Leading education and social research
Institute of Education
University of London

Department of Quantitative Social Science

University access for disadvantaged children: A comparison across English speaking countries

John Jerrim
Anna Vignoles
Ross Finnie

DoQSS Working Paper No. 12-11
December 2012

Disclaimer

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

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

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

University access for disadvantaged children: A comparison across English speaking countries

John Jerrim[1], Anna Vignoles[2] and Ross Finnie[3]

Abstract

In this paper we consider whether certain countries are particularly adept (or particularly poor) at getting children from disadvantaged homes to study for a bachelor's degree. A series of university access models are estimated for four English speaking countries (England, Canada, Australia and the United States) which include controls for comparable measures of academic achievement at age 15. We not only consider access to any university but also admission to a 'selective' institution. Our results suggest that socio-economic differences in university access are more pronounced in England and Canada than Australia and the United States, and that cross-national variation in the socio-economic gap remains even once we take account of differences in academic achievement. We discuss the implications of our findings for the creation of more socially mobile societies.

JEL classification: I20, I21, I28.

Keywords: University access, educational inequality, social mobility, PISA.

[1] Department of Quantitative Social Science, Institute of Education, University of London (J.Jerrim@ioe.ac.uk)
[2] Faculty of Education, University of Cambridge (av404@cam.ac.uk)
[3] Graduate School of Public and International Affairs, University of Ottawa (rfinnie@uottawa.ca)

Acknowledgements

Jerrim has been partly funded by the British Academy and the ESRC. He also wishes to acknowledge the invaluable support of the PATHWAYS postdoctoral programme. Vignoles time has been partly funded by the Nuffield Foundation. Helpful comments have been received from participants at the Sutton Trust social mobility summit, the 2012 IWAAE conference, 2012 LLAKES conference and seminars at the Institute of Education and Michigan State University. Particular thanks go to Stephen Childs for research assistance with the YITS data and to Barbara Schneider for facilitating Jerrim's visit to MSU.

1. Introduction

Young people from disadvantaged backgrounds are less likely to enter a well-paid professional job than their more advantaged peers. This holds true throughout the developed world (D'Addio 2007, Ermisch et al 2012). Yet it has also been suggested that in certain countries the relationship between family background and labour market outcomes is particularly strong. The intergenerational income elasticity (the association between fathers' and sons' earnings) is high in the United States by international standards (Blanden 2011, Jantti et al 2006). The same is arguably true of the United Kingdom (Blanden et al 2005), though this is a contentious issue (Gorard 2008, Goldthorpe 2012, Saunders 2012). In contrast, Canada and Australia are thought to be much more socially mobile according to this measure (Blanden 2011). A small but rapidly growing literature is attempting to explain why this is the case (Bradbury et al 2012, Blanden et al 2012, Jerrim 2012, Magnuson et al 2012). In this paper we make an important contribution to this literature by investigating the link between family background, academic achievement in school and university participation across England, Canada, Australia and the United States (see Appendix A for key characteristics of these countries).

The simple framework of Haveman and Wolfe (1995) illustrates the important role tertiary education plays in the intergenerational transmission of advantage – see Figure 1. Family background influences children's outcomes in three broad stages. Heredity and parental investments (time and goods input) combine to create large socio-economic differences in cognitive ability even on entry into school. Socio-economic background then further interacts with school quality and peers, potentially widening the socio-economic difference in achievement by the mid-teenage years (stage 1). Young people then decide whether to enter university, which subject to study and which institution to attend (stage 2). Socio-economic gaps in the decision of whether to go to university will emerge at this point due to (i) disadvantaged children's weaker academic preparation and (ii) other non-academic constraints upon their choices (e.g. credit constraints, risk aversion, lack of information, peer pressure, lack of aspiration). Young people then enter the labour market in stage 3, where those who graduate from university (who are more likely to be from affluent backgrounds) receive a significant wage premium. Universities are therefore one of the key mechanisms by which affluent parents pass on their socio-economic status (and ability to generate high earnings) to the next generation.

Figure 1

Access to university thus differs across socio-economic groups for three broad reasons:

i. Differences in academic achievement
ii. Factors that constrain students' choices above and beyond their academic ability (credit constraints, financial support, risk aversion);
iii. Other non-academic factors including student aspirations and possible discrimination in university admissions.

Yet the extent to which the first of these factors is able to explain socio-economic differences in university admission rates is a controversial issue that has been the subject of much recent academic and public policy debate. Leading economists (Cunha et al 2006) argue that the reason why rich and poor children follow different pathways as young adults is not due to factors affecting young people at the point they decide to go to university or not (e.g. credit constraints or debt aversion), but rather by what happens much earlier in their life. This reflects a growing belief amongst economists that inequality in university access largely reflects differences in prior achievement (Cameron and Heckman 2001, Chowdry et al 2012, Ermisch and Del Bono 2012) and that other factors (including those listed above) are of less significance. However, others disagree. A significant amount of sociological work continues to stress the importance of factors other than scholastic attainment that constrain young people's choices (Jackson et al 2007) and argues that intervention later in life (e.g. financial support to complete higher education) may be effective in reducing intergenerational inequalities. The resolution of this debate seems to hinge upon one key question – to what extent can socio-economic differences in university access be explained by differences in prior achievement?

Evidence from the Anglophone countries on this matter is somewhat mixed. Chowdry et al (2012), Ermisch and Del Bono (2012) and Jackson et al (2007) all find 'raw' socio-economic gaps in educational transitions at age 16 and 18 to be large in England (around 40 percentage points). But Ermisch and Del Bono (2012) state that there is 'virtually no relationship' between family background and university access once age 16 academic achievement has been controlled. In contrast, Chowdry et al (2012) find that a sizeable and statistically significant difference remains (approximately 12 percentage points) even once academic ability at age 16 is taken into account (though this is reduced to just 5 percentage points when grades at age 18 have also been controlled). Jackson et al (2007) go a step further, suggesting that up to half the socio-economic gap in teenagers' educational decisions in England is due to factors other than academic ability (Jackson et al 2007). Cameron and Heckman (2001) and Carneiro and Heckman (2002) focus on the role of credit constraints in the United States for a sample of young adults entering university in the 1980s. They find prior academic achievement to be a greater barrier to disadvantaged children's prospects of entering higher education. However Belley and Lochner (2007) argue that the direct effect of family income, and credit constraints in particular, have become substantially more important in recent years (see also Duncan and Murnane 2012). Using Canadian data, Finnie and Mueller (2008) find that the association between parental education and university participation decreases by 50% once high school grades have been taken into account; yet they also note that the 'effect' of family income remains unchanged. In a follow-up paper, Finnie (2012) stresses the importance of the culture children are exposed to throughout early childhood, with a suggestion that this factor is more important than credit constraints. In Australia, Le and Miller (2005) argue that 'equity-based scholarships or university fee rebates [need] to be provided to Year 12 graduates' in order to address the socio-economic imbalance in university education. In response Cardak and Ryan (2009) suggest that, conditional upon school achievement at age 18, disadvantaged Australian students are just as likely to attend university as their more fortunate peers. They thus argue that raising disadvantaged children's low school achievement is actually the more appropriate policy response.

The aim of this paper is to provide further clarity on this matter for Australia, Canada, England and the United States. Specifically, we shall provide comparable evidence on the link between family background, academic achievement in secondary school and access to university for these four large English speaking countries. We begin by investigating the 'raw' socio-economic gap in university access to establish whether any country is particularly adept (or poor) at limiting the association between family background and university attendance. We then investigate the extent to which these gaps can be explained by prior achievement of pupils (point i above). We use two sets of prior achievement measures (a) cognitive test scores measured at age 15 (PISA test scores) and (b) a host of school achievement measures recorded up to age 18. With these data we can demonstrate whether family background is still associated with university entry even amongst young people who are eligible and equally well qualified to attend.

We recognise that all forms of higher education may not be of equal value; labour market opportunities, dropout rates and wage returns vary by both subject and institution (Black

2006, Hoekstra 2009, Powdthavee and Vignoles 2009, Chevalier and Conlon 2003). Hence we not only consider access to a bachelor's degree (broadly defined), but also admission to more 'selective' universities and enrolment in a science, technology, engineering or mathematics (STEM) degree. Our results show that:

- The association between family background and university access is notably stronger in England and Canada than in Australia and the United States.
- Academic achievement up to age 18 can explain approximately 80% of the socio-economic university access gap in England, 60% in the United States, 55% in Australia and 30% in Canada.
- School-level factors (including school peer effects) explain only a very small amount of the socio-economic gap in university entrance, over and above their potential influence on age 15 academic achievement.
- In all four countries, children from affluent backgrounds remain significantly more likely to attend a selective institution than their middle and low SES peers, even after school achievement up to age 18 has been taken into account.

The paper now proceeds as follows. Section 2 provides an overview of key differences in the higher education systems across the countries considered. The data and econometric methodology are then described in section 3. Results follow in section 4, with conclusions and policy recommendations in section 5.

2. University systems

University systems differ markedly across the four countries that we consider. A set of key indicators can be found in Table 1. Notice how young people face different 'non-academic' constraints to university participation across these four countries (e.g. up-front costs, access to finance, investment risks, peer pressures, lack of information). Perhaps the most obvious example is cost. Partly due to its large private sector, annual tuition costs are almost three times higher in the United States ($11,605) than England ($4,731), Canada ($3,774) and Australia ($4,369). Moreover, bachelor's degree courses take longer to complete in the US (typically four years) than in some other countries (e.g. three years in England), further increasing the total tuition price of university as well as the opportunity costs[4]. Yet the countries also differ in terms of financial support, and thus their ability to limit the role of credit constraints and risk aversion for poorer students. The United States is the most generous country in terms of the proportion of the population receiving (non-repayable) scholarships and grants (65%). But England and Australia have a sophisticated system of income contingent public loans which greatly helps offset the risks associated with human capital investment (Chapman and Ryan 2005). For instance, in Australia the public loan covers the upfront cost of study, with graduates paying back a certain percentage (somewhere between 4% and 8%) once their earnings exceed a certain threshold (this was approximately $US 47,000 in 2010)[5]. Repayments are thus much more strongly linked with ability to pay in England and Australia than in Canada and the United States.

School – to – university transitions also differ across the countries. In England young people can decide to leave full time education at age 16. Those aiming for university continue in full time education for a further two years, with university offers based largely upon *predicted* grades in national examinations. Supply is also constrained in England, with a limited number of places available in different higher education institutions. There is, in contrast, a single educational transition point in Canada and the United States (at age 18), with a well-developed two-tier tertiary education system (made up of two and four year degrees awarded by community colleges and universities respectively). In Australia, the compulsory school leaving age varies by state/territory and the university admission process is also generally centralised at state or territory level, with entry determined by school grades ('ENTER' scores)[6]. Like England, use of SAT / ACT style cognitive tests are limited to a few subjects and institutions.

[4] There is also significant heterogeneity in tuition costs by institution in the United States, while in England tuition for the period in question was essentially set at a flat fee (i.e. studying Economics at Oxford cost the same as studying Art at a 'modern' community institution). In Australia, there is significant variation in tuition fees by subject (but little by institution). Specifically, there are four price bands, with tuition fees ranging from $US 4,500 for government priority areas (e.g. maths and science) up to $US 9,000 for courses including Medicine, Law and Economics.

[5] Between 2006 and 2012, English students did not start repaying university loans until their earnings exceeded approximately $US 25,000. They then repaid 9% of all earnings over this amount until the debt was repaid. Any outstanding amount was written off after 25 years. A new system has been introduced as of September 2012, with repayments not starting until individuals earn around $US 35,000, with debt written off 30 years after graduation.

[6] 'ENTER' stands for Equivalent National Tertiary Entrance Rank and 'ATAR' for Australian Tertiary Admission Rank. Young people are assigned a certain percentile rank based upon the subjects they have studied and their performance in those subjects in school exams.

3. Data and methodology

For each of the four countries we have access to a longitudinal dataset which contains information on respondents' family background, academic achievement and post-secondary school destinations. These data have a high degree of comparability, with each being a nationally representative sample of the youth population. The datasets we analyse are:

- The Longitudinal Study of Australian Youth (LSAY 2003 - Australia)
- The Youth in Transition Study Cohort A (YITS 2000 - Canada)
- The Longitudinal Study of Young People in England (LSYPE 2004 – England)
- The Educational Longitudinal Study (ELS 2002 – United States).

3.1 Sample design and response rates

The Canadian and Australian datasets are longitudinal follow-ups of the Programme for International Student Assessment (PISA) 2000 or 2003 cohort. Annual (Australia) or bi-annual (Canada) follow-ups have been conducted. ELS (United States) began by interviewing a cohort of 16 year olds in 2002, with longitudinal follow-ups at ages 18 and 20. The LSYPE (England) started by surveying 14 year olds in 2004 with annual follow-ups until age 20. In each study, schools were selected as the primary sampling unit, with probability proportional to size, and pupils randomly chosen from within (35 pupils in Australia and Canada, 26 in the US and approximately 33 in England)[7]. Of the original sample members, 78% (US), 64% (Australia), 55% (Canada) and 55% (England) remain in the sample up to age 20. To try and correct for this attrition, we apply the longitudinal weights provided. Sample sizes are 12,575 (US), 9,446 (Canada), 7,715 (England) and 6,536 (Australia).

3.2 Measurement of family background

Family background is measured using the highest level of education achieved by either parent. Parental education is a key determinant of the financial resources available within a household, and the time and goods parents invest in their offspring (Haveman and Wolfe 1995, Leibowitz 1974). It is widely used in social stratification and cross-national research

[7] We account for the clustered sample design by either making the appropriate adjustment to the estimated standard errors (using the 'svy' survey command in STATA) or by including a school level fixed effect.

(Smeeding et al 2011, Ermisch et al 2012) and is the most readily comparable indicator of family background at our disposal.

In England, Canada and the US, parents were asked directly about their educational attainment, while in Australia children acted as proxy respondents. We have converted responses into International Standard Classification of Education (ISCED) categories which are designed to compare educational attainment across countries. However, we recognise that national qualifications do not always fit easily into this framework. In an attempt to overcome this difficulty we aggregate ISCED levels into the following three broad categories (phrases in brackets refer to the US):

- 'Low' education = ISCED 0 – 2 [less than high school]
- 'Medium' education = ISCED 3 – 5b [high school to associate degree]
- 'High' education = ISCED 5a/6 [bachelor's degree and higher]

A similar combination of ISCED categories has been widely used in academic and public policy research (e.g. Ermisch and Del Bono 2012, Blanden et al 2012, Eurostat 2009, OECD 2012). The distribution of the highest parental education variable can be found in Table 2[8]. The spread of respondents across the ISCED levels is quite similar across countries, but with notably fewer individuals in the top category in England and Canada than in Australia and the US.

Table 2

3.3 Academic achievement in secondary school

A key goal of this paper is to establish the extent to which socio-economic gaps in university participation can be explained by the accumulation of academic skill. But at what age should these skills be measured? One possibility is just before university entry (e.g. age 18). This would have the advantage of truly removing differences in prior achievement as a potential explanation as to why university access differs between socio-economic groups (and thus reveal whether the higher education system is meritocratic at the point of entry). A drawback, however, is that such skill measures are potentially endogenous; the decision of whether to

[8] We have compared the distribution of mothers' educational attainment from the longitudinal datasets to the distribution of qualifications held by 35 – 55 year old women drawn from OECD (2012) and other sources of information (e.g. Chowdry 2012 in the case of England). The agreement between these different sources of information is reasonable.

enter higher education may have already been made, which could impact upon young people's motivation at school and their final school grades[9]. Alternatively, one could control for children's skills at a younger age (e.g. age 15) when such endogeneity is likely to be less of a problem. The drawback, of course, is that family background can continue to influence achievement beyond this point. As both approaches have strengths and weaknesses, we estimate a series of university access models controlling for prior achievement measures at both age 15 and age 18 (see section 3.5).

A major limitation of existing cross-national longitudinal research is the lack of comparable information on young people's academic skill. We attempt to overcome this problem by using a common, cross-nationally comparable measure of children's reading and maths ability based upon the OECD's PISA framework. The Canadian and Australian data are longitudinal follow-ups of children who sat the PISA test[10]. American children sat reading and maths tests as part of ELS which, critically, included some questions from PISA. The survey organisers have used this fact to estimate PISA reading and maths test scores using equipercentile equating (see Ingels et al 2005, pages 37 – 41, for further information). English and maths test scores are also available in the LSYPE (England), but refer to national exam performance at age 14 (key stage 3 tests). However, Micklewright and Schnepf (2006) and Micklewright et al (2012) show that the correlation between these key stage 3 national exam and PISA test scores in England is high (e.g. 0.70 for reading and almost 0.85 for maths). These papers also provide detailed regression models illustrating how children's PISA maths and reading scores map onto their national exam performance at key stage 3. Again the crucial point is that the R^2 underlying these regressions is high. This suggests we can use the former to predict the latter, and can thus produce a measure of PISA achievement that is consistent with that available in the other countries. We use results from Micklewright and Schnepf (2006) and Micklewright et al (2012) to develop such a predictive model (using a methodology with similarities to a two sample two stage least squares approach – see Angrist and Krueger 1992 and **Inoue and Solon 2010)**. Further details can be found in

[9] Moreover, in countries like England, young people have to decide whether to continue on to upper secondary education or enter the labour market at age 16. One may view this as the first step in a series of sequential decisions that combine to determine university attendance, causing measurement of acquired skill after this age to be potentially endogenous.

[10] Children's answers to the test questions were summarized by the survey organizers using an 'item-response model', producing five 'plausible values'. These are five different estimates of children's 'true' reading ability at age 15. The first plausible value is used here. Substantive findings remain intact when other plausible values are used instead. See page 129 of OECD (2009) for further information on using just one plausible value in analysis of the PISA data.

Appendix B[11]. One limitation of this approach, however, is that national exam score information is missing for private school pupils in England (7% of the LSYPE sample). This means that results for England shall be based upon state school pupils only[12]. We have investigated how the exclusion of private school pupils in England influences our results, and have found that the substantive conclusions drawn remain largely unchanged (further details available upon request)[13].

Full details of the age 18 prior attainment measures are presented in Table 3. Many of these qualifications are country specific and hence not as comparable across countries as PISA test scores. For instance, detailed information on school grades are available in England, whereas a combination of SAT / ACT scores, grade point averages and results on a cognitive math test are available in the US. However, these age 18 test scores are typically those that are taken into account when universities make their admission decisions, and so we argue it is also informative to use them in our analysis as indicators of students' ability to progress to university in the country concerned.

Table 3

3.4 Bachelor's degree, 'selective' institutions and STEM subjects

Individuals are classified as university entrants if they ever enrolled in a bachelor's degree course up to age 20 (in the US this includes young people who enrolled in a four year college but *excludes* those studying for an associate's degree)[14] We focus upon this particular qualification as it is a standard, well-known and comparable level of educational attainment across the countries considered. Age 20 is the latest point that all of our four datasets have followed children up to[15]. The implication of this, however, is that we are only able to consider socio-economic differences in university access shortly after completion of upper secondary school, and not eventual graduation rates (or participation in later adult life). We

[11] Whilst the strong correlations found by Micklewright et al (2012) gives us some confidence in the approach we have taken, we must acknowledge that the proxy PISA test scores we have created may be measured with some error. If one assumes that this measurement error is 'classical', the coefficient on the PISA parameter estimates will be downwardly biased, while those for family background will be upwardly biased.

[12] The datasets for Canada, Australia and the US include both state and private school pupils (see Appendix A).

[13] Specifically, we estimate test scores for private school children using multiple imputation. Typically, the low SES and high SES parameter estimates increase by approximately 0.10 of a standard deviation in absolute magnitude (both in the base specification and when PISA test scores have been controlled).

[14] Whilst we recognise that some two-year college students may go on to complete a four year degree, upgrade rates remain relatively low. For instance, Long and Kurlaender (2009) find that only 26% on community college students in Ohio graduate with a bachelor's degree, nine years after they began their study.

[15] In Canada the fourth survey wave has been used, when respondents were age 21. However, estimates are largely unchanged when using the third survey wave (when respondents were age 19) instead.

are conscious that graduation rates do vary across countries, with the US having a particularly high drop-out rate, and that we cannot allow for this in our work. This is one of the reasons why we also investigate entry into selective universities, where drop-out rates are much lower than for the university sector as a whole – see Powdthavee and Vignoles (2009). However, defining 'selective' institutions is not a trivial task. We take a pragmatic approach and use the following pre-defined groups:

England = 'Russell Group' institutions (www.russellgroup.ac.uk/our-universities.aspx)

Australia = 'Group of Eight' institutions (www.go8.edu.au/go8-members/go8-member-profiles)

Canada = 'U15' institutions (://rd-review.ca/eic/site/033.nsf/vwapj/sub198.pdf/$file/sub198.pdf)

United States = 'Highly selective' (Carnegie classification) institutions.

In England, Australia and Canada these are self-selected alliances of research intensive institutions, whilst the US categorisation is based on SAT / ACT scores of entrants. Using these classifications, approximately equal proportions of the age 20 population attended a selective university in each country (16% Canada, 13% US, 12% Australia and 10% England). We have tested the robustness of our results to different definitions of 'selective', with substantive conclusions largely unchanged[16].

STEM subjects include biomedical sciences, computer science, engineering, health professions, mathematics and physical sciences. Our decision to focus on this particular grouping is that (a) STEM degrees are often associated with high wage returns (see Black et al 2003, O'Leary and Sloane 2005), (b) there is policy interest in increasing STEM graduates and (c) it is a comparable grouping across countries.

[16] We have used a variety of alternative definitions for 'selective' universities. This includes average PISA test scores of entrants and the use of an international university ranking (The Times World 500 universities). Substantive conclusions remained largely unchanged when one of these alternate definitions is used.

3.5 Econometric specification

A series of logistic regression models are estimated. Explanatory factors include a set of basic controls (e.g. gender and language spoken at home), parental education dummy variables, children's cognitive skills at age 15 and academic achievement at age 18. Formally the model is specified:

$$log\left(\frac{\pi(E_{ij})}{1-\pi(E_{ij})}\right) = \alpha + \beta.S_i + \gamma.P_i + \delta.G_i + \varphi.C_i + \mu_j \quad \forall \text{ k}$$

Where:

$\Pi(E_{ij})$ = Probability of enrolment for child i in school j (E = 1 if they do, E = 0 otherwise)

S = A vector of parental education dummy variables (reference: ISCED level 3 – 5b or intermediate level of education)

P = A vector of age 15 (PISA or equivalent) test scores

C = A vector of control variables (dummy variables for gender and language spoken at home)

G = A vector of age 18 academic achievement measures

μ = A school level fixed effect

i = Child i

j = School j

k = Country k

We use three outcome variables. First, a binary response which takes a value of one if the child enters university and zero if they do not. Second, a response variable which has a value of one if the sample member attended a selective university and zero otherwise. Third, the binary outcome takes a value of one if the child is enrolled in a STEM degree and zero otherwise.

Four specifications of the above model are estimated. In specification 1 only parental education dummies and basic controls are included. The β coefficients thus reveal the overall socio-economic gap in university entrance rates, capturing all the channels by which family background influences university attendance (from in-uteri experiences to the educational decisions made in the late teenage years). In the second specification PISA test scores are included. The estimated β coefficients will now capture socio-economic differences in university entrance rates for young people with the same level of measured cognitive skill at age 15. In our third specification we include a school level fixed effect. These estimates reveal whether school-level factors explain any of the remaining SES gap in university attendance, above and beyond schools possible influence upon young people's cognitive skills at age 15. This school effect might include differences across schools in the information provided to pupils regarding post-secondary options and peer effects, hence it cannot be interpreted as an indicator of school quality per se. Unfortunately we are unable to estimate a school fixed effect model for our selective institution models due to perfect *collinearity (i.e. either all students or no students from a given socio-economic group within a given school attend a selective university)*. Finally in specification 4 we restrict the sample to only those young people who hold the qualifications required to start a bachelor's degree, and include a full set of school achievement controls up to age 18 (note the school fixed effect is removed from these estimates)[17]. This will reveal whether a socio-economic gap in university access remains amongst children who are eligible and equally well qualified to attend.

All parameter estimates shall be presented in differences in log-odds. This measure is more attractive than alternatives like the odds ratio and marginal effect (predicted probabilities) as they are not sensitive to the point on the logistic distribution on which they are estimated, and are therefore not influenced by differences between countries in the absolute proportion of children who enter university. However, appreciating that this metric is rather cumbersome to interpret, we also present predicted probabilities in the text to aid interpretation. These probabilities are based upon estimates from linear probability models following the same specification as presented in section 3[18].

[17] For instance, in the US the sample is restricted to high school graduates only, with controls included for (i) PISA test scores quintiles (ii) SAT / ACT quintiles (iii) grade point average in the 12th grade and (iv) cognitive math scores at age 18.

[18] Note that, being linear, estimated marginal effects from the linear probability model do not depend upon the level of the probability

4. Results

This section summarises the main findings from our university access models. Tables including all parameter estimates can be found in Appendix C.

4.1 Access to a bachelor's degree

Figure 2 illustrates the socio-economic gap in university access from the four model specifications discussed above. The light grey segments of the bars illustrate differences in university access between the low SES (ISCED 0 - 2) and middle SES (ISCED 3 – 5b) groups. The dark grey segments refer to the middle SES – high SES (ISCED 5A/6) comparison[19]. The thin black lines represent the estimated 90% confidence intervals.

Figure 2

Panel A presents results from our base specification (basic controls only). There is a statistically significant difference in the university participation rate between the low and middle SES groups in each country. This gap is of substantial magnitude in Canada (1.2 log – odds or 20 percentage points), England and the United States (approximately 0.85 log-odds or 15 percentage points)[20]. Interestingly, the gap is significantly smaller (at the 1% level) in Australia (0.25 log – odds or just 7 percentage points). Thus disadvantaged children are relatively more likely to start a bachelor's degree course in Australia than in the other countries.

Turning to the middle – high SES comparison, differences are substantial (always more than 1.0 log odd) and significantly different from 0 at the 1% level. The gap is particularly big in England (1.5 log odds or 35 percentage points), meaning that high SES children are approximately two and a half times more likely to enter university than a young person from an 'average' family background. The analogous figures are 1.3 log – odds or 31 percentage points in Canada and the US (England is significantly different to both at the 5% level) and 1.15 log-odds or 26 percentage points in Australia (which is significantly different at the 5% level to England and the US). Bringing these results together, the overall socio-economic gap (i.e. the difference between the high and low SES groups) is notably bigger in

[19] Thus the total length of the bars represents the overall socio-economic gaps at age 20 (i.e. the difference between the low SES and high SES groups).
[20] These probabilities are based upon estimates from linear probability models (following the same specification as presented in section 2).

Canada (2.45 log – odds) and England (2.35 log – odds) than the United States (2.15 log – odds) and, particularly, Australia (1.50 log – odds). Thus, taking the system as a whole, Australia succeeds in achieving the smallest gap in the likelihood of going to university between higher and lower SES children.

Panel B presents estimates controlling for PISA test scores. The previous substantial difference between the low SES and middle SES groups has been greatly reduced in England and the United States (from 0.85 log – odds to approximately 0.20 and 0.35 respectively), modestly in Canada (from 1.15 to 0.85 log - odds), but with virtually no change in Australia (0.36 log – odds in specification 1 to 0.35 in specification 2). Although the difference between low and middle SES children is still statistically significant in each country, the magnitude is now reasonably small (roughly 5 percentage points) in three of the countries considered (the exception is Canada where the gap remains around 20 percentage points). Thus the reason why disadvantaged children are less likely to go to university than a child from an 'average' background is largely due to differences in the skills that have been acquired earlier in life (before age 15). It is also notable that differences in age 15 skills explain most of the cross-national variation observed in previous estimates. University participation amongst the low and middle SES groups is more equal in Australia than in England and the US due to factors that take hold before age 15 and *not* differences in the design of the tertiary education system per se. This has important implications for public policy. If the goal of British and American policymakers is to raise university participation amongst the most disadvantaged children, enhancing their academic achievement is imperative. Focusing on the design of the higher education system alone (e.g. entry pathways, tuition fees, financial support, credit and information constraints, entry criteria) is unlikely to be enough, at least on the basis of this evidence.

The middle – high SES gap also declined once PISA test scores are included in the model. The gap declined from approximately 1.5 to 1.0 log – odd in England, 1.4 to 1.1 in Canada, 1.3 to 1.0 in the US, and 1.15 to 0.75 in Australia. Academic attainment at age 15 thus accounts for one third of the difference in degree enrolment between these groups in England and Australia, and around a quarter in Canada and the US. Yet even after age 15 achievement has been taken into account, a large SES gap (and cross-national variation) remains. For instance, high SES children in England, Canada and the US are still twice as likely to enter university as their middle SES peers (and roughly 1.7 times as likely in Australia). Again, this has important implications for public policy. Firstly, the high rates of university access amongst the most advantaged group cannot solely be attributed to their superior academic skill at age 15. Secondly, high SES children are significantly more likely to enter university than both low SES children and those from ‘average’ backgrounds. This would imply that we need policies designed to close the university participation gap between the ‘elite’ and the ‘rest’ of the population as much as the gap between the poor and the ‘rest’.

A school fixed effect is included in Panel C. Interestingly the estimated SES parameters hardly change. The middle – high gap declines by just 0.01 log - odds in England, 0.05 in the United States, and only slightly more (0.10 log-odds) in Canada and Australia. This suggests that schools currently play a relatively minor role in explaining SES differences in university access (net of their influence upon age 15 cognitive skills). This is a powerful result. It suggests that even when children attend the same school and have similar levels of achievement at age 15, those from the middle SES group are still less likely to go to university than their high SES peers. Hence the SES gap in degree enrolment is not simply caused by poorer students attending lower quality schools or schools that do not help their students apply to go to university. The implication of this finding is also, for example, that school peer effects (e.g. disadvantaged young people not going to university due to peer pressure) do not seem to be an important factor beyond their possible influence upon age 15 achievement.

In specification 4 the sample is restricted to those eligible to enter university, with a wide range of academic achievement scores up to age 18 included in the model. The difference between the bottom and middle groups is no longer statistically significant at conventional thresholds (in any country other than Canada). This supports our claim that in three of the countries, disadvantaged children’s low level of academic achievement at school is the primary cause of their relatively low levels of university participation as compared to children from an ‘average’ background. The middle – high SES gap is also smaller once we allow for age 18 achievement measures, and it is reduced by more in some countries than in others. For instance, compared to panel B/C, estimates of the SES gap are reduced by roughly 5% in Canada, 20% in Australia and the United States and by approximately 50% in England (from 0.95 to 0.45 log odds). This is perhaps unsurprising, given that young people in England have to decide whether to stay in education at *two* key transition points (age 16 and age 18). Hence many children in England do not complete schooling up to age 18 and thus do not hold the qualifications needed to access a bachelor’s degree. Therefore the raw socio-economic gaps in university participation are larger in England initially, but once we allow for entry qualification at age 18, they are reduced to a more modest level. One could argue that the English system is more meritocratic since one’s achievement and qualification level at age 18 is the main driver of university participation rather than one's socio-economic status. Alternatively one could make the point that the socio-economic differences in university participation are larger in England because the education system gets increasingly selective even before university entry[21].

To complete this section, we summarise findings regarding our central research question – to what extent can academic achievement explain the socio-economic gap in university participation in each of the four countries? The estimated difference in university access between the top and bottom SES groups is reduced by 35% in Canada, 55% in Australia, 60% in the United States and 80% in England between specification 1 (basic controls only) and specification 4 (full set of achievement controls up to age 18)[22]. Thus academic achievement up to age 18 accounts for most of the socio-economic gap in university access in three of the four countries considered. Yet it is notable that non-trivial differences in university participation by family background remain (e.g. roughly ten percentage points between the most and least advantaged groups in Australia and six percentage points in England), even amongst the subset of young people who are eligible and equally well qualified to attend. With regard to specific countries, although socio-economic gaps in university access are particularly large in England, this is mostly driven by differences in academic preparation for university. Equalising school achievement (and completion of secondary education up to age 18) between rich and poor should thus be this country's most pressing policy concern. In contrast, previous work has shown that the association between family background and teenage achievement is weak in Canada by international standards (Jerrim 2012, Jerrim and Micklewright 2011). Yet this does not stop socio-economic differences in university access in this country being particularly large. This perhaps suggests that in Canada, in contrast to England, policies to equalise school achievement will only be effective at raising the university participation rate if they are accompanied by initiatives helping to promote access at the point of entry.

[21] Another possible explanation is that the age 18 achievement scores used for England contain less noise than those for other countries.
[22] These figures refer to percentage changes in the log – odds.

4.2 Selective institutions

Figure 3 presents results from the second set of university access models, focusing upon participation at selective universities. Panel A shows the results from the model that includes all students and just basic controls. Socio-economic differences in the likelihood of attending a selective institution are large in all four countries, particularly the gap between those from middle and high SES backgrounds. We conclude that access to these elite institutions is highly socially graded in all of the countries that we consider. Previous research has found that qualifications from these institutions offer economic rewards above and beyond those from a 'typical' bachelor's degree (Chevalier and Conlon 2003, Hoekstra 2009). Hence it is a concern that Figure 3 panel A demonstrates how young people from advantaged homes are the main beneficiaries of this labour market premium.

Perhaps a better way to consider the SES gap in the likelihood of a student attending a selective institution is to focus only on those who enrol in university in the first place. The remaining panels show results when the sample is restricted to only those young people who enter university. Figure 3 panel B thus illustrates the socio-economic gap in entry to a selective institution, *conditional upon* university attendance. Notice how the difference between the low and middle SES groups is now small and statistically insignificant in England, Canada and Australia; conditional upon going to university, disadvantaged children are just as likely to enter a selective institution as a young person from an 'average' background. This gap is much larger, and statistically different from 0 at the 5% level, in the US (0.6 log – odds or 7 percentage points). Yet caution is needed when interpreting these results. Confidence intervals are wide, reflecting the fact that relatively few disadvantaged children remain in the sample now that it has been restricted to university attendees only. This is revealing in itself; it suggests that the major issue currently facing disadvantaged young people is access to university in general and not specifically about admission to selective institutions. Due in part to this reduction in sample size, one can not reject the null hypothesis that the low – middle SES gap is equal to 0 at the 5% level in any of the remaining model specifications (panels B and D).

We turn to the middle – high SES comparison from Panel B. The middle-high SES gap is larger in England and the United States (1.0 log – odd) than in Australia (0.75 log – odds) and Canada (0.60 log – odds) and differences in the gap across countries are statistically significant at the 10% threshold. Nevertheless, in all four countries the middle-high SES gap is substantial (around 15 percentage points in Australia and Canada and 20 percentage points in England and the United States). This suggests that not only are high SES children more likely to go to university, but that they are more likely to go to a selective institution conditional upon their higher rate of attendance.

Is this socio-economic gap in selective university access simply a reflection of high SES children's superior cognitive ability and higher school grades? Estimates in Panel C (PISA controls) and Panel D (PISA controls plus age 18 school grades) suggest that this is only part of the explanation. For instance, in the US estimates decline from 1.0 (panel B) to 0.75 when PISA test scores are controlled (panel C). This then falls to 0.60 when age 18 school grades are included in the model (panel D). A similar pattern occurs in England and Australia. But a non-trivial difference between young people from high SES backgrounds and the other two groups remain. In the previous section, we demonstrated how high SES children in England are six percentage points more likely to enter university as their middle SES peers, even once a host of achievement measures have been controlled. Figure 3 Panel D illustrates that, conditional upon this already greater likelihood of going to university, children from advantaged backgrounds are then a further eight percentage points more likely to attend a selective institution (having controlled for academic achievement).

4.3 STEM courses

We now turn to the issue of socio-economic differences in subject choice, focusing upon whether disadvantaged children are less likely to undertake a STEM qualification than their advantaged peers. All estimates are conditional upon university enrolment, and thus reveal the extent to which high and low SES children take different subjects having made the decision to enter university. Results are presented in Table 4. In Australia and the United States, estimates are small and statistically indistinguishable from zero at conventional thresholds for all model specifications. The parameter estimate for the middle – high SES comparison in England is statistically significant in the base specification, but not once PISA math and reading test scores have been included in the model. Similarly, the low – middle parameter estimate is significant at the 5% level in the base specification for Canada. However this reduces to significance at the 10% level once PISA test scores have been included in the model and becomes insignificant when the full range of achievement measures have been controlled.

Table 4

The conclusion we therefore reach is that, conditional upon university entry, there is little evidence of socio-economic inequality in this particular dimension of subject choice. In additional analysis (not presented for brevity) we find that the same holds true for other possible subject groupings, including differences in access to university courses that offer high economic returns. This is consistent with Chowdry et al (2012) who, using population level administrative data for England, found little association between material deprivation and university subject choice. Our contribution has been to show that this result seems to hold true across Anglophone countries.

5. Conclusions

Recent international comparisons of intergenerational income mobility (the link between fathers' and sons' earnings) have suggested that the US and (arguably) the UK are less socially mobile than Australia and Canada (Blanden 2011). Although this is a contentious issue (Gorard 2008, Goldthorpe 2012) a growing number of academics are now investigating how the link between family background and children's skills varies across these English-speaking countries. This interest has stemmed from the widely held view that socio-economic differences in skill development maybe one of the key drivers of intergenerational income persistence. Most of these studies have focused on the extent of social grading of educational achievement scores in early or middle childhood (Bradbury et al 2012, Blanden et al 2012, Jerrim 2012). This paper makes an important contribution by considering how the link between family background, achievement in secondary school and access to university compares across these four countries[23]. Our focus has been on the extent to which low and middle SES children's lower rates of university participation can be explained by their weaker scholastic achievement prior to university entry. This can inform the debate about whether improving the achievement of more disadvantaged children in the school system should be a priority rather than policy reform at the point of entry into tertiary education (e.g. credit constraints). This is a crucial policy question given the difficult decisions governments

[23] In this paper we have only been able to investigate socio-economic differences in university participation. Yet university persistence and graduation rates also vary by family background, with non-completion a particular problem in certain countries, such as the US (see OECD 2008). Information on university completion is likely to become available in future waves of the data we have analysed, and is an important area for future research.

are facing about which part of the education system they should spend their increasingly scarce resources on.

Our analysis suggests that socio-economic gaps in university participation are substantial in all four countries. For instance, young people from disadvantaged backgrounds in England are five times less likely to enter university than their more advantaged peers. Yet we also find evidence of cross-national variation, with the rich – poor divide being more pronounced in England and Canada than Australia and (to a certain extent) the United States. We can thus conclude that the system as a whole in Australia appears to create more equality in university participation rates across rich and poor children. However, similarities across these four countries are perhaps as striking as any differences, with the same broad patterns holding across each of these English – speaking countries. School achievement is found to be an important determinant of the university access gap in all four countries, with estimates of the SES gap reduced by 80% in England, 60% in the US, 55% in Australia and 35% in Canada once academic ability up to age 18 has been controlled. We conclude that SES differences in children's achievement at primary and secondary school are the major cause of the large differences in university participation rates by SES that we observe, particularly in England and the US. This is consistent with the seminal work of Cunha et al (2006) which suggested that inadequate resources invested throughout childhood is the primary cause of disadvantaged children's low university participation rates. However, it is important to recognise that sizeable and statistically significant socio-economic differences in university participation rates remain in all four countries, even amongst young people who are eligible and equally well qualified to attend. Indeed, even after conditioning upon university attendance and a range of school achievement scores up to age 18, young people from affluent backgrounds remain twice as likely to enter a 'selective' higher education institution as their less fortunate peers (across all four countries). We thus concur with Jackson et al (2007) when they state *'ability that is demonstrated by children from less advantaged backgrounds in their earlier academic careers is still often not exploited as fully as it could be at later stages'*.

What do these findings imply for public policy? Firstly, as we find the SES gap at the top (i.e. between high and middle SES pupils) is substantial in all four countries, this implies that interventions should seek to reduce differences in the likelihood of going to university between the most advantaged children and the rest of the student body, rather than focusing exclusively on young people from the most disadvantaged homes. Secondly, the key role of prior academic achievement suggests that initiatives designed to boost school performance (rather than lowering the costs of university through bursaries and fee waivers) will be pivotal to reducing socio-economic inequality in university participation, particularly in England and the US. Given that earlier intervention is thought to be more cost effective than later intervention (Cunha et al 2006) we argue that this is where the vast majority of governments' 'widening access' funds should be spent.

Finally, although raising school achievement is vitally important, policymakers should not lose sight of the under-representation of low and middle class students at selective higher education institutions (particularly as we continue to find low participation rates amongst low and middle SES children, even when they have the academic ability and school grades to gain entry). Despite the prevailing focus on early years policy, additional interventions are still needed at this later stage. But these must be cost effective, to ensure sufficient resource is allocated to improving children's achievement in school (and the pre – school years). One possible option that is likely to be relatively low cost is to encourage universities to use 'contextual' information (e.g. family background) when admitting prospective students. The fact that low and high SES pupils do not start from the same place, financially and academically, needs to be acknowledged by universities and incorporated into their decision making about admissions and bursaries. Although some countries (e.g. the US) are well advanced in their use of such information, others are not (e.g. England). Our recommendation is that the use of such 'contextual' information in the university admission process should become more widespread, alongside concerted action to narrow socio-economic gaps in pupil achievement at school.

References:

Angrist, J. and Krueger, A. (1992) 'The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples', *Journal of the American Statistical Association*, 87(418): 328 – 336.

Belley, P. and Lochner, L. (2007) 'The Changing Role of Family Income and Ability in Determining Educational Achievement', Journal of Human Capital*, 1(1): 37 – 89.*

Black, D., Sanders, S. and Taylor, L. (2003) 'The *Economic Reward for Studying Economics'*, *Economic Inquiry*, 41(3): 365–377.

Black, D. (2006) 'Estimating the Returns to College Quality with Multiple Proxies for Quality', *Journal of Labor Economics*, 24(3): 701 – 728.

Bradbury, B., Corak, M, Waldfogel, J. and Washbrook, E. (2012) 'Inequality during the Early Years: Child Outcomes and Readiness to Learn in Australia, Canada, United Kingdom and United States.' In John Ermisch, Markus Jantti and Timothy Smeeding (eds) '*Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage',* New York: Russell Sage Foundation.

Blanden, J., Gregg, P. and Machin, S. (2005) 'Intergenerational Mobility in Europe and North America', London: Centre for Economic Performance.

Blanden, J. (2011) 'Cross-national Rankings of Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology', *Journal of Economic Surveys*, doi: 10.1111/j.1467-6419.2011.00690.x.

Blanden, J., Katz, I., and Redmond, G. (2012) 'Persistent Inequality? A Comparison of the Impact of Family Background on Children's Outcomes in the U.K. and Australia.' In John Ermisch, Markus Jantti and Timothy Smeeding (eds) *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage*, New York: Russell Sage Foundation.

Cameron, S. and Heckman, J. (2001) 'The Dynamics of Educational Attainment for Black, Hispanic, and White Males', *Journal of Political Economy*, 109(3): 455 – 499.

Cardak, B. and Ryan, C. (2009) 'Participation in Higher Education in Australia: Equity and Access', *The Economic Record*, 85(271), 433 – 448.

Carneiro, P. and Heckman, J. (2002) 'The Evidence on Credit Constraints in Post-Secondary Schooling', *Economic Journal*, 112(482): 705 – 734.

Chapman, B. and Ryan, C. (2005) 'The Access Implications of Income Contingent Charges for Higher Education: Lessons from Australia', *Economics of Education Review*, 24(5): 491 – 512.

Chevalier, A. and Conlon, G. (2003), 'Does it Pay to Attend a Prestigious University?', Centre for the Economics of Education (CEE) Discussion Paper number 33. Accessed 18th October 2012 from http://cee.lse.ac.uk/ceedps/ceedp22.pdf

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2012) 'Widening Participation in Higher Education: Analysis Using Linked Administrative Data', *Journal of the Royal Statistical Society Series A*, doi: 10.1111/j.1467-985X.2012.01043.x.

Cunha, F., Heckman, J. and Lochner, L. (2006) 'Interpreting the Evidence on Life Cycle Skill Formation'. In Eric Hanushek and Finis Welch (eds) *Handbook of the Economics of Education*, Amsterdam: Holland North.

D'Addio, A. (2007) 'Intergenerational Transmission of Disadvantage: Mobility or Immobility Across Generations?', OECD Social, Employment and Migration working paper 52, Paris: OECD.

Dearden, L., Fitzsimons, E. and Wyness, G. (2011) 'The Impact of Tuition Fees and Support on University Participation in the UK', Institute for Fiscal Studies working paper. Access 26th November 2012 from http://www.ifs.org.uk/wps/wp1117.pdf

Duncan, G. and Murnane, R. (2012) *Whither Opportunity: Rising Inequality, Schools, and Children's Life Changes.* New York: Russell Sage Foundation.

Ermisch, J. and Del Bono, E. (2012) 'Inequality in Achievements during Adolescence' in J. Ermisch, M. Jantti and T. Smeeding (eds) *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage*, New York: Russell Sage Foundation.

Ermisch, J., Jantti, M. and Smeeding, T. (2012) *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage*, New York: Russell Sage Foundation

EUROSTAT (2009) 'The Bologna Process in Higher Education in Europe: Key Indicators on the Social Dimension and Mobility', doi: 10.2785/32096.

Finnie, R. and Mueller, R. (2008) 'The Effects of Family Income, Parental Education and Other Background Factors on Access to Post-Secondary Education in Canada: Evidence from the YITS', MESA Research Paper 2008-2. Accessed 25th October 2012 from http://higheredstrategy.com/mesa/pub/pdf/MESA_Finnie_Mueller.pdf

Finnie, R. (2012) 'Access to Post-Secondary Education: The Importance of Family Culture', Children and Youth Service Review, 34(6): 1161-1170.

Goldthorpe, J. (2012) 'Understanding – and Misunderstanding – Social Mobility in Britain: The Entry of the Economists, the Confusion of Politicians and the Limits of Educational Policy', University of Oxford Department of Social Policy and Intervention working paper 1/2012. Accessed 18th October 2012 from www.spi.ox.ac.uk/fileadmin/.../Goldthorpe_Social_Mob_paper.pdf

Gorard, S. (2008) 'A Reconsideration of Rates of Social Mobility in Britain', *British Journal of Sociology of Education*, 29(3): 317 – 324.

Haveman, R. and Wolfe, B. (1995) 'The Determinants of Children's Attainments: A Review of Methods and Findings.' *Journal of Economic Literature* 33(4): 1829-1878.

Hoekstra, M. (2009) 'The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach', *The Review of Economics and Statistics,* Volume 91 (4): 717 – 724.

Ingels, S., Pratt, D., Rogers, J., Siegel, P. and Stutts, E. (2005) 'Education Longitudinal Study of 2002: Base-year to First Follow-up Data file Documentation', US Department of Education. Washington, DC: National Center for Education Statistics.

Inoue, A. and Solon, G. (2010) 'Two-Sample Instrumental Variables Estimators', *The Review of Economics and Statistics,* 92(3): 557-561.

Jackson, M., Erikson, R. Goldthorpe, J. and Yaish, M. (2007) 'Primary and Secondary Effects in Class Differentials in Educational Attainment: The Transition to A-level Courses in England and Wales', *Acta Sociologica,* 50(3): 211-229.

Jantti, M. Bratsberg, B. Roed, K. Raaum, O. Naylor, R. Osterbacka, E. Bjorklund, A. and Erikson, T. (2006) 'American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic countries, the United Kingdom and the United States', IZA Discussion Paper 1938.

Jerrim, J. (2012) 'The Socio-economic Gradient in Teenagers' Literacy Skills: How Does England Compare to Other Countries?', *Fiscal Studies,* 33(2): 159 - 184.

Jerrim, J. & Micklewright, J. (2011) 'Children's Education and Parents' Socio-economic Status: Distinguishing the Impact of Mothers and Fathers.' In Robert Erikson, Markus Jantti and Timothy Smeeding (eds) *Persistence, Privilege, and Parenting: The Comparative Study of Intergenerational Mobility,* New York: Russell Sage Foundation.

Le, A. and Miller, P. (2005) 'Participation in Higher Education: Equity and Access', *The Economic Record,* 81(253): 152 – 165.

Leibowitz, A. (1974) 'Home Investments in Children', *Journal of Political Economy,* 82(2,II): S111-31.

Long, B. and Kurlaender, M. (2009) 'Do Community Colleges Provide a Viable Pathway to a Baccalaureate Degree?', Educational Evaluation and Policy Analysis, 31(1): 30 – 53.

Magnuson, K., Waldfogel, J. and Washbrook, E. (2012) 'The Development of SES Gradients in Skills during the School Years: Evidence from the U.S. and U.K.' In John Ermisch, Markus Jantti and Timothy Smeeding (eds) *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage,* New York: Russell Sage Foundation.

Micklewright, J. and Schnepf, S. (2006) 'Response Bias in England in PISA 2000 and 2003', Department for Education and Skills Research Report 771. Accessed 18th October 2012 from www.education.gov.uk/publications/eOrderingDownload/RR771.pdf.

Micklewright, J., Schnepf, S. and Skinner, C. (2010) 'Non-response Biases in Surveys of School Children: The Case of the English PISA Samples', DoQSS working paper. Access 26th November 2012 from http://ideas.repec.org/p/qss/dqsswp/1004.html.

Micklewright, J., Schnepf, S. and Skinner, C. (2012) 'Non-response Biases in Surveys of School Children: The Case of the English PISA Samples', *Journal of the Royal Statistical Society Series A*, 175(4): 915 – 938.

Milburn, A. (2012) 'Social Mobility in Higher Education', GMT interview, accessed 17th October 2012 www.hecsu.ac.uk/assets/assets/documents/Alan_Milburn_Spring_2012.pdf.

OECD (2008) 'Education at a Glance 2008: OECD indicators', OECD: Paris.

OECD (2009) 'PISA Data Analysis Manual: SPSS Second Edition', OECD: Paris.

OECD (2012) 'Education at a Glance 2012: OECD Indicators', OECD: Paris.

O'Leary, N. and Sloane, P. (2005) 'The Return to a University Education in Great Britain', National Institute Economic Review, 193(3): 75 – 89.

Powdthavee, N. and Vignoles, A. (2009), 'The Socioeconomic Gap in University Dropouts', *The B.E. Journal of Economic Analysis & Policy*, 9(1): article 19.

Saunders, P. (2012) *Social Mobility Delusions*, London: Civitas.

Smeeding, T. Erikson, R. and Jantti, M. (2011) *Persistence, Privilege and Parenting: The Comparative Study of Intergenerational Mobility*, New York: Russell Sage Foundation.

Table 1. Higher education across Anglophone countries

	Source	US	England	Canada	Australia
Educational expenditure					
% of GDP spent on tertiary education	OECD (EAG)	1.3	0.8	1.8	1.1
Bachelor's degree Enrolment					
% of population starting bachelor's degree by age 20	Author calculation	45	37	43	39
% of population obtaining bachelor's degree (all ages)	OECD (EAG)	50	48	36	38
Non-completion rate (% of entrants)	OECD (EAG - 2008)	44	21	25+	28
% of enrolments by foreign students	OECD (EAG)	3	18	7	22
% tertiary students rolled in private universities	OECD (EAG)	32	0	0	3
University tuition fees					
Average annual tuition fees *public* institutions ($US)	OECD (EAG)	6,312	4,731*	3,774	4,222
Average annual tuition fees *private* institutions ($US)	OECD (EAG)	22,852	-	-	9,112
Average tuition fee all students ($US)	Author calculation	11,605	4,731*	3,774	4,369
Average length of bachelor's degree course (years)		4	3	3-4^	3 – 4^
Tuition cost of a bachelor's degree ($US)	Author calculation	46,419	14,193	15,096	17,475
University scholarships					
% of pupils receiving grant / scholarship	OECD (EAG)	65	58	-	8
% of pupils receiving public loans	OECD (EAG)	50	87	-	81
% NOT receiving loan, scholarship or grant	OECD (EAG)	24	6	-	19

Notes:

1 EAG stands for Education at a Glance.

2 Tuition costs have been converted into US dollars by the OECD using purchasing power parity.

3 * Figures refer to pre September 2012.

4 + Refers to Québec only.

5 ^ Degree length varies by subject in Canada and honours degrees are 4 years in Australia.

Table 2. Distribution of highest level of parental education across countries

	Australia	England	United States	Canada
ISCED 0 - 2	13	12	7	9
ISCED 3 - 5B	47	68	56	63
ISCED 5A / 6	40	20	38	27

Notes:

1 Figures refer to column percentages
2 ISCED 0 – 2 = Below high school; ISCED 3 – 5B = High school to associates degree; ISCED 5A /6 = Bachelor’s degree or higher.

Table 3. School achievement grades within the longitudinal datasets

England	United States	Australia	Canada
Completed age 18 schooling	High school graduate	High school graduate	High school graduate
A* - C GCSE maths	GPA in grade 12	Tertiary entry rank quintile	GPA high school
A* - C GCSE English	Carnegie units taken		GPA maths
Key stage 4 total points	SAT quintile (or equivalent ACT)		GPA reading
Key stage 5 total points	Age 18 maths test quintile		
A-Level grades			

Table 4. Access to a STEM qualification (conditional upon university participation)

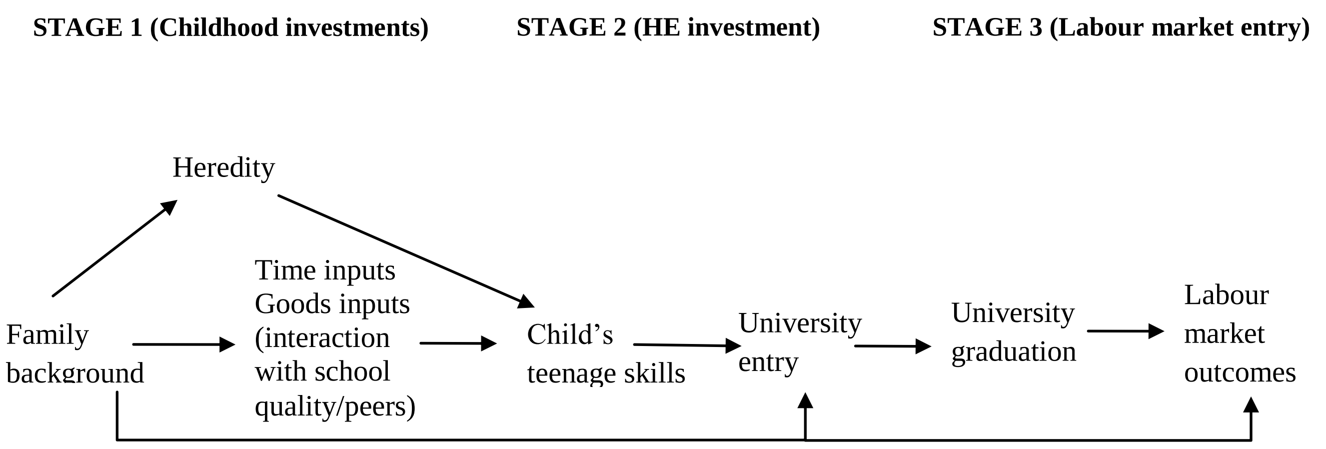
		Canada		England		USA		Australia	
	Comparison	**Log - odds**	**SE**	**Log - odds**	**SE**	**Log – odds**	**SE**	**Log - odds**	**SE**
Basic controls	Low vs Middle	-0.74*	0.34	0.20	0.16	-0.06	0.27	0.29	0.26
	Middle vs High	0.18	0.11	0.37*	0.17	0.11	0.08	0.01	0.12
PISA controls	Low vs Middle	-0.56**	0.35	0.15	0.16	0.02	0.28	0.26	0.24
	Middle vs High	0.08	0.11	0.20	0.17	0.06	0.08	-0.09	0.12
Full achievement controls	Low vs Middle	-0.28	0.40	0.11	0.17	0.00	0.28	0.28	0.24
	Middle vs High	0.12	0.13	0.13	0.18	0.05	0.08	-0.11	0.12

Notes:

1 * and ** indicate statistical significance at the 5% and 10% thresholds.

2 ‘Comparison’ column refers to differences between SES groups (e.g. “low vs middle” refers to the difference between low and middle SES groups)

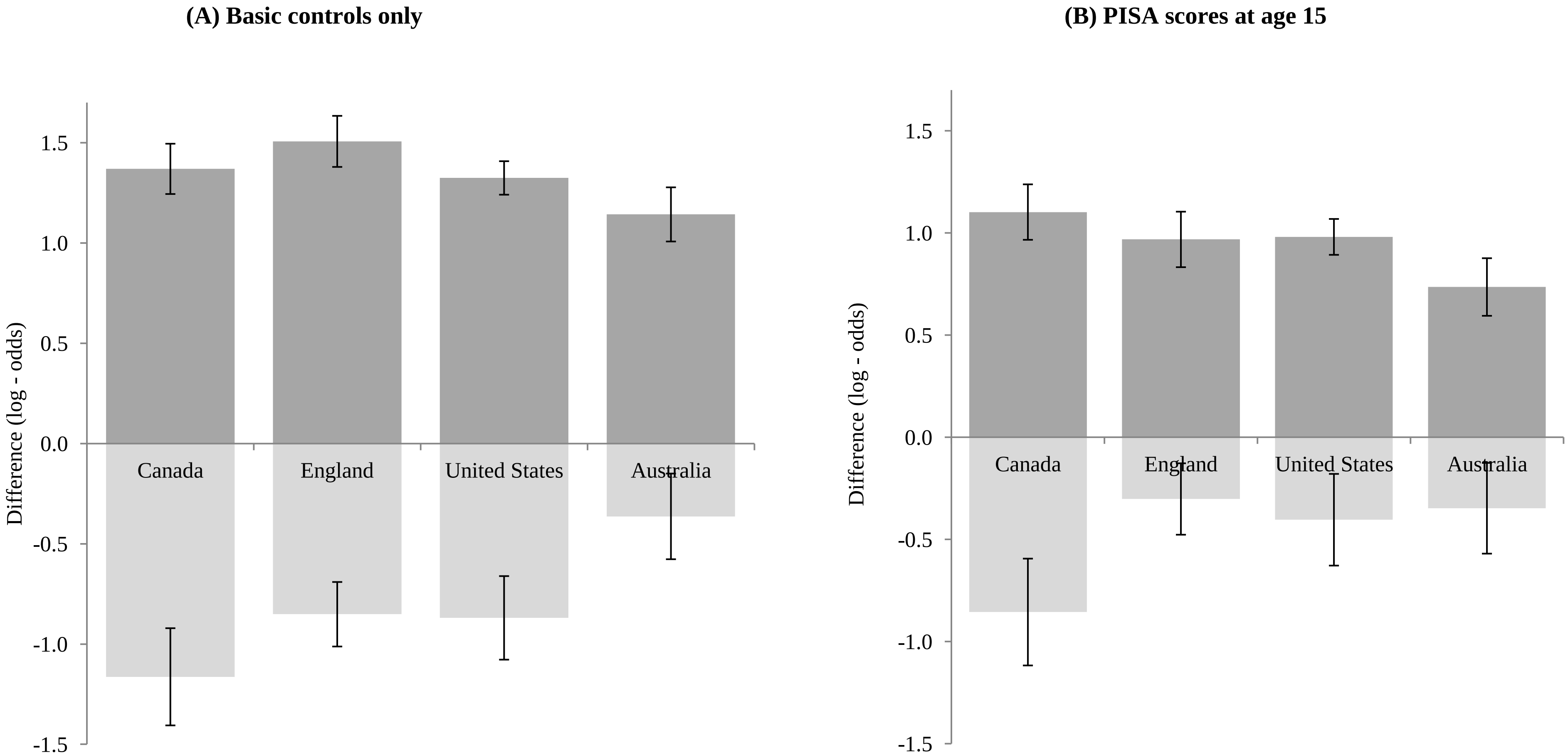
Figure 1. A framework of intergenerational persistence



Notes

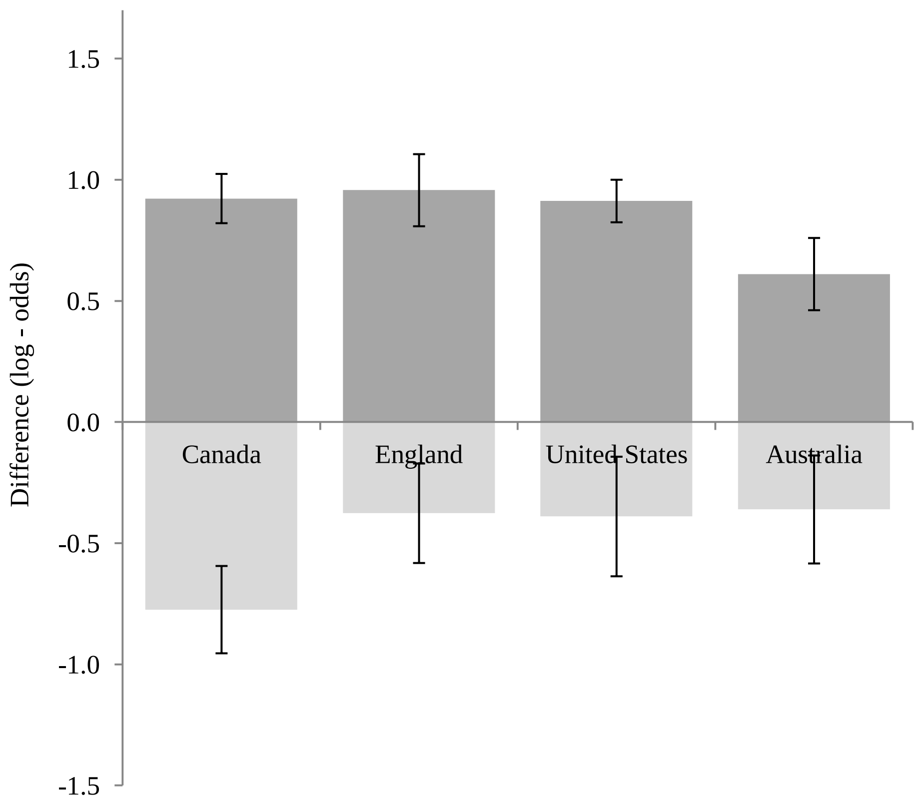
1 Source: Adapted from Haveman and Wolfe (1995, figure 1).

Figure 2. The socio – economic gap in college participation across Anglophone countries



Notes: Figures for England refer to state school pupils only. The light grey segment of the bars illustrates the difference between ISCED 0 – 2 and ISCED 3 – 5B groups. Dark grey segments refer to the difference between ISCED 3 – 5B and ISCED 5A /6 groups. Thin black lines running through the centre are the estimated 90% confidence intervals.

(C) School fixed-effects



(D) School grades age 18

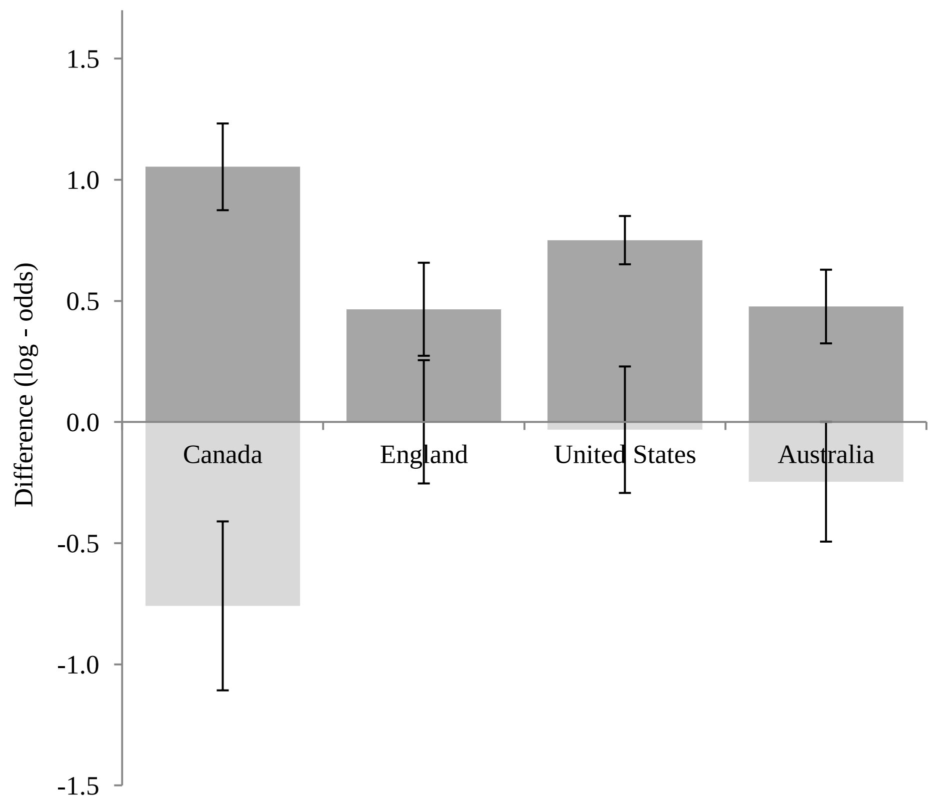
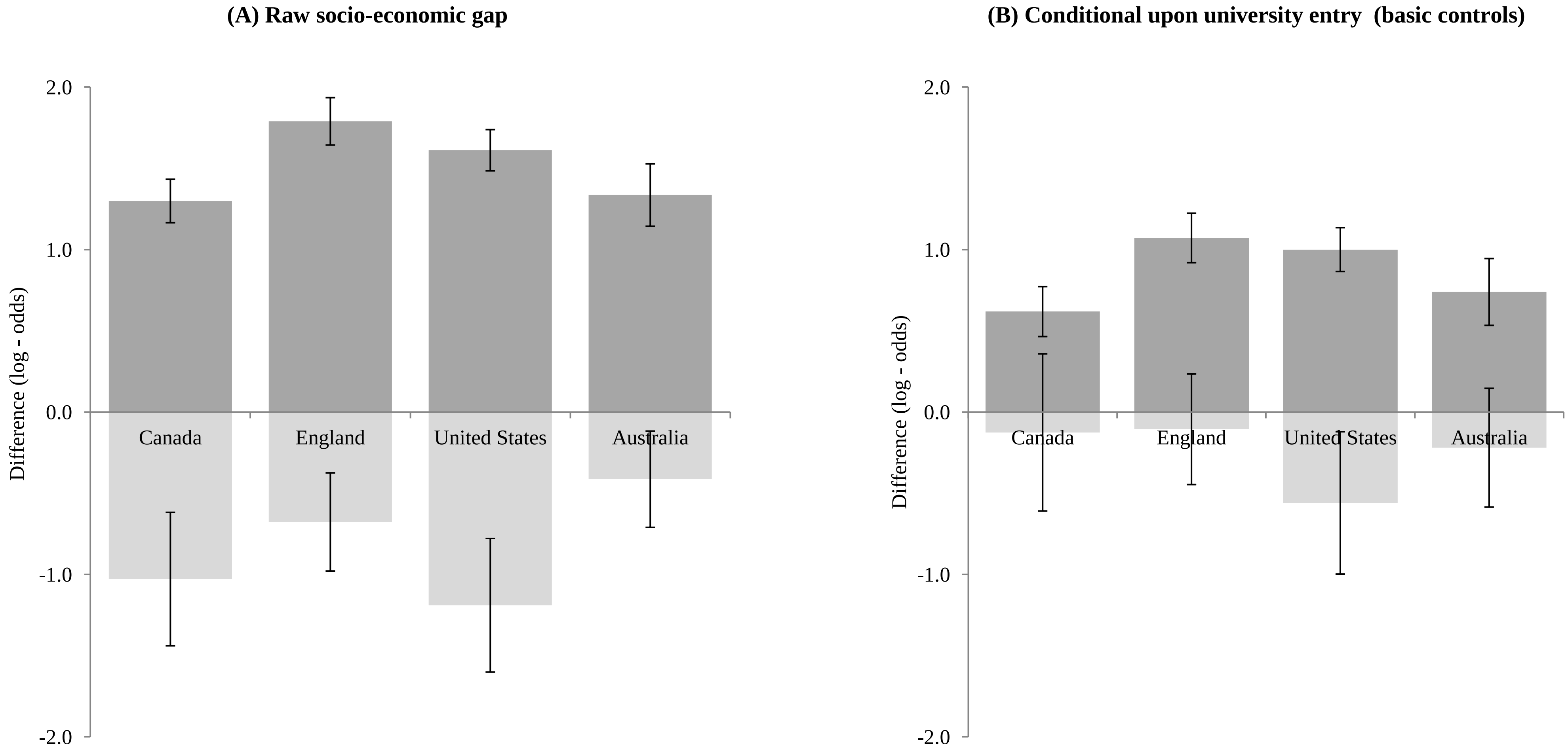
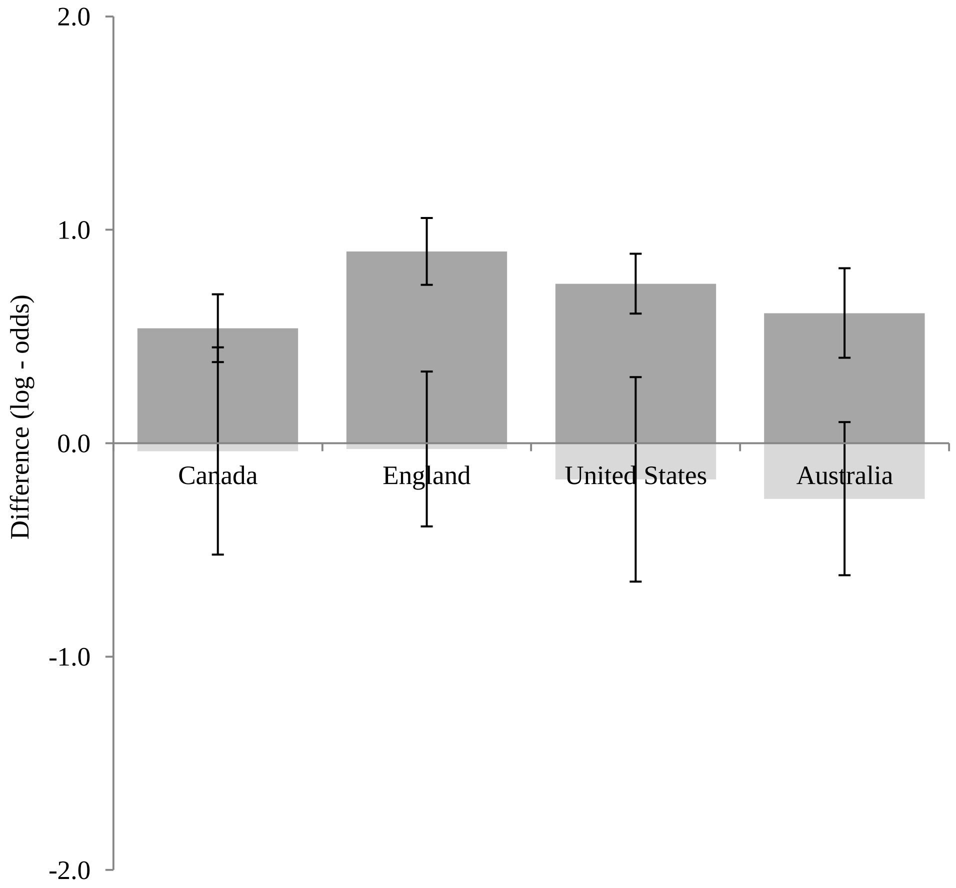


Figure 3. The socio – economic gap in entry to a *selective* higher education institution

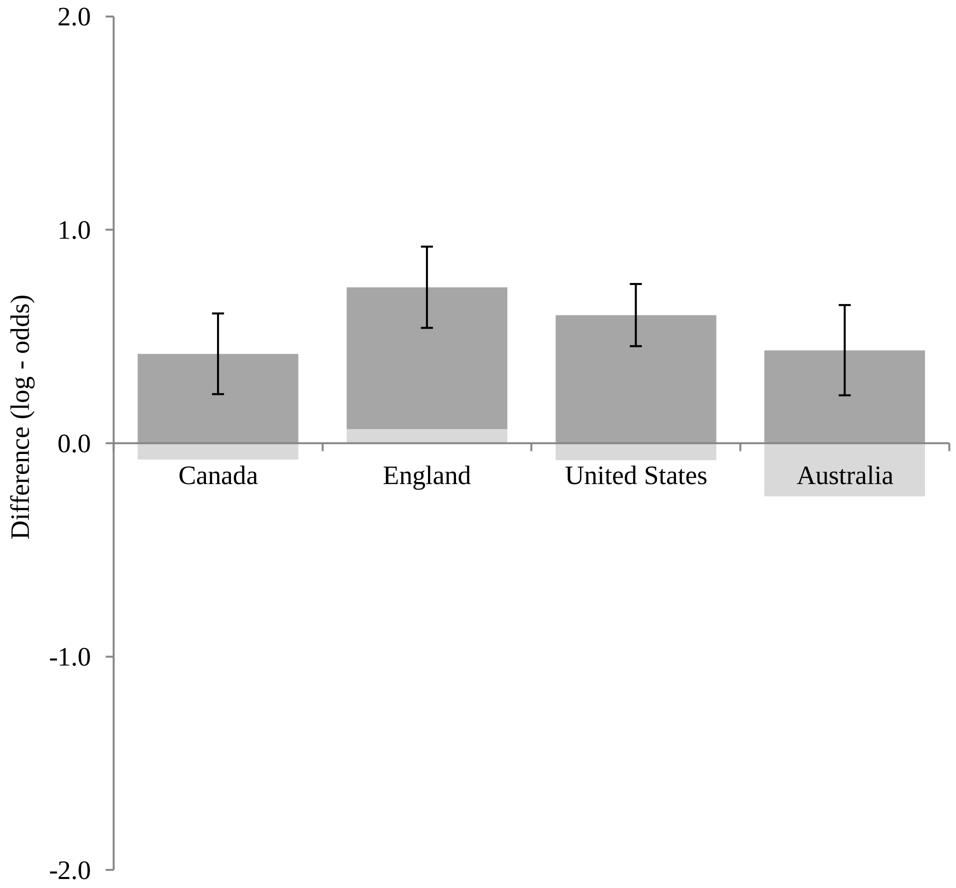


Notes: Figures for England refer to state school pupils only. The light grey segment of the bars illustrates the difference between ISCED 0 – 2 and ISCED 3 – 5B groups. Dark grey segments refer to the difference between ISCED 3 – 5B and ISCED 5A /6 groups. Thin black lines running through the centre are the estimated 90% confidence intervals. Estimates in Panel A are based upon the full sample and includes only basic controls (gender and language spoken at home). In panel B the datasets have been restricted to university graduates only. PISA test scores are then controlled for in panel C, with achievement scores at age 18 also included in panel D.

(C) PISA test scores



(D) School grades



Appendix A. Table 1. Key economic statistics across Anglophone countries

	Source	US	England	Canada	Australia
Economy					
GDP per capita ($US 000)	OECD	46.5	35.8	39.1	40.8
Population size (million)	OECD	307	53	34.1	22.3
% of 20 – 24 year olds NEET	OECD (EAG)	16	16	14	12
Poverty, inequality and social mobility					
Intergenerational income elasticity	Blanden (2011)	0.41	0.37	0.23	0.25
Income inequality (Gini coefficient)	OECD	0.38	0.34	0.32	0.30
% of children living in poverty	OECD	22	13	15	14
Educational achievement (PISA test scores)					
PISA reading rank in 2009	PISA 2009	17th	25th	6th	9th
Mean PISA reading test score in 2009	PISA 2009	500	494	524	515
Standard deviation of PISA reading test score	PISA 2009	97	95	90	99
SES gap in reading ability at 15 (years of schooling)	Jerrim (2012)	2.6	2.3	1.7	2.3
% private school children	Author	4.0	7.0	7.0	17.0

Notes:

1 Figures are taken from various sources. EAG stands for Education at a Glance.
2 Countries with a high figure for the intergeneration income elasticity are the least socially mobile.
3 Tuition costs have been converted into US dollars by the OECD using purchasing power parity.

Appendix B. Prediction of PISA test scores for the LSYPE sample

In this Appendix we provide further discussion of the test score data for England. This will focus upon how we estimate English children's PISA maths and reading scores from their performance on national exams.

In contrast to children in the other three countries, LSYPE sample members have not sat a PISA-style cognitive assessment. Rather information is available on their performance on national exams taken at age 14 ('Key Stage 3') and age 16 ('Key Stage 4' / 'GCSE'). The tests children take in England at age 14 examine their academic ability in three areas: English, maths and science (note the similarity to the PISA domains). The average mark children achieve across these three subjects is known as their Key Stage 3 average score. This information is contained within the LSYPE dataset for all children who attended a state school (roughly 93% of the English school population)[24]. National exams taken at age 16 typically involve assessment in around 10 subjects and lead to nationally recognised qualifications (the General Certificate of Secondary Education or 'GCSE'). This information has also been linked into the LSYPE from administrative records (for both private and state school pupils). With regards to the GCSE information, we focus upon five summary measures:

(a) A dummy variable indicating whether the child achieved at least five 'good' grades (A*-C) across all the GCSE exams they sat (this is a common performance threshold used in England).
(b) The total number of 'good' GCSE grades (A*-C) the child achieved
(c) The total points scored across all GCSE subjects taken (key stage 4 total points score)
(d) Capped GCSE points score (the marks the child achieved in their best eight exams)
(e) Average GCSE points score (total points scored divided by the number of subjects the child sat exams in).

Two recent studies (Micklewright and Schnepf 2006 and Micklewright et al 2012) have used a restricted version of the PISA 2000 and 2003 England datasets (which have been linked to population level administrative records) to investigate the relationship between children's age 14/16 exam performance and their PISA test scores. Crucially for our purposes, the authors report an estimated correlation between children's PISA maths test

[24] Information on private school pupils is not held in the administrative records and is this unavailable for roughly 7% of children who attend such institutions.

performance and average key stage 3 test scores of around 0.80. This implies a high degree of consistency between the key stage 3 test scores and PISA test scores. Micklewright and co-authors then estimate a series of regression models with the aim of predicting children's PISA test scores from their performance on national exams (their aim in doing so was to create a series of response weights in order to adjust for non-response). In Micklewright and Schnepf (2006) two bivariate models are estimated, where children's PISA scores were simply regressed upon average key stage 3 points scores (page 95, Table 5.6a). Note that, even in this very simple model, roughly two-thirds of the variance in PISA test scores can be explained ($R^2 \approx 0.65$).

We use the results from Micklewright and Schnepf (2006) to inform our first prediction of LSYPE sample members PISA test scores. This is only possible, however, for state school pupils. The following prediction equation is used:

$$\widehat{\mathrm{PISA}} = \hat{\alpha} + \widehat{\beta.}\mathrm{KS3} + \mathrm{error}$$

Where, in the case of reading:

$\hat{\alpha}$ = 118.05, $\hat{\beta}$ = 11.07, KS3 = children's average key stage 3 test score ($R^2 = 0.65$)

And from maths:

$\hat{\alpha}$ = 102.3, $\hat{\beta}$ = 11.54 , KS3 = children's average key stage 3 test score ($R^2 = 0.69$)

We assume that the error term (the difference between PISA and KS3 test scores) is normally distributed with a mean on zero and standard deviation σ:

error ~ N (0 , σ).

To incorporate this error into children's predicted PISA test scores, we take a random draw from a simulated normal distribution. The mean of this simulated distribution is zero, while we set σ so that the standard deviation of our predicted PISA test score approximates that for the actual PISA 2003 cohort in England (around 95 test points).

We also generate a second set of predicted PISA test scores for state school children in England. We follow a similar process to that described above, with the main difference being that the underlying prediction model is richer (in terms of the number of variables used). Specifically we now turn to the regression model presented in Table 9 of Micklewright et al (2012), which includes children's GCSE performance (along with some other variables) as additional covariates. Note again the high R^2 of around 0.7 achieved in their model. Our new prediction equations are as follows:

$$\widehat{\text{PISA read}} = \hat{\alpha} + \widehat{\beta_1}.\text{Male} + \widehat{\beta_2}.\text{KS3} + \widehat{\beta_3}.\text{KS3}_{\text{Miss}} + \widehat{\beta_4}.\text{KS4}_{\text{GG}} + \widehat{\beta_5}.\text{KS4}_{\text{APS}} + \widehat{\beta_6}.\text{KS4}_{\text{CPS}} + \widehat{\beta_7}.\text{KS4}_{\text{TPS}} + \widehat{\beta_8}.\text{FSM}_{\text{Miss}} + \text{error}$$

Where:

PISA read = PISA reading test score

Male = A binary indicator of whether the child was male (coded as 1) or female (coded as 0)

KS3 = Children's average key stage 3 score

KS3_Miss = A binary indicator of whether children's key stage 3 score was missing

$KS4_{GG}$ = A binary indicator of whether the child achieved 5 good GCSE's (1=yes , 0 = no)

$KS4_{APS}$ = Children's key stage 4 average points score (total points / number of entries)

$KS4_{CPS}$ = Children's capped key stage 4 points score (best eight grades)

$KS4_{TPS}$ = Children's total key stage 4 points score

FSM_{MISS} = A binary indicator of whether the Free School Meals variable (a state benefit in England) was missing

and $R^2 = 0.68$

$$\widehat{\text{PISA math}} = \hat{\alpha} + \widehat{\gamma_1}.\text{Male} + \widehat{\gamma_2}.\text{KS3} + \widehat{\gamma_3}.\text{KS3}_{\text{Miss}} + \widehat{\gamma_4}.\text{KS4}_{\text{NGG}} + \widehat{\gamma_5}.\text{KS4}_{\text{APS}} + \widehat{\gamma_6}.\text{KS4}_{\text{CPS}} + \widehat{\gamma_7}.\text{KS4}_{\text{TPS}} + \widehat{\gamma_8}.\text{Private\_Sch} + \text{error}$$

Where:

$KS4_{NGG}$ = The number of good grades (A*-C) the child achieved in their key stage 4 exams

Private_Sch = Whether the child attended a private school

$R^2 = 0.70$

All other variables are defined as above.

Predictions are calculated as before. Estimates of the α and β parameters come directly from Table 9 of Micklewright et al (2012), with covariate values plugged in for each of the LSYPE sample members that attended state school. The error term is once more assumed to be random noise, and is incorporated into our predictions by taking a random draw from a simulated normal distribution.

A comparison between (predicted) LSYPE and actual PISA (2006) test scores

Clearly the approach we have taken requires a strong correlation between an individual's key stage 3 test scores and their PISA test scores, as was observed by Micklewright and Schnepf (2006) and Micklewright et al (2012). We argue in the main body of the paper that, given the high correlation between the two sets of test scores as reported in those papers, our strategy is reasonable.

We now discuss the results of our predictive model. Specifically we examine whether the distribution of the predicted PISA test scores for the LSYPE sample is consistent with those observed for English children who have actually taken the PISA exam. With the predictive model we have estimated, and the assumptions we have made, unsurprisingly our predicted PISA test scores are distributed similarly to actual PISA test scores. Specifically, we compare the predicted LSYPE PISA test score distribution to the actual PISA 2006 England test score distribution. We have chosen to compare our prediction to the 2006 wave for two reasons. Firstly, information from PISA 2006 has played no role in our prediction of test scores for the LSYPE sample members. This is important as we wish to validate our predictions against a completely external source[25]. Secondly, this particular wave of PISA covers children of approximately the same age as those in the LSYPE cohort (PISA 2006 children were born between September 1990 and August 1991, while LSYPE children were born between September 1989 and August 1990).

Kernel density estimates of the actual PISA 2006 reading test scores for England (solid black line) and our predictions for the LSYPE sample (the dashed red line refers to our first method and the blue dotted line the second) can be found in Appendix Figure 1. All results refer to estimates with the LSYPE or PISA 2006 survey weights applied.

Appendix Figure 1

One can see that the three distributions largely overlap, suggesting that our predicted PISA test scores for the LSYPE cohort are reasonably consistent with those of children who actually sat the PISA 2006 exam. Moreover, we find that the difference in LSYPE (predicted)

[25] This thus rules out the use of the PISA 2003 sample, as this was the source used by Micklewright et al (which we in turn use in our prediction model).

and PISA 2006 (actual) average test scores is rather small. Specifically, the mean reading score for PISA 2006 pupils is 495 compared to a predicted mean for the LSYPE cohort of 492.

In Appendix Figure 2 we turn to gender differences in reading test scores. Again, predicted LSYPE and actual PISA 2006 test score distributions largely overlap for both males and females. Meanwhile average predicted reading test scores for boys in the LSYPE stands at 479 compared to 481 for boys who actually sat the PISA 2006 assessment. This difference of just two points is both small and statistically insignificant at conventional thresholds. Similar results hold for girls.

Appendix Figure 2

Finally, in Appendix Figure 3 we compare the LSYPE (predicted) and PISA 2006 (actual) test score distributions for children with different levels of parental education. Some caution is required when interpreting these results as the information on mother's and father's education has been collected in different ways across the two studies. Specifically, in PISA 2006 parental education is based upon child reports while in the LSYPE this information is drawn directly from parents. If children's reports are prone to measurement error, then this could also cause the actual PISA 2006 reading test score distribution (for a given level of parental education) to differ from the predicted LSYPE PISA test score distribution.

Appendix Figure 3

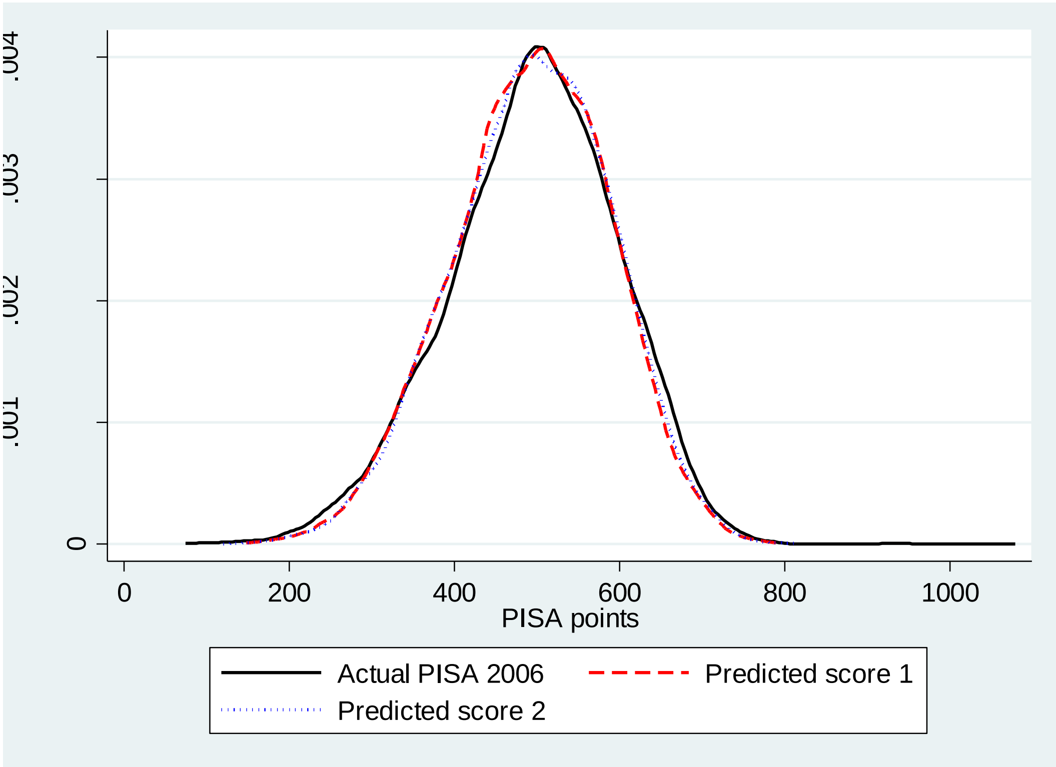
Panel (a) and (b) illustrate our findings for the ISCED level 0 to 2 (less than secondary school) and ISCED level 3 (high school) groups. In both cases there is still a great deal of consistency between the LSYPE (predicted) and PISA 2006 (actual) test scores. Note, in particular, that the solid black, dashed red and dotted blue lines all continue to overlap (to a great extent). Moreover, for the ISCED level 0 – 2 group we find that the means differ by just two points, with the spread being only slightly greater for the LSYPE predicted values than for the actual scores of the PISA 2006 cohort (a standard deviation of 93 compared to 91). Average predicted and actual test scores for ISCED level 3 children are also similar (there is a difference in the means of just three test points) with the spread being almost identical (the standard deviation is 93 points for both).

Turning to panel C, one can see that a similar story holds for the ISCED level 4/5B comparison – the three distributions overlap, with the measures of average and spread being similar for the LSYPE (predicted) and PISA 2006 (actual) reading test scores[26]. There is, however, a slightly more noticeable difference when one turns to children from the most advantaged homes (ISCED level 5A and above) presented in panel d. Firstly, the median predicted test score for the LSYPE sample stands at 550 compared to 541 in the actual PISA 2006 sweep (a difference of nine points). The analogous figures for the mean are 541 and 530 points (a difference of eleven points). It is important to note, however, that these differences are not statistically significant at the 5%, and thus one cannot reject the null hypothesis that this is simply a reflection of sampling variation. Perhaps the more noticeable difference is, however, in the spread of the data. Notice how the LSYPE (predicted) reading test score distribution for the ISCED 5A group tends to be narrower and more clustered around the centre than that for the (actual) PISA 2006 cohort. This is reflected by the standard deviation being 94 in the former and 104 in the latter.

This should not, however, detract from the general message of this Appendix. Based on an underlying strong correlation between key stage 3 and PIS test scores (Micklewright and Schnepf 2006; Micklewright et al 2012), our model predicts PISA scores for the LSYPE sample that are generally in-line with those of the actual PISA 2006 cohort.

[26] There is again no evidence of a statistically significant difference in average test scores (at any of the conventional thresholds) between the two.

Appendix Figure 1. Distribution of predicted PISA reading test scores for the LSYPE sample compared to the distribution of actual reading test scores for the PISA 2006 wave

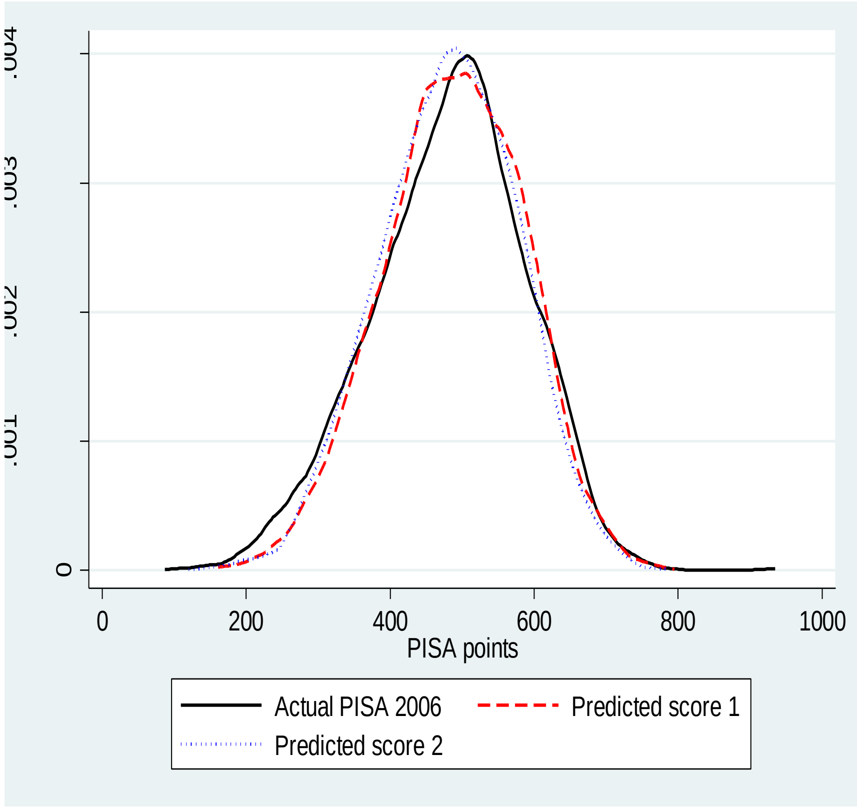


Notes:

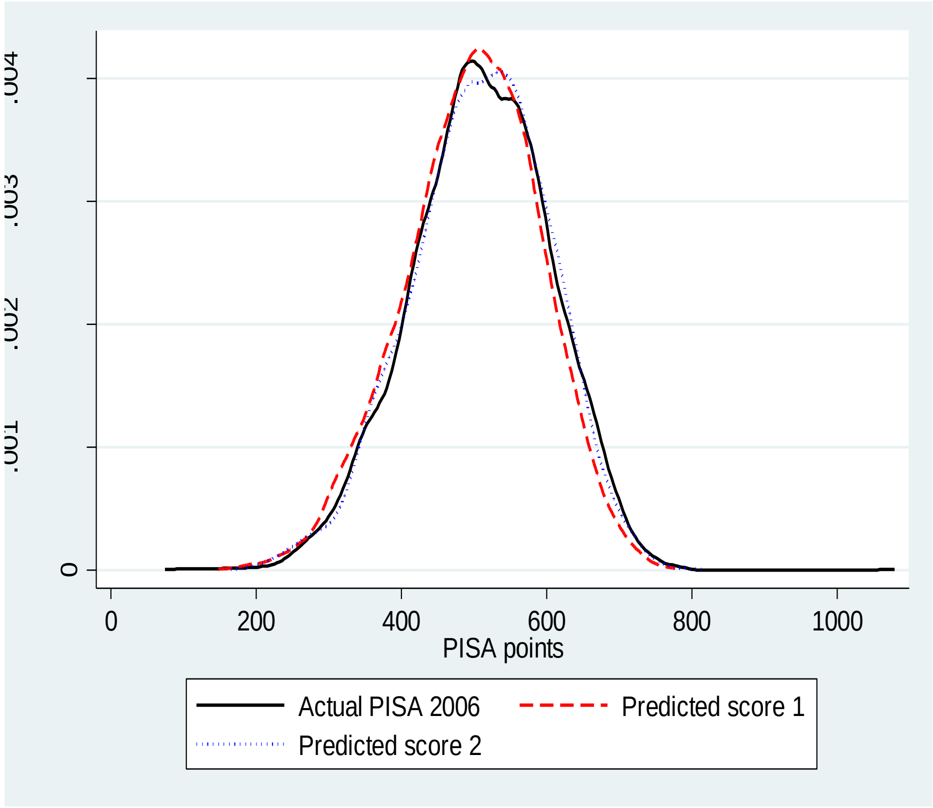
The solid black solid line refers to the actual PISA 2006 reading test score distribution. The dashed red line is the estimated PISA reading score for the LSYPE sample using our first prediction method (based solely upon their Key Stage 3 exams average points score). The dotted blue line is the estimated PISA reading score for the LSYPE sample using our second prediction method (based upon children's Key Stage 3 scores, Key Stage 4 scores and other auxiliary information).

Appendix Figure 2. Distribution of predicted PISA reading test scores for the LSYPE sample compared to the distribution of actual reading test scores for the PISA 2006 wave – Males and Females

Males

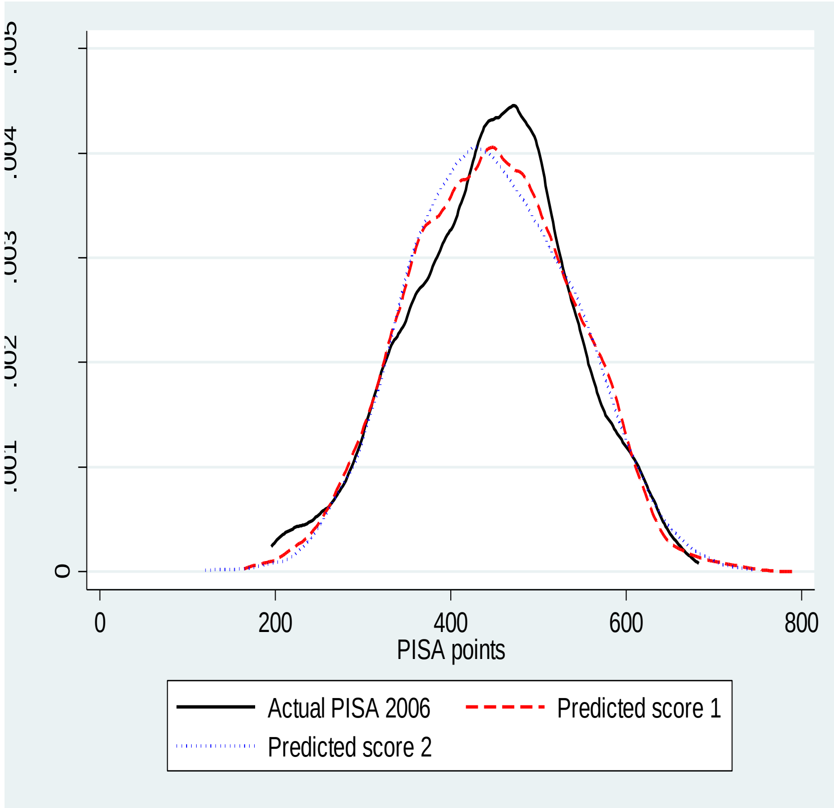


Females

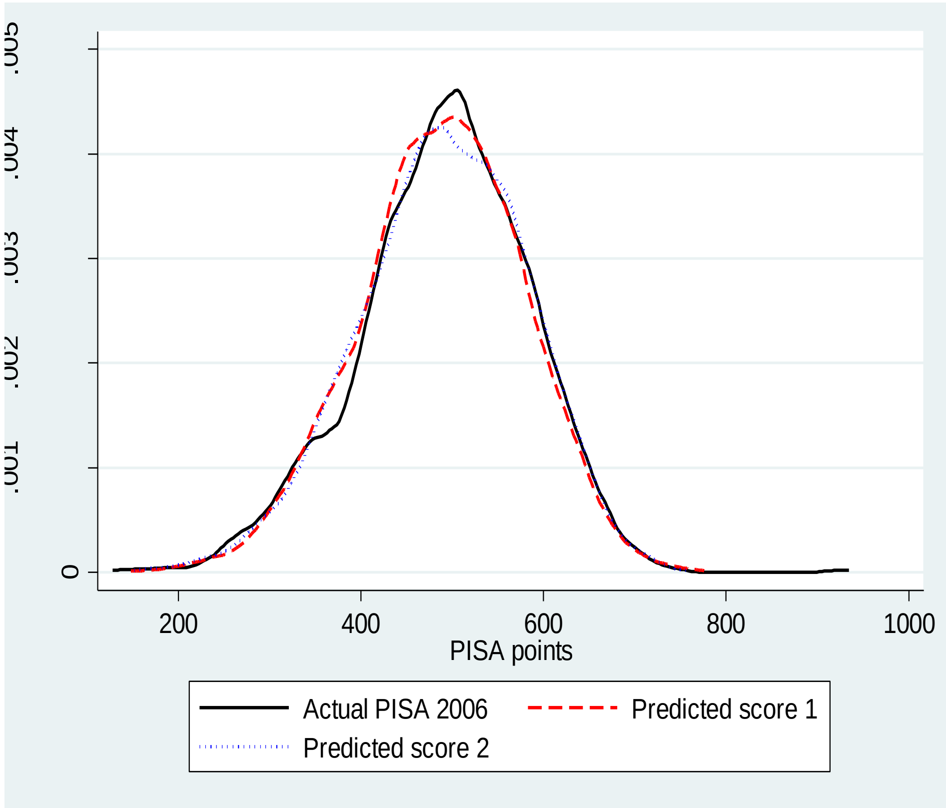


Appendix Figure 3. Distribution of predicted PISA reading test scores for the LSYPE sample compared to the distribution of actual reading test scores for the PISA 2006 wave – by parental education level

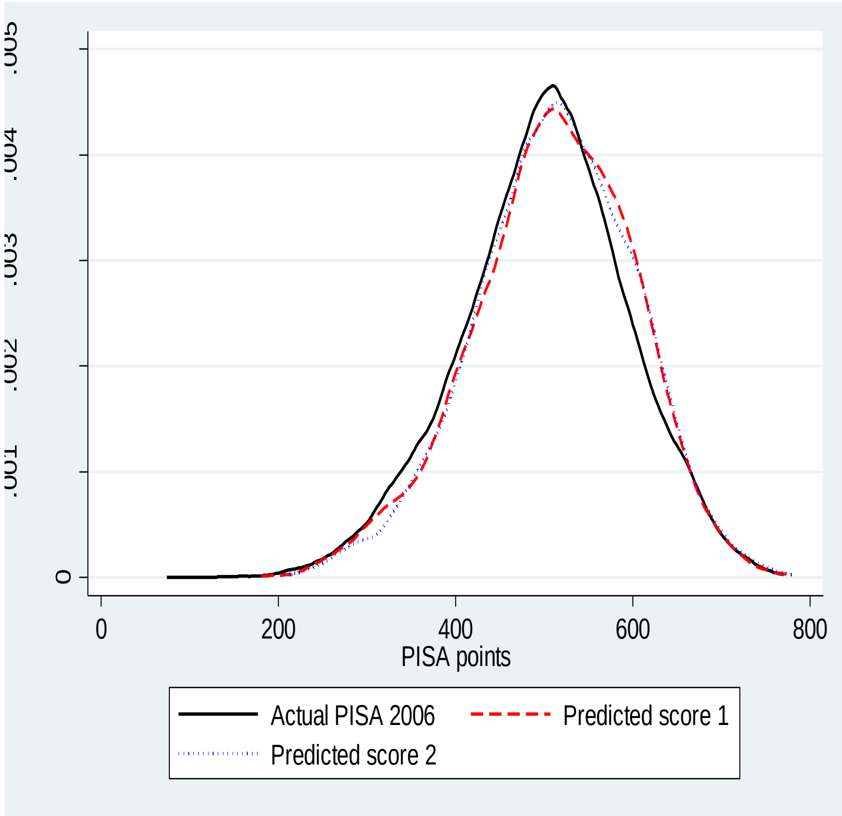
(a) ISCED level 0 - 2



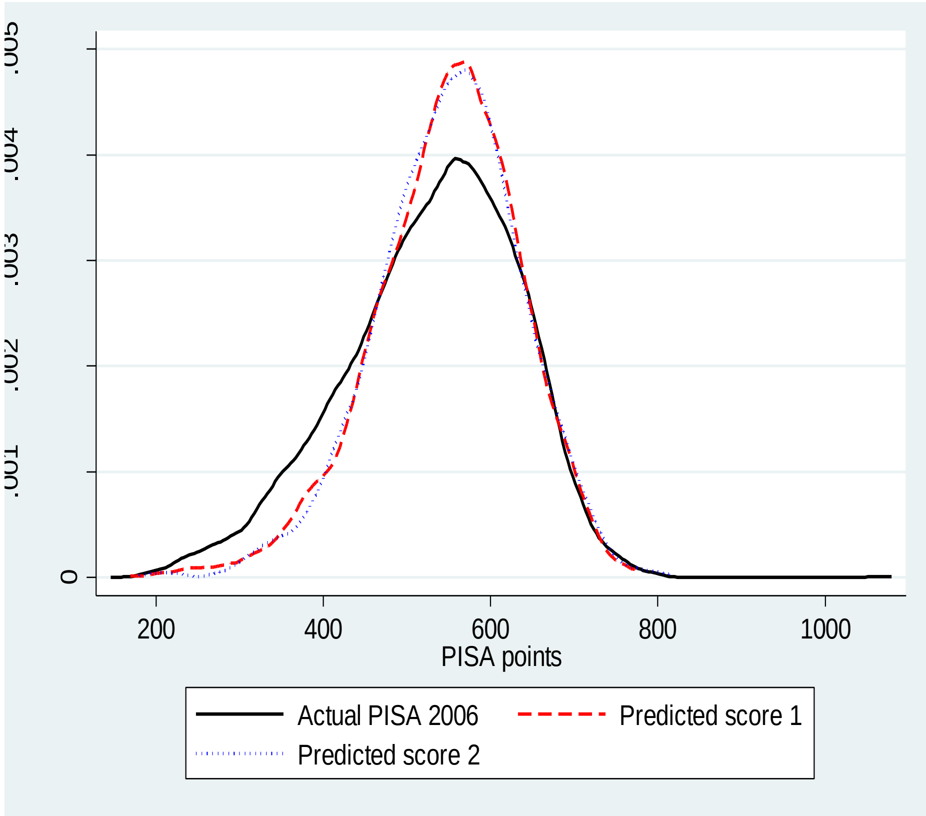
(b) ISCED level 3



© ISCED 4 / 5B



(d) ISCED level 5A/6



Appendix C. Parameter estimates

Appendix Table 1a. Parameter estimates for bachelor's degree models (specifications 1 and 2)

		Australia		Canada		England		United States	
		Beta	**SE**	**Beta**	**SE**	**Beta**	**SE**	**Beta**	**SE**
Specification 1 (Unconditional)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.363	0.129	-1.163	0.147	-0.851	0.097	-0.869	0.126
	ISCED level 5A +	1.143	0.082	1.370	0.076	1.507	0.077	1.325	0.051
	Gender (Ref: Male)								
	Female	0.611	0.084	0.780	0.065	0.367	0.061	0.329	0.046
	Language at home (Ref: Native)								
	Non-native	0.888	0.132	-0.009	0.174	1.284	0.102	-1.112	0.160
	Constant	-1.253	0.083	-1.035	0.057	-1.057	0.056	-0.818	0.047
	n	**6536**		**9446**		**7715**		**12575**	
Specification 2 (PISA test scores controlled)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.346	0.133	-0.855	0.158	-0.304	0.107	-0.409	0.137
	ISCED level 5A +	0.753	0.085	1.102	0.082	0.958	0.084	1.001	0.053
	Gender (Ref: Male)								
	Female	0.665	0.088	0.642	0.073	0.366	0.065	0.433	0.053
	Language at home (Ref: Native)								
	Non-native	1.425	0.140	0.295	0.209	1.703	0.119	-0.362	0.179
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	0.423	0.235	0.371	0.161	0.734	0.134	0.467	0.106
	Third Quintile	0.890	0.227	0.368	0.163	1.107	0.135	0.991	0.102
	Fourth Quintile	1.408	0.229	0.444	0.179	1.673	0.132	1.482	0.111
	Top Quintile	2.003	0.237	0.815	0.184	2.278	0.143	1.987	0.130
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	0.897	0.215	0.768	0.173	0.981	0.138	0.395	0.098
	Third Quintile	1.035	0.187	1.523	0.184	1.397	0.134	0.660	0.104
	Fourth Quintile	1.462	0.221	1.945	0.190	1.857	0.134	0.847	0.110
	Top Quintile	1.694	0.241	2.603	0.201	2.405	0.138	1.316	0.124
	Constant	-3.313	0.176	-2.958	0.155	-3.826	0.150	-2.488	0.107
	n	**6536**		**9446**		**7715**		**12575**	

Appendix Table 1b. Parameter estimates for bachelor's degree models (specifications 3 and 4)

		Australia		Canada		England		United States	
		Beta	**SE**	**Beta**	**SE**	**Beta**	**SE**	**Beta**	**SE**
Specification 3 (School fixed effect)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.348	0.136	-0.774	0.109	-0.411	0.123	-0.396	0.152
	ISCED level 5A +	0.624	0.089	0.923	0.062	0.913	0.091	0.925	0.065
	Gender (Ref: Male)								
	Female	0.716	0.089	0.464	0.052	0.429	0.073	0.529	0.064
	Language at home (Ref: Native)								
	Non-native	1.317	0.191	0.164	0.161	1.631	0.152	-0.538	0.215
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	0.481	0.213	-	-	0.816	0.150	0.497	0.121
	Third Quintile	1.062	0.203	-	-	1.188	0.149	1.126	0.120
	Fourth Quintile	1.573	0.208	-	-	1.779	0.149	1.661	0.129
	Top Quintile	2.199	0.222	-	-	2.380	0.159	2.278	0.158
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	0.919	0.211	0.928	0.101	1.066	0.150	0.466	0.114
	Third Quintile	0.987	0.206	1.799	0.100	1.439	0.147	0.719	0.116
	Fourth Quintile	1.405	0.213	2.279	0.101	1.957	0.149	0.938	0.123
	Top Quintile	1.587	0.233	3.146	0.108	2.534	0.157	1.450	0.140
	School fixed effect	Yes		Yes		Yes		Yes	
	n	**6473**		**10034**		**7439**		**11974**	
Specification 4 (School grades controlled)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.246	0.150	-0.759	0.211	0.045	0.154	-0.028	0.160
	ISCED level 5A +	0.477	0.092	1.054	0.108	0.460	0.119	0.760	0.061
	Gender (Ref: Male)								
	Female	0.605	0.100	0.409	0.094	0.018	0.093	0.100	0.066
	Language at home (Ref: Native)								
	Non-native	1.021	0.171	0.114	0.263	1.113	0.161	-0.162	0.195
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	0.168	0.259	0.210	0.183	0.114	0.209	-0.100	0.145
	Third Quintile	0.434	0.253	0.242	0.177	0.222	0.202	-0.190	0.157
	Fourth Quintile	0.629	0.252	0.281	0.200	0.307	0.210	-0.178	0.169
	Top Quintile	0.968	0.263	0.424	0.211	0.246	0.223	-0.292	0.198
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	0.459	0.227	0.629	0.203	0.258	0.232	0.238	0.116
	Third Quintile	0.546	0.205	1.232	0.207	0.211	0.225	0.366	0.118
	Fourth Quintile	0.781	0.229	1.527	0.219	0.285	0.234	0.273	0.132
	Top Quintile	0.863	0.249	1.965	0.228	0.423	0.238	0.438	0.147
	Grades	Yes		Yes		Yes		Yes	
	Constant	-1.793	0.225	-0.679	0.261	13.515	0.947	-5.921	1.273
	n	**5541**		**7196**		**4248**		**11570**	

Notes: Parameter estimates refer to log-odds ('logits'). Sample sizes are slightly reduced in the school fixed effect model (compared to in specification 1 and 2) due to perfect multi-collinearity. The number of observations falls in the 'school grades' estimates as the samples are being restricted to only those young people who complete education up to age 18 (e.g. high school graduates in the US) and who are thus eligible to complete university.

Appendix Table 2a. Parameter estimates for selective institution models (specifications 1 and 2)

		Australia		Canada		England		United States	
		Beta	**SE**	**Beta**	**SE**	**Beta**	**SE**	**Beta**	**SE**
Specification 1 (Unconditional)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.41	0.18	-1.03	0.25	-0.68	0.18	-1.28	0.25
	ISCED level 5A +	1.34	0.12	1.30	0.08	1.79	0.09	1.61	0.08
	Gender (Ref: Male)								
	Female	0.17	0.13	0.45	0.09	0.14	0.11	0.20	0.06
	Language at home (Ref: Native)								
	Non-native	0.81	0.16	0.37	0.21	0.41	0.15	0.01	0.11
	Constant	-2.83	0.14	-2.22	0.08	-2.98	0.10	-2.82	0.12
	n	**6472**		**9446**		**7715**		**12575**	
Specification 2 (Conditional on university entry)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.22	0.22	-0.13	0.29	-0.11	0.21	-0.77	0.27
	ISCED level 5A +	0.74	0.12	0.62	0.09	1.07	0.09	1.01	0.08
	Gender (Ref: Male)								
	Female	-0.23	0.12	-0.04	0.10	-0.08	0.11	0.02	0.07
	Language at home (Ref: Native)								
	Non-native	0.46	0.15	0.50	0.24	-0.28	0.17	-0.34	0.12
	Constant	-1.20	0.14	-0.60	0.09	-1.52	0.10	-1.19	0.12
	n	**2905**		**4539**		**3426**		**6384**	

Appendix Table 2b. Parameter estimates for selective institution models (specifications 3 and 4)

		Australia		Canada		England		United States	
		Beta	SE	Beta	SE	Beta	SE	Beta	SE
Specification 3 (PISA test scores)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.26	0.22	-0.04	0.29	0.02	0.22	-0.17	0.30
	ISCED level 5A +	0.61	0.13	0.54	0.10	0.85	0.09	0.75	0.09
	Gender (Ref: Male)								
	Female	-0.13	0.12	-0.09	0.10	-0.06	0.11	0.15	0.08
	Language at home (Ref: Native)								
	Non-native	0.65	0.14	0.65	0.26	0.04	0.17	-0.03	0.27
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	1.26	0.46	-	-	0.38	0.53	0.45	0.39
	Third Quintile	1.30	0.47	-	-	0.71	0.50	1.21	0.38
	Fourth Quintile	1.60	0.49	-	-	1.23	0.48	1.64	0.38
	Top Quintile	2.08	0.49	-	-	1.69	0.48	2.32	0.38
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	-0.17	0.39	0.04	0.36	0.28	0.48	-0.15	0.30
	Third Quintile	-0.36	0.39	0.02	0.32	0.54	0.46	-0.03	0.31
	Fourth Quintile	0.06	0.38	0.37	0.31	0.90	0.48	0.40	0.28
	Top Quintile	0.16	0.39	0.86	0.31	1.35	0.47	0.75	0.29
	Constant	-2.86	0.50	-1.02	0.31	-3.74	0.78	-3.52	0.37
	n	**2905**		**4539**		**3426**		**6384**	
Specification 4 (School grades controlled)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	-0.25	0.25	-0.08	0.42	0.07	0.24	-0.08	0.32
	ISCED level 5A +	0.44	0.13	0.42	0.11	0.67	0.10	0.61	0.09
	Gender (Ref: Male)								
	Female	-0.26	0.11	-0.11	0.13	-0.10	0.11	0.19	0.08
	Language at home (Ref: Native)								
	Non-native	0.62	0.15	0.44	0.30	-0.03	0.19	0.04	0.29
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	1.21	0.47	-	-	0.52	0.54	0.03	0.51
	Third Quintile	1.16	0.49	-	-	0.78	0.52	0.47	0.51
	Fourth Quintile	1.13	0.51	-	-	1.10	0.50	0.51	0.52
	Top Quintile	1.40	0.50	-	-	1.27	0.50	0.67	0.53
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	-0.27	0.41	-0.01	0.40	0.31	0.49	-0.13	0.29
	Third Quintile	-0.46	0.39	-0.04	0.36	0.48	0.47	-0.15	0.31
	Fourth Quintile	-0.16	0.38	-0.01	0.34	0.77	0.50	0.09	0.28
	Top Quintile	-0.14	0.40	0.48	0.35	0.94	0.48	0.08	0.30
	Grades	Yes		Yes		Yes		Yes	
	Constant	-2.14	0.51	-0.03	0.38	-4.13	0.83	-4.33	0.39
	n	**2871**		**3555**		**3426**		**6383**	

Notes: Parameter estimates refer to log-odds (‘logits’).

Appendix Table 3. Parameter estimates STEM qualifications models

		Australia		Canada		England		United States	
		Beta	**SE**	**Beta**	**SE**	**Beta**	**SE**	**Beta**	**SE**
Specification 1 (Unconditional)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	0.292	0.262	-0.744	0.336	0.195	0.164	-0.057	0.273
	ISCED level 5A +	0.008	0.116	0.176	0.106	0.373	0.172	0.107	0.079
	Gender (Ref: Male)								
	Female	-0.699	0.110	-0.879	0.103	-0.429	0.081	-0.136	0.075
	Language at home (Ref: Native)								
	Non-native	0.321	0.139	0.441	0.290	0.274	0.138	0.174	0.286
	Constant	-0.222	0.103	-0.705	0.099	-0.565	0.167	-1.235	0.074
	n	**2905**		**4188**		**3412**		**6571**	
Specification 2 (PISA test scores controlled)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	0.259	0.241	-0.561	0.348	0.144	0.169	0.016	0.280
	ISCED level 5A +	-0.090	0.119	0.082	0.107	0.204	0.180	0.057	0.080
	Gender (Ref: Male)								
	Female	-0.567	0.109	-0.817	0.112	-0.413	0.082	-0.085	0.075
	Language at home (Ref: Native)								
	Non-native	0.411	0.151	0.621	0.294	0.436	0.143	0.162	0.290
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	0.421	0.332	-0.235	0.360	0.016	0.260	-0.099	0.205
	Third Quintile	0.655	0.335	-0.034	0.365	-0.016	0.253	0.107	0.206
	Fourth Quintile	0.964	0.373	0.342	0.381	0.417	0.245	0.145	0.211
	Top Quintile	1.329	0.345	0.626	0.387	0.550	0.248	0.611	0.207
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	-0.212	0.327	-0.144	0.408	0.301	0.294	0.147	0.203
	Third Quintile	-0.028	0.318	0.159	0.400	0.449	0.280	0.174	0.204
	Fourth Quintile	-0.149	0.322	0.438	0.399	0.618	0.276	-0.076	0.212
	Top Quintile	-0.307	0.337	0.435	0.417	0.652	0.283	-0.199	0.213
	Constant	-1.009	0.318	-1.391	0.411	-1.397	0.359	-1.478	0.209
	n	**2905**		**4188**		**3412**		**6571**	
Specification 4 (School grades controlled)	**Parental Education (Ref: ISCED 3 - 5b)**								
	ISCED 0 - 2	0.284	0.244	-0.282	0.396	0.108	0.169	-0.003	0.279
	ISCED level 5A +	-0.111	0.120	0.116	0.127	0.131	0.181	0.050	0.081
	Gender (Ref: Male)								
	Female	-0.584	0.112	-0.745	0.137	-0.393	0.082	-0.089	0.075
	Language at home (Ref: Native)								
	Non-native	0.405	0.154	0.578	0.271	0.440	0.144	0.155	0.290
	Maths test score (Ref: Bottom Quintile)								
	Second Quintile	0.423	0.333	-0.243	0.409	-0.137	0.269	-0.099	0.210
	Third Quintile	0.658	0.334	-0.203	0.427	-0.249	0.264	0.107	0.208
	Fourth Quintile	0.936	0.370	0.073	0.431	0.137	0.256	0.141	0.213
	Top Quintile	1.262	0.343	0.182	0.440	0.226	0.261	0.607	0.210
	Read test score (Ref: Bottom Quintile)								
	Second Quintile	-0.197	0.327	-0.291	0.436	0.232	0.300	0.123	0.205
	Third Quintile	-0.003	0.321	0.277	0.432	0.322	0.294	0.130	0.206
	Fourth Quintile	-0.149	0.325	0.554	0.441	0.437	0.293	-0.115	0.213
	Top Quintile	-0.318	0.343	0.570	0.460	0.437	0.301	-0.235	0.215
	Grades	YES		YES		YES		YES	
	Constant	-0.958	0.348	-0.484	0.503	-1.397	0.359	-1.515	0.244
	n	**2871**		**3465**		**3412**		**6521**	