Department of Quantitative Social Science

Disadvantaged children's "low" educational expectations: Are the US and UK really so different to other industrialized nations?

John Jerrim

DoQSS Working Paper No. 11-04
June 2011

Leading education and social research
Institute of Education
University of London

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.

Disadvantaged children's "low" educational expectations: Are the US and UK really so different to other industrialized nations?

John Jerrim[*†]

Abstract. In most countries, children from disadvantaged backgrounds are under-represented amongst the undergraduate population. One explanation is that they do not see higher education as a realistic goal; that it is 'not for the likes of them'. In this paper, I use the Programme for International Assessment data to investigate whether 15 year olds from disadvantaged backgrounds are less likely to expect to complete university than their advantaged peers. I explore this issue across the OECD nations, though paying particular attention to the US and UK. My results suggest that children from less fortunate families are not as likely to make early plans for university as their affluent peers. Yet the extent to which these findings differ across countries is rather modest, with little evidence to suggest that the UK stands out from other members of the OECD. The US, on the other hand, appears to be a nation where the relationship between socio-economic background and the expectation of completing higher education is comparatively weak.

JEL classification: I21, I28, J62.

Keywords: Higher Education, University Access, Educational Expectations, PISA.

[*]Department of Quantitative Social Science, Institute of Education, University of London, 20 Bedford Way London, WC1H 0AL. E-mail: (J.Jerrim@ioe.ac.uk)

[†]Thanks is given to the ESRC for providing funding for this project. I would particularly like to thank John Micklewright, Patrick Sturgis, Sandra McNally and Sylke Schnepf for their helpful comments, and the hospitality of the Institute for Research on Poverty at the University of Wisconsin, Madison, on a visit while this paper was produced

"Children who grow up in inferior environments may expect less of themselves and may not fully develop their academic potential because they see little hope for ever being able to complete college or use their schooling in any effective way"

Cameron and Heckman (1999), Financing College Tuition: Government Policies & Educational Priorities, page 76-124

Educational attainment has risen dramatically across the developed world over the past 15 to 20 years, with particularly strong growth in university participation. Yet despite this rising trend, access to tertiary education remains unequal. Children with well-educated, affluent parents are still over-represented in higher education, with relatively limited opportunities for those from disadvantaged backgrounds. This issue has taken on particular prominence either side of the Atlantic, where worries mount over equality of educational opportunity and social mobility (see Blanden and Machin (2004), Machin and Vignoles (2004)). Consequently, British and American governments have introduced policies to increase the number of disadvantaged children entering higher education. Ensuring children hold high expectations and aim for university from an early age is seen as a crucial step towards reaching this goal. In other words, there is a belief that future educational plans made during adolescence have a significant impact on later academic attainment, and a concern that poor children's low expectations may be limiting their opportunities to succeed.

Initiatives to "raise expectations" have thus become a prominent feature of educational policy in these countries, including the "Gear-up" and "I have a dream" programmes in the US and the "Aim Higher" and "Gifted and Talented" schemes in the UK. It has also become a hot topic of academic debate, with an increasing number of studies by economists (Chowdry et al 2009, Chevalier et al 2009 and Emmerson et al 2005), sociologists (Reynolds and Pemberton 2001, Morgan 2005) and social psychologists (Schoon 2010, Patton and Creed 2007, Gottfredson 2002), all of which investigate adolescents' educational expectations within a single national setting (typically either the US or UK). I contribute to this literature by placing the link between family background and children's educational expectations (which is often the focus of such studies) into a comparative context. Specifically, this paper shall investigate whether the socio-economic gap in children's educational expectations is particularly 'big' in the UK and US compared to other developed nations, and whether this is greater than one would anticipate given these countries level of educational inequality.

I begin in section 2 by reviewing the relevant literature. This includes a discussion of how children's expectations differ to their aspirations, and how this concept is linked to their eventual educational attainment. Section 3 describes the Programme for International Student Assessment (PISA) data on 15 year old children that I analyse and my empirical methodology. During sections 4 and 5 I discuss the results. I conclude in section 6 with a discussion of how my findings may inform educational policy in countries that encourage disadvantaged children to 'aim higher', like the UK.

2. Existing literature and research questions

Friedman and Friedman (1980), amongst others, have argued that all young adults who can benefit from university should have access to the returns it offers, regardless of their family background. One reason is that this may lead to a more equitable society. Yet it is also important for economic efficiency. Labour is a scarce resource that needs to be allocated appropriately, but the brightest children may be excluded from the best jobs if they are unable to 'fully develop their academic potential'. However, as noted in the introduction, disadvantaged children tend to be under-represented amongst the undergraduate population and in the most prestigious jobs (see Sutton Trust 2009). As suggested by Cameron and Heckman (1999) at the start of this paper, some disadvantaged children may perceive there to be a lack of opportunity to complete higher education which stops them from applying or, in the words of Shields and Mohan (2008), believe that university is 'not for the likes of them'[1]. Consequently, policymakers in the US and UK have introduced a series of programmes to raise disadvantaged children's expectations of being able to obtain a tertiary level qualification.

At this point, it is important to draw a distinction between children's "expectations" and their "aspirations". The former implies a realistic assessment of future outcomes, while the latter reflects children's hopes and dreams (Gutman and Akerman 2008). So if a child *expects* to obtain a university qualification, they truly believe that they will go on to complete this level

[1] For instance a report by the Sutton Trust (2008), a UK based charity, states: 'exam grades on their own will not necessarily lead to university if young people do not have a high level of expectation and make ill-informed decisions'

of education. It is this concept that I attempt to explore in this paper. However, one must consider whether 15 year old children (the age group that I study) are able to make such realistic assessments of the future. Drawing from the developmental literature, Gottfredson (2002) notes that, around age 14, children are beginning to recognize the need for compromise in their educational and occupational goals. Likewise, Gutman and Akerman (2008) suggest that at this age young people 'relinquish their most preferred choices and settle for more acceptable, available choices', recognising the external constraints that they face. From a different prospective, Morgan (1998) finds that adolescents' educational expectations are not 'irrational fantasies'; rather, they are grounded in logical thinking, and vary with the marginal costs and benefits associated with such continued schooling. Hence there is evidence which suggests that young adults are able to distinguish between their aspirations and expectations. Accurately capturing such details in a social survey is, however, another matter. I shall further elaborate on this point when discussing the PISA data in the following section.

It is also important to make clear that the value of any scheme that attempts to 'raise disadvantaged children's expectations' is based on the assumption that this will have a causal influence on their later behaviour and attainment. A conceptual model to illustrate this relationship is set out in Figure 1, drawing upon the work of Chowdry et al (2009).

Figure 1 about here

This framework recognizes the multi-dimensional nature of family background, based around measures of parental education, occupation (socio-economic status) and income. The authors then specify four “transition mechanisms” (schools, neighbourhoods, parental attitudes, material resources) by which family background influences children’s attitudes, behaviour, beliefs (including their aspirations for the future) and outcomes at age 14. The main focus of this paper is, however, on the next stage of the model – the transition from adolescence to young adulthood (i.e. from age 14 to 18). During this period, children may change their attitudes, behaviours and beliefs about the future, which, in turn, alters their academic trajectory[2]. Based on the work of Gottfredson (2002), I propose that one key development between these ages is that children begin to recognize the external constraints that they face,

[2] This framework also recognises that family circumstances and parental characteristics will continue to play a role.

and thus start to develop *expectations* about their future (regarding, in particular, higher education). These expectations then become the key behavioral "transmission mechanism" that encourage greater effort and investment in school and less "risky" behaviour (drinking, drug use and early sexual activity) between ages 14 and 18 which, in turn, leads to higher educational attainment.

It is important to recognise, however, that this is not a static relationship; children will continually revise these expectations, based on their on-going attainment. Indeed, it is likely that higher expectations lead to higher attainment, which leads to continued high expectations, and so forth. Yet, as one can not identify the exact age at which such feedback begins, it has proven to be methodologically challenging to estimate the extent to which one factor is driving the other. Nevertheless, several authors have explored the association between these variables, with some attempt to address the direction of causality. For instance, Khoo and Ainley (2005) investigate the educational plans and achievement of a sample of Australian teenagers. Estimating a structural equation model, they show that children's expectations are strongly associated with their later outcomes, even after controlling for a host of potentially confounding factors. In a similar manner, Reynolds and Pemberton (2001) find that expecting to go to university at age 15 is almost a prequisite for actual later attendance in the US; they show that less than 3% of children who do not expect to go to university actually obtain a degree by the time they turn 30. Likewise, Morgan (2004) uses a regression based path analysis to investigate whether educational expectations held during the mid-teens determines entry into post-compulsory schooling in the US. In turn, he finds evidence of a strong and statistically significant association. Of course, economists may express concerns about the potential endogeneity of expectations in any regression based set-up, particularly due to omitted variable bias. Consequently, Morgan (2004) shows that expectations remain a highly significant predictor of later outcomes using an instrumental variable analysis. However he also recognizes the difficulties of identifying such models in this set-up. Brown et al (2004) use similar methods to Morgan, and find a strong and highly significant relationship between children's expectations and later attainment in the UK. Using panel data, with measurement of young adults' educational expectations at several ages, Morgan (2005) finds university plans are serially correlated across time. He suggests that this is consistent with an underlying dynamic causal relationship between expectations and attainment as described above. Similarly, Chowdry et al (2009) find that a number of disadvantaged children in England stop believing that they will enter university between ages

14 and 16, and that these teenagers subsequently make less academic progress than their peers who maintain high expectations. In a wider context, Cowan (2009) investigates the relationship between American teenagers' educational expectations and their chances of engaging in risky behaviour. Using an instrumental variable analysis, he finds that 'anticipated schooling has an effect on behaviour above and beyond the effect of realized schooling' and thus that raising children's expectations of completing university may prove to be an inexpensive way of reducing their tobacco, marijuana and alcohol consumption. Finally, one may have concerns about the negative consequences for young people if their expectations are not met. Reynolds and Baird (2010), however, find no long term emotional costs of "shooting for the stars".

Given the above, any difference between advantaged and disadvantaged children's educational expectations will lead to a division in their behaviour, attainment and eventual graduation rates. Indeed, the framework set out in Figure 1 suggests that such a divergence in beliefs may well occur; expectations are assumed to have six primary determinants (schools, neighbourhoods, parental attitudes, family resources, childhood attitudes and prior attainment) all of which are associated with family background. For instance, advantaged children will tend to go to better schools, where teachers may build their children's academic confidence and emphasise their ability to complete this level of study. Similarly, it may be that only well-educated parents stress the wider benefits of learning (meeting new people, broadening horizons, growing up) to their offspring, who become driven towards higher education as a result. Availability of resources will also determine children's expectations; those from less fortunate households may believe they are credit constrained and thus do not have the necessary finance to complete higher education. There may also be peer and role model influences, both in school and the wider community, where disadvantaged children do not see university as a realistic goal because they do not know any adult who has completed higher education and have few friends who believe they can achieve the same. Attitudes may also be transferred between generations, such as ambition and work ethic, which could influence children's educational plans via the extent they are willing to stretch themselves in the future. Finally, as expectations involve the recognition of external constraints, they will be tempered by children's pre-existing skill, with large socio-economic differences in academic achievement already evident at age 14 (see Hanushek and Woessmann (2010) for a survey of the international evidence).

The analysis I undertake in this paper is motivated by the theoretical framework and empirical analysis described above, which suggests that children's expectations have an important influence on their later academic attainment, and that there are likely to be large differences in these expectations between socio-economic groups. For instance, the work of Chowdry et al (2010) suggests that expectations regarding higher education differ substantially between advantaged and disadvantaged teenagers in the UK, and that this makes the biggest contribution to the widening of the socio-economic attainment gap towards the end of compulsory schooling. However, few have considered whether the difference in educational expectations between rich and poor is bigger in some countries than in others. Yet there are numerous reasons to suspect that this may be the case. For instance, the paragraph above described how family background is linked to children's expectations (e.g. through schools, resources, parental attitudes etc). Yet the strength of such associations are likely to differ across nations. For instance, the US and UK are known for having high levels of income inequality (see OECD 2008) and quite segregated schooling systems compared to other nations (see Jenkins et al 2006). Similarly, both have comparatively high costs of university tuition, which may mean disadvantaged British and American teenagers are more likely to feel credit constrained than their peers in other countries. Such factors may consequently lead to these countries experiencing particularly large gaps in educational expectations between socio-economic groups.

In my first research question I consider this issue, exploring the size of the socio-economic gap in children's educational expectations across the OECD, with a focus on results for the US and UK. Although concern has been expressed about the difference between advantaged and disadvantaged children's educational expectations in these countries, I have never seen it put into a comparative perspective. It is therefore difficult to know if the socio-economic gap in expectations is especially 'big' within these countries, and whether this is a bigger "problem" here than other parts of the developed world. Hence I ask:

Research Question 1. What is the absolute size of the gap between advantaged and disadvantaged children's expectations of completing university? Is this gap particularly large in the US and UK compared to other members of the OECD?

Of course, these countries also differ in terms of educational inequality; the gap between rich and poor teenagers test scores is greater in some countries than in others. Schutz et al (2008), for instance, compares the relationship between family background and children's test scores across a range of developed countries. They show that the association is particularly strong in England, Scotland and the US compared to elsewhere. In other words, these countries seem to suffer high levels of educational inequality, and hence one might also expect there to be especially large socio-economic differences in teenagers' educational plans. My second research question considers this possibility and explores whether countries with a big rich-poor gap in teenagers test scores are also the ones with a big rich-poor gap in children's educational plans. One particular focus of this analysis will be whether the expectation gap in the US and UK is bigger than one would predict given their level of educational inequality, or if these countries manage to 'buck the trend' (i.e achieve a smaller gap in young peoples' educational expectations than one would predict given their level of educational inequality). In summary:

RQ 2. Is the socio-economic gap in children's educational expectations particularly big in the UK and US, given their relatively high levels of educational inequality?

Indeed, given the arguments made above, one might argue that differences in test scores at age 15 are entirely responsible for the socio-economic gap in children's educational expectations; the only reason why advantaged 15 year olds are more likely to expect entry into university than their disadvantaged peers is that they have developed superior cognitive skills (e.g. "outcomes at age 14" in Figure 1) by this point in time. On the other hand, sizeable differences may remain even after controlling for academic skill measured at age 15. That is to say that disadvantaged 15 year olds may be less likely to expect a university education than their wealthy peers, even if they score equally well on assessments nearing the end of compulsory schooling. I again consider whether this is a specific problem to the US and UK, or if the situation here is comparable to other parts of the developed world. My final research question is therefore:

Research Question 3. Do the higher educational expectations of advantaged children only reflect their higher test scores at age 15? After controlling for this factor, is the socio-economic gap in the UK and US particularly large in comparison to other members of the OECD?

In answering the question above, I am not able to make a causal statement about the relationship between children's test scores, socio-economic status and their expectations. As laid out in Figure 1, children are assumed to begin making firm educational plans between ages 14 and 16, yet the exact point in time is almost impossible to identify. It could be that children start thinking seriously about university from a younger age than I can measure (e.g. 14), which has already had an impact upon their motivation at school, and is thus reflected in their scores on the PISA test (taken at age 15)[3]. In other words, the process of educational expectations influencing motivation and behaviour has already begun, causing age 15 test scores to be endogenous in the models that I estimate. This may be a particularly big issue in countries like England where children have to make educational decisions at a young age, and who receive regular updates on their ability through test performance. Hence this set of results needs to be treated with caution, and interpreted simply as the socio-economic gap in plans to enter higher education amongst children who manage to score the same on the PISA tests.

To summerise, this paper has one central aim – to place the socio-economic gap in children's higher educational plans in the US and UK into a comparative context. In other words, is the difference between "advantaged" and "disadvantaged" children's educational expectations greater in these countries than elsewhere? Section 3 now turns to the PISA data that I use to explore this topic, with the results from my analysis to follow in sections 4 and 5.

3. Data

The data I use are drawn from the 2003 round of the Programme for International Student Assessment (PISA); a study of 15 year-olds' cognitive skills held every three years. Although 46 countries took part, I restrict my analysis to 33 industralised nations[4]. In each country, a minimum of 150 schools were included in the sample, selected with probability proportional to size. Thirty students were then randomly selected from within. Average response rates of both schools (90%) and pupils (90%) were high, though this varies moderately between countries[5]. Further details are available in the PISA 2003 technical report (OECD 2004b). A

[3] In other words expectations at prior time points (that I am unable to control for) are confounding the relationships that I estimate.
[4] Here I treat the constituent parts of the United Kingdom (England, Scotland and Northern Ireland) as separate countries. Likewise, I separate Flemish from French Belgium.
[5] The lowest of which was England, at 64% for schools and 77% for pupils . Micklewright et al (2010)

set of sampling weights are also provided by the survey organisers that tries to correct for the unit non-response. The achieved sample size, across the 33 countries I consider, is 224,094.

As part of the study, children were asked to complete a questionnaire. This included the question:

"Which of the following do you **expect** to complete" [emphasis in original question]

Lower secondary education (Middle or junior high school)
Upper Secondary education (High school)
Post-secondary non-tertiary (Vocational/technical certificate after high school)
Tertiary "Type b" education (Associate's degree)
Tertiary "Type a" education of higher (Bachelors degree or higher)

Country specific options were provided in the questionnaire. The phrases in brackets illustrate these for the US. The primary outcome I analyse in this paper is whether the child ticked the top category (bachelors degree or higher). Response rates to this question were very high. Table 1 shows that almost 99% of children responded, from a low of 93% in France to a high of 100% in Poland. Consequently, I exclude the few (1%) observations where educational expectations are missing[6].

Table 1 about here

Recall that my concern in this paper is children's expectations (realistic assessments of their future), *not* their aspirations (idealistic goals). As noted in section 2, the developmental literature suggests that by the time of the PISA study (approaching age 16), children typically separate one concept from the other. Indeed, there has been work in the sociological literature that compares children's expectations to their aspirations around this age (see Patton and Creed (2007)). Such studies usually distinguish between the two concepts by altering and emphasizing the operative word (e.g. asking children what they would "<u>like</u>" to do, and then what they "<u>expect</u>"). Yet, to my knowledge, there has been little work on the validation of

investigate this non-response, and create an alternative set of responses weights (as opposed to those provided in the dataset by the survey organisers) to try and correct for bias in the estimates. They show that the UK only moves one place in the PISA ranking of children's test scores once these weights have been applied.

[6]These observations are not a random selection from the population. Rather they tend to be children with lower test scores, who also do not have complete information on family background.

such questions in quantitative surveys. In particular, there seems scant evidence of whether such subtle phrases are able to elicit the appropriate information from respondents. Unfortunately, the question asked in PISA shares much of the same criticism. It emphasizes the word **"expect"** using bold, underlined letters, yet provides children with no further instruction. Hence this is the only guide they have towards reporting their expectations rather than their aspirations. Whether such subtle wording can be adequately translated into other languages, as required in this cross-national analysis, is a further concern.

If this question is actually capturing children's *aspirations*, then the proportion reporting that they *"expect"* to complete university will be significantly higher than current graduation rates[7]. If this only occurs in certain countries, then these nations will out-lie from the rest. Indeed, if it is a translational issue that is causing this problem, language will be a common theme amongst these outlying nations. I search for such patterns in Figure 2. Specifically, in each country I compare the proportion of children who expect to obtain a degree (that I have calculated from the OECD PISA data) with actual graduation rates drawn from OECD (2009). The 45 degree line is where the proportion of children expecting to complete university equals actual graduation rates.

Generally, Figure 2 suggests that responses are not out of touch with reality; 45% of OECD children expect to complete university against actual graduation rates of 40%. Indeed, several countries, including England, sit below the 45 degree line; the proportion of children expecting to enter university is *below* actual completion rates. I do note, however, that there are some countries where one may have concerns. For instance the proportion of Canadian and American children "expecting" to obtain a bachelor's degree is significantly higher than actual graduation rates. However, Reynolds and Pemberton (2001) point out that there are high drop out rates from university (at least in the US), and as such the proportion of US children expecting to complete university are at least in-line with current *entrance* rates[8]. It is also interesting to see that the proportion of children expecting to complete university varies quite substantially across the English speaking countries, suggesting that this cross-national

[7] Of course, such a finding may just reflect that children are not very good at predicting the future (the question does capture children's expectations, it is just that these expectations are inaccurate). Nevertheless, if this pattern occurs consistently across all nations then one may question whether this is actually capturing adolescents' expectations

[8] Reynolds and Pemberton (2001) noted a similar finding when using the American NLSY 97 sample. OECD (2009) suggests that 65% of US school leavers enter university, very similar to the number I find expecting to obtain a degree. Unfortunately this information is not available for Canada.

variation is not simply due to a difference in language. Another feature of Figure 2 is the low correlation between the proportion of young people expecting to obtain a degree and current graduation rates (r = 0.02). Hence there is no suggestion that countries with a greater proportion of the population completing university are also the ones where children are more likely to expect to complete higher education.

Figure 2 about here

Nevertheless, despite concerns with some countries, the overall pattern of response is quite encouraging, and generally seems to be consistent with a measure of children's expectations. Hence these data do seem to be of value in answering the research questions I set out in section 2. Yet I am unable to investigate (and thus rule out) other potential problems regarding measurement error. For instance, it might be that advantaged children have a tendency to report their expectations and disadvantaged children their aspirations, and that this particular response pattern varies across OECD countries. Likewise, I advise caution in interpreting results for less developed members of the OECD like Turkey, Greece and Mexico, where "expectations" often seem to be out of touch with reality.

I now turn my attention to the variables that I use to distinguish between children from "advantaged" and "disadvantaged" backgrounds. In this paper, I view such concepts as multi-dimensional, reflecting a whole host of factors that one can not capture in a single variable (e.g. income). Rather, I define "advantaged" and "disadvantaged" on the basis of three factors – parental education, parental occupation and the home learning environment –following the theoretical framework of Chowdry et al (2009) set out in Figure 1.

I start by describing the importance (and measurement) of parental education. As noted in section 2, parents with more schooling might place greater emphasis on their children's education, or instill a taste for learning in their off-spring. Likewise, parents may be able to provide more information and encouragement about going to university if they hold a tertiary qualification themselves, and perhaps act as educational role models. Parental education will also be a key factor driving household income and children's cognitive development. This is therefore a key distinguishing feature between children from advantaged and disadvantaged backgrounds.

Information on parental education was captured through the sampled children as part of the PISA background questionnaire. Specifically, children were asked to report the level of education their mother and father completed at school and what type of tertiary qualifications they hold[9]. Schlutz (2005) and Jerrim and Micklewright (forthcoming) investigate possible measurement error in such reports using the PISA 2006 wave, where parents and children were asked separately to report mother's and father's level of education[10]. They gave the same category in two thirds of cases, though this was notably higher (around 86% of occasions) when the parents held a degree.

These responses were then recoded by the survey organisers into ISCED levels of education, a measure designed by UNESCO to aid cross-national analysis (though some differences in definitions across countries may remain - Steedman (2001)). The highest ISCED level achieved by either parent is then used to create the "highest parental education" variable. Appendix Table 1 shows how this is distributed across each of the OECD nations, including a "missing" category where this information is unavailable (typically 5-10% of cases)[11].
As sample sizes become small for certain groups, I recode this information into three broad categories (low, medium and high) following a similar re-categorisation in the Luxemburg Income Study (a well known dataset often used in cross-national research). I define high as holding a tertiary qualification (ISCED 5B or 5A+), medium as post-secondary schooling but no experience of higher education (ISCED 4) and low as completed secondary schooling or less (ISCED 0-3). This broad categorisation also helps to ensure that I have a sufficient number of observations within the "advantaged" and "disadvantaged" groups that I define later in this section.

[9] Note that children were instructed to report this information for their *mother and father like figures*. Consequently, children living in a household with a complex family structure, for instance with a step-mother or step-father, may not be reporting the education of their biological parents. I have experimented with including a variable that captures this in my analysis. However, I have chosen to exclude it as the estimated effect was usually small and statistically insignificant.

[10] The parental questionnaire that contains this information was an "international option" in 2006. This information is therefore only available for 11 countries, and has relatively high rates of non-response. I could not use the 2006 data for this analysis as it did not contain a question on children's educational expectations for the US or UK.

[11] These children are not a random sample from the population. Rather, they disproportionately come from children who performed poorly on the PISA test and come from less well-off families. One may argue that this could be driving some of the cross-national differences observed. Given the relatively small non-response in the majority of countries, I do not attempt any correction for this issue. However, I do include a "missing" dummy variable in all subsequent regressions to ensure these children are not dropped from the analysis.

The second measure of family background that I use to distinguish between advantaged and disadvantaged children is parental occupation. This variable is probably the best proxy available for household income and financial resources (which are unfortunately not collected as part of the PISA study) that play an important role in the development of children's educational expectations as laid out in section 2. Parental occupation will also pick up relevant aspects of social class, such as the societies, cultures and communities the child has grown up in.

As with parental education, information on mother's and father's occupation was collected directly from the sampled children. Specifically, they were asked the title of their mother's and father's main job and a description of the type of work this involves. Responses were coded by the survey organisers into four digit ISCO codes (the International Labour Organisation's occupational classification), which assigns the reported occupation to one of over 300 categories. Schulz (2005) investigates the potential measurement error in this data using the 2006 PISA field trial. He found that parents and children reported the same occupational group (defined in terms of the major ISCO groups) on roughly seven out of ten occasions. My experimentations with the final 2006 PISA sample revealed similar results.

Children's responses were then coded by the PISA survey organisers into the quasi-continuous ISEI index of occupational status, designed by Ganzeboom et al (1992). This index assigns each occupation a score between 16 and 90, depending upon the relevant "inputs" (educational level required) and "outputs" (the salary commanded) from that particular job[12]. Hence this is an objective occupational scale which is designed to be correlated with income. Moreover, Ganzeboom et al (1992) specifically designed this scale to aid the type of cross-national analysis I undertake in this paper, and have thus attempted to validate it as a measure of socio-economic status across a large number of developed countries (although some still question aspects of its validity – see Bukodi, Dex and Goldthorpe (forthcoming)). Nevertheless, the ISEI index remains an attractive measure against the possible alternatives (such as an aggregation of the 4 digit ISCO codes into the 9 major occupational groups). Summary statistics for the distribution of this variable across the

[12] The OECD describes: "The index captures the attributes of occupations that convert parents' education into income. The index was derived by the optimal scaling of occupation groups to maximise the indirect effect of education on income through occupation and to minimise the direct effect of education on income, net of occupation (both effects being net of age)."

OECD countries can be found in Appendix Table 2. It is interesting to note that the distribution of the ISEI index generally seems to be quite similar across countries, with very little cross-national variation in the 10th and 90th percentiles, and only slightly more at the 25th and 75th percentiles.

The third variable that I use to measure family background is children's reports of the number of books at home. It has been argued that this is correlated with a number of aspects of family background including parental education, household income and social origin (Ammermueller and Pischke 2009, Schuetz et al 2008). Yet the same authors also suggest that it picks up factors like the value parents place on their children's education and the encouragement they provide with regards to schooling. Likewise, the PISA survey organisers argue it is a measure of the 'home educational resources' available to the child. Hence this picks up such residual aspects of family background that are not fully captured within my measures of parental education and occupation, but are nevertheless likely to be important in the development of children's expectations.

This information was also reported by the participating children in the background questionnaire. Specifically, they were asked 'how many books are there in your home' (excluding magazines, newspapers and textbooks) with six possible options. However, the two bottom and two top categories contain a rather sparse number of observations. Thus I combine the bottom (0-10, 11-25, 26-100), and top (101-200, 201-500, above 500) three fields to form low and high groups, along with a 'missing' category, following a strategy similar to that used by the survey organisers (see OECD 2004c page 283). As with my re-categorisation of parental education, this also helps to ensure that I have sufficient observations within the "advantaged" and "disadvantaged" groups that I shall define in the following paragraph. The distribution of this variable across the OECD countries can be found in Appendix Table 3.

In the next section, I use the aforementioned variables in a logistic regression model of children's educational expectations. In all models, I also control for gender and whether the child was a first or second generation immigrant (as this group may be under different pressure from their family to complete higher education)[13]. Likewise, in all estimations I

[13] Children had to answer three questions regarding whether they, their mother or their father was born outside the country that they are taking the test in. I define a child as an "immigrant" if they answer yes to any of these

include 'missing' categories (dummy variables) to ensure children are not dropped from the analysis when pieces of information are unavailable. Thus the final form of this model is:

$$\log\left(\frac{\Pi(\boldsymbol{E}_{ij})}{1-\Pi(\boldsymbol{E}_{ij})}\right) = \alpha + \beta_1.\boldsymbol{Sex}_i + \beta_2.\boldsymbol{I}_i + \beta_3.\boldsymbol{SES}_i + \beta_4.\boldsymbol{I}_i * \boldsymbol{SES}_i$$

Where:

$\Pi(E_{ij})$ = Probability of the child expecting to graduate from university, where

E= 1 if the child expects to complete university, 0 otherwise

Sex = A binary indicator of the child's gender (0 = female, 1 =male).

I= Whether the child is a first or second generation immigrant (0 = Native , 1 = Immigrant)

SES = A vector of variables capturing the child's socio-economic background. This includes:

- Highest parental education – A set of two dummy variables, one referring to "some post-secondary education but no tertiary" (medium) the other "tertiary and above" (high). (Ref: "compulsory schooling or less" – i.e. low)
- Number of books in the home - A single dummy variable referring to whether there are "more than 100 books" (high) in the family home (Ref: "Less than 100 books" - low)
- Highest parental occupation measured on the ISEI scale - Entered as a piecewise linear term with knots at the 10th, 25th, 50th, 75th and 90th percentile of the national ISEI distribution

I then use this model to generate predictions of how likely a hypothetical child with given characteristics is to expect to complete university. Specifically, I create these predictions for:

---

three questions.

1. *An "advantaged" child*

Defined as:

- Either of their parents holds a tertiary qualification ("high" parental education)
- There are over 100 books in the family home ("high" books)
- The highest occupation of their parents sits at 75th percentile of the national ISEI distribution
- Country native
- Female

2. *A "disadvantaged" child*

Who I define as:

- Neither parent has completed any post-compulsory schooling ("low" parental education)
- There are less than 100 books in the family home ("low" books)
- The highest occupation of their parents sits at 25th percentile of the national ISEI distribution
- Country native
- Female

I then calculate the difference between these two predictions in order to compare the expectations of "advantaged" and "disadvantaged" groups[14].

Note that the definition of "advantage" and "disadvantage" that I use in this paper is multi-dimensional. Specifically, the prediction for the "advantaged" group refers to those children with "multiple advantages". So not only is at least one of these children's parents working in a high-level occupation, but also they have good access to educational resources and at least one of their mother or father is university educated. The second group can be thought of as "multiply disadvantaged" (i.e. low parental occupation, low parental education, poor access to home educational resources) following a similar logic. In section 5 I shall use an alternative definition of family background, and of "advantaged" and "disadvantaged" groups, to test the robustness of my results

[14] See Appendix Table 4 for the proportion of children who are defined as "advantaged" and "disadvantaged" using this definition

The final variable that needs to be described, which forms an integral part of my second and third research questions, is the PISA measure of children's academic skill. As part of the PISA 2003 study, children (aged 15) sat a two hour test. The PISA consortia claim that this measures children's 'functional ability' (how well they can use the concepts examined in 'real life' situations) in three domains (reading, maths and science). In 2003, maths was assigned as the major domain, where the vast majority of questions children were asked were on this topic. All test questions were explicitly designed with cross-national comparability in mind. Answers were summarized by the survey organizers into a single score for each of the three domains using an 'item-response model'; the intuition being that true skill in each subject is unobserved, and must be estimated from the answers to the test. Consequently, five 'plausible values' are generated for each pupil, estimating their true proficiency in each subject. These scores were scaled by the survey organizers to have a mean (across all OECD countries) of 500 points and standard deviation of 100. Throughout my analysis, I use the first of these plausible values for the maths domain[15]. The correlation between test scales is high ($r \approx 0.8$ between maths and reading, and the same between maths and science), with little change in my substantive results when I use the reading and science scales instead[16]. The distribution of this variable across all the countries I consider can be found in OECD (2004a).

Throughout this paper, I shall present results mainly in terms of 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 expect to complete higher education. I illustrate this point in Table 2. The second and third columns present the proportion of children expecting to complete university depending on whether either of their parents holds a degree. Column 4 provides the marginal effect (the percentage point difference between columns 2 and 3) while column 5 illustrates the difference in terms of the log odds.

Table 2 about here

[15] I experimented using the other plausible values, and by running five separate models and averaging the estimated coefficients and standard errors. Results are very similar to those presented.

[16] Note that only around half the children within each country actually answer questions in each of "minor" PISA domains (reading and science). Scores are estimated by the study organisers for the remaining children using a Rasch modelling approach.

Comparisons between countries can look very different depending on which measure is used. Take England, one of my countries of interest, and Korea. The difference between the second and third column is quite similar in terms of the log-odds (1.69 in Korea to 1.73 in England), but very different when considering the marginal effect (22 percentage points compared to 39)[17]. This is being driven by the fact that, across the population, Korean children are generally more optimistic about their prospects of completing university than those in England (77% expect to obtain a degree in the former, compared to 29% in the latter). As my concern in this paper is the expectations of disadvantaged children *relative* to their advantaged peers, I prefer the log-odds as it abstracts from the *absolute* proportion of the population believing that they will complete higher education. However, appreciating that this metric is rather cumbersome to interpret, I also occasionally present predicted probabilities to assist the reader's understanding.

4. Results

I shall now present results for the research questions set out in section 2. Parameter estimates and an illustration of my predictions for England and the US can be found in Appendix 1, with those for other countries available upon request.

Figure 3 illustrates the gap between advantaged and disadvantaged children's expectations of completing higher education[18]. In all countries, disadvantaged children are less likely to expect entry into university than their more affluent peers. This gap is generally big (around two and two and a half log odds in most countries) and is always significantly different from zero at the one percent level. To put this into perspective, if a hypothetical disadvantaged child had a 50% chance of expecting to complete university, the probability for their identical (but advantaged) peer would be closer to 90%[19].

Figure 3 about here

[17] Across all countries, the estimated correlation between the marginal effect and log odds is 0.78.
[18] A list of country abbreviations can be found in the left hand column of Table 1.
[19] These probabilities were calculated using the formula: probability = exp(log[odds]) /(1+exp(log[odds])). Log odds of 0 correspond to a probability of 50%. Log odds of 2.5correspond to a probability of 92%. Hence, in this hypothetical example, a difference of 2.5 log odds (i.e. the difference between advantaged and disadvantaged groups) leads to a 42% difference in the probability.

Focusing firstly on the results for England, the difference in log odds between advantaged and disadvantaged children's expectations is roughly 2.4; the 10th largest estimate in the OECD. Larger gaps are found in Germany, Switzerland and Austria; countries where access to university is restricted to children on the appropriate educational "track"[20]. Dustman (2004) shows that this "track" is strongly associated with family background, hence one would anticipate there to be large differences between advantaged and disadvantaged children's expectations in these countries. Likewise, in a number of Eastern European countries (e.g. Hungary, Slovakia) the log-odds seems large when compared to other nations. This should not been surprising, as several studies (see Shavit and Blossfeld 1993) have shown that these countries have a large degree of educational polarisation. England is also ranked higher than the other Anglo-Saxon countries. This includes Scotland and Northern Ireland (the other constituent parts of the UK) where the estimated difference in log odds is roughly equal to the OECD average (around 2.2). However the 95% confidence interval (the thin black line running through the centre of each bar) suggests that caution is required when interpreting this result. One can not reject the null hypothesis that England is significantly different to Scotland or Northern Ireland at any of the conventional threshold. Indeed, Table 3 shows that the socio-economic gap in England is only significantly stronger than only two countries at the 5% level (Finland and the US) and a further one at each of the 5% and 10% levels (Turkey and Portugal). To summerise, there is little evidence that England stands out in comparison to other parts of the UK (Scotland and Northern Ireland) or generally amongst the OECD.

Indeed, the general suggestion of Figure 3 is that cross-national variation in the socio-economic expectation gap is rather modest – although there are some interesting outliers. The US, for instance, immediately stands out in Figure 3 - though perhaps not in the direction one might initially expect. The socio-economic gap here on the log-odds scale is around 1.6; the second *smallest* out of the 33 countries. Moreover, this result is not just a matter of sampling variation; Table 3 shows that results for the US are statistically different to 11 other nations at the 1% level and a further 5 and 4 at the 5% and 10% levels. Hence it seems that the absolute gap between advantaged and disadvantaged children's educational expectations is actually quite "small" in the US – at least when compared to other members of the OECD.

[20] In these countries, children are sorted into different schools by their level of ability at a young age (known as "tracking").

Is this difference in educational expectations greater or smaller than we would expect given each countries level of educational inequality? The answer to this question can be found in Figure 4. Specifically, I re-estimate exactly the same model as set out in section 3 – but now using Ordinary Least Squares (OLS) with children's age 15 test scores as the response. This provides a measure of inequality in educational achievement that can be found on the x-axis. The y-axis, on the other hand, provides the socio-economic gap in children's educational expectations (as just presented in Figure 3). Running through the centre of the graph is a regression line, which represents the difference in educational expectations one would predict in a country given its level of educational inequality.

Figure 4

One of the most notable features of Figure 4 is the relatively strong correlation (approximately 0.7) between educational inequality and the socio-economic gap in university plans. This is in stark contrast to Figure 2 where I demonstrated that the correlation between the proportion of children expecting to go to university and the proportion actually attended was close to zero. In any case, Figure 4 indicates that England, Scotland and Northern Ireland all sit very closely to the estimated regression line. Hence, despite policy concern in the UK that disadvantaged children are a lot less likely to expect to complete higher education than their advantaged peers, the international evidence suggests this is not particularly more than one would anticipate given its level of educational inequality. The US, on the other hand, again emerges as an intriguing outlier. It is the furthest *below* the regression line out of any of the OECD countries. That is, the difference between advantaged and disadvantaged children's educational expectations in the US is much lower than one would predict given its level of educational inequality. One can, of course, come up with several possible explanations for why this might be. For instance, the US is a country with a diverse ethnic composition and a large number of young people of afro-Caribbean dissent. Such minority groups are known to generally hold high educational expectations, but also suffer from a large "expectation-attainment gap" (i.e. their expectations rarely match with reality - see Gutman and Akerman for a discussion of this literature). Alternatively, one might argue that policies to raise disadvantaged children's expectations of completing higher education are long established in the US, and that the comparatively narrow gap between rich and poor is essentially illustrating such initiatives success. It is beyond the scope of this paper, and indeed the data available, to try and distinguish between these explanations. But attempting to explain why

such a pattern emerges in the US (and why it stands out from other countries) is an interesting possibility for future research.

A final feature of Figure 4 that is worth highlighting is the positions taken up by countries that "track" children into different schools from an early age. Note how these countries (shown in circles) tend to sit above the line and mostly to the right (with the exception of Switzerland). These are countries where educational inequalities tend to be high, but also where the difference between advantaged and disadvantaged children's educational expectations is greater than one might predict. One cautious interpretation of this result is that, along with tracking being associated with educational inequality (as suggested in studies by Hanushek and Woessmann 2006 and Ammermueller 2006), it may also be related to the size of the socio-economic gap in adolescents' expectations for the future.
In Table 4, I investigate whether England and the US still stand out in the international ranking of SES differences in educational expectations if I now *control* for differences in children's age 15 test scores (i.e. when I add children's score on the PISA test as an additional right hand side variable to the model that was presented in section 3). I of course recognise the potential endogeniety of this variable, and that estimates from these models are not causal; they should hence be treated simply as conditional associations found in the data. This limitation does not, however, restrict my ability to address the simple question I am asking – does there remain a socio-economic gap in educational expectations amongst children who score the same on tests at age 15, and do the cross-national patterns remain similar to those I presented before? Results can be found in Table 4.

Table 4 about here

The US is again placed towards the top of the table, while England is around the middle. The socio-economic gap in the latter is now significantly stronger than in only one other OECD countries at the five percent level (Finland), and a further two at the ten percent level (US and Mexico). Once again, there is little difference compared to Scotland and Northern Ireland, and all three countries sit broadly in line with the OECD average of around 1.7 log-odds. Consequently, after controlling for test scores, there remains no evidence that the difference between the educational expectations of advantaged and disadvantaged children is unusually large in the UK. On the other hand, the association between family background and children's

expectations of completing university remains significantly weaker in the US than in other OECD countries. Five countries are significantly different at the 1% level, with a further seven at the 5% level and six more at the 10% level. This, once more, includes countries within the United Kingdom, several from Scandinavia (e.g. Sweden, Norway) and those which assign children to different educational tracks (e.g. Germany, Switzerland and Austria). Hence, there remains a suggestion that the socio-economic gap in educational expectations is comparatively small in the US.

5. Robustness of results

In this section, I use a different measure of family background to test the robustness of the aforementioned results. I do so for two reasons. Firstly, using my initial definition of "advantage" and "disadvantage", a different proportion of the population in each country is being assigned to these groups. For example, Appendix Table 4 shows that 9% of English children were defined as "disadvantaged" under the definition I described in section 3, compared to 3% of those from Norway and roughly 20% in Portugal, Poland and Turkey. Hence I have thus far presented results where a more extreme proportion of the population has been defined as "advantaged" / "disadvantaged" in some countries than in others. I will now re-perform my analysis, using a different measure of family background, which will allow me to compare the most advantaged and the least advantaged quarter of the population within each country (i.e. I will now investigate whether my results hold when defining the same *relative* proportion of children in each country as "advantaged" and "disadvantaged"). Secondly, the definitions and measures that I used in the previous section were just one (quite specific) way to divide children into advantaged and disadvantaged groups. It is important to test whether my results and substantive conclusions hold under possible alternatives.

To do so, I turn to the "Economic, Social and Cultural Status" (ESCS) index; a continuous measure of family background (scaled to be mean 0 and standard deviation 1 across the OECD countries) that is produced by the PISA survey organisers and is contained within the provided dataset. Specifically, the survey organisers have produced a weighted average, via a principal component analysis, of three variables (highest level of parental education, parental occupation, and availability of items in the family home) to generate a measure of children's socio-economic status. The first two of these variables (parental education and occupation) are as described in the earlier data section (although parental education has now been

converted by the survey organiser into a linear term reflecting ‘years of schooling’ – see OECD 2004: 308 for details). The “availability of home possessions” is itself an index (from another principal components analysis) based upon children's reports of whether they have various items (e.g. computers, works of art, number of books) in their family home. According to OECD (2004), this provides an approximate measure of household wealth. Further details on this measure and its construction can be found in OECD (2004b), while Schulz (2005) investigates its properties (reporting reasonable levels of internal consistency and stability).

This index has several attractions as an alternative measure of family background. Firstly, it continues to capture the multi-dimensional nature (education, occupation, income/wealth) of “advantage”. Secondly, as this variable is continuous, I can easily widen the proportion of children contained within my definition of advantaged and disadvantaged groups (to, for instance, the top and bottom quartile). Also note that, by using this measure, I can ensure that the same relative proportion of the population is defined as advantaged and disadvantaged in each of the OECD nations. Yet this variable also has a number of limitations. As it is created via a principal components analysis, it is somewhat difficult to interpret. There is also likely to be some information loss from suppressing various measures into one, all-encompassing, continuous index. One may also have some doubts over the validity of using household items as a measure of family wealth. Whether a child grows up in a home with a dishwasher or works of art will to some extent reflect parental preferences, and thus may provide little insight into whether they truly come from an advantaged or disadvantaged background. Similarly, one may question the cross-national comparability of such measures[21]. Yet, despite these limitations, this remains an attractive alternative measure of socio-economic status due to its flexible nature. The distribution of this index across OECD countries can be found in OECD (2004a).

[21]Indeed, it is for these reasons that I did not use this variable or include such information in my initial definition of advantaged and disadvantaged groups (described in section 3).

I proceed by dividing this variable into four equal groups (separately for each country) and defining:

'Advantaged' = top quartile group of the national ESCS distribution
'Disadvantaged' = bottom quartile group of the national ESCS distribution

I then use this information in a regression model as a set of dummy variables, with the bottom quartile ('disadvantaged') as the reference group. Formally, this model is specified:

$$\log\left(\frac{\Pi(\boldsymbol{E_{ij}})}{1-\Pi(\boldsymbol{E_{ij}})}\right) = \boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1.\boldsymbol{Sex}_i + \boldsymbol{\beta}_2.\boldsymbol{I}_i + \boldsymbol{\beta_3}.\boldsymbol{SES}_i + \boldsymbol{\beta_4}.\boldsymbol{I}_i * \boldsymbol{SES}_i$$

where:

SES = A vector of three dummy variables reflecting advantage, based upon quartiles the ESCS measure of family background described above (reference = bottom quartile)

All other variables are as described in section 3.

The estimated coefficient for the "top quartile" dummy variable captures the difference between "advantaged" and "disadvantaged" children's educational expectations, and can be found on the y-axis of Figure 5. I then re-run the model above, but using OLS regression with children's scores on the PISA maths test as the response. This then provides a measure of educational inequality within each country, which can be found on the x-axis. One may think of these results as analogous to those presented previously in Figure 4, but now using this alternative definition of "advantaged" and "disadvantaged" groups.

Figure 5

Generally, these results are largely consistent with those that I presented before, although there is slightly more variation around the estimated regression line (the correlation is now roughly 0.55 down from 0.70 for the estimates presented in section 4). Nevertheless, England and Scotland still sit very closely to this predicted line – again suggesting that the difference in advantaged and disadvantaged children's educational expectations is not greater than one

would anticipate given their level of educational inequality. Likewise, it still seems to be the case that countries which "track" children into different types of school from an early age tend to sit above the 45 degree (although not by particularly large magnitudes in most cases). The US, on the other hand, is slightly less of an outlier than before; it does not stick out quite so much from all other developed countries. Yet, it is still quite some distance from the fitted line, and there remains a strong suggestion that the socio-economic gap in children's educational expectations is smaller than one would predict given the level of educational inequality in the US. Hence the overall message is one that tally's with the preceding section, with my substantive conclusions largely unchanged.

6. Summary and conclusions

There is a concern in many countries that children from disadvantaged backgrounds are under-represented amongst the undergraduate population. In particular, policy-makers are worried that some young adults who could benefit from higher education decide not to seek out the returns that this investment can offer. One explanation for this is that many disadvantaged young people in these countries see university as "not for the likes of them". As such, there has been increasing concern regarding the socio-economic gap in children's expectations of completing higher education. "Widening access" schemes that try to address this issue have thus become common across the developed world. A particular feature of such programmes in the US and UK is that they explicitly aim to raise disadvantaged children's expectations of completing university. It is claimed that, in raising the proportion of disadvantaged students expecting to complete higher education, such policies will reduce the socio-economic divide in tertiary graduation rates.

By considering the size of this expectation gap in the US and UK, and how it compares to other developed countries, this paper makes an important contribution to the existing literature. I have shown that, although there are large differences between advantaged and disadvantaged children's educational plans, this holds true across all countries in the developed world. There is little evidence that the UK stands out when compared to other members of the OECD, and that the socio-economic gap between "advantaged" and "disadvantaged" children's educational plans is roughly what one would predict given this country's level of educational inequality. There is, on the other hand, quite a strong suggestion that the socio-economic gap in educational expectations is atypically small in the

US.

It is of course important to note the limitations of this study and, in doing so, aspects of the wider literature. To begin, I remind the reader that none of the estimates I have presented should not be treated as causal. Rather this paper has simply attempted to give the socio-economic gap in children's educational expectations, which is of great academic and political interest in many countries, a comparative context. On a related issue, I have undertaken this research based on the assumption that adolescents' educational expectations have an important influence on their later behaviour and schooling attainment. Although there is evidence supporting this from a broad range of disciplines, including sociology, social psychology and economics, further work in this area still needs to be done. In particular, future research should focus on untangling the relationship between these variables and whether such associations vary across different national settings. Finally, all my analyses and subsequent inferences are based on the assumption that children are reporting their educational *expectations,* rather than their *aspirations,* in the PISA survey. Although general patterns within the data are consistent with this view, formal validation of this type of question would represent a significant forward step for the wider literature.

One should, nevertheless, not lose sight of the contribution this paper makes to the wider literature. Single national studies are always likely to find a socio-economic gap in attitudes and expectations towards higher education, producing differences that (at first sight) tend to be rather striking. But before deciding whether this expectation gap is "big" or not, and if a country has a particular "problem" in this area, it needs to be put into context. With reference to the US and UK, although there is a significant socio-economic gap in children's educational expectations, it is not atypically big compared to other developed nations. In fact, given its level of educational inequality, the socio-economic gap in university expectations in the US is rather small. This does not mean that policymakers should stop their investment in this area. However, it is important that they understand that socio-economic differences in attitudes and expectations regarding higher education is not an area where the UK is doing particularly badly (at least when this data was collected in 2003)[22]. Similarly, for

[22] It is worth bearing in mind that since this data was collected in 2003, higher education finance in the UK has undergone a substantial change. It is difficult to say whether higher tuition fees (introduced in 2005 and to be extended in 2012) may alter this result. For example, one could argue that higher tuition fees may make disadvantaged children feel more credit constrained, hence lowering their expectations towards HE compared to their advantaged peers. The UK has unfortunately not collected data on educational expectations in subsequent

policymakers and academics in the US, it is worth bearing in mind that getting disadvantaged young people to aim high is something which they already do comparatively well.

PISA survey waves (i.e. past 2003), meaning I am unable to investigate this interesting possibility.

References:

Ammermueller, A. (2006) "Educational Opportunities and the Role of Institutions", ZEW Discussion Paper No. 05-44, Mannheim.

Ammermueller A & Pischke J-S (2009) "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study", Journal of Labor Economics , volume 27, p 315-48

Blanden, J. & Machin, S. (2004) "Educational Inequality and the Expansion of UK Higher Education", Scottish Journal of Political Economy, Special Issue on the Economics of Education, volume 51, p 230–49

Brown, S. Sessions, J. & Taylor, K. (2004) "What Will I Be When I Grow Pp? An Analysis of Childhood Expectations and Career Outcomes", Working Paper No. 05/2, Department of Economics, University of Leicester

Bukodi, E. Dex, S. & Goldthorpe, J. (forthcoming) "The Conceptualisation and Measurement of Occupational Hierarchies: A Review, a Proposal and Some Illustrative Analysis"

Cameron, S. & Heckman. J. (1999) "Can Tuition Combat Rising Wage Inequality?", In Kosters, M. (ed), Financing College Tuition: Government Policies & Educational Priorities

Chevalier, A. Gibbons, S. Thorpe, A. Snell, M. & Hoskins, S. (2009) "Students' Academic Self-Perception", Economics of Education Review, volume 28, p 716-727

Chowdry, H. Crawford, C. Dearden, L. Goodman, A. & Vignoles, A. (2008) "Widening Participation in Higher Education: Analysis using Linked Administrative Data", Institute for Fiscal Studies report

Chowdry, H. Crawford, C. & Goodman, A. (2009) "Drivers and Barriers to Educational Success Evidence from the Longitudinal Study of Young People in England", DCSF research Report RR102

Cowan, B. (2009) "Forward-Thinking Teens: The Effects of College Costs on Adolescent Risky Behavior", Universiy of Madison, Wisconsin job market paper

Chowdry, H. Crawford, C. , Dearden L., Joyce, R., Sibieta, L., Sylva, K. Washbrook, E. (2010) "Poorer children's education attainment: how important are attitudes and behaviour?"

Dustmann, C (2004) "Parental Background, Secondary School Track Choice, and Wages," Oxford Economic Papers, Oxford University Press, volume 56, p 209-230

Emmerson, C. Frayne, C. McNally, S. & Silva, O. (2005) "Evaluation of Aimhigher: Excellence Challenge: The Early Impact of Aimhigher: Excellence Challenge on Pre-16 Outcomes: An Economic Evaluation", Department for Education and Skills Research Report 652

Friedman, M. & Friedman, R. (1980) "Free to Choose"

Ganzeboom, H. De Graaf, P. & Treiman, D. (1992) "A Standard International Socio-Economic Index of Occupational Status", Social Science Research , volume 21, p 1-56

Gottfredson, S. (2002) "Gottfredson's Theory of Circumscription, Compromise, and Self Creation", In D. Brown (Ed.) Career Choice and Development , 4th edition, p 85-148

Gutman, L. & Akerman, R. (2008) "Determinants of Aspirations", Centre for Research on the Wider Benefits of Learning Research Report 27

Hanushek, E. & Woessmann, L. (2006) "Does Educational Tracking Affect Performance and Inequality? Differences in Differences Evidence across Countries", Economic Journal, volume 116, pp C63-C76.

Hanushek, E. & Woessmann, L. (2010) "The Economics of International Differences in Educational Achievement", IZA Discussion Paper 4925

Jenkins, S. Micklewright, J. & Schnepf, S. (2006) "Social Segregation in Secondary Schools: How Does England Compare with Other Countries?", IZA Discussion Papers 1959, Institute for the Study of Labor (IZA).

Jerrim, J & Micklewright, J (forthcoming) "Children's Education and Parents' Socio-economic Status: Distinguishing the Impact of Mothers and Fathers"

Khoo, S. & Ainley, J. (2005) "Attitudes, Intentions and Participation". LSAY Research Report 41

Machin, S. & Vignoles, A. (2004) "Educational Inequality: The Widening Socioeconomic Gap", Fiscal Studies, volume 25, p 107–28

Micklewright, J. Schnepf, S. & Skinner, C. (2010) "Non-response Biases in Surveys of School Children: The Case of the English PISA samples", S3RI Applications & Policy Series working paper A10/01

Morgan, S. (1998) "Adolescent Educational Expectations. Rationalized, Fantasized, or Both?" Rationality and Society, volume 10, p 131–162

Morgan, S. (2004) "Methodologist as Arbitrator: Five Methods for Black-White Differences in the Causal Effect of Expectations on Attainment", Sociological Methods and Research, volume 33, p 3–53

Morgan, S. (2005) "On the Edge of Commitment: Educational Attainment and Race in the United States", Stanford University Press

Organisation for Economic Co-operation and Development (2004a) "Learning for Tomorrow's World – First Results from PISA 2003", Paris: Organisation for Economic Co-operation and Development

Organisation for Economic Co-operation and Development (2004b) "PISA 2003 Technical Report", Paris: Organisation for Economic Co-operation and Development

Organisation for Economic Co-operation and Development (2004c) "PISA 2003 data analysis manuals for SAS and SPSS users", Paris: Organisation for Economic Co-operation and Development

Organisation for Economic Co-operation and Development (2008) "Growing unequal? : Income distribution and poverty in OECD countries" , Paris: Organisation for Economic Co-operation and Development

Organisation for Economic Co-operation and Development (2009) "Education At a Glance 2009: OECD Indicators", Paris: Organisation for Economic Co-operation and Development

Patton, W. & Creed, P. (2007) "The Relationship between Career Variables and Occupational Aspirations and Expectations for Australian High School Adolescents", Journal of Career Development, volume 34, p 127-148

Reynolds, J. & Pemberton, J. (2001) "Rising College Expectations among Youth in the United States: A Comparison of 15 and 16 Year Olds in the 1979 and 1997 NLSY", Journal of Human Resources, volume 36, p 703-726

Reynolds, J. & Baird, C. (2010) "Is There a Downside to Shooting for the Stars? Unrealized Educational Expectations and Symptoms of Depression", American Sociological Review, volume 75, issue 1, pp 151–172

Schoon, I. (2010) "Planning for the future. Changing education expectations in three British cohorts", Historical Social Research, volume 35, issue 2, p 99-119

Shavit, Y. & Blossfeld, H. (1993) "Persistent Inequality: Changing Educational Attainment in Thirteen Countries", Boulder, Colorado: Westview Press

Schütz G., Ursprung, H. & Wössmann, L. (2008) "Education policy and equality of opportunity", Kyklos, volume 61, p 279–308

Schulz, W. (2005) "Measuring the Socio-economic Background of Students' Achievement in PISA 2000 and PISA 2003", Paper presented at the Annual American Educational Research Association (AERA), San Francisco

Shields, C. & Mohan, J. (2008) "High-Quality Education for All Students: Putting Social Justice at its Heart", Teacher Development, volume 12, p 289-300.

Steedman, H. & McIntosh, S. (2001) "Measuring Low Skills in Europe: How Useful is the ISCED Framework?", Oxford Economic Papers, volume 53, p 564-581.

Sutton Trust (2008), "Increasing Higher Education Participation Amongst Disadvantaged Young People and Schools in Poor Communities", Sutton Trust report to the National Council for Educational Excellence

Sutton Trust (2009) "Sutton Trust submission to the House of Commons Children, Schools and Families Committee on Social Mobility and Education and Access to the Professions"

Table 1. Sample sizes and missing expectations data across the OECD countries

Country	Starting sample size	Complete educational expectation data	% Missing expectations data
Poland (POL)	4,383	4,381	0.0
Finland (FIN)	5,796	5,793	0.1
Italy (ITA)	11,639	11,631	0.1
Japan (JAP)	4,707	4,700	0.1
Korea (KOR)	5,444	5,440	0.1
Spain (ESP)	10,791	10,776	0.1
Turkey (TURK)	4,855	4,852	0.1
Hungary (HUN)	4,765	4,756	0.2
Slovakia (SLOV)	7,346	7,328	0.2
Greece (GRE)	4,627	4,613	0.3
Portugal (PORT)	4,608	4,594	0.3
Switzerland (SWZ)	8,420	8,393	0.3
Sweden (SWE)	4,624	4,605	0.4
Australia (AUS)	12,551	12,492	0.5
Mexico (MEX)	29,983	29,845	0.5
Denmark (DEN)	4,218	4,191	0.6
Scotland (SCO)	2,723	2,707	0.6
USA (USA)	5,456	5,419	0.7
Austria (AUT)	4,597	4,558	0.8
Iceland (ICE)	3,350	3,324	0.8
Ireland (IRE)	3,880	3,848	0.8
Luxembourg (LUX)	3,923	3,892	0.8
Northern Ireland (NI)	2,853	2,829	0.8
Belgium(French) (BELFREN)	2,958	2,931	0.9
Norway (NOR)	4,064	4,023	1.0
New Zealand (NZ)	4,511	4,447	1.4
Netherlands (NLD)	3,992	3,902	2.3
Belgium(Flemish) (BELFLEM)	5,838	5,696	2.4
England (ENG)	3,959	3,817	3.6
Czech Republic (CZE)	6,320	6,076	3.9
Germany (GER)	4,660	4,457	4.4
Canada (CAN)	27,953	26,707	4.5
France (FRA)	4,300	3,997	7.0
TOTAL	**224,094**	**221,020**	**1.4**

Notes:
1 Missing data refers to item non-response only. Details on unit non-response can be found in the OECD (2004b) Technical Report.
2 Data sorted by the percentage of missing observations

Table 2. Children's expectations of completing university, depending on whether either of their parents' holds a bachelor's degree

	% Expecting to complete university if *either* their mother or father holds a degree	% Expecting to complete university if *neither* their mother or father holds a degree	Marginal Effect	Difference in log-odds
Mexico	69	53	16	0.68
Netherlands	56	31	25	1.04
Finland	67	41	25	1.07
Belgium(French)	48	23	25	1.13
France	56	28	28	1.19
Portugal	73	45	28	1.20
Sweden	50	23	27	1.21
Canada	80	54	26	1.23
Ireland	76	48	29	1.23
Greece	81	55	25	1.25
Norway	45	19	26	1.25
Italy	72	41	31	1.31
USA	82	55	28	1.32
Denmark	47	19	28	1.33
Scotland	73	41	33	1.36
Japan	69	36	34	1.38
Spain	74	41	33	1.41
Australia	80	49	30	1.43
Turkey	93	75	18	1.49
Belgium(Flemish)	59	24	35	1.52
Switzerland	39	12	27	1.55
New Zealand	70	32	38	1.60
Luxembourg	71	33	39	1.60
Slovakia	73	35	38	1.61
Northern Ireland	64	26	37	1.62
Iceland	64	26	38	1.62
Korea	93	71	22	1.69
England	60	21	39	1.73
Austria	59	20	40	1.75
Germany	44	12	32	1.75
Czech Republic	73	31	42	1.79
Poland	67	23	44	1.92
Hungary	87	41	45	2.26
OECD average	**70**	**40**	**30**	**1.25**

Notes:
1 The column labeled 'marginal effect' illustrates the percentage point difference between children's expectations. This is the difference between the first two columns. Conversely, the final column illustrates the difference between the same figures, but in terms of the log-odds.
2 The final row, labeled OECD average, refers to when one combines data across all 33 OECD countries considered.
3 Countries sorted by the difference in terms of log-odds

Table 3. Difference between the expectations of advantaged and disadvantaged children

Country	Log odds	SE	Sig diff from Eng?	Sig diff from USA?
Fin	1.41	0.15	***	-
USA	1.59	0.18	***	-
Turk	1.70	0.25	**	-
Port	1.93	0.14	*	-
Fra	1.94	0.23	-	-
Ita	1.96	0.16	-	-
Ice	1.97	0.18	-	-
Can	1.99	0.13	-	*
Gre	1.99	0.23	-	-
Lux	2.01	0.31	-	-
NZ	2.05	0.25	-	-
Nld	2.05	0.25	-	-
NI	2.11	0.26	-	*
Ire	2.11	0.17	-	**
Pol	2.14	0.17	-	**
Den	2.15	0.26	-	*
Swe	2.20	0.20	-	**
Sco	2.22	0.30	-	*
Nor	2.23	0.24	-	**
BelFren	2.24	0.32	-	*
Esp	2.36	0.14	-	***
Cze	2.39	0.18	-	***
Eng	2.43	0.25	-	***
Swz	2.43	0.36	-	**
Aus	2.47	0.22	-	***
Aut	2.56	0.28	-	***
BelFlem	2.62	0.23	-	***
Slov	2.81	0.19	-	***
Hun	3.09	0.23	*	***
Germany	3.20	0.39	*	***
Hungary	3.21	0.37	*	***
Japan	3.31	0.00	***	***

Table 4. Difference between the expectations of advantaged and disadvantaged children, after controlling for age 15 test scores

Country	Log odds	SE	Sig diff from Eng?	Sig diff from USA?
Fin	1.095	0.151	**	-
USA	1.190	0.182	*	-
Mex	1.202	0.153	*	-
Nld	1.232	0.256	-	-
Turk	1.275	0.250	-	-
Port	1.416	0.147	-	-
Fra	1.437	0.245	-	-
NI	1.442	0.294	-	-
NZ	1.524	0.247	-	-
Ice	1.557	0.192	-	-
Lux	1.567	0.344	-	-
Ire	1.600	0.186	-	-
Can	1.636	0.139	-	*
Den	1.641	0.267	-	-
Cze	1.654	0.189	-	*
Pol	1.689	0.180	-	**
Sco	1.692	0.309	-	-
Ita	1.695	0.154	-	**
Esp	1.704	0.149	-	**
BelFren	1.738	0.331	-	-
Nor	1.740	0.246	-	*
Gre	1.743	0.253	-	*
Eng	1.752	0.264	-	*
Aus	1.760	0.241	-	*
Swe	1.835	0.196	-	**
Aut	1.871	0.264	-	**
Slov	1.968	0.184	-	***
BelFlem	1.979	0.242	-	***
Ger	1.984	0.300	-	**
Swz	1.992	0.359	-	**
Hun	2.116	0.237	-	***

Notes:
1 The final two columns illustrate whether the estimated difference in log odds are significantly different to those in England and the US. *, ** and *** indicate a statistically significant difference at the 10%, 5% and 1% level.
2 Countries sorted by the difference in expectations of advantaged and disadvantaged groups
3 Statistical significance calculated using a two sample t-test assuming independent samples are drawn between countries

Figure 1. A model linking family background to children's aspirations, expectations and outcomes, based upon Chowdry et al (2009)

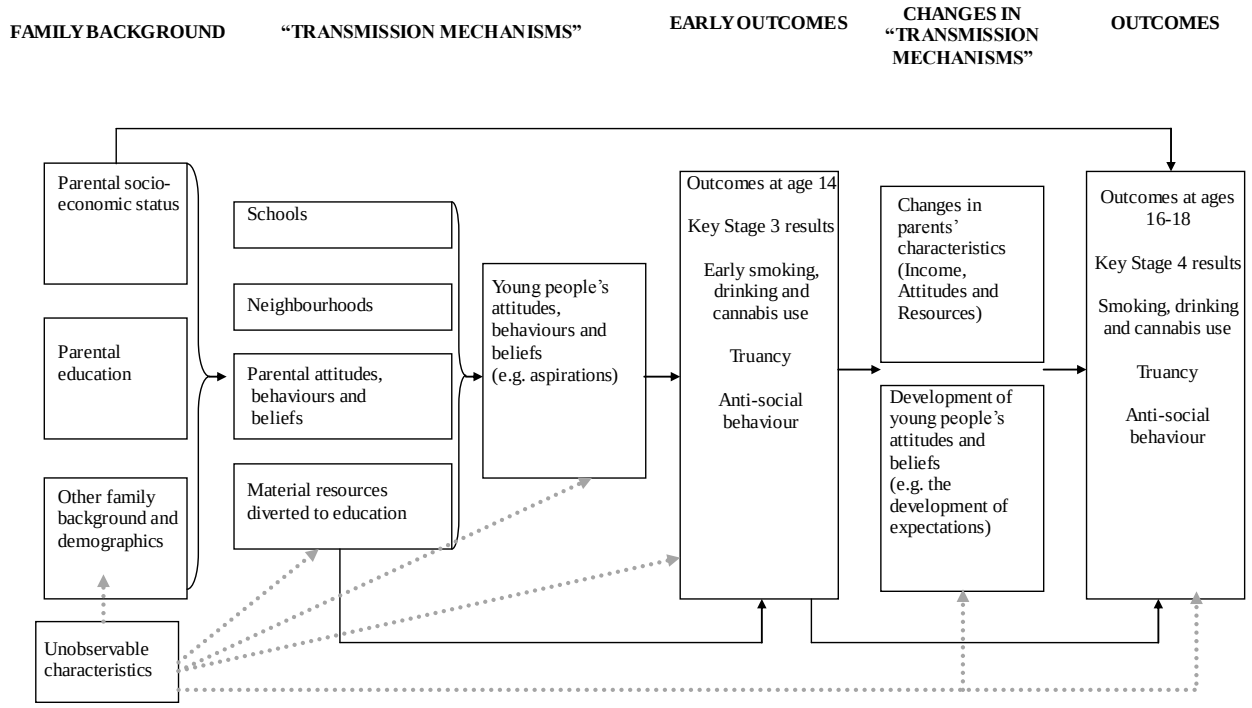
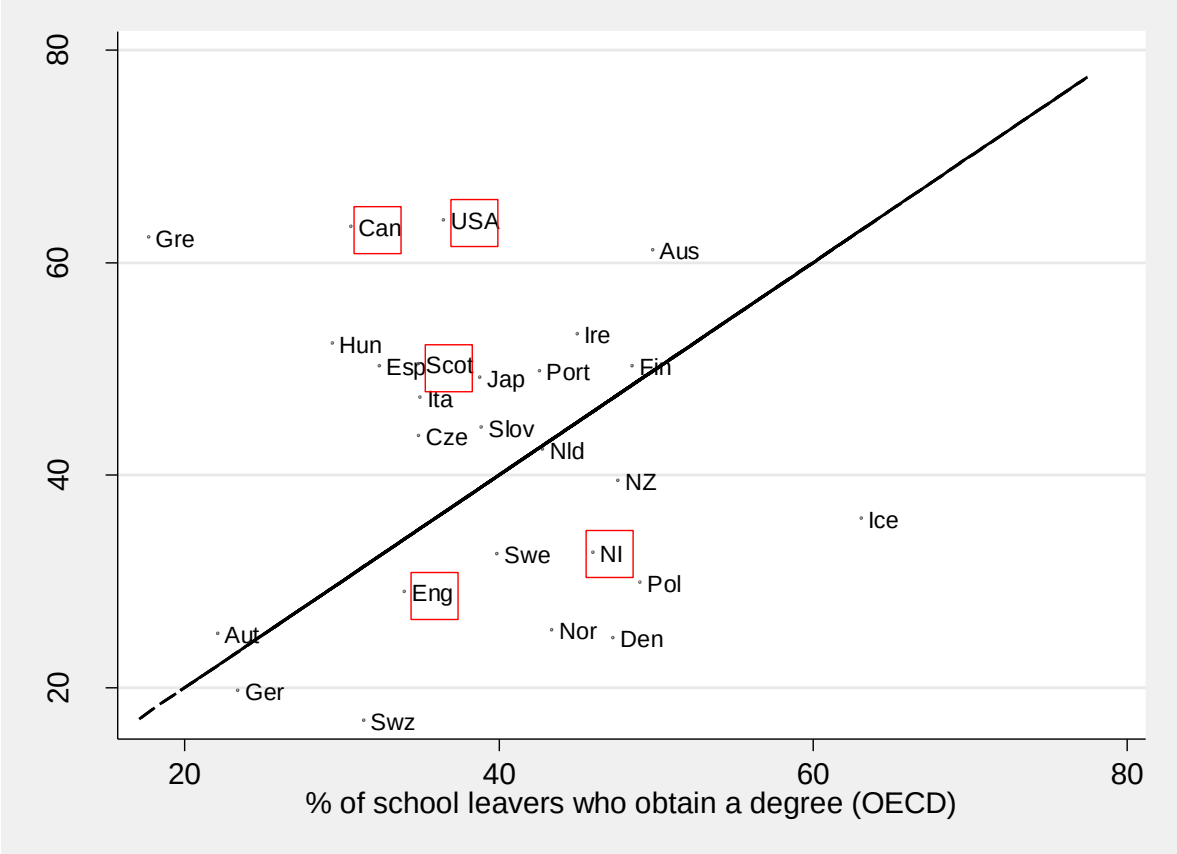


Figure 2. Proportion of children expecting to obtain a degree versus actual graduation rates



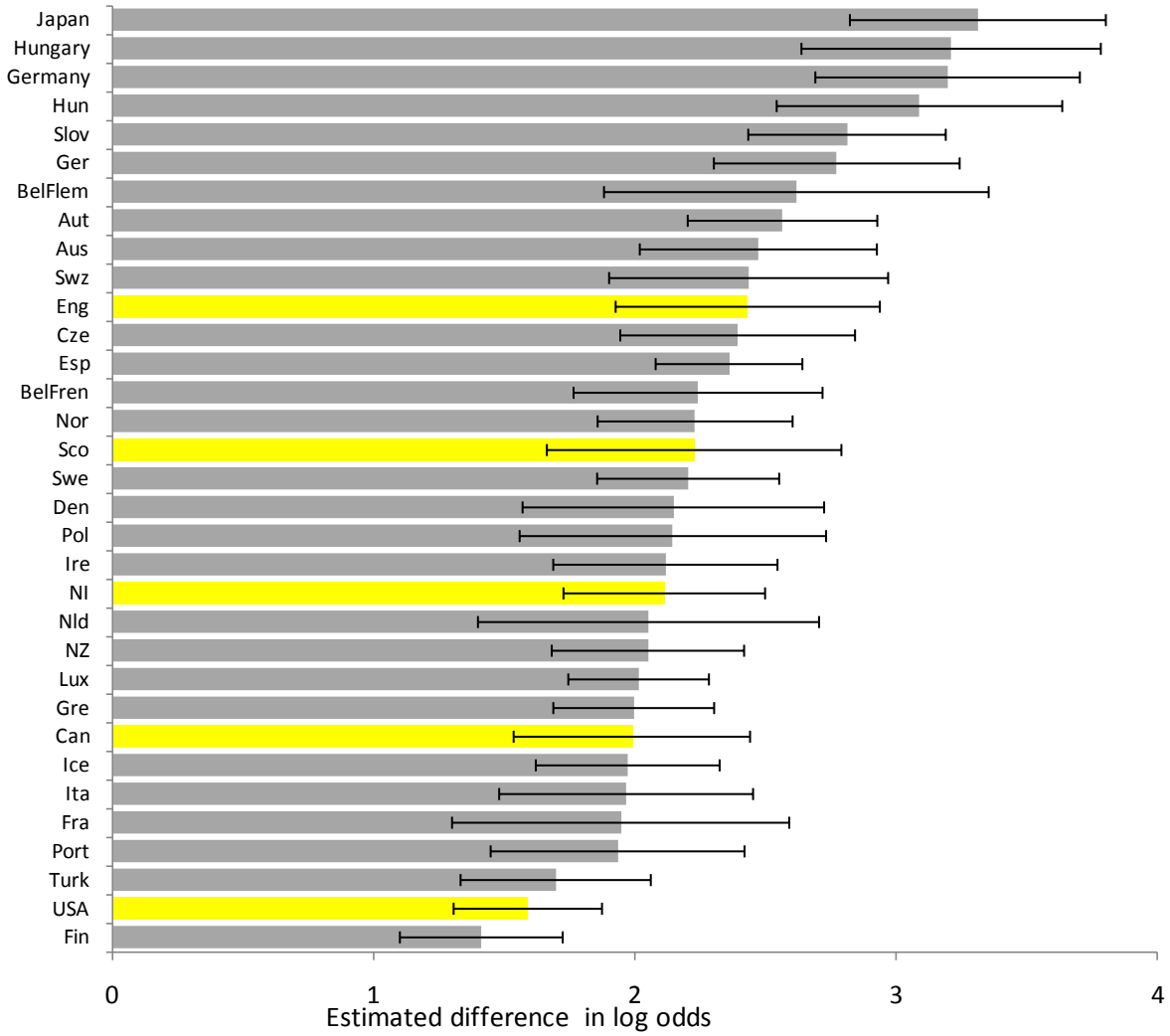
Notes:
1 Data on the % of school leavers who obtain a degree (x-axis) has been drawn from OECD Education At A Glance Report (2009), Table A3.2, page 74. This refers to net graduation rates (*i.e.* as the sum of age-specific graduation rates). See Annex 1 of OECD (2009) for further details. Information on Mexico, Luxemburg, Korea, France and Belgium not available in this data. Data on the proportion of children who expect to obtain a degree (y-axis) is based on my calculations from the PISA 2003 data.

2 Data is only available for the UK as a whole (not separately for England, Scotland and Northern Ireland) in the OECD (2009) report. Hence I use data on higher education participation for these countries Data taken from:
http://www.dcsf.gov.uk/rsgateway/DB/SFR/s000716/SFR10_2007v1.pdf for England
http://www.scotland.gov.uk/Publications/2009/11/20112425/4 for Scotland
http://www.delni.gov.uk/he-api0607.pdf for Northern Ireland

3 To calculate statistical significance I have compared the proportion of children expecting to complete university (drawn from the PISA data) to the OECD 'Education At A Glance' figures of actual graduation rates. I assume the latter refer to the population, hence conduct a one sample test of the PISA figures against these values.

4 Distance from the 45 degree line (where average expectations = actual graduation rates) are statistically significant in all countries at the 1% level except Germany, Austria, Netherlands, Slovakia, Czech Republic and Finland

Figure 3. Estimated difference between advantaged and disadvantaged children's plans to complete higher education (based on model 1)



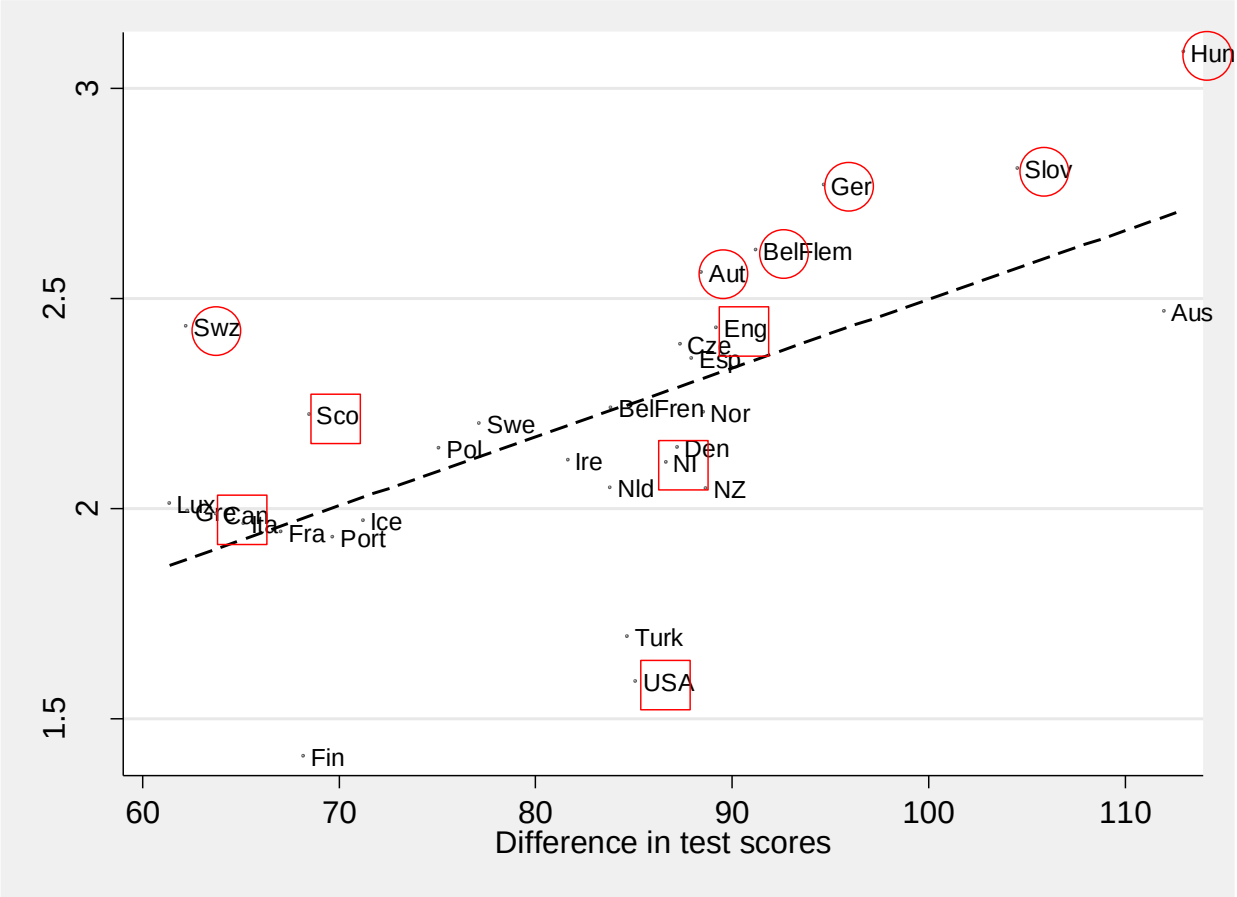
Notes:

1 Results are based upon predictions from the regression model that I describe in section 3. These predictions are based upon measures of highest parental education, highest parental occupation and number of books in the home. Other controls include gender, immigrant status and an interaction between immigrant status and the three measures of advantage listed above (note that children's PISA test scores are NOT included as a covariate).

2 The thick, solid bars represent the difference between advantaged and disadvantaged children's expectations of completing university, as measured in log-odds. The thin black line at the ends of these bars illustrates the 95% confidence interval of this estimate.

3 Country names corresponding to abbreviations can be found in the first column of Table 1

Figure 4. Estimated difference between advantaged and disadvantaged children's plans to complete higher education versus difference in their scores in the PISA test



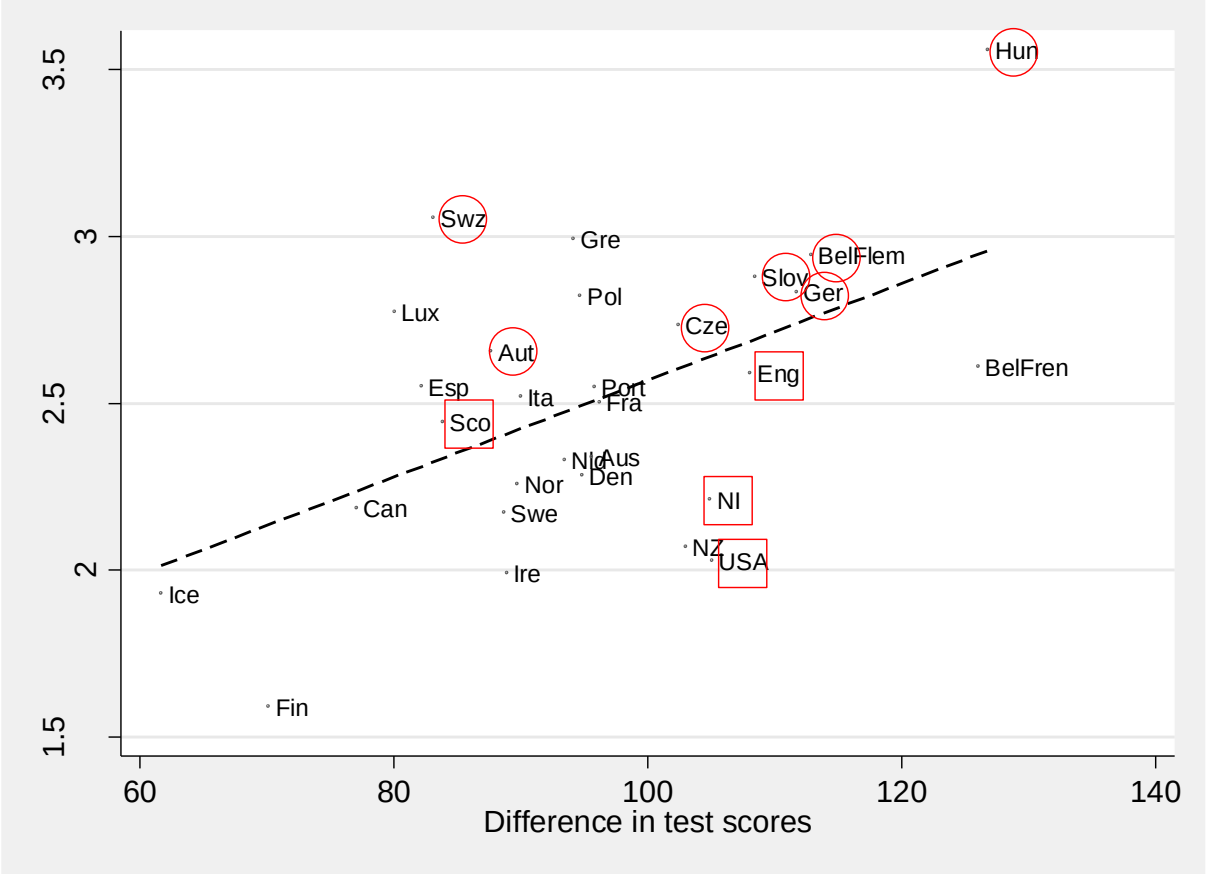
Notes:
1 Results on the y-axis are based upon predictions from the logistic regression model of children's educational expectations that I describe in section 3. Specifically, this is the difference in educational expectations between advantaged and disadvantaged groups in terms of log-odds. These predictions are based upon measures of highest parental education, highest parental occupation and number of books in the home. Appendix 1 provides further details. Other controls include gender, immigrant status and an interaction between immigrant status and the three measures of advantage listed above. Note that these results are drawn from the first model specification and so do not include information on the children's PISA test scores as a right hand side variable.

2 Results on the x-axis are based upon predictions from an OLS regression model of children's age 15 PISA maths test score. The covariates in this model are exactly the same as those that enter the logistic regression of educational expectations described in note 1 above. Figures refer to the number of points difference between advantaged and disadvantaged groups on the PISA test. For example, advantaged children score roughly 85 points more on average than their disadvantaged peers in the US (note: 100 PISA points =1standard deviation).

3 Country names corresponding to abbreviations can be found in the first column of Table 1

4 Squares refer to the countries of particular interest in this paper. Circles indicate tracking countries that sit above the line of best fit

Figure 5. Robustness test using national ESCS quartiles



Notes:
1 Results on the y-axis are based upon predictions from the logistic regression model of children's educational expectations that I describe in section 5. Specifically, this is the difference in educational expectations between advantaged and disadvantaged groups in terms of log-odds. These predictions are based upon the difference between the top and bottom national quartile of the ESCS measure of family background. Other controls include gender, immigrant status and an interaction between immigrant status and the three measures of advantage listed above. Note that these results are drawn from the first model specification and so do not include information on the children's PISA test scores as a right hand side variable.

2 Results on the x-axis are based upon predictions from an OLS regression model of children's age 15 PISA maths test score. The covariates in this model are exactly the same as those that enter the logistic regression of educational expectations described in note 1 above. Figures refer to the number of points difference between advantaged and disadvantaged groups on the PISA test (e.g. advantaged children score roughly 85 points more on average than their disadvantaged peers in the US). A 100 PISA points difference refers to a 1 international standard deviation change.

3 Country names corresponding to abbreviations can be found in the first column of Table 1

4 Squares refer to the countries of particular interest in this paper. Circles indicate tracking countries that sit above the line of best fit

Appendix 1. Calculation of socio-economic expectation gap in section 4

In section 4 I present the difference between the educational expectations of a hypothetical "advantaged" and "disadvantaged" child (in terms of log-odds). Recall that these hypothetical children differ in terms of highest parental education, highest parental occupation and number of books in the home as described in section 3.

Parental education and books in the home are dummy variables, where the reference group refers to characteristics of the disadvantaged child (less than 100 books, neither parent completed any more then compulsory schooling). One must sum the relevant coefficients (those on the "high" books and "high" education dummies) to form this part of the prediction of the difference between the hypothetical advantaged and disadvantaged children.

Adding in the contribution of parental occupation to these predictions is a little trickier. Recall that parental occupation is based upon the continuous ISEI index. Also recall that I define my hypothetical "advantaged" child as having a parent at the (national) 75^{th} percentile of this continuous index, while a "disadvantaged" child is defined as having highest parental occupation at the (national) 25^{th} percentile. Hence to add this into the predicted difference between "advantaged" and "disadvantaged" groups, one needs to know:

(1) How many ISEI points there are between the 25^{th} and 75^{th} national percentile
(2) How a one point increase in ISEI changes a child's expectations

Note that BOTH of these factors may differ across countries.

If the ISEI index had been entered as a single, simple linear term, one would simply multiply (1)*(2) above to calculate the contribution parental occupation makes to the difference between advantaged and disadvantaged groups. However, I have entered the ISEI index as a series of piece-wise linear components, with a knot at the 50^{th} percentile. Hence one must break point (1) and point (2) above into two further components:

(1a) How many ISEI points there are between the 25th and 50th percentile of the national ISEI distribution

(1b) How many ISEI points there are between the 50th and 75th percentile of the national ISEI distribution

(2a) How a one point change in ISEI between the 25th and 50th national percentile alters a child's expectations

(2b) How a one point change in ISEI between the 50th and 75th national percentile alters a child's expectations

Now, to calculate the contribution of occupation to the advantaged-disadvantaged expectation gap, once must sum {(1a*2a) + (1b*2b)}. Information on 1a and 1b (i.e. percentiles of the ISEI distribution) can be easily calculated from Appendix Table 2 (1a is, for example, simple the 50th percentile minus the 25th percentile- for England this is 51 minus 35 which equals 16). Details on 2a and 2b are provided in the Appendix Table 5 for the US and Appendix Table 6 for England (e.g. under the label "occupation spline 26-50th percentile coefficient" for 2b).

To summerise, to get the difference between the expectations of advantaged and disadvantaged children for models 1-3, one must sum[23]:

Parental education "high" coefficient
+
Books in the home "high" coefficient
+
Occupation spline 26-50th percentile coefficient * Number of ISEI points between 25th and 50th percentile
+
Occupation spline 51-75th percentile coefficient * Number of ISEI points between 50th and 75th percentile

[23] The coefficients in the tables below refer to a one point increase in the ISEI index over the relevant range (e.g. so the coefficient for "Occupation spline 26-50th percentile" refers to how the log-odds of expecting to go to university change with a one point increase in the ISEI scale that occurs between the 26th and 50th national percentile).

A worked example for England is given below (and correspond to results presented in section 4 – e.g. Figure 4)

	Educational expectations (no test score control)	**Test scores**
Coefficient on High (over 100) Books (Ref: Low)	0.86	40.3
+		
Coefficient on High (tertiary) Parental Education (Ref: Low)	0.70	7.24
+		
Occupation spline 26-50th percentile coefficient * Number of ISEI points between 25th and 50th percentile	0.027*16	1.31*16
+		
Occupation spline 51-75th percentile coefficient * Number of ISEI points between 51th and 75th percentile	0.029*15	1.39*15
=		
Difference (in log odds) between advantaged and disadvantaged children's expectations	2.43	89.3

1 Source: Author's calculations using PISA 2003 data. Sample size = 3,817.

Appendix Table 1. Distribution of highest parental education across OECD countries

	% None	% ISCED 1	% ISCED 2	% ISCED 3B or 3C	% ISCED 3A	% ISCED 4	% ISCED 5B	% ISCED 5A +	% Missing
Turkey	4	31	20	1	23	0	7	13	0
Austria	0	1	5	34	9	6	29	14	2
Poland	1	0	2	21	42	12	7	15	0
Northern Ireland	0	1	11	24	4	19	17	17	7
Portugal	19	20	16	3	15	0	7	17	2
Mexico	8	18	25	3	13	0	14	19	1
Switzerland	2	2	18	23	7	6	21	19	3
Ireland	1	5	10	0	17	27	20	19	2
England	1	1	6	22	5	20	18	20	9
Denmark	1	0	8	8	12	10	36	20	4
Italy	0	2	22	5	16	19	13	20	1
New Zealand	3	1	5	13	9	18	20	20	12
Luxembourg	4	9	2	6	8	13	24	20	13
Germany	5	1	8	19	5	15	15	23	10
Norway	0	0	3	4	6	22	36	24	4
Hungary	0	0	7	20	16	24	7	24	1
Slovakia	1	0	2	14	38	17	3	25	1
France	2	2	12	22	21	0	11	25	6
Iceland	0	2	10	9	11	26	14	26	2
Spain	3	18	7	2	16	10	13	27	5
Greece	0	8	12	4	18	16	13	27	0
Czech Republic	0	0	1	21	37	7	2	28	4
Korea	2	5	14	11	31	0	7	30	1
Scotland	3	0	5	17	14	0	23	30	8
Belgium(Flemish)	1	2	4	4	15	17	20	30	7
USA	1	1	4	0	29	16	13	34	3
Canada	0	1	4	0	19	15	21	34	6
Finland	0	3	7	0	21	3	30	36	1
Belgium(French)	2	3	4	4	12	10	21	36	8
Sweden	2	1	7	7	21	0	21	37	5
Australia	1	1	11	2	16	13	13	39	3
Japan	0	3	3	6	30	0	17	41	0
Netherlands	1	4	10	0	6	27	0	45	7
OECD	**3**	**6**	**10**	**8**	**17**	**11**	**16**	**26**	**4**

Notes:
1 Data sorted by the percentage of children who reported either parent as holding an ISCED level 5A+ qualification
2 Figures refer to row percentages.
3 ISCED level 0 refers to no formal school, level 1 is equivalent to primary education only, level 2 is lower secondary education, level 3B/3C refers to basic vocational education, 3A is upper secondary education, level 4 is post secondary education (either short vocational courses of preparation for tertiary education), level 5B is specialised vocational education, level 5A is a university education (bachelors degree), while level 6 refers to doctorates.

Appendix Table 2. Distribution of highest parental occupation (ISEI index) across countries

	Percentile							
	10th	**25th**	**50th**	**75th**	**90th**	**mean**	**SD**	**% Missing**
Mexico	24	28	33	54	69	42	19	5
Turkey	23	29	45	49	66	42	15	12
Portugal	26	30	39	51	69	43	16	3
Poland	23	33	43	53	67	45	15	2
Spain	25	30	43	54	70	45	17	4
Greece	26	31	46	56	69	46	17	6
Korea	29	37	45	51	69	46	13	3
Austria	27	34	45	56	69	47	16	4
Hungary	30	38	45	56	69	48	15	6
Ireland	29	34	49	57	69	48	16	4
Italy	29	34	49	56	70	48	16	2
Luxembourg	29	34	50	56	69	48	17	4
Switzerland	29	34	48	55	69	48	16	3
Northern Ireland	29	34	48	59	69	48	17	6
Denmark	29	38	51	57	69	49	15	3
France	29	34	51	59	70	49	17	4
Belgium(French)	29	37	51	67	70	50	17	6
Germany	30	38	51	59	70	50	16	9
Japan	33	38	45	55	69	50	15	11
Slovakia	30	37	50	64	69	50	16	4
England	29	35	51	66	70	50	17	7
Belgium(Flemish)	29	35	51	66	70	51	17	5
Canada	29	38	51	65	69	51	16	7
Finland	29	34	51	67	71	51	17	1
Sweden	30	38	51	66	70	51	16	3
Scotland	30	40	51	66	70	51	16	4
Czech Republic	33	42	51	64	69	52	15	4
Netherlands	30	39	51	67	70	52	16	7
New Zealand	29	40	51	66	69	52	16	16
Australia	30	43	52	69	69	53	16	5
Iceland	29	43	53	67	71	54	17	2
USA	30	40	56	67	71	54	16	6
Norway	34	43	53	69	71	55	15	3
OECD	**28**	**34**	**49**	**59**	**69**	**48**	**17**	**5**

Notes:
1 Data refers to points on the ISEI scale of occupational status, as described in section 3. On this scale, higher values indicate a more prestigious occupation.
2 Countries sorted by mean ISEI score

Appendix Table 3. Distribution of the number of books in the home across OECD countries

	% 0-100 books	% Over 100 books	% Missing
Mexico	86	10	4
Turkey	79	18	3
Portugal	68	31	2
Greece	65	34	2
Belgium(Flemish)	59	38	4
Scotland	58	40	2
Northern Ireland	58	40	2
USA	58	40	2
Ireland	58	40	2
France	57	41	2
Poland	58	41	1
Switzerland	56	41	2
Netherlands	55	42	3
Austria	56	42	2
Luxembourg	54	44	2
Italy	54	45	1
Japan	54	45	1
Denmark	52	45	3
Slovakia	54	45	1
England	50	45	5
Belgium(French)	51	46	4
Finland	52	47	1
Germany	46	48	6
Canada	43	49	8
Korea	51	49	0
New Zealand	47	51	3
Spain	47	52	1
Australia	41	56	2
Sweden	40	58	2
Hungary	41	58	1
Norway	36	61	2
Iceland	36	63	2
Czech Republic	33	63	4
OECD	**55**	**42**	**3**

Notes:
1 Data refers to row percentages
2 Data sorted by % over 100 books

Appendix Table 4. The proportion of children define as "advantaged" / "disadvantaged"

	Main analysis		Robustness analysis	
	% Disadvantaged"	% Advantaged	% Disadvantaged"	% Advantaged
Norway	3	19	25	25
New Zealand	6	16	25	25
Belgium(French)	7	17	25	25
Iceland	7	15	25	25
Netherlands	8	14	25	25
Australia	8	16	25	25
England	9	13	25	25
Belgium(Flemish)	9	12	25	25
Sweden	9	17	25	25
Denmark	9	15	25	25
Finland	9	17	25	25
Germany	9	15	25	25
Japan	10	15	25	25
Czech Republic	10	19	25	25
Hungary	10	17	25	25
Scotland	11	15	25	25
Canada	11	13	25	25
USA	11	12	25	25
Spain	11	17	25	25
Ireland	11	13	25	25
Luxembourg	11	16	25	25
Italy	12	14	25	25
Slovakia	12	13	25	25
Austria	13	13	25	25
Northern Ireland	14	13	25	25
Greece	14	13	25	25
Korea	16	14	25	25
Switzerland	16	12	25	25
France	16	12	25	25
Mexico	17	5	25	25
Turkey	19	7	25	25
Poland	19	12	25	25
Portugal	20	11	25	25

Appendix Table 5. Parameter estimates US

Model of ..	Educational Expectations				Test Scores	
	No test score control		With test score control			
	Beta	SE	Beta	SE	Beta	SE
Gender (Ref: Girl)						
Boy	-0.32	0.07	-0.38	0.07	8.53	2.37
Books (Ref: Low)						
High	0.58	0.08	0.32	0.08	48.31	3.29
Missing	-0.06	0.32	-0.03	0.31	-8.72	14.70
Parental Education (Ref: Low)						
Medium	0.13	0.11	0.15	0.12	-2.97	4.37
High	0.93	0.08	0.94	0.09	7.15	3.45
Missing	0.04	0.29	0.21	0.29	-29.86	9.95
Occupation spline 0-10th percentile	0.009	0.028	-0.010	0.029	3.544	1.398
Occupation spline 11-25th percentile	0.035	0.014	0.028	0.014	1.201	0.598
Occupation spline 26-50th percentile	0.002	0.012	-0.005	0.012	1.303	0.498
Occupation spline 50-75th percentile	0.005	0.013	-0.002	0.013	1.153	0.515
Occupation spline 76-90th percentile	0.139	0.057	0.092	0.058	8.019	2.015
Occupation spline 91-100th percentile	-0.015	0.016	-0.005	0.016	-1.374	0.529
Immigrant Status (Ref: Native)						
Immigrant	-0.43	1.49	-0.74	1.46	65.17	54.19
Books*Immigrant						
High Books, Immigrant	0.58	0.08	-0.14	0.21	1.13	7.04
Missing Books, Immigrant	-0.06	0.32	-0.80	0.69	-8.72	14.70
Parental Education*Immigrant						
Medium Education, Immigrant	0.20	0.23	0.16	0.24	10.09	9.57
High Education, Immigrant	0.17	0.21	0.16	0.22	1.70	7.56
Missing Education, Immigrant	-0.44	0.59	-0.60	0.62	25.13	21.24
Occupation 0-10 * Immigrant	0.03	0.05	0.04	0.05	-2.60	1.93
Occupation 11-25 * Immigrant	0.00	0.03	0.00	0.03	0.06	1.09
Occupation 26-50 * Immigrant	0.01	0.02	0.01	0.02	0.92	0.99
Occupation 51-75 * Immigrant	-0.05	0.03	-0.04	0.03	-2.42	1.29
Occupation 76-90 * Immigrant	0.16	0.13	0.14	0.13	7.26	5.39
Occupation 91-100 * Immigrant	-0.04	0.04	-0.04	0.04	-0.07	1.08
Constant	-0.66	0.77	-2.24	1.16	325.29	40.55
Test scores Controlled	No		Yes		NA	
Regression technique	Logistic		Logistic		OLS	
Response variable	Educational Expectations		Educational Expectations		PISA Maths test score	

Calculation of the advantaged-disadvantaged gap			
High Parental Ed	0.93	0.94	7.15
High Books	0.58	0.32	48.31
Occupation 26-50th percentile * 13	0.02	-0.06	16.94
Occupation 51-75th percentile * 11	0.06	-0.02	12.68
Advantaged - disadvantaged gap	**1.59**	**1.19**	**85.08**

Appendix Table 6. Parameter estimates England

Model of ..	Educational Expectations				Test Scores	
	No test score control		With test score control			
	Beta	SE	Beta	SE	Beta	SE
Gender (Ref: Girl)						
Boy	-0.46	0.10	-0.60	0.10	6.10	4.00
Books (Ref: Low)						
High	0.86	0.11	0.49	0.11	40.33	3.43
Missing	0.35	0.53	0.41	0.50	-11.96	18.93
Parental Education (Ref: Low)						
Medium	-0.09	0.16	0.06	0.16	-11.35	4.46
High	0.70	0.12	0.71	0.13	7.24	4.27
Missing	-0.53	0.36	-0.18	0.36	-34.37	7.55
Occupation spline 0-10th percentile	0.07	0.03	0.08	0.04	0.65	0.87
Occupation spline 11-25th percentile	0.00	0.05	-0.03	0.05	2.34	1.38
Occupation spline 26-50th percentile	0.03	0.01	0.01	0.01	1.31	0.39
Occupation spline 50-75th percentile	0.03	0.02	0.02	0.02	1.39	0.62
Occupation spline 76-90th percentile	0.04	0.05	0.00	0.05	3.90	1.71
Occupation spline 91-100th percentile	0.02	0.01	0.02	0.02	-0.26	0.69
Immigrant Status (Ref: Native)						
Immigrant	4.54	2.19	7.24	2.41	-172.45	45.75
Books*Immigrant						
High Books, Immigrant	-0.46	0.22	-0.71	0.23	40.33	3.43
Missing Books, Immigrant	0.03	0.91	0.58	1.16	-11.96	18.93
Parental Education*Immigrant						
Medium Education, Immigrant	-0.01	0.33	0.07	0.38	-10.70	10.35
High Education, Immigrant	0.13	0.25	0.21	0.28	0.10	8.75
Missing Education, Immigrant	0.56	0.62	0.26	0.66	29.71	16.37
Occupation 0-10 * Immigrant	-0.11	0.08	-0.20	0.09	6.17	1.77
Occupation 11-25 * Immigrant	0.10	0.11	0.08	0.13	2.96	3.50
Occupation 26-50 * Immigrant	-0.06	0.03	-0.04	0.03	-2.00	1.01
Occupation 51-75 * Immigrant	-0.03	0.03	-0.03	0.03	-0.98	1.35
Occupation 76-90 * Immigrant	0.00	0.09	-0.01	0.10	1.48	3.80
Occupation 91-100 * Immigrant	-0.02	0.03	-0.03	0.03	0.58	1.29
Constant	-4.01	0.90	-14.62	3.23	436.30	22.59
Test scores Controlled	No		Yes		NA	
Regression technique	Logistic		Logistic		OLS	
Response variable	Educational Expectations		Educational Expectations		PISA Maths test score	

Calculation of the advantaged-disadvantaged gap			
High Parental Ed	0.70	0.71	7.24
High Books	0.86	0.49	40.33
Occupation 26-50th percentile * 16	0.44	0.22	20.91
Occupation 51-75th percentile * 15	0.44	0.34	20.81
Advantaged - disadvantaged gap	**2.43**	**1.76**	**89.29**