,* Oxford University Press, Oxford.

CEDEFOP (2008) *Initial Vocational Education and Training (IVET) in Europe: Review*, European Centre for the Development of Vocational Training, Thessaloniki.

Chong, T. (2014) 'Vocational education in Singapore: meritocracy and hidden narratives', *Discourse: Studies in the Cultural Politics of Education,* 35 (95), 637-648.

Dolton, P. & Marcenaro-Gutierrez, O. D. (2011) 'If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance', *Economic Policy,* 26 (65), 5-55.

Desjardins, R. & Ederer, P. (2015) 'Socio-demographic and practice-oriented factors related to proficiency in problem solving: A lifelong learning perspective', *International Journal of Lifelong Education,* 34 (4), 468-486.

Dumas, A., Méhaut, P., & Olympio, N. (2013) 'From Upper Secondary to Further Education: European Models of Post-Compulsory Learning' In: Janmaat J.G., Duru-Bellat M., Green A., Méhaut P. (eds) *The Dynamics and Social Outcomes of Education Systems. Education, Economy and Society,* Palgrave Macmillan, London.

Esping-Andersen, G. (1990) *The Three Worlds of Welfare Capitalism*, Princeton University Press, New Jersey.

Evans, K. & Kersh (2006) '*Adults Learning In, For and Through the Workplace: the Significance of Biography and Experience',* Paper presented at the British Educational Research Association Annual Conference, University of Warwick, 6-9 September 2006.

Green, A. (2003) 'The many faces of lifelong learning: Recent education policy trends in Europe', *Journal of Education Policy,* 17 (6), 611–626.

Green, A. (2013) *Education and State Formation: Europe, East Asia and the USA*. Revised and Extended Second Edition, Palgrave, Basingstoke.

Green, A., Green, F. and Pensiero, N. (2014) *Why are Literacy and Numeracy Skills in England so Unequal? Evidence from the OECD's Survey of Adult Skills and other International Surveys,* Research Paper 27, Centre for Learning and Life Chances in Knowledge Economies, Institute of Education, London.

Green, A., Green, F. and Pensiero, N. (2015) 'Cross-country variation in adult skills inequality: why are skill levels and opportunities so unequal in anglophone countries?', *Comparative Education Review*, 59 (4), 595-618.

Green, A. & Janmaat, J.-G. (2011) *Regimes of Social Cohesion: Societies and the Crisis of Globalisation*, Palgrave, Basingstoke.

Green, A., Mostafa, T. & Preston, J. (2010) *The Chimera of Competitiveness: Varieties of Capitalism and the Economic Crisis*, Research Paper 8, LLAKES Centre, Institute of Education, London.

Green, A. & Pensiero, N. (2016) ‘The effects of upper secondary education and training systems on skills inequality. A quasi-cohort analysis using PISA 2000 and the OECD Survey of Adult Skills’, *British Education Research Journal,* 45 (5), 756-779.

Greenwood, C., Jenkins, A., & Vignoles, A. (2007) *The Returns to Qualifications in England: Updating the Evidence Base on Level 2 and Level 3 Vocational Qualifications,* Centre for the Economics of Education, Discussion Paper 89.

Greinert, W. (2004) ‘European vocational training ‘systems' - some thoughts on the theoretical context of their historical development’, *European Journal of Vocational Training*, 32, 18-25.

Hall, P. & Soskice, D. (eds) (2001) *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*, Oxford University Press, Oxford.

Hanushek, E. A., Piopiunik, M. & Wiederhold, S. (2019) ‘The value of smarter teachers: international evidence on teacher cognitive skills and student performance’, *Journal of Human Resources,* 54 (4), 857-899.

Hanushek, E. A., & Rivkin, S. G. (2007) 'Pay, working conditions, and teacher quality', *Future Child*, 17 (1), 69-86.

Hanushek, E. A. & Woessmann, L. (2006) ‘Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries’, *Economic Journal,* Royal Economic Society, 116 (510), 63-76.

Hanushek, E. and Woessmann, L. (2009) *Schooling, Cognitive Skills and the Latin American Growth Puzzle,* Working Paper 15066, National Bureau of Economic Research, Cambridge, M.A.

Hanushek, E. and Woesmann, L. (2010) *The Economics of International Differences in Educational Achievement*, Working Paper 15949, National Bureau of Economic Research Cambridge, M.A.

Hanushek, E. & Woessmann, L. (2017) ‘School resources and student achievement: A review of cross-country economic research’, In: *Cognitive Abilities and Educational Outcomes,* 149-171, Springer.

Heath, A. & Sullivan, A. (2011) ‘Introduction: The democratisation of upper secondary education?’, *Oxford Review of Education*, 37 (2), 123-138.

Hodgen, J., Pepper, D., Sturman, L., & Ruddock, G. (2010) *Is the UK an Outlier,* Nuffield Foundation, London.

Hogan, D. (2011) *Why is Singapore' School System so Successful and is it a Model for the West*? *The Conversation*. Available at:
https://theconversation.com/why-is-singapores-school-system-so-successful-and-is-it-a-model-for-the-west-22917

Horn, D. (2009) ‘Age of selection counts: A cross-country analysis of educational institutions’, *Educational Research and Evaluation, 15*(4), 343–366.

Jackson, M. (ed.) (2013) *Determined to Succeed? Performance versus Choice in Educational Attainment*, Stanford University Press, Stanford.

Janmaat J.G., Duru-Bellat M., Green A., Méhaut P. (eds) *The Dynamics and Social Outcomes of Education Systems. Education, Economy and Society*, Palgrave Macmillan, London.

Jerrim, J. & Vignoles, A. (2015) *The Causal Effect of East Asian 'Mastery' Teaching Methods on English Children's Mathematics Skills*, Working Paper No. 15-05, Department of Quantitative Social Science, UCL Institute of Education.

Keep, E. & Mayhew, K. (2014) 'Inequality – 'wicked problems', labour market outcomes and the search for silver bullets', *Oxford Review of Education*, 40 (6), 764–781.

Koike, K. & Inoki, T. (eds) (1990) *Skill Formation in Japan and Southeast Asia*, University of Tokyo Press, Tokyo.

Landman, T. & Carvalho, E. (2017) *Issues and Methods in Comparative Politics: An Introduction*, Routledge, London.

Lasonen, J. & Young, M. (1998) *Strategies for Achieving Parity of Esteem in European Upper Secondary Education. Final Report on the Project*, Institute for Educational Research, University of Jyvaskyla, Finland.

Lave, J. & Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*, Cambridge University Press, Cambridge.

Law, S. S. (2015) *A Breakthrough in Vocational and Technical Education: The Singapore Story*, World Scientific Publishing Company, Singapore.

Liang, G., & Akiba, M. (2017) 'Teachers' working conditions: a cross-national analysis using the OECD TALIS and PISA data', In: Le Tendre, K & Akiba, A. (eds) *International Handbook of Teacher Quality and Policy*, 388-402, Routledge, London.

Lucas, S. R. (2001) 'Effectively maintained inequality: education transitions, track mobility, and social background effects', *American Journal of Sociology*, 106 (6), 1642-1690.

Lynn, R. (1988) *Educational Achievement in Japan. Lessons for the West*, Macmillan, London.

Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008) 'Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes', *Child Development Perspectives, 2* (2), 99–106.

Maurice, M., Sellier, F. & Silvestre, J-J. (1986) *The Social Foundations of Industrial Power: A Comparison of France and Germany*, MIT Press, Cambridge, M. A.

McLean, M. (1990) *Britain and a Single Market Europe. Prospects for a Common School Curriculum*, Kogan Page, London.

Méhaut P. (2013) ‘Formal or actual convergence? Three cases of hybridisation’, In: Janmaat J.G., Duru-Bellat M., Green A. & Méhaut P. (eds) *The Dynamics and Social Outcomes of Education Systems. Education, Economy and Society.* Palgrave Macmillan, London.

Micklewright, J. & Schnepf, S.V. (2007) *Who Gives for Overseas Development,* IZA Discussion Papers 3057, *Institute of Labor Economics* (IZA).

National Centre on Education and the Economy (NCEE) (undated), *South Korea Teacher and Principal Quality.* Available at:
https://ncee.org/what-we-do/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/south-korea-teacher-and principal-quality/

Nir, A. E. & Naphcha, M. (2007) 'Teachers’ salaries in public education: Between myth and Fact’, *International Journal of Educational Management*, 21(4), 315–328.

Norris, N., Asplund, R., MacDonald, B., Schostak, J. & Zamorski, B. (1996) *An Independent Evaluation of Comprehensive Curriculum Reform in Finland*, National Board of Education, Helsinki.

Oates, T. (2010), *Could do Better: Using International Comparisons to Refine the National Curriculum in England.* Available at:
https://www.cambridgeassessment.org.uk/Images/112281-could-do-better-using-international-comparisons-to-refine-the-national-curriculum-in-england.pdf

OECD (1985) *Education and Training Beyond Basic Schooling*, OECD, Paris.

OECD (2001) *Education at a Glance* (2001), OECD, Paris.

OECD (2004) *Completing the Foundation for Lifelong Learning – An OECD Survey of Upper Secondary Schools*, OECD, Paris.

OECD (2008) *Education at a Glance*, OECD, Paris.

OECD (2010) *PISA 2009 Results: Overcoming Social Background. Equity in Learning Opportunities and Outcomes (Vol. 11*), OECD, Paris.

OECD (2011) *Education at a Glance 2011: OECD Indicators*, OECD, Paris.

OECD (2012) *Education at a Glance 2012*, OECD, Paris.

OECD (2013) *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices*, vol. IV, OECD, Paris.

OECD (2013b) *The Survey of Adult Skills: Reader’s Companion*, OECD, Paris.

OECD & Statistics Canada (2000) *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*, OECD, Paris.

Ost, B., & Schiman, J. C. (2017) ‘Workload and teacher absence’, *Economics of Education Review*, *57*, 20-30.

Pensiero, N. and Green, A. (2018) 'The effects of post-compulsory education and training systems on literacy and numeracy skills: a comparative analysis using PISA 2000 and the 2011 Survey of Adult Skills, *European Journal of Education*, 1–16.

Pontusson, J. (2008) *Social Democracy, Varieties of Capitalism and Globalization: Sketch of Paper*. Available at:
http://www.people.fas.harvard.edu/~ces/conferences/left/left_papers/SocDemo05-08.pdf

Prais, S.J. (2003) 'Cautions on OECD's recent educational survey (PISA)', *Oxford Review of Education*, 29 (2).

Raffe, D., Brannen, K., Fairgrieve, J. & Martin, C. (2001) 'Participation, inclusiveness, academic drift and parity of esteem: A comparison of post-compulsory education and training in England, Wales, Scotland and Northern Ireland', *Oxford Review of Education*, 27 (2), 173-203.

Raffe, D., Howieson, C., Spours, K. & Young, M. (1998) 'The unification of post-compulsory education: towards a conceptual framework', *British Journal of Educational Studies*, 46 (2), 169-187.

Raftery, A. E., & Hout, M. (1993) 'Maximally maintained inequality: expansion, reform, and opportunity in Irish education 1921-75', *Sociology of Education*, 41-62.

Rivas, A. (2015) *Latin America after PISA: Lessons Learned about Education in Seven Countries. Executive Summary*, CIPPEC-Natura-Instituto Natura, Buenos Aires.

Rose, A. B. & Sika, J. O. (2019) 'Determining influence of teacher's workload on academic performance in secondary schools, Suba sub-county Kenya', *Advances in Social Sciences Research Journal*, 6 (3), 287-295.

Rubenson, K. (1987) 'Participation in recurrent education: A research review' In: Schuetze, H. G., & Istance, D. (eds), *Recurrent Education Revisited: Patterns of Participation and Financing*, Almqvist and Wiksell, Stockholm.

Sahlberg, P. (2015) *Finnish Lessons 2.0,* Teachers College Press, Columbia University: NY.

Salverda, W., Nolan, B., Checchi, D., Marx, I., McKnight, A., Tóth, I. & van de Werfhorst, H. (2014) *Changing Inequalities in Rich Countries: Analytical and Comparative Perspectives*, Oxford University Press.

Schimdt, W. H. and and Prawat, R. S. (2006) 'Curriculum coherence and national control of education: Issue or non-issue?', *Journal of Curriculum Studies,* 38 (6), 641-658.

Schleicher, A. (2011) 'Lessons from the World on Effective Teaching and Learning Environments,' *Journal of Teacher Education*, 62 (2), 202–221.

Schneider, S. L. and Tieben, N. (2011) 'A healthy sorting machine? Social inequality in the transition to upper secondary education in Germany', *Oxford Review of Education*, 37 (2), 139-166.

Schütz, G., Ursprung, H.W. & Woessmann, L. (2008) ‘Education policy and equality of opportunity’, *Kyklos*, 61, 279-308.

Sinnema, C. (2016) ‘The ebb and flow of curricular autonomy: balance between local freedom and national prescription in curricula’, In: *The SAGE Handbook of Curriculum, Pedagogy and Assessment,* SAGE Publications, London.

Solga, H., Protsch, C., Ebner, C. & Brzinsky-Fay, C. (2014) *The German vocational education and training system: its institutional configuration, strengths and challenges*. Discussion Paper, University of Cologne and Federal Institute for Vocational Education and Training, Bonn.

Steedman, H. (2001) *Benchmarking Apprenticeship: UK and Continental Europe Compared*. Report for London, Centre for Economic Performance, London School of Economics, London.

Steedman, H. (2010) *The State of the Apprenticeship in 2010*, Centre for Economic Performance, London School of Economics, London.

Toropova, A., Myrberg, E., & Johansson, S. (2020) ‘Teacher job satisfaction: the importance of school working conditions and teacher characteristics’, *Educational Review*, 1-27.

UNESCO (2020) *Global Education Monitoring Report 2020*, UNESCO, Paris.

Van de Werfhorst, H. G. & Mifs, J. (2010) ‘Achievement inequality and the institutional structure of educational systems: A comparative perspective’, *Annual Review of Sociology,* 36 (1), 407-428.

Van Houtte, M. & Demanet, J. (2012) ‘Self-esteem of academic and vocational students: Does within-School tracking sharpen the difference?’, *Acta Sociologica.* 55(1), 73–89.

Van Houtte, M. and Demanet, J. (2012) ‘Teachers' attitudes and students' opposition. School misconduct as a reaction to teachers' diminished effort and affect’, *Teaching and Teacher Education,* 28 (6), 860–869.

Verdier E. (2013) ‘Lifelong learning regimes versus vocational education and training systems in Europe: The growing hybridisation of national models’, In: Janmaat J.G., Duru-Bellat M., Green A., & Méhaut P. (eds) *The Dynamics and Social Outcomes of Education Systems. Education, Economy and Society,* Palgrave Macmillan, London.

Winkler, D. R. (1993) ‘Fiscal decentralization and accountability in education: Experiences in four countries,’ In: Hannaway, J. & Carnoy, M., *Decentralization and School Improvement: Can We Fulfil the Promise,* Wiley, New York, 102-134.

Woessmann, L. (2003) ‘Schooling resources, educational institutions and student performance: the international evidence’, *Oxford Bulletin of Economics and Statistics*, 65, 117-170.

Woessmann, L. (2005) *Public-Private Partnerships in Schooling: Cross-Country Evidence on their Effectiveness in Providing Cognitive Skills,* PEPG Research Paper 05-09, Program on Education Policy and Governance, Harvard University, Cambridge, MA.

Woessmann, L. (2005b) 'The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA', *Education Economics*, 13 (2), pp. 143-169.

Woessmann L. (2008) 'How equal are educational opportunities? Family background and student achievement in Europe and the US', *Betriebswirtsch,* 78, 45–70.

Woessmann, L. (2008b) 'Efficiency and equity of European education and training policies', *International Tax Public Finance,* 15, 199–230.

Young, M. and Muller, J. (2010) 'Three educational scenarios for the future: lessons from the sociology of knowledge', *European Journal of Education*, 45, 11-27.

Young, M., Lambert, D. & Roberts, C. (2014) *Knowledge and the Future School: Curriculum and Social Justice,* Bloombury, London.

Zongyi Deng & Gopinathan, S. (2016) 'PISA and high performing education systems: explaining Singapore's education success', *Comparative Education*, 52 (4), 449-472.

.

Appendices

Table A1: Descriptive statistics for outcomes by country/territory

Country or territory	Country code	System type*	PISA outcomes (age 15): PISA reading scores (mean)	PISA Maths scores (mean)	Inequality in PISA reading scores (Gini	Inequality in PISA Maths scores (Gini coefficient)	Gap in average PISA reading score between students with high and low parental	Gap in average PISA Maths score between students with high	PIAAC outcomes (age 18 to 20): PIAAC literacy scores (mean)	PIAAC numeracy scores (mean)	Inequality in PIAAC literacy scores (Gini coefficient)	Inequality in PIAAC numeracy scores (Gini coefficient)
Australia	AU	Type 4	512.89	519.91	0.10	0.10	55.18	51.10	291.33	276.18	0.09	0.10
Austria	AT	Type 3	490.19	505.48	0.12	0.11	82.68	88.28	276.04	276.73	0.08	0.09
Canada (English)	CA-en	Type 2a	527.01	527.01	0.10	0.09	74.16	52.53	277.86	265.94	0.10	0.12
Canada (French)	CA-fr	Type 2a	527.01	527.01	0.10	0.09	74.16	52.53	268.61	268.48	0.09	0.09
Chile	CL	Type 4	449.37	421.06	0.10	0.11	72.64	73.22	238.39	223.36	0.11	0.13
Czech Republic	CZ	Type 1	482.72	509.86	0.13	0.11	116.90	146.78	283.41	281.20	0.07	0.08
Denmark	DK	Type 1b	494.48	513.03	0.10	0.09	74.34	71.53	275.36	269.54	0.08	0.09
England (UK)[1]	UK-en	Type 4	495.56	495.19	0.12	0.10	73.77	63.67	269.37	255.64	0.09	0.10
Estonia	EE	Type 1	500.75	514.58	0.10	0.09	52.47	58.39	285.84	278.24	0.08	0.08
Finland	FI	Type 1b	546.87	548.36	0.08	0.08	37.98	39.49	293.77	283.25	0.08	0.09
Flanders (Belgium)	BE-fl	Type 1	521.99	543.44	0.11	0.10	118.80	108.04	281.90	278.94	0.08	0.09
France	FR	Type 1	487.71	495.54	0.12	0.11	71.32	78.02	270.74	258.03	0.09	0.11
Germany	DE	Type 3	494.94	503.79	0.13	0.11	100.84	85.28	276.48	271.81	0.09	0.10
Greece	GR	Type 1	482.78	466.10	0.11	0.11	66.83	60.08	259.16	252.56	0.09	0.09
Ireland	IE	Type 4	517.31	501.47	.	.	.	.	270.90	255.73	0.08	0.10

Israel	IL		Type 4		473.99	446.86	0.13	0.13	112.15	120.06		261.46	251.42	0.11	0.13
Italy	IT		Type 1		468.52	461.69	0.13	0.12	45.16	44.54		263.83	254.64	0.08	0.10
Japan	JP		Type 1		497.96	523.10	0.12	0.10	93.46	84.97		298.79	281.84	0.06	0.08
Korea	KR		Type 1		556.02	547.46	0.09	0.10	43.81	63.21		294.02	283.55	0.06	0.07
Lithuania	LT		Type 1		468.44	476.60	0.10	0.10	76.17	90.39		279.86	285.22	0.08	0.08
Netherlands	NL		Type 1b		506.75	530.65	.	.	.	.		291.38	284.21	0.08	0.09
New Zealand	NZ		Type 4		520.88	519.30	0.11	0.10	73.99	78.16		281.38	268.97	0.08	0.10
N. Ireland (UK)	UK-ni		Type 4		489.32	490.50	0.12	0.11	79.59	55.89		280.73	271.66	0.08	0.09
Norway	NO		Type 2b		484.29	489.85	0.12	0.11	90.62	77.11		273.65	268.87	0.08	0.09
Poland	PL		Type 1		507.64	495.43	0.11	0.10	121.50	110.63		282.81	270.34	0.08	0.09
Russia	RU		Type 1		439.86	475.68	0.12	0.11	67.20	76.57		272.30	268.94	0.08	0.08
Singapore[2]	SG		Type 1b		525.90	562.02	0.10	0.11	78.46	81.11		291.18	291.73	0.08	0.08
Slovak Republic	SK		Type 1		466.35	492.11	0.13	0.11	158.42	175.85		278.07	279.72	0.08	0.09
Slovenia	SI		Type 1		483.08	501.47	0.11	0.11	92.15	99.74		274.99	274.24	0.08	0.09
Spain	ES		Type 4		460.83	479.96	0.11	0.10	45.24	50.26		264.26	255.30	0.08	0.09
Sweden	SE		Type 2b		507.31	502.36	0.11	0.10	56.47	51.18		284.36	277.39	0.08	0.09
United States	US		Type 2a		497.50	474.35	0.00	0.11	0.00	81.92		271.03	248.14	0.08	0.11

***Note**

Systems types:

Type 1: Differentiated systems with dedicated upper secondary

Type 1b: Differentiated systems with dedicated upper secondary and school-based apprenticeships

Type 2a: Comprehensive systems (North American model)

Type 2b: Comprehensive systems (Nordic model)

Type 3: Dual systems

Type 4: Mixed systems

Table A2a: Descriptive statistics for systems-characteristic indicators by country/territory

Country or territory	**Country code**		**Duration and timing of upper secondary education**					**Financial and human resources**					
			Theoretical duration of upper secondary education (years)	Theoretical age at completion of upper secondary education	Age at first selection	Educational stage of first selection		Average starting salary for teachers at upper secondary level (US$ PPP)	Average expenditure per FTE students at upper secondary level (US$ PPP)	Average expenditure per FTE students at upper secondary level (% per capita GDP)	Average student-teacher ratio at upper secondary level	Average number of teaching hours per year at upper secondary level	Teacher workload (student-hours per year)
Australia	AU		2	18	16*†	On entry to upper secondary		34321.37	9965.843	24.42528	12.022	803.2593	9656.78
Austria	AT		4	18	10*†	On entry to lower secondary		32680.05	12389.946	30.65999	10.13933	588.6	5968.01
Canada (English)	CA-en		3	18	16*†	During upper secondary[6]		34588.31	11316.83	28.1965	13.85	744.1676	10306.72
Canada (French)	CA-fr		2	17	16*†	On entry to upper secondary		34588.31	11316.83	28.1965	13.85	744.1676	10306.72
Chile	CL		2[3]	18	16†	On entry to upper secondary		17941.41	3119.0959	18.01742	26.08949	1086.888	28356.35
Czech Republic	CZ		4[4]	19	11*†	On entry to lower secondary		15533.45	6244.0142	24.61773	12.08275	617.4	7459.89
Denmark	DK		3	19	16*†	On entry to upper secondary		44640.11	11913.813	29.34423	13.09	377	4934.93
England (UK)[1]	UK-en		2	18	16*†	On entry to upper secondary		30204.19	10452.15	29.42746	15.6	703	10966.80
Estonia	EE		3	19	16‡	On entry to upper secondary		11875.61	6834.0061	34.0115	16.58726	577.5	9579.14

Finland	FI		3	19	16*†	On entry to upper secondary		32276.16	7911.7432	21.95906	17.05877	553.014	9433.74
Flanders (Belgium)	BE-fl		4	18	12*†	On entry to lower secondary		38938.84	13019.816	30.1792	10	630	6300.00
France	FR		3	18	15*†	On entry to upper secondary		27420.37	12873.886	37.42999	9.667581	631.87	6108.65
Germany	DE		3	19	10*†	On entry to lower secondary		53962.67	7656.8315	17.5546	13.24182	713.2475	9444.70
Greece	GR		3	18	15*†	On entry to upper secondary		26583.21	6382.6128	25.33783	8.068	415.089	3348.94
Ireland	IE		2	18	15*†	On entry to upper secondary		32600.85	11264.662	27.47449	13.946	734.8	10247.52
Israel	IL		3	18	15*†	On entry to upper secondary		13994.66	4013.0188	14.05032	11.04056	521.2745	5755.16
Italy	IT		5	19	14*†	On entry to upper secondary		29121.71	8646.2344	26.92714	12.07926	630	7609.94
Japan	JP		3	18	15*†	On entry to upper secondary		25453.62	10064.162	28.56061	12.15208	500.3465	6080.25
Korea	KR		3	18	14*†	On entry to upper secondary		26670.08	9477.2195	32.87416	16.52279	616	10178.04
Lithuania	LT		2	19	14‡	On entry to lower secondary		.	4422.7041	22.33132	8.036	610.2	4903.57
Netherlands	NL		3	18	12*†	On entry to lower secondary		38001.34	11750.35	28.1908	18.602	750	13951.50
New Zealand	NZ		3	18	16*†	During upper secondary[6]		28535.18	9007.2312	30.39965	14.35273	760	10908.07
N. Ireland (UK)	UK-ni		2	18	11	On entry to lower secondary		30204.19	10452.15	29.42746	14.7	.	.
Norway	NO		3	19	16*†	During upper secondary[6]		35990.91	14844.685	33.11661	9.432906	522.5	4928.69
Poland	PL		3	19	15*	On entry to upper secondary		12118.71	5529.6607	27.60165	12.05448	494.1	5956.12

Russian Federation	RU		2	18	15*	On entry to upper secondary		.	6420.5366	38.19516	11.31828	507	5738.37
Singapore[2]	SG		3	18[5]	12†	On entry to lower secondary		.	11035.767[7]	14.6472[8]	11.29151[9]	.	.
Slovak Republic	SK		4	19	11*†	On entry to lower secondary		11027.94	4500.5186	19.40412	14.59157	623.7	9100.76
Slovenia	SI		4	19	15*	On entry to upper secondary		26690.19	7472.1915	28.03908	14.30919	632.7	9053.42
Spain	ES		2	18	16*†	On entry to upper secondary		42324.69	10305.948	32.64113	9.572668	692.55	6629.55
Sweden	SE		3	19	16*†	During upper secondary[6]		30649.93	10496.706	26.74239	13.12136	.	.
United States	US		3	18	16*†	During upper secondary[6]		37266.83	13045.486	28.0258	14.98368	1051.2	15750.84

[1] upper secondary education in England (UK) and Northern Ireland (UK) refers to 'A' levels and equivalent programmes, which typically start at age 16 and last two years

[2] upper secondary education in Singapore refers to programmes at junior colleges and polytechnics, which typically start at age 16 and last two or three years

[3] in Chile, there is no formal delineation between lower- and upper-secondary education although the final 2 years are often regarded as being ‘upper secondary’

[4] in Czech Republic, academic programmes last for 4 years, whilst vocational programmes last for between 2 and 4 years

[5] in Singapore, junior college programmes finish at age 17 or 18; polytechnic courses usually finish at age 19, or later for those undertaking military service.

[6] Students tracked 'during upper secondary' include those in comprehensive systems who select vocational programmes within a comprehensive school system (i.e. different programmes offered within the same institutions)

[7] Figure calculated from MOE figures and converted from SG$ using PPP for consumption (https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm)

[8] Figure calculated from MOE figures using GDP in PPP USD (https://data.oecd.org/gdp/gross-domestic-product-gdp.htm)

[9] Figure calculated from MOE numbers of students and teachers in upper secondary education

Table A2b: Descriptive statistics for systems-characteristic indicators by country/territory (cont.)

		Educational provision				Upper secondary outcomes		PIAAC indicators (age 20 to 24)			
Country or territory	**Country code**	Number of school types	Standardised curriculum score	Number of programmes offered to 15 year-olds	Share of upper secondary students enrolled on vocational programmes (%)	Upper secondary completion rate (%)	Cohort years of upper secondary (Duration of upper secondary * Upper secondary	Gap in average PIAAC literacy score between students with high and low	Gap in average PIAAC numeracy score between students with high and low	Vocational Gap in average PIAAC literacy score	Vocational Gap in average PIAAC numeracy score
Australia	AU	5	0	1	57.75974	.	.	.	.	.	.
Austria	AT	6	2	4	76.77422	85.0	340	45.33	48.32	-1.79	3.27
Canada (English)	CA-en	2	1	1	7.48646	.	.	28.99	29.06	25.19	26.89
Canada (French)	CA-fr	2	1	1	7.48646	.	.	33.18	35.52	-12.23	-3.91
Chile	CL	3	2	1	35.90461	89.0	.	39.69	40.18	-22.33	-31.13
Czech Republic	CZ	6	4	6	72.60926	90.0	.	20.14	29.41	-29.81	-32.63
Denmark	DK	4	3	1	51.50747	73.0	219	27.82	31.26	-35.54	-32.38
England (UK)[1]	UK-en	6	0	1	23.79127	77.0	154	48.12	55.60	.	.
Estonia	EE	2	4	1	34.18363	.	.	35.89	39.25	.	.
Finland	FI	2	4	1	56.20564	88.0	264	42.01	47.65	-25.71	-32.05

Flanders (Belgium)	BE-fl		6	3	4	72.98175		92.0	368		44.78	46.96	.	.
France	FR		4	2	3	44.33764		89.0	267		21.60	26.40	-41.48	-48.50
Germany	DE		6	2	4	51.49402		75.0	225		38.09	42.27	-23.95	-23.64
Greece	GR		3	4	2	30.69278		88.0	264		32.20	31.23	-28.76	-35.05
Ireland	IE		4	0	4	35.90711		91.0	182		25.56	27.61	-26.89	-29.97
Israel	IL		5	.	2	38.19294		90.0	270		40.33	46.91	.	.
Italy	IT		6	4	4	59.98679		66.0	330		24.63	25.14	-34.15	-32.86
Japan	JP		3	4	2	23.5024		92.0	276		-9.60	-16.82	.	.
Korea	KR		3	4	3	23.67696		97.0	291		25.39	21.90	-16.56	-19.23
Lithuania	LT		2	4	5	29.8232		94.0	188		16.62	33.44	-23.12	-27.83
Netherlands	NL		3	2	7	67.90686		80.0	240		23.57	24.58	-36.29	-38.77
New Zealand	NZ		5	0	1	30.14357		83.0	249		34.63	35.03	-24.14	-22.24
Northern Ireland (UK)	UK-ni		5	0	1	23.79127		73.0	146		44.15	47.61	.	.
Norway	NO		2	4	1	53.86086		76.0	228		46.11	51.89	-36.75	-37.98
Poland	PL		2	4	1	53.28378		91.0	273		33.47	35.00	-29.78	-33.09
Russian Federation	RU		2	2	3	51.80285		96.0	192		31.40	33.94	-17.61	-14.47
Singapore[2]	SG		3	1	4	72.75942		98.0	294		22.26	24.00	.	.
Slovak Republic	SK		4	.	5	70.78732		89.0	356		50.95	64.70	-22.60	-24.70
Slovenia	SI		4	3	3	59.41485		86.0	344		17.24	20.92	-40.09	-42.33
Spain	ES		4	2	1	44.57024		62.0	124		21.82	20.36	-26.35	-25.15
Sweden	SE		1	4	1	59.5674		87.0	261		33.68	30.02	.	.
United States	US		3	4	1	.		90.0	270		24.24	30.26	-28.23	-19.19