

RESEARCH

Open Access



How long is the shadow? The relationships of family background to selected adult outcomes: results from PIAAC

Henry Braun* 

*Correspondence:
braunh@bc.edu
Boston College, Chestnut Hill,
MA, USA

Abstract

Background: Ongoing interest in the relationships between family background and adult outcomes is motivated by concerns regarding the intergenerational transmission of advantage/disadvantage. Currently all countries are far from achieving the ideal that all individuals, irrespective of their starting points or their demographic characteristics, are able to accumulate sufficient human capital so that they can achieve success in the workplace and fulfill their responsibilities as family members and as citizens. This study quantifies the length of the shadow cast by family background and personal characteristics on an individual's prospects in the labor market. It also examines the extent to which these relationships are mediated by factors more proximal to labor market entry.

Methods: This study uses data for 21 OECD countries from the first round of PIAAC. It employs descriptive statistics, correlations and logistic regression. Two dichotomous variables are derived from each country's national annual income distribution: $Q1 = 1$ if the individual's income is in the first (lower) quartile and $Q4 = 1$ if the individual works full-time and whose income is in the fourth (upper) quartile. For each country, a nested sequence of logistic regression models are fit to ascertain the role of more proximal factors in mediating the impact of family background and demographic characteristics on these labor market outcomes.

Results: The patterns of relationships are qualitatively similar across the 21 countries, although the estimated associations vary greatly in strength. Parental education accounts for substantial proportions of the variation in respondents' Educational Attainment and Cognitive Skills. In most countries, children born to parents with lower levels of education have less than a fifty-fifty chance of exceeding that level. Family background is strongly associated with income, but the relationship is largely mediated by Educational Attainment and Cognitive Skills. Females and younger individuals have much higher odds of being in the lower quartile and much lower odds of reaching the upper quartile, even after adjusting for other variables. The magnitudes of these adjusted odds are concerning.

Conclusions: Family background and gender cast a long shadow on individuals' life prospects. Countries vary greatly in their success in mitigating these disadvantages. Formulating effective policies will depend on understanding a complex set of dynamics that surely differ among countries.

Keywords: PIAAC, Parental education, Cognitive Skills, Educational Attainment, Annual income, Logistic regression

Background

Every nation confronts the challenge of preparing its people to thrive in an increasingly competitive, global economic system. A good part of that preparation involves building the stock of human capital, which comprises Cognitive Skills and knowledge, along with interpersonal skills such as collaboration and teamwork, and character traits such as motivation, persistence, reliability and self-discipline. (Kirsch et al. 2016). Ideally, all individuals, irrespective of their family background or their demographic characteristics, should be able to accumulate sufficient human capital so that they not only can achieve success in the workplace, but also fulfill their responsibilities as family members and as citizens.

In principle, governments would like to monitor the full stock of human capital. However, to this point, only certain Cognitive Skills have been assessed systematically and in a manner that facilitates both intra-national and international comparisons. With regard to those skills, interest centers on documenting their overall distribution, as well as the differences in the distributions across sub-populations defined by combinations of such factors as gender, race/ethnicity, location and family of origin socio-economic status.

International comparisons provide: (i) Transparency to education systems that is frequently not otherwise available; (ii) Independent benchmarks that jurisdictions can use to set their own targets; (iii) Models that can be adapted to help meet the targets (Singer et al. 2018). Lack of expected progress on key national indicators, such as means of cognitive skill distributions, as well as (apparently) unfavorable comparisons based on international assessments, can serve as an impetus for policy changes, although much depends on aspects of the local context (Ritzen 2013). In the United States, for example, there are concerns with the persistence of skill gaps among groups defined by race/ethnicity, as well as the growing skill gaps associated with differences in the socio-economic status of the family of origin (Reardon 2011; Reardon and Portilla 2016). Though much of the discussion centers on changes over time, evidence from cross-sectional international comparisons is also relevant and informative.

There is ongoing interest across the Organization for Economic Cooperation and Development (OECD) in the relationships between family background on the one hand, and a range of adult outcomes on the other. This interest is motivated, in large part, by concerns regarding the intergenerational transmission of advantage/disadvantage and its impact on intergenerational social mobility.

This study investigates some key aspects of the intergenerational transmission of advantage/disadvantage: specifically, the cross-national patterns of statistical associations between indicators of family background and (i) measures of Cognitive Skills and markers of Educational Attainment; (ii) indicators of labor market success. Furthermore, the study documents the extent to which these associations are mediated by individual characteristics that are more proximal to entry into the labor market. In the process, it also examines the relationships of certain demographic characteristics and those same indicators of labor market success. In metaphorical terms, the study attempts to quantify

the length of the shadow cast by family background (and personal characteristics) on an individual's prospects in the labor market.

The study draws on first round data from the Programme for the International Assessment of Adult Competencies (PIAAC). PIAAC is the third in a sequence of periodic surveys conducted under the auspices of the OECD to examine the relationships among adults' demographic and family characteristics, Cognitive Skills, Educational Attainment, work experiences and labor market outcomes. PIAAC is a household survey that draws nationally representative samples of adults ages 16–65. The utility of PIAAC is that it offers a common framework for comparing patterns of relationships across countries and, in particular, the contribution(s) of differences in family background, Cognitive Skills, and Educational Attainment in accounting for the variation in labor market outcomes. Unlike the administrative databases that are often used to address similar questions, PIAAC provides direct measures of foundational skills such as literacy and numeracy. We employ a combination of descriptive statistics, correlations and logistic regression to examine these relationships, as well as to convey the cross-country patterns in the results.

The target population for this study comprises adults ages 25–54 in 21 countries. Family background is here represented by Parental Education and Books in the Home.¹ The labor market outcomes constitute two dichotomous variables indicating whether the respondent is in the upper quartile or in the lower quartile of the national annual income distribution.

Briefly, we show that not only does family background indeed cast a long shadow on adult outcomes, but also that gender casts one that is equally long. These results bear directly on the intergenerational transmission of advantage and disadvantage. More specifically, we find that across all countries:

- (i) Parental Education has a strong association with Educational Attainment, but the strength varies very substantially across countries.
- (ii) The well-known relationship between Parental Education and Cognitive Skills persists even when controlling for Educational Attainment, albeit in a weaker form.
- (iii) Countries also vary substantially with regard to the probability that adults' Educational Attainment equals or exceeds that of their parents'.
- (iv) In all countries, the strongest predictors of the probability that an adult achieves (or exceeds) specified levels of proficiency in both Literacy and Numeracy are Educational Attainment and Parental Education.
- (v) Family Background (Parental Education and/or Books in the Home) is strongly associated with both wage-related labor market outcomes, but the strength varies across countries and is largely mediated by Educational Attainment and Cognitive Skills.

¹ Parental Education is a relatively simple indicator for a complex construct, family socio-economic background. Thus, the apparent strength of its relationship with a criterion will tend to be enhanced by those unmeasured facets of the construct that are positively (or negatively) correlated with both variables, but attenuated by those unmeasured facets that are positively correlated with one variable and negatively correlated with the other. More comprehensive studies of the impact of family background on children's development and adult outcomes are available. See, for example, Alexander et al. (2014).

- (vi) Gender and age are strongly associated with both wage-related labor market outcomes, but the strength varies substantially across countries. These associations are NOT mediated by either Cognitive Skills or Educational Attainment—or even by Occupational Category.

Thus, adults from more disadvantaged family backgrounds or are female, or younger, are considerably less likely to achieve the upper quartile of the annual income distribution and considerably more likely to be in the lower quartile of that distribution. This is the case in all the countries studied, though some have been more successful than others in reducing these inequities. These findings substantially extend the results reported in OECD (2013b) and OECD (2016).

The article is organized as follows. The next section provides some background and a brief review of the literature, followed by a section describing the data and methods employed. The next two sections present results: the first offers findings related to Cognitive Skills and the second findings related to labor market outcomes. The final sections comprise a summary and conclusions.

Literature review

As noted above, there is ongoing global interest in the extent to which education and other social services are able to mitigate the influence of differences in family background on children's success in adult life. Accordingly, many international studies have been utilized to address this question. In the main, they have shown that there are strong statistical associations between the socio-demographic characteristics of individuals' families of origin and their academic achievements, as well as a range of adult outcomes (Ermisch et al. 2012).² In this regard, the OECD has provided invaluable resources. Most recently, it has published statistical reports based on PIAAC data (OECD 2013b, OECD, 2016) that contain a broad range of relevant analyses.

For example, according to the OECD (2013b, Fig. 3.1, <http://dx.doi.org/10.1787/888932900821>) differences in certain characteristics (Age, Nativity/Language, Educational Attainment, Parental Education, and Occupation) are all strongly associated with differences in Literacy scores. The results for Parental Education are particularly revealing. The focus is on the average difference in mean Literacy scores between respondents who had at least one parent with tertiary education and respondents with neither parent having attained upper secondary. For the 21 countries in the study, the differences range from about 30 points (Estonia) to more than 50 points (United States). After adjustment for a broad constellation of factors, the differences range from about 11 points (Japan, Estonia) to about 29 points (United States).

The OECD (2013b, Fig. 3.12, <http://dx.doi.org/10.1787/888932901068>) offers further evidence of the long shadow of parental education. It displays the adjusted odds ratios for scoring at Level 2 or below on the Literacy scale (i.e. below the commonly accepted threshold of proficiency) for individuals characterized by both Parental Education and (their own) Educational Attainment.³ The odds ratios for individuals with the lowest

² OECD (2013b, p. 111–112) offers a compelling rationale for why family socio-economic background (proxied by Parental Education) should be strongly associated with individuals' Cognitive Skills.

³ The reference group is "Both respondent's and parents' education is at least upper secondary". The term *adjusted* refers to the fact that there are other variables in the model generating the odds ratios. Statistical adjustment is with respect to Age, Gender, Occupation, and Nativity/Language background.

level of Educational Attainment and whose parents were also at the lowest level of Educational Attainment ranged from about 3 (Estonia, Cyprus) to more than 10 (United States.). Thus, other things being equal, individuals in the United States at the lowest level of Educational Attainment with parents also at the lowest level of Educational Attainment, were more than 10 times more likely to score in the lowest levels of the Literacy scale than individuals whose Educational Attainment was at least upper secondary, as was that of their parents.

However, the adjusted odds ratios for individuals at the lowest level of Educational Attainment, but who had at least one parent who attained upper secondary, range from about 2 (many countries) to 4 (England/N. Ireland), with most countries falling between 2 and 3. The adjusted odds ratio for the United States is about 2.5. Thus, holding the individual's Educational Attainment constant (at the lowest level), differences in Parental Education are still strongly associated with differences in the probability that the individual falls in the lower levels of the Literacy distribution.

Some investigators have employed data from earlier OECD studies. For example, Park and Kyei (2011) used data from IALS (1994–1998) to investigate the relationships of Educational Attainment with both prose and document literacy. The focal population was adults aged 26–35. Across all 19 countries studied, they found substantial skills gaps between those individuals in the highest categories of Educational Attainment (ISCED 5–7) and those in the lowest categories (ISCED 0–2). The sizes of the gaps varied considerably across countries.⁴

The article by Massing and Schneider (2017) examines a different facet of the relationship between Cognitive Skills and Educational Attainment. One important finding was that literacy levels varied substantially within ISCED categories across countries, even when controlling for other factors related to literacy levels. This indicates that the goal of harmonization of education levels across countries has been imperfectly achieved and implies that studies involving education levels, particularly those more coarse-grained than the ISCED levels, should be done on a country-by-country basis.

There is also substantial evidence for the strong relationship between Cognitive Skills and labor market outcomes. For example, Fig. 6.4 of OECD (2013b, <http://dx.doi.org/10.1787/888932902493>) displays for each country the 25th, 50th, and 75th percentiles of the wage distributions by literacy proficiency level.⁵ In every country, the wage distributions shift to the right with increasing proficiency. However, there is also considerable overlap in the wage distributions across literacy proficiency levels. The United States appears to have the least overlap, particularly between Level 3 and Level 4/5.

An important question is the strength of the relationship between Cognitive Skills and labor market outcomes when other relevant variables are included as predictors in the model. It must be borne in mind that the statistical relationships among these variables embody complex dynamics and that in a cross-sectional observational study such as PIAAC it is impossible to disentangle the causal mechanisms at play.

⁴ Park and Kyei also investigated the extent to which, at the national level, differences in school-related factors could account for difference in gaps.

⁵ A slightly updated version of this figure can be found in Figure 5.3 of OECD (2016).

With this caveat, it is of interest to consider Fig. 6.7 (OECD 2013b, <http://dx.doi.org/10.1787/888932902550>). It displays the percentage differences in wages associated with a one standard deviation difference in Years of Education and a one standard deviation difference in literacy scores.⁶ For Years of Education, the percentage differences in wages range from about 8% (Sweden) to about 25% (Poland). The United States stands at 23%. For Literacy scores, the percentage differences in wages range from about 4% (Italy) to about 14% (England/N. Ireland). The United States stands at 12%. In all countries, the percentage differences associated with Years of Education are much greater than those associated with Literacy scores. For both variables, the results vary substantially across countries.

There has been longstanding interest among economists in delving into the relationships among Cognitive Skills and labor market outcomes. A common obstacle has been the lack of direct measures of individuals' skills that could be linked to databases that contain information on their labor market outcomes. As a fall back, economists have relied on years of schooling or indicators of Educational Attainment.

There have been some exceptions, see for example, Murnane et al. (2000) and the references therein. Murnane et al. (2000) used two United States longitudinal data sets to examine the relationships between math skills measured at the end of high school and wages, in one case, at age 27 and, in the other, age 31. For males and females separately, they fit a number of regression models that also included race/ethnicity and Educational Attainment. The results were somewhat mixed, but the general conclusion was that math skills did have a statistically significant relationship with wages even when Educational Attainment was included in the model. At the same time, differences in wages across Educational Attainment levels were very substantial: An increase of one standard deviation in math scores was associated with a wage increase that was typically less than a quarter of the difference between those with a college degree and those with a high school diploma. The authors also adduced evidence that the effect of math skills on wages was partially mediated through Educational Attainment. A similar, but more extensive, study was conducted by Lin et al. (2016).

Holzer and Lerman (2015) used PIAAC data for the United States to examine the relationships between Cognitive Skills (literacy, numeracy, problem solving in technology-rich environments [PS-TRE]) and labor market outcomes. They studied respondents aged 25–65. Literacy and numeracy scores were segmented into three categories: low proficient (levels 2 and below), proficient (level 3), and highly proficient (levels 4 and 5). For problem solving in technology-rich environments the corresponding categories were: low proficient (level 1 and below), proficient (level 2), and highly proficient (level 3).

They found that all three skills were strongly and positively associated with categories of Educational Attainment. Moreover, controlling for Educational Attainment (3 categories), both Literacy and Numeracy were strongly associated with earnings—not surprising in view of the high correlation (0.85) between Literacy and Numeracy. They also carried out a regression of the natural logarithm of earnings on demographics (Gender,

⁶ Years of education has a s.d. of 3.05 and literacy has a s.d. of 45.8.

Age, and Nativity) and the three skills, with the base categories being Low Proficiency. The results showed that, for each skill, the regression coefficients for both included skill levels were significantly different from zero. Notably, the stronger association was with Numeracy and the weaker one with Literacy. When Education Attainment was added to the model, the regression coefficients were reduced by as much as a factor of 2.⁷ The coefficients for the Literacy were no longer significant.

Economists have employed PIAAC to conduct international comparisons of the relationships between skills and Educational Attainment, on the one hand, and wage-related outcomes on the other. For example, Hanushek et al. (2013) used data from the first round of PIAAC to investigate the relationships between the natural logarithm of wages and various individual characteristics, especially Cognitive Skills. Their sample was restricted to respondents working at least 30 h per week. In all countries, they found that even with gender, years of experience and years of schooling in the model, the coefficient of Numeracy was statistically significant and substantively important. The same was the case for Literacy and PS-TRE, although the coefficients for the latter were substantially smaller than those for Literacy that, in turn, were similar to those for Numeracy.

Cappellari et al. (2015) focused on how well Numeracy and Years of Education could account for the variation in different wage outcomes. They compared results from ordinary least squares (OLS) and instrumental variables (IV) approaches in order to explore issues of endogeneity.⁸ Inasmuch as one of their goals was to relate patterns derived from PIAAC to education policy choices and governance features, they linked the PIAAC data to other databases with the requisite information. Consequently, they were restricted to using the data from only 13 European countries. In their analyses, they pooled the data from the 13 countries but incorporated country fixed effects in their models, along with variables representing a range of respondent characteristics.

Overall, they found that Years of Education was a stronger predictor of wage outcomes than Numeracy. For example, for the dichotomous outcome indicating whether the respondent's income is in the top quartile of the income distribution, the IV-estimated coefficient of the (standardized) education variable is more than twice as large as the IV-estimated coefficient of the (standardized) Numeracy variable.⁹ As we shall see, these findings, as well as those from the other studies cited, are broadly consistent with ours.¹⁰

In sum, the literature has illuminated the relationships between various individual characteristics on the one hand, and educational and labor market outcomes on the other. However, what is lacking is a systematic, temporally coherent study of the relationships among these characteristics and the outcomes of interest. This exploratory study aims to partially fill this gap. Moreover, the approach adopted here is somewhat different from that carried out by most economists. The latter tend to focus on the relationships between Cognitive Skills (or Years of Education or Educational Attainment)

⁷ Holzer and Lerman also investigated the relationship of each occupational category to earnings, with various controls in the regression model.

⁸ In order to apply the IV approach, the authors made use of another data base that contained longitudinal information on countries' education policies.

⁹ A reviewer has pointed out that the different results for Years of Education and for Numeracy could be due, in part, to differences in the magnitudes of the measurement error associated with the two variables.

¹⁰ There are a number of differences between their analytic strategy and ours, aside from their use of IV. In particular, they employ data from the full age distribution, use the average of the plausible values as the predictor variable, and appear to use least squares (rather than logistic regression) for fitting a model with a dichotomous outcome.

and labor market outcomes. Demographic characteristics and family background indicators are treated as confounding factors to be controlled for. By contrast, we begin with family background and demographic characteristics and attempt to trace the pathways through which they impact those same labor market outcomes.

Data and methods

PIAAC differs from previous rounds of international surveys of adult competencies in a number of important ways. Foremost is the introduction of computer administration of the assessment and the background survey.¹¹ Taking advantage of computer delivery, the design employed in the administration of the cognitive items was a multi-stage, stage-wise adaptive design. For further information, see OECD (2013a) and Kirsch and Lennon (2017). In point of fact, the assessment was actually administered in two modes: by computer or, for those who were deemed unable or unwilling to be assessed by computer, by paper-and-pencil. In the computer administration, respondents were assessed in Literacy and Numeracy, albeit with variable degrees of precision. A small proportion of respondents were also assessed in Problem Solving in Technology-rich Environments.¹² In the paper-and-pencil administration, respondents were assessed in Literacy and Numeracy only.¹³ All respondents completed an extensive background questionnaire.

We employ data from the first round of PIAAC, for which data collection took place during the period August 1, 2011 to November 21, 2012.¹⁴ Twenty-four countries participated in the first round of PIAAC. Two countries, Russian Federation and Cyprus, are not members of the OECD, while Australia, which is an OECD member, does not have data on the public use files and so could not be included in the analysis. Accordingly, as noted earlier, we consider the remaining 21 OECD countries. See Appendix 1 for a list of the countries comprising this subset, as well as the abbreviations used in this study.¹⁵ PIAAC sampled adults ages 16–65 at the time of the survey. For reporting purposes respondents were placed in one of five age categories: 16–24, 25–34, 35–44, 45–54, 55–65. This study focuses on the middle three categories (i.e., ages 25–54), in as much as many individuals in the youngest category are engaged in education/training, while many of those in the oldest category are in their final years of labor force participation or may be out of the labor force entirely.

We employ the country-level data files produced by the PIAAC consortium. Each country's data file is organized by individual respondent, with data on cognitive outcomes, background characteristics, education and training, labor market outcomes, and sampling-related quantities. Further information on PIAAC design, development, implementation, and post-administration processing (including output data) can be found in OECD (2013a).

¹¹ The design employed in the administration of the cognitive items is an adaptive, multi-stage design.

¹² PS-TRE is a new domain of assessment made possible by computer administration.

¹³ Respondents who demonstrate very low levels of literacy on the initial (core) module are routed to an instrument that assesses fundamental components of literacy.

¹⁴ The actual data collection period varied from jurisdiction to jurisdiction. For more detailed information see Table 10.1 in OECD (2013a).

¹⁵ Data for the United Kingdom (GBR) was based on samples from England and Northern Ireland only.

The criteria of interest in this study comprise Cognitive Skills, Educational Attainment, and two wage-related labor market outcomes. The Cognitive Skills are Literacy and Numeracy.¹⁶ They are defined as follows (OECD 2013b; Chapter 2):

Literacy

The ability to understand, evaluate, use and engage with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential.

Numeracy

The ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life.

Direct measurement of each cognitive skill is accomplished through administration of a specially designed set of instruments, such that each respondent only receives a subset of the item pool developed for that skill.¹⁷ This design improves overall construct representation but complicates the psychometric analysis. The skill values associated with a respondent are estimated by applying a country-specific, IRT/latent regression model that combines information from the respondent's answers to the cognitive items administered to her with the respondent's background information (von Davier et al. 2007; OECD 2013a). While the IRT parameters are estimated from the entire international sample, the specific country population parameters are estimated from the national samples. The output of the model consists of a set of ten plausible values (PV) for each cognitive skill for each respondent.¹⁸ The PVs are represented as numerical values on a scale of 0–500. It is these sets of PV that are subject to secondary analyses of the type reported here.¹⁹ For further information regarding PV, consult Sect. 5 of OECD (2013a) or Braun and von Davier (2017).

The two labor market outcomes examined in this study are calculated from the country-level annual income distributions.²⁰ The computation of an individual's annual income is described in (OECD 2013a, Ch. 20.4). Briefly, respondents were given the option of reporting earnings on whatever time scale that was most convenient or familiar to them. For those who did not want to report precise amounts, there was the option of reporting in broad categories. Both respondents on salary and those who were self-employed were included. Given the range of reporting formats, a great deal of work (transformations) was done to place the responses on a common scale. Ultimately, estimated monthly earnings for salaried workers (including bonuses) and

¹⁶ PIAAC also collected data on a third skill, problem solving in a technology-rich environment [PS-TRE]. We do not examine this skill in the present study.

¹⁷ More detail can be found in OECD (2013a) (Chapter 2).

¹⁸ "To increase the accuracy of the assessment of the uncertainty associated with the estimates of the population parameters of interest, PIAAC uses plausible values (PVs)—which are multiple imputations—drawn from an [estimated] posterior distribution [obtained by] combining the IRT scaling of the cognitive items with a latent regression model using information from the BQ ... in a population model." (OECD 2013a; Chapter 17, p1). "The 'plausible value' methodology correctly accounts for error (or uncertainty) at the individual level by using multiple imputed proficiency values (plausible values) rather than assuming that this type of uncertainty is zero." (OECD 2013a; Chapter 17, p2).

¹⁹ Respondents taking the computer administration are presented with two full modules that contain items linked to either one or two Cognitive Skills. However, through the latent regression model all respondents are associated with sets of PV for all three skills.

²⁰ Derived from variable YearlyIncPR from the PIAAC data files for public use.

for the self-employed were calculated, combined and then converted to annual earnings. After correcting for purchasing power parity, the distributions were represented in terms of both deciles and quartiles.

In deriving the national income distributions for this study, the PIAAC contractors decided to exclude individuals with zero incomes. Respondents with missing data were also excluded. Of course, the proportions of individuals reporting no income varies across countries and should be taken into account when interpreting the results. Appendix 2 contains the country-specific exclusion percentages. They range from 15% (Sweden) to 42% (Spain), with a median of 29.6%. The value for the United States is 28.2%.

The percentiles employed in subsequent analyses are derived from these national income distributions. We defined a dichotomous variable, Q1, to take the value 1 if the respondent's annual income was at or below the 25th percentile (lower quartile) with respect to the national income distribution and 0 otherwise. With regard to the upper end of the national income distribution, a decision was made to restrict attention to those individuals reporting they were working full time, defined as working 30 or more hours per week. This decision was motivated by the simple fact that weekly earnings are the product of the hourly wage and the number of hours worked. Thus, focusing on full-time workers enhances the utility of earnings as a criterion variable in models that are intended to elucidate the relationships between earnings and various possible explanatory factors. For example, if one of these factors (e.g. Gender) is correlated with both hourly earnings and hours worked, then there would be ambiguity in the interpretation of the gender coefficient in the wage equation. That said, this restriction does limit the generalizability of the results obtained from the reduced sample.

With this caveat in mind, we defined a second dichotomous variable for full-time workers, Q4, to take the value 1 if the respondent's annual income was at or above the 75th percentile (upper quartile) with respect to the national income distribution. Note that since the national annual income distributions were based on data from the full national samples (only excluding those respondents with zero or missing incomes), the proportions falling in the lower and upper quartiles for the analytical samples may differ substantially from 0.25.

The implications for sample sizes are as follows: The total sample size in the age categories selected for this study is 88,818. Of those, there were 65,082 respondents who had both paid income in the previous 12 months and valid income data. Disaggregated by country, it is these respondents who were assigned a "1" or "0" depending on whether their annual income was located in the lower quartile. Finally, of the 65,082 respondents there were 31,876 who reported working full-time (30 h per week, or more). They comprise the full-time (FT) sample. Disaggregated by country, it is these respondents who were assigned a "1" or "0" depending on whether their annual incomes were located in the upper quartile.

In addition to these four variables, the study investigates a number of possible explanatory factors derived from information collected through the BQ. These factors are listed in Table 1. For discrete factors the table displays the number of categories

Table 1 Factors employed in analyses, the number of categories corresponding to each factor and the reference group used in regression analyses (if appropriate)

Factor name	Number of categories	Reference group
Age	3	45–54
Gender	2	Male
Parental education	3	At least one parent with BA or BA+
Books in the home	3	101 books or more
Literacy and numeracy literacy and numeracy	Scale scores 2	N/A Score in level 2 or below
Educational Attainment	4	BA or BA+

and the reference group used in the analyses. The categories corresponding to factors other than Age are described below:

- *Gender*: male and female.
- *Parental Education*: neither parent attained upper secondary education; at least one parent attained upper secondary education; at least one parent attained tertiary education.
- *Books in the Home*: 25 books or less, 26–100 books, 101 books or more.
- *Educational Attainment*: below secondary education; secondary education; beyond secondary education but not tertiary education; tertiary education.

In earlier iterations of model fitting, we also introduced two other factors. The first, Nativity/Language, refers both to the respondent's country of birth and mother tongue. It is represented by four categories: Native born/Native language; Native born/Foreign language; Foreign born/Native language; Foreign born/Foreign language. We found that in six countries the sample sizes in some cells were so small that the model couldn't be fit. In the other countries, the associated regression coefficients were generally not significant and eliminating this factor did not materially degrade the fit. The second, Years of Experience, refers to years of paid work during one's lifetime (top-coded at 35). Again, introducing this factor did not meaningfully improve the fit. Accordingly, we have not included these two factors in the reported results.

For certain reporting purposes of the OECD, the scales for Literacy and Numeracy have been segmented into 6 ordered levels. Group performance can then be described by the percentages of the population falling in each level. The boundary between Levels 2 and 3 (occurring at 276 scale score points) typically is taken as the demarcation between those with adequate skills (i.e. proficient) and those without them.²¹ For further information on these levels, see OECD (2013b, Chapter 2). In this study, when Literacy and Numeracy serve as the criterion, they are dichotomized into indicator variables that take the value 1 if the PV scale score is at 276 or above, and 0 otherwise.²²

Because the wage-related criterion (dependent) variables are dichotomous, we employed logistic regression to examine their relationships to the explanatory factors.

²¹ Reasonably proficient is taken to mean that they possess the minimal skill levels generally required to succeed in the 21st century.

²² The process generates ten PV for each indicator variable.

Since the data associated with each respondent includes sets of ten PV for each cognitive skill, as well as the sampling weights that reflect the probability sampling design that was employed, it is strongly recommended to take these into account in a logistic regression analysis. We accomplished this by employing the IES's, IDB Analyzer (IEA 2015).²³

The basic logistic regression model takes the following form: Suppose Y is a random quantity that can take the values 1 or 0, denoting group membership. Let P_i denote the probability that individual i has the value $Y_i = 1$. Then

$$\ln\left(\frac{P_i}{1 - P_i}\right) = \beta X_i + \gamma Z_i$$

where X_i = vector of demographic characteristics (Age, Gender, ParEd, Book), Z_i = vector of acquired characteristics (Cognitive Skills, Education Attainment), β and γ = vectors of fixed coefficients to be estimated. (Note: In the baseline model, $Z_i = \emptyset$).

It is common with logistic regression to report results in terms of the estimated coefficients of the explanatory variables. These are difficult to interpret directly, however, and should be transformed into odds ratios and probabilities. Readers should also bear in mind that results from different studies are often not directly comparable as they differ in the filters used to construct the analytic databases. The filters can involve combinations of age category and employment status, as well as other factors.

When presenting the fitted coefficients of the logistic regressions we follow general practice and indicate whether the coefficients are significant at the 0.05 or 0.01 levels. The estimated standard errors do take into account uncertainty due to both measurement error and sampling error. However, given the number of models that are fit, as well as the number of predictors in each model, these indications of significance should be taken simply as guides to potentially interesting results and not be over-interpreted.²⁴

To complement the presentation of the estimated coefficients for the logistic regression models, we employ a statistic, D , the *coefficient of determination* (Tjur 2008). This statistic possesses distinct advantages over the more commonly used goodness-of-fit statistics for logistic regression, including the fact that it can be interpreted in a manner analogous to that of the R^2 statistic in ordinary least squares regression.

In the first part of this study, we examine: (i) the relationships of Parental Education to Educational Attainment, as measured by diplomas or degrees achieved, (ii) the probabilities that children's Educational Attainment exceeds that of their parents, and (iii) the relationships of Parental Education to both Literacy and Numeracy, holding Educational Attainment constant. The findings should command considerable policy interest. Ideally, education ought to reduce early differences in Cognitive Skills, including those correlated with differences in family background—with greater reductions associated with greater Years of Education and/or higher levels of Educational Attainment.

One expects substantial heterogeneity across countries in both sets of relationships. Indeed, the strength of the residual relationships between Parental Education and Cognitive Skills within Educational Attainment strata can serve as an additional

²³ The original version of the software could only carry out ordinary regression analysis. An extension completed in 2014 enabled logistic regression to be carried out as well.

²⁴ This caution reflects the problem of multiplicity, which arises whenever multiple significance tests are conducted on a dataset. Strategies to control Type I error rates are available but, for this exploratory study, would take us too far afield.

indicator of the extent to which public investment has compensated for differences in family resources to “level the playing field of opportunity” (Kirsch et al. 2016; Braun 2016; Smeeding 2016). At the same time, we must be mindful that parental education is only one of many facets of family socio-economic background that influence both the development of Cognitive Skills and the capacity of the education system to reduce those early differences.

In the second part of the study, the goal is to extend the cited results in OECD (2013b). To this end, we examine the relationships between the two indicators of family background, as well as Age and Gender, and the two wage-related variables, Q1 and Q4. We also investigate the extent to which those relationships are mediated by more proximal variables such as Cognitive Skills and Educational Attainment. As this is an exploratory study, we do not begin with a formal set of hypotheses to be tested.

Results: Parental Education, Educational Attainment and cognitive outcomes

We begin by quantifying the relationship between Parental Education and Educational Attainment within each of the participating countries, employing a well-known rank correlation coefficient, denoted by *gamma*, often used as a measure of the association between factors represented by ordered categories.²⁵ Lower correlations signal that the country’s educational system, presumably in conjunction with other social policies, is more successful in promoting educational mobility, with higher correlations signaling less success in mitigating the influence of family background on human capital development.²⁶ The weighted sample results, along with approximate standard errors, are presented in Table 2, column A. The range of gammas is 0.30–0.76, surprisingly large. The lowest, but still moderate, correlations are found in Finland, Norway, and Estonia, while the largest, very strong, correlations are found in the Slovak Republic, Italy, Poland and the Czech Republic. The value for the United States is 0.54, slightly greater than the median of 0.51.

Of course, correlations do not tell the whole story of intergenerational education mobility. There is also increasing interest in directly comparing individuals’ Educational Attainment to that of their parents.²⁷ With PIAAC data we can compute two relevant statistics for each country: (i) The probability that an individual’s educational level **equals or exceeds** that of her most educated parent and (ii) For an individual whose most educated parent’s education lies in one of the two lower categories, the probability that her education level **exceeds** that of her most educated parent. The unweighted sample results are contained in Table 2, columns B and C. Note that the denominators for the two statistics differ. In the first case, the denominator consists of the entire population (i.e. yielding an unconditional probability) while in the second case, the denominator consists of the subset of the population with Parental Education in the lower two categories (i.e. yielding a conditional probability).

²⁵ Gamma is one of a large number of measures of association proposed by Goodman and Kruskal (1954).

²⁶ We computed gamma coefficients separately for male and females. The country rankings are very similar to those in Table 2, Column A.

²⁷ The use of both correlations and probabilities to describe intergenerational education mobility is not novel. See for example de Broucker and Underwood (1998).

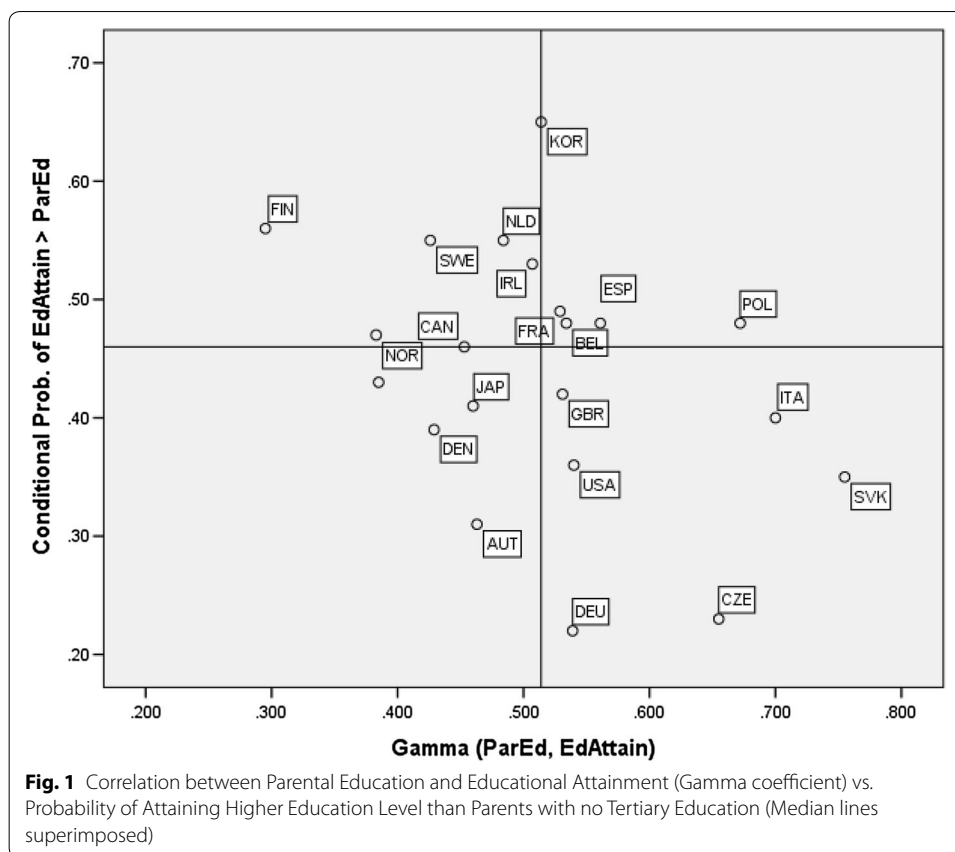
Table 2 Association between Parental Education and Educational Attainment (Gamma coefficients), Probability of Attaining Comparable/Higher Education Level than Parents and Probability of Attaining Higher Education Level than Parents

Country (ranked by [A])	Gamma for all (asymptotic standard error) [A]	Prob. of attaining comparable/ higher education level ^a [B]	Prob. of attaining higher education level ^b [C]
SVK	0.76 (.017)	0.92	0.35
ITA	0.70 (.019)	0.95	0.40
POL	0.67 (.018)	0.94	0.48
CZE	0.66 (.023)	0.88	0.23
ESP	0.56 (.019)	0.92	0.48
USA	0.54 (.020)	0.78	0.36
DEU	0.54 (.021)	0.74	0.22
BEL	0.53 (.019)	0.82	0.48
GBR	0.53 (.015)	0.82	0.42
FRA	0.53 (.018)	0.88	0.49
KOR	0.51 (.016)	0.92	0.65
IRL	0.51 (.017)	0.87	0.53
NLD	0.48 (.021)	0.86	0.55
AUT	0.46 (.021)	0.81	0.31
JPN	0.46 (.020)	0.79	0.41
CAN	0.45 (.009)	0.78	0.46
DEN	0.43 (.018)	0.73	0.39
SWE	0.43 (.022)	0.76	0.55
EST	0.39 (.017)	0.75	0.43
NOR	0.38 (.021)	0.78	0.47
FIN	0.30 (.023)	0.87	0.56
Median	0.52	0.82	0.46

^a Probability that an individual's Educational Attainment equals or exceeds the educational level of the most educated parent

^b Probability that an individual's Educational Attainment exceeds the educational level of the most educated parent, conditional on neither parent having tertiary education

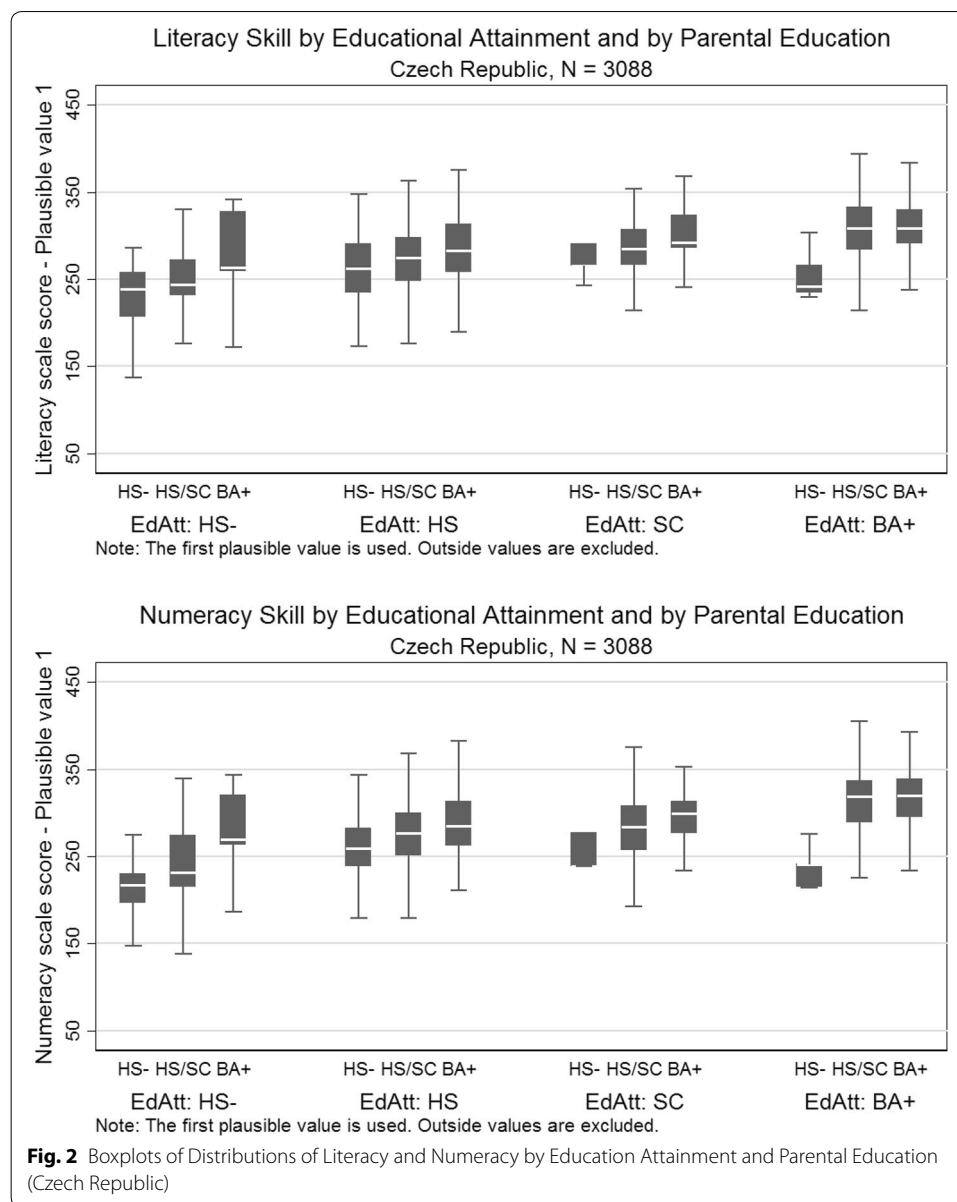
With regard to (i), the probabilities range from 0.73 to 0.95, with a median of 0.82. Thus, even in countries with the lowest probabilities (Denmark, Germany, Estonia), only about one-quarter of the population in the focal age groups lost ground with respect to Educational Attainment. At the high end (Italy, Poland, Spain, Korea, Slovak Republic) fewer than ten percent lost ground. With regard to (ii), the probabilities range from 0.22 (Germany) to 0.65 (Korea). The median is 0.46, while about half the countries have values between 0.39 and 0.49. The value of 0.36 for the United States places it in the bottom quartile. Again, higher probabilities signal countries that have been more successful in mitigating the effects of parental background—at least with respect to Educational Attainment. Generally speaking, countries with higher probabilities for (i) and (ii) had distributions of parental education that were stochastically smaller than the OECD average (i.e., parents typically had lower levels of Educational Attainment than was the case



in the OECD overall). The low value for Germany may be due to the multiplicity of alternate routes (e.g. apprenticeships) into tertiary education or the workforce.²⁸

Figure 1 displays the scatter plot of the gamma coefficients (Table 2, Column A) against the conditional probabilities (Table 2, Column C). There is a strong negative slope indicating, as one might expect, high-gamma countries tend to have lower conditional probabilities of individuals' Educational Attainment exceeding that of their (lower educated) parents, and low-gamma countries tend to have higher conditional probabilities of individuals' Educational Attainment exceeding that of their (lower educated) parents. There are, however, some apparent outliers. For example, Austria has a gamma of 0.46, just below the median, but the third lowest conditional probability. On the other hand, there are number of countries (e.g. Czech Republic, Poland, Italy and Slovak Republic) that have high gammas and low conditional probabilities, but high overall probabilities of individuals' Educational Attainment equaling or exceeding that of their parents (Column B). Further elucidation of these patterns would require consideration of the distributions of parental Educational Attainment, as well as an historical analysis of countries' education policies over two generations—well beyond the scope of this study.

²⁸ This point illustrates the importance of considering the country-specific context in explaining deviations from general patterns of results.



Parental Education is strongly positively associated with an individual's Literacy and Numeracy scores (See for example, OECD 2013b, p. 113). These results demonstrate that Parental Education is also strongly positively associated with Educational Attainment. We now take the analysis a step further by examining the relationship between Parental Education and Literacy and Numeracy, separately, within each of the four strata defined by respondents' Educational Attainment. The results are illustrated (Fig. 2) by the data from the Czech Republic, where the patterns are particularly strong for both Literacy and Numeracy. Figure 2 displays the distributions of the skill (either Numeracy or Literacy) for each combination of Parental Education (3 levels) and Educational Attainment (4 levels). The distributions are represented by box-and-whisker plots, which visually highlight five percentiles (10, 25, 50, 75, 90) and the two adjacent values.

Consider, for example, the lowest level of Educational Attainment (respondent did not complete secondary school). The difference in median Numeracy scores between respondents with the highest level of Parental Education and those with the lowest level is 61 points. For Literacy the comparable difference is 33 points. At the highest level of Educational Attainment (respondent obtained a bachelor's degree or higher), the difference in median Numeracy scores between respondents with the highest level of Parental Education and those with the lowest level is 67 points. For Literacy the comparable difference is 35 points.

In most countries, within each Educational Attainment category, there are systematic differences in the distributions of Literacy and Numeracy across levels of Parental Education, particularly between the lowest and highest levels. The differences are generally larger the lower the level of Educational Attainment, though there are a few exceptions. Interestingly, Germany and Sweden have similar ranges for the distributions of Literacy and Numeracy for each Educational Attainment category. However, although there is a moderate association with Parental Education within attainment categories in Germany, there is almost no association in Sweden. Box plots for Germany, Sweden, and the United States are contained in Appendix 3.

The results suggest that in most countries children coming from more poorly educated families suffer from a double disadvantage: They are not only less likely to attain higher levels of education, but also less likely to achieve adequate Cognitive Skills than their peers at the same level of Educational Attainment. As we shall see below, this double disadvantage is detrimental to their labor market success, as measured by wages.

In a preliminary set of analyses, we estimated for each country a logistic regression prediction model for which the criterion was a high score in both Literacy and Numeracy. Specifically, the dichotomous criterion equaled 1, if and only if the respondent achieved Level 3 or above on both the Literacy and Numeracy scales, and 0 otherwise. Such individuals are generally considered to have the foundational skills needed to achieve some degree of success in the modern world (Kirsch (2001), pp. 39–44). The explanatory factors were Age, Gender, Parental Education, and Educational Attainment. We found that across all countries the strongest predictors were the different levels of Educational Attainment followed by the lowest level of Parental Education. The coefficients of Age and Gender were not significant. Thus, it appears that the relationship between Parental Education and Literacy and Numeracy is largely, but not entirely, mediated through Educational Attainment.

Results: Parental Education and labor market outcomes

We now turn to the relationships of different individual-level characteristics to the labor market outcomes represented by Q1 and Q4. Since income distributions typically have an extreme right skew, it is common in the economics literature to use the natural log of income as the criterion in a single, ordinary least squares regression. The advantage of the present approach is that it enables us to see if there are markedly different patterns in the relationships between income and individual characteristics at the two tails of the

Table 3 Order of introduction of factors into logistic regression models

	Factors
Block 1	Age, gender, parental education, books in the home
Block 2	Literacy and Numeracy
Block 3	Educational Attainment

income distributions.²⁹ Apparently, this is the motivation in Cappellari et al. (2015) to employ Q1 and Q4 as dependent variables in a set of auxiliary analyses. Although logistic regression is not as efficient in its use of data as ordinary regression, the number of observations available in these analyses are sufficient to yield reasonably accurate estimates of the regression parameters.

In this set of analyses, initial interest centers on the strength of the association between two indicators of family background (Parental Education and Books in the Home) and Q1 and Q4. We note that the country-level correlations (gamma coefficients) between these two indicators range from 0.44 (Belgium, Japan) to 0.74 (Italy), suggesting that it is reasonable to include both indicators in the regression. By fitting a sequence of nested models, we are able to study, for each country, how the magnitudes of the regression coefficients for the two indicators change with the introduction of Cognitive Skills and Educational Attainment. Of course, changes in the coefficients of other explanatory factors are also of interest.

For both analyses, the factors displayed in Table 1 were introduced successively in blocks as indicated in the Table 3. Block 1 comprises standard demographic characteristics and family background indicators, while Block 2 comprises a single composite measure of Cognitive Skills; that is, the Literacy scale scores and the Numeracy scale scores were summed and standardized to form a single predictor. As it is expected to be the strongest predictor, Educational Attainment was introduced next (Block 3) so as not to mask the other relationships of interest.

Findings for Q4 analyses

In another set of preliminary analyses (not detailed here) we employed an analytic sample, with the only filters being that the respondent had a valid, non-zero income. A striking result was the very large, negative coefficients (Models 1, 2, and 3) associated with Gender in all countries. The straightforward interpretation is that even controlling for the other factors in the model, being female was associated with much lower odds of reaching the upper quartile of the income distribution. It seemed plausible that, to some extent, this could be due to the different proportions of males and females working full-time. Accordingly, we constructed a second analytic sample, the FT sample described above, for subsequent analyses.

The shift to the FT analytic sample forced us to drop Canada from the Q4 analyses as Canada did not collect information on full-time work. Overall, this shift resulted in

²⁹ Some investigators introduce certain interactions in the predictor set, but they are not likely to pick up these sorts of differences.

a loss of 29.3% of the data from the remaining 20 countries. The reductions by gender were 17.4% for males and 39.8% for females. These gender-related differences are of independent interest. As we have argued above, for some policy questions the FT sample should be a better basis for making gender comparisons.

Since we do not have full information on a respondent's history of labor market participation, we cannot account for differences by gender in the number of years spent in part-time work, or being out of the labor force entirely, factors that can also influence current annual income. Consequently, gender-related differences in these (and other) unobserved characteristics are confounded with the characteristics that are represented in the model. Since women are generally more likely to have worked part-time, or stopped out entirely, during their prime working years, the estimates of gender disadvantage obtained here may well be somewhat larger than would be the case if the variables corresponding to these other factors were included in the model.

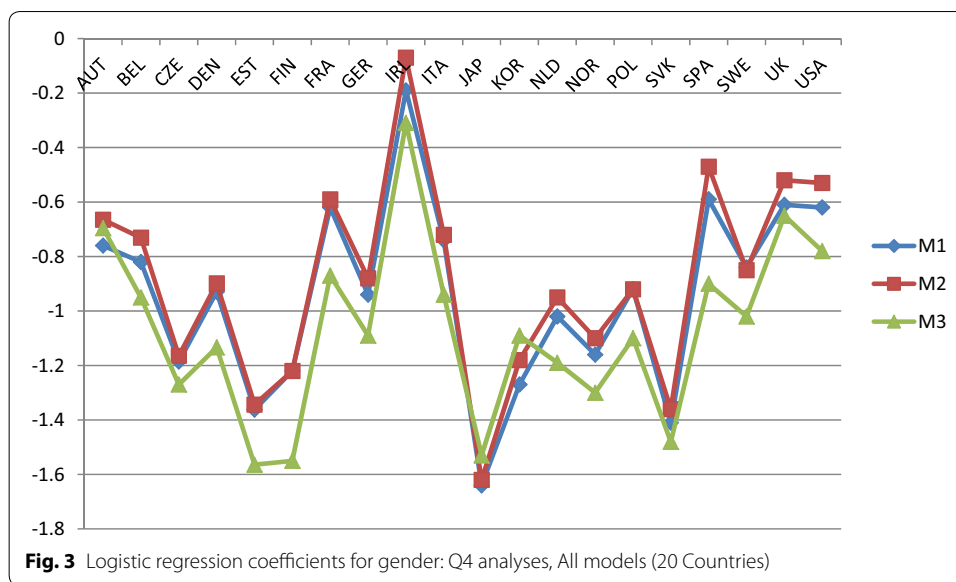
It bears mentioning that the magnitudes of the gender coefficients, though still very large, are substantially smaller in the set of analyses using the FT sample in comparison to those in the set of preliminary analyses referred to above. For example, the Netherlands has one of the largest gender coefficients. In Model 3, the gender coefficient declines from -2.16 to -1.19 , corresponding to odds ratios of 0.12 and 0.30, respectively. An odds ratio of 0.12 means that holding constant all other variables in the model, males are about 8.5 times more likely than females to achieve the upper quartile. The corresponding interpretation of the odds ratio of 0.30 is that for the FT sample the male advantage is "only" a factor of 3.3. With one of the smallest gender coefficients among the 20 countries, the decline in the U.S. is only from -0.91 to -0.78 , corresponding to odds ratios of 0.40 and 0.46, respectively. The male advantage in terms of odds ratios is thus reduced from a factor of 2.48 to 2.18.

In what follows, we discuss the upper quartile (Q4) results for the FT sample. Before turning to these results, we note that for each country the quality of the fit improves with the inclusion of additional blocks of predictors. In particular, the values of D for Model 3 range from 0.06 (Slovak Republic) to 0.30 (Germany), with a median value of 0.21.³⁰ The full set of D statistics is in Appendix 4.

Overall, we find for almost all countries, controlling for other factors in the model, significantly lower odds of reaching the upper quartile are associated with being female, being younger, or having a lower level of Educational Attainment. Higher Cognitive Skills are associated with significantly greater odds of reaching the upper quartile. All country-level results are contained in Appendix 5, to which we now turn.

With respect to the focal question, the results from fitting Model 1, which also includes Gender and Age as predictors, indicate that in all countries the coefficients associated with lower Parental Education and/or fewer Books in the Home are negative and significantly different from zero. (Recall that a negative coefficient corresponds to an odds ratio less than one relative to the reference category, while a positive coefficient corresponds to an odds ratio greater than one relative to the reference category.) Thus, these

³⁰ The low D values for the Slovak Republic are likely due to the small number of sample respondents falling in Q4.



family background characteristics are strong predictors of reaching the upper quartile of the income distribution.

The addition of Cognitive Skills (Model 2) somewhat reduces the magnitudes of the coefficients for Family Background, but they usually remain statistically significant. However, with the addition of Educational Attainment (Model 3), the coefficients retain significance only in the Czech Republic, Estonia, France, Japan, Great Britain and the United States—albeit with magnitudes further reduced. It appears, then, that the long shadow of family background typically acts through the more proximal factors comprising Cognitive Skills and, especially, Educational Attainment.

Turning to demographic characteristics, in Model 1 the coefficients of the variables representing Gender and Age are large, negative, and strongly significant. Their magnitudes are not much diminished, if at all, by the addition of other predictors (Models 2 and 3). In part, these results are expected since in most work settings younger workers do not have sufficient seniority to move into the high end of the income distribution and females' incomes typically trail males' incomes. What is somewhat unexpected are the magnitudes of the coefficients for Gender and their robustness to the inclusion of Cognitive Skills and Educational Attainment. This robustness is evident in Fig. 3, which displays for 20 countries the coefficients for Gender for Models 1, 2, and 3.

To illustrate the range of gender disadvantage, consider the results for Model 3. The greatest disadvantage occurs in Japan, with a coefficient of -1.56 corresponding to an odds ratio of 0.21. The least disadvantage occurs in Ireland, with a coefficient of -0.31 corresponding to an odds ratio of 0.73. Thus, in Japan, other things being equal, the odds of males reaching the upper quartile are nearly five times greater than those for females. In Ireland the male advantage is only about 1.4 times. The median male advantage for the countries studies is about 3 in terms of odds ratios.

Cognitive Skills enter Model 2 as significant predictors of the criterion, but the introduction of Educational Attainment reduces the magnitudes of the corresponding coefficients by about 25% on the logit scale, though they remain statistically significant

and substantively important. Note that the coefficients of Cognitive Skills cannot be interpreted in terms of odds ratios, as the variable is continuous and not categorical. It is possible, nonetheless, to make some meaningful comparisons between predictors. Reviewing the results for Model 3, we see that the impact on the log odds of a one standard deviation reduction in Cognitive Skills is typically about one-fifth to one-third as large as the shift from the highest to the lowest category of Educational Attainment.

The coefficients of the indicators for categories of Educational Attainment are all large, negative, and strongly significant.³¹ In particular, the coefficients for the lowest level of Educational Attainment correspond to odds that are very substantially lower than those of the reference category. Here, the United States is an outlier. In Model 3, the coefficient for the lowest category of Educational Attainment is -3.01 corresponding to an odds ratio of 0.05. The interpretation is that, other things being equal, the odds of individuals in the highest category of Educational Attainment reaching the upper quartile are 20 times greater than the odds for those in the lowest category.

In a supplemental analysis (not shown here), Model 3 was augmented by a set of indicators representing seven Occupational Categories. In most countries, this addition led to only slight improvements to the fit, although most of the (partial) regression coefficients for the indicators associated with the different occupational categories were statistically significant.³² The coefficients for Gender and Age were not much changed, while the coefficients for Cognitive Skills are somewhat reduced but still remained statistically significant. On the other hand, the coefficients for Educational Attainment were reduced across the board by about half on the logit scale, suggesting substantial confounding between Occupational Category and Educational Attainment. A plausible interpretation is that individuals with more education earn more because they are able to get jobs in higher paying occupations than individuals with less education.³³

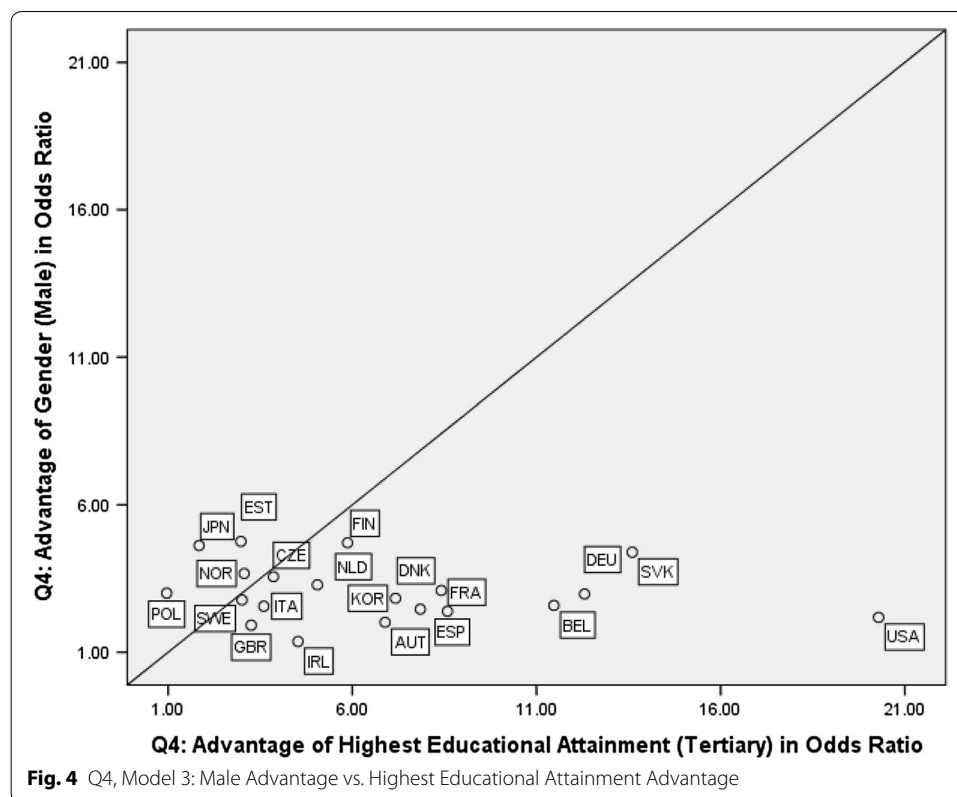
Figure 4 compares the advantage for males (relative to females) with the advantage for those in the highest category of Educational Attainment (relative to those in the lowest category). Advantage is represented by odds ratios based on the partial regression coefficients that have been adjusted for all the other variables included in Model 3. For most countries, not having finished secondary school appears to be a greater disadvantage than being female. Particularly extreme examples are the United States, Slovak Republic, Germany and Belgium. Estonia, Japan, and Poland are exceptions, with gender being a greater disadvantage. It is evident that in every country poorly educated women are extremely unlikely to reach the upper quartile.

When analogous analyses are carried out separately by gender there are substantial differences between males and females in some countries in the coefficients for Age and Educational Attainment. However, there are no systematic patterns. On the other hand, in half the countries Cognitive Skills are a stronger predictor for females, while in a quarter of the countries they are a stronger predictor for males.

³¹ In 10 countries the sample size in the lowest category of Educational Attainment is around 50 or below. In 5 of those countries, the estimation of the corresponding regression coefficients is unstable.

³² In most countries the sample sizes for some of the categories are small and the corresponding estimates are unstable.

³³ The author thanks the referee for this interpretation.



Findings for the lower quartile (Q1)

Before turning to the logistic regression results, we discuss the D statistics displayed in Appendix 6. In comparison to the D statistics in Appendix 4, the model fit statistics for Q1 are considerably smaller than those for Q4. This may be due, in part, to the exclusion of respondents with zero incomes. For Model 3, the median D value is 0.13. (For Q4, the median D value for Model 3 is 0.22). Here, Canada has the poorest fit ($D=0.016$) and Ireland the best ($D=0.29$). The weak fit statistics argue for caution in interpreting the regression results, which are presented in Appendix 7.

The patterns in the coefficients are generally similar to those obtained for Q4. Parental Education and Books in the Home are significant in almost all countries in Model 1, but in Model 3, they are significant in only about half the countries with magnitudes much reduced from those in Model 1. By contrast, the logistic regression coefficients for Gender are large and positive in Model 1 and remain so in Model 3. Thus, controlling for the other variables in the model, females have considerably higher odds than males of being in Q1. Indeed, in about half the countries, the (adjusted) disadvantage associated with being female is greater than the (adjusted) disadvantage associated with being even in the lowest category of Educational Attainment. This pattern is different from the one in Q4, where in about three-quarters of the countries the disadvantage associated with being in the lowest category of Educational Attainment is greater than the disadvantage associated with being female.

More specifically, Japan is something of an outlier with a (partial) regression coefficient for Gender ($b=2.40$) that is considerably higher than that for the Netherlands (ranked

second, $b = 1.88$) and for South Korea (ranked third, $b = 1.72$). The corresponding odds ratios are [11.0, 6.6 and 5.6, respectively]. Clearly, these values represent very high levels of gender-related earnings inequality. Countries with the smallest regression coefficients for gender are Canada, Denmark and the United States ($b = 0.79, 0.80$ and 0.81 , respectively), all corresponding to odds ratios of approximately 2.2. Although smaller, these odds ratios are still substantial. It is worth noting that the United States is an outlier in that the disadvantage associated with all categories of Educational Attainment (below the reference category) are greater than that associated with being female.

Age appears as a strong predictor in almost all countries in Model 1. As one might expect, in comparison to the reference group (aged 45–54), the younger age groups generally have significantly higher odds of being in Q1. In Model 3, the youngest cohort (aged 25–34) has significantly higher odds of being in Q1 in 14 out of 21 countries.³⁴ For the middle cohort (aged 35–44) the coefficients are significant in seven countries, with smaller magnitudes.

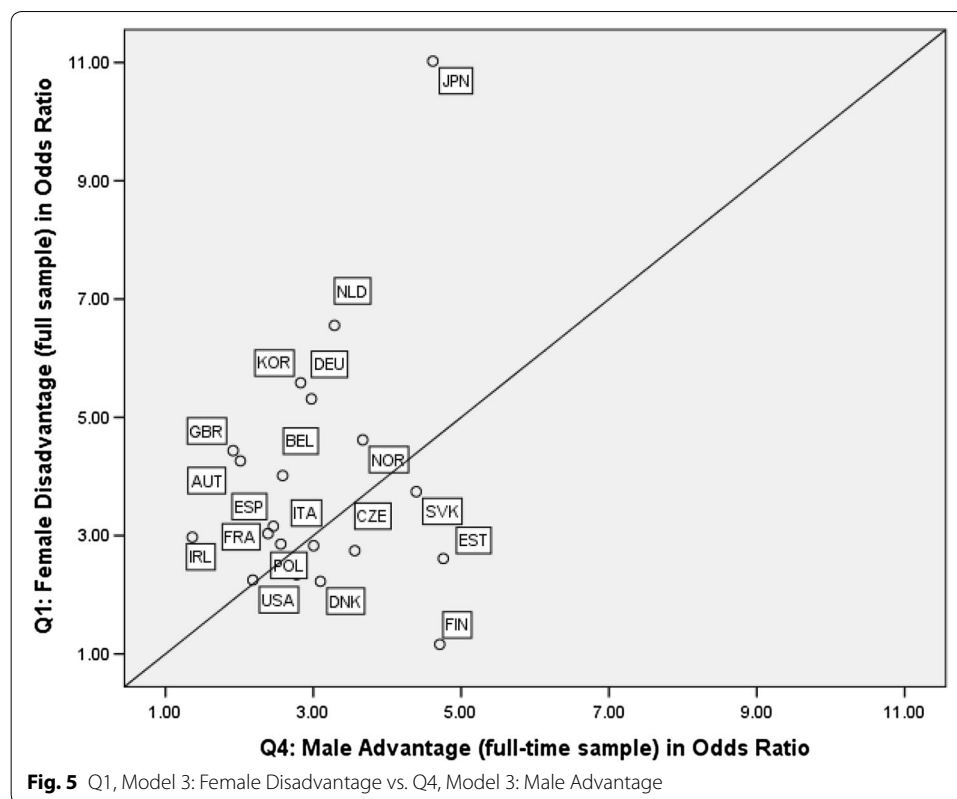
The regression coefficients for the cognitive skill variable are negative, indicating that higher proficiency is associated with lower odds of being in Q1. They are statistically significant in all countries except Canada and the U.S. The countries with the strongest results are Denmark, Norway, Belgium, and Great Britain ($b = -0.49, -0.45, -0.44, -0.43$, respectively). The impact on the log odds of a one standard deviation increase in Cognitive Skills is typically about one-fifth to one-third as large as the shift from the lowest to the highest category of Educational Attainment.

Finally, others things being equal, individuals with lower levels of Educational Attainment have greater odds of being in the lower quartile in comparison to individuals with at least a Bachelor's degree. The odds ratios are greatest in Poland, with the three lower attainment categories having odds ratios of 8.7, 4.3 and 3.1, respectively. In the United States the odds ratios corresponding to the three lower attainment categories are 5.5, 4.1 and 3.8, respectively.

Interesting differences arise in the gender-specific analyses. For males, the D statistics for Model 1 are uniformly close to zero indicating that demographic and background variables have very little explanatory power. Even with Model 3, the D statistics are very small (median = 0.045), indicating that the addition of Cognitive Skills and Educational Attainment does not add much explanatory power. Accordingly, we do not comment on the fitted models.

On the other hand, for females the D-statistics are similar to those found for models fit to all the data (median = 0.12). Comparing the Model 3 results for females to those for the full sample (Appendix 7), we find that the regression coefficients in the analyses for females only are typically larger in absolute magnitude than the corresponding regression coefficients in the analyses for all respondents. That is, females being in the youngest age group, or the lowest Educational Attainment category, or having lower Cognitive Skills corresponds to a greater probability of being in the lower quartile in comparison to the results for all respondents, with Gender included as an explanatory variable. For

³⁴ Estonia is the only country in which younger cohorts have lower odds of being in Q1. The relationship becomes insignificant in gender-specific analyses. South Korea is the only country in which younger cohorts have lower odds of being in Q1 in both gender-specific analyses.

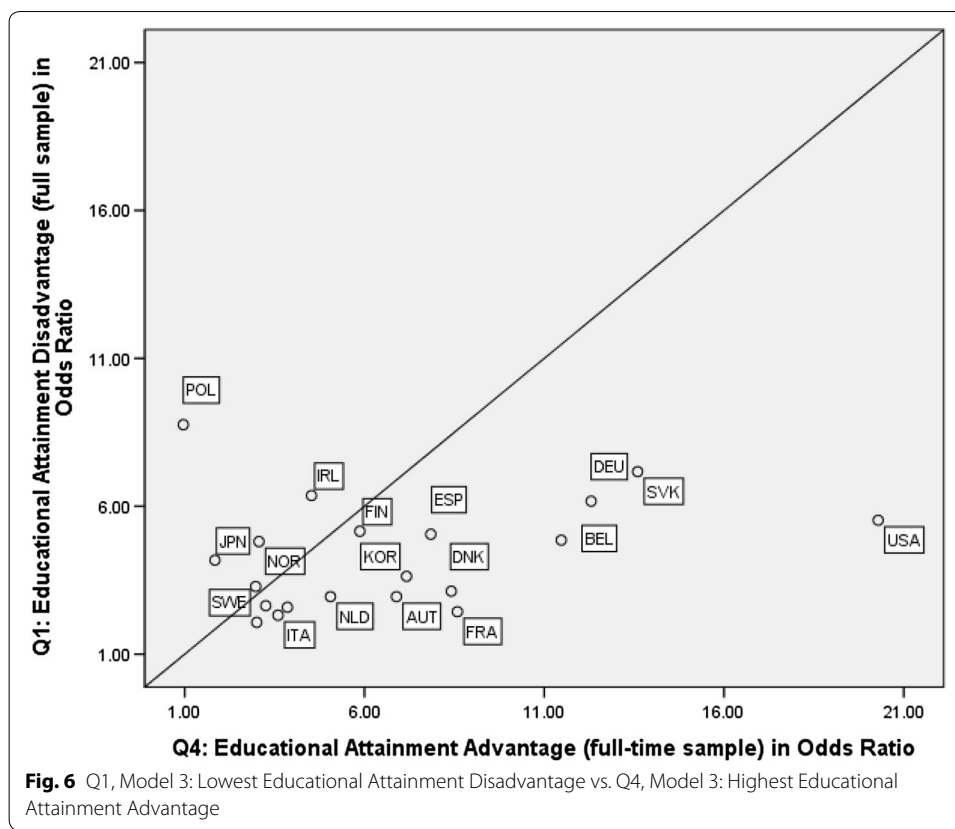


females, the disadvantage associated with the youngest age category is greatest in Finland, Poland, and Spain; with lower Cognitive Skills it is greatest in Korea, Ireland, the Netherlands, and Germany; with the lowest category of Educational Attainment it is greatest in Japan, Poland, and Belgium.

We conclude this section with two graphical displays. Figure 5 plots for each country the odds ratios of the gender disadvantage in Q1 and Q4. In order to constrain the plot to the first quadrant, the Y-axis represents the (adjusted) odds ratio for females being in Q1, but the X-axis represents the (adjusted) odds ratio for males being in Q4. (The latter is just the inverse of the female disadvantage.) In both cases, the odds ratios are derived from the estimated partial regression coefficients that have been adjusted for the other variables included in Model 3.

Countries where the gender disadvantage is approximately the same in both quartiles will fall on or near the 45° line. This is the case for most countries, with a few notable exceptions. In the Netherlands, Korea, and Germany, the female disadvantage is much greater for Q1 than for Q4, while the opposite is true for Estonia and Finland. In addition, Japan is an obvious outlier. The female disadvantage is extremely large for both quartiles, but much greater for the lower quartile. The magnitudes of these odds ratios are particularly striking—and concerning—given that they are based on partial regression coefficients and not on simple probability ratios.

Figure 6 is analogous to Fig. 5. The Y-axis represents the odds ratio for individuals in the lowest Educational Attainment category being in Q1 and the X-axis represents the odds ratio for individuals in the highest Educational Attainment category being in



Q4. Again, in both cases, the odds ratios are adjusted for the other variables included in Model 3. We see here that, with the exception of Poland, for most countries Educational Attainment matters more for Q4 than for Q1. The odds ratios for Q4 are particularly large in the United States, Slovak Republic, Germany, and Belgium.

Summary and limitations

This study contributes to the elucidation of the relationships of Parental Education to the foundational Cognitive Skills of Literacy and Numeracy, as well as the relationships between two indicators of family background to two wage-related labor market outcomes. The populations of interest comprise individuals aged 25–54 in 21 OECD countries. An earlier report (OECD 2013b) established the strong association between Parental Education and those skills for all participating jurisdictions. Here we have shown that in all 21 countries, Parental Education is also strongly associated with Educational Attainment. Neither set of results is particularly surprising—except, perhaps, that these relationships are clearly evident even in those countries, such as Denmark, France and Sweden, that have invested heavily in reducing the impact of children’s family background on their adult outcomes.

Turning to the question of intergenerational progress in Educational Attainment, we documented the fact that in all countries no more than a quarter (and typically far fewer) of the population has lost ground relative to their parents. For those whose parents’

Educational Attainment were in the lowest two categories, the probability of exceeding their parents' levels were typically between one-third and one-half.

We then examined the relationship between Parental Education and Literacy and, separately, Numeracy, while controlling for Educational Attainment. Again, we found a fairly consistent pattern across countries; namely, higher levels of Parental Education are associated with higher skill levels within each Educational Attainment stratum. Generally, the association is stronger at lower levels of Educational Attainment and more pronounced in the contrast between the lowest and highest levels of Parental Education.

Thus, complementing the long-established finding that the association between Parental Education and Cognitive Skills is largely captured through Educational Attainment, it appears that in many countries a substantial residual disadvantage accrues to those individuals whose parents are in the lowest educational category. At the same time, the finding that this residual disadvantage tends to be greater at lower levels of Educational Attainment indicates that, overall, more education does help to mitigate differences in learning outcomes associated with family background. Undoubtedly, there are multiple sources of this residual disadvantage. Most likely it reflects differences in the nature and intensity of interactions within the family, as well as differences in the quality of the education received among individuals in the same Educational Attainment category. Those differences may further impact the kind of jobs obtained, the level of cognitive demand in those jobs, as well as the opportunities to obtain further training.

Turning to the two wage-related outcomes, the analyses led to somewhat different narratives. With regard to the upper quartile criterion, for full-time workers, Parental Education and Books in the Home were generally statistically and substantively significant in Model 1 (baseline) but were much reduced in magnitude, and usually no longer statistically significant in Model 3 (baseline + Cognitive Skills + Educational Attainment). This finding is consistent with the hypothesis that the relationships between Family Background and adult outcomes are mediated by more proximal individual characteristics.

One set of results were quite unexpected: The disadvantages associated with being female, or being in the youngest age category, were very large, even after controlling for other factors. In fact, the coefficients for Gender and Age obtained in Model 1 generally did not change systematically with the addition of other explanatory variables (Models 2 and 3).

Higher levels of Cognitive Skills were strongly associated with higher odds of reaching the upper quartile of the income distribution. However, the addition of Educational Attainment resulted in substantial reductions in the regression coefficients of the cognitive skill variable. Indeed, Educational Attainment was the strongest predictor, with significant disadvantage associated with lower levels of attainment. Roughly speaking, the lower odds associated with being female, being in the age category 25–34, or with having only a secondary education, were about the same order of magnitude.

Turning to the results for the lower quartile, for all workers with non-zero incomes, we found that model fit was substantially lower than that for the upper quartile, with the fit for males being particularly poor. This phenomenon merits further investigation.

In comparison to the findings for the upper quartile, the disadvantage associated with Family Background was somewhat more robust to the addition of other factors and Gender was a relatively stronger predictor of disadvantage. With regard to Cognitive Skills,

the regression coefficients in Model 3 were statistically significant ($p < 0.01$) in almost all countries. In absolute magnitude, however, they were smaller than the corresponding coefficients in Model 3 for the upper quartile.

As noted earlier, the findings for the lower quartile should be interpreted with some caution in view of the poorer model fit. At the same time, the differences in the regimes that appear to be operating in the two tails of the income distribution raise the possibility that the standard econometric models employed in these settings may be badly misspecified. Further investigation of this conjecture is called for.

With regard to the United States, the results presented in OECD (2013b) indicate that it is somewhat of an outlier with regard to the size of the gaps associated with differences in cognitive proficiency, as well as the gradients of cognitive proficiency with respect to various covariates. In this study, we also found that the United States was an outlier: (1) Differences in Literacy and Numeracy (by Parental Education) within Educational Attainment strata were among the largest in the study, and (2) With regard to reaching the upper quartile, the difference in the coefficients of the lowest and second lowest Educational Attainment categories, corresponding to 1.5 logits, was the largest in the study. At the same time, the coefficients of the other variables were not particularly exceptional.

PIAAC shares many of the limitations of all cross-sectional studies. For example, respondents' Cognitive Skills are measured contemporaneously and are likely influenced to some degree by experiences subsequent to the completion of their formal education. This introduces endogeneity into the model leading to estimation bias. Further, there are no measures of respondents' socio-emotional skills, skills that have been shown to be predictive of labor market success (Levin 2013) and, arguably of increasing importance (Deming 2017). Thirdly, the database contains full information on labor market experiences only for the previous 12 months. Thus, systematic differences in earlier work histories (e.g. working part-time, stopping out, etc.) among focal groups complicates interpretations of group differences. The highest category of Educational Attainment includes all individuals with a tertiary degree or higher. In the United States there is evidence of a divergence in income between those with a four-year degree and those with more advanced credentials (Autor 2014). With larger sample sizes it would have been instructive to examine that divergence in all countries.

Two further cautions: With only two indicators directly related to family background, we surely have an incomplete representation of that construct. Finally, as noted earlier, there is substantial missing data on income in all countries and the range of the proportions of missing data, 15–42%, is considerable. Although the similarity across countries in the patterns of relationships between income quartile location and the explanatory variables is reassuring, comparisons of the magnitudes of the logistic regression coefficients are likely sensitive to differences in the amounts of missing data.

There are also well-known limitations to the policy implications of findings from cross-sectional data because of the difficulty in supporting causal interpretations of those findings. A number of authors have attempted to circumvent these difficulties either through methodologies for combining data from repeated cross-sectional surveys (Gustafsson 2013) or through linking different surveys in order to follow specific cohorts (Kaplan 2009; Kaplan and McCarty 2013; Chmielewski 2015). See the review

by Chmielewski and Dhuey (2017), as well as recent work by Hampf et al. (2017) and by Bind and Rubin (2017). Hampf et al. investigated a number of different approaches to establishing the causal linkage between skills and wages, concluding there was compelling evidence for such an inference. Bind and Rubin explored different strategies for embedding an observational study in a hypothetical randomized study, with the intent of comparing the resulting causal estimates across studies. These methodological developments represent important directions for strengthening the policy relevance of analyses of cross-sectional large-scale assessments surveys.

Conclusions

The overall conclusion is that Family Background, when represented by Parental Education and Books in the Home, indeed casts a long shadow on individuals' life prospects: Across all countries studied, Parental Education accounts for substantial proportions of the variation in respondents' Educational Attainment, as well as some of the differences in Cognitive Skills within Educational Attainment strata. These results complement the well-established associations between Parental Education and Cognitive Skills, such as Literacy and Numeracy. Turning to income, Family Background does help to predict who will be in the upper and lower quartiles of the national annual income distribution. However, as one might expect, the strength of the direct association is much attenuated by the addition of relevant intermediate factors (i.e. Cognitive Skills and Educational Attainment) that are more proximal in time to entry into the labor market. Being female greatly diminishes the odds of reaching the upper quartile and greatly enhances the odds of being in the lower quartile, even controlling for other factors. The strength and robustness of the gender advantage, favoring males, is somewhat surprising and should be of concern to all participating countries.

As always, appropriate policy-relevant interpretations of the statistical findings for a particular country depend to a great extent on knowledge of historical trends in economic activity, government policies, as well as considerations of culture, demography and the like (Cappellari et al. 2015). Formulating effective policies to counter the disadvantage associated with Family Background and Gender will depend on understanding a complex set of dynamics that differ markedly among countries and even regions within countries. Thus, in some countries the gender disadvantage may be largely due to the occupational choices women make in attempting to find an appropriate work-life balance. In other countries, it may be mostly due to structural bias in the labor market. In yet others, a combination of these, and perhaps other factors, offer the best explanation. A recent review (<https://www.vox.com/2018/2/19/17018380/gender-wage-gap-childcare-penalty>) cites a number of studies of the gender gap in wages, including one of Danish data that suggests that the gap widens with the birth of the first child.

With respect to the long shadow of family background in the United States, there is evidence of substantial variation by location in both economic mobility and the inter-generational transmission of advantage or disadvantage (Moretti 2013; Chetty et al. 2014; Rothstein 2017). Empirical studies suggest that this variation can be explained, to some degree, by differences by location in factors such as the distribution of socio-economic status, marriage patterns, labor markets, and the interactions among them (Rothstein 2017; Khatiwada and Sum 2016). Of course, macro-economic forces and policies

also play an important role in determining who succeeds in the labor market (Bernstein 2016).

Thus, aggregate comparative results can only suggest interesting directions for further study to inform policy decisions (Carnoy 2015; Singer and Braun 2018; Singer et al. 2018). To take full advantage of these findings, it is essential to conduct in-depth, longitudinal investigations of the context and circumstances pertaining to each country.³⁵ That said, PIAAC is unique in enabling the linkage of individuals' family backgrounds, Cognitive Skills, and life/work experiences to a variety of adult outcomes. Consequently, PIAAC and its successors will serve as indispensable tools in attempts to identify and understand the key relationships that are essential foundations for evidence-informed policy formulation.

Abbreviations

ALL: Adult Literacy and Lifeskills; BQ: background questionnaire; FT: full-time; IALS: International Adult Literacy Survey; IDB: international database; IEA: International Association for the Evaluation of Educational Achievement; IRT: item response theory; IV: instrumental variables; LMO: labor market outcomes; OECD: Organization for Economic Co-operation and Development; OLS: ordinary least squares; PIAAC: Programme for the International Assessment of Adult Competencies; PS-TRE: problem solving in technology rich environments; PV: plausible values.

Authors' contributions

HB conceived, designed and implemented the study. Data analyses were conducted by graduate assistants under the direct supervision of the author. The study manuscript is the sole work of the author. The author read and approved the final manuscript.

Acknowledgements

The author extends his appreciation to the OECD for its support through the T.J. Alexander Fellowship, as well as its hospitality at the initiation of this work. He also benefitted from the comments and advice from Irwin Kirsch, Richard Murnane, William Thorne, Matthias von Davier, Kentaro Yamamoto, the editor and an anonymous referee. Excellent research assistance was provided by Shiya Yi, Yiran Chen, and Maria Baez Cruz. Any errors or omissions are the responsibility of the author.

Competing interests

The author declares he has no competing interests.

Availability of data and materials

All data and materials are publicly available.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

The work reported here was supported in part by a T. J. Alexander Fellowship from the Organization for Economic Cooperation and Development.

Appendix 1

See Table 4.

³⁵ For a discussion of the importance of conducting longitudinal policy analyses along with longitudinal data collection, see (Braun et al. 2006).

Table 4 Country names and abbreviations

Country abbr.	Country name
AUT	Austria
BEL	Belgium
CAN	Canada
CZE	Czech Republic
DEU	Germany
DNK	Denmark
ESP	Spain
EST	Estonia
FIN	Finland
FRA	France
GBR	United Kingdom
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
NLD	Netherlands
NOR	Norway
POL	Poland
SVK	Slovak Republic
SWE	Sweden
USA	United States

Appendix 2

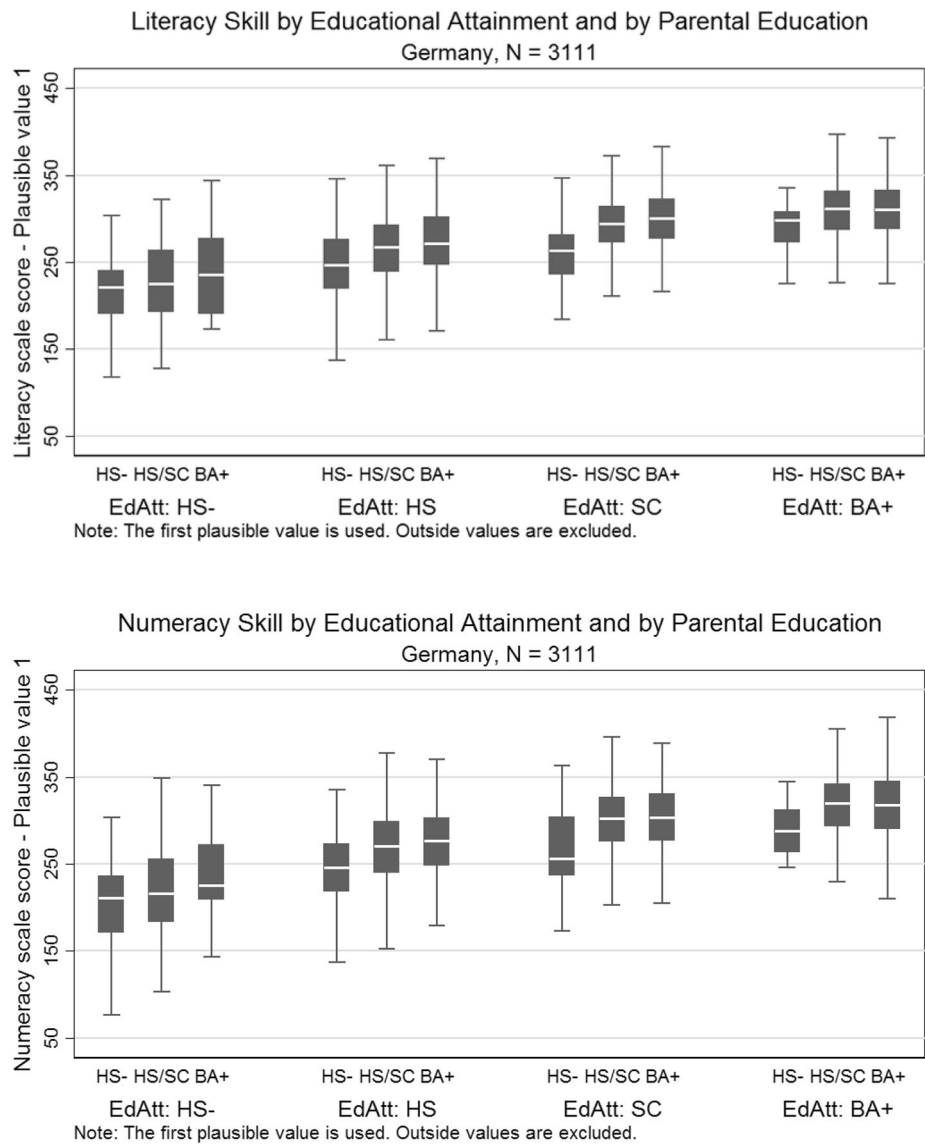
See Table 5.

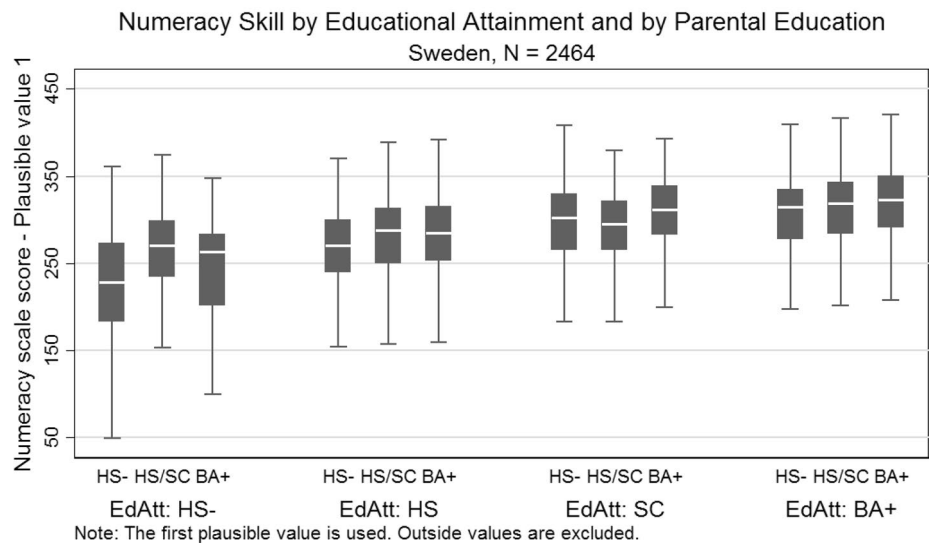
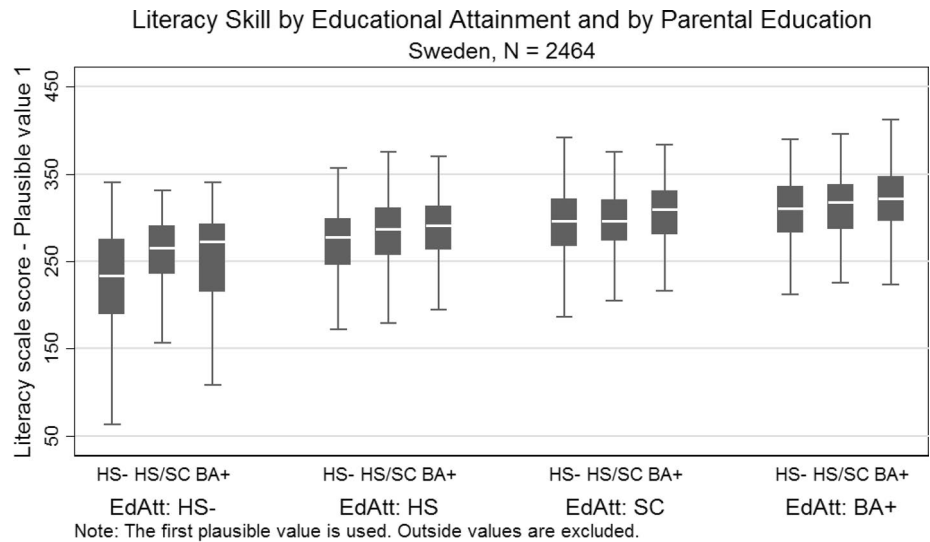
Table 5 Percentage of missing/not inferred/not stated earning information by country (unweighted)

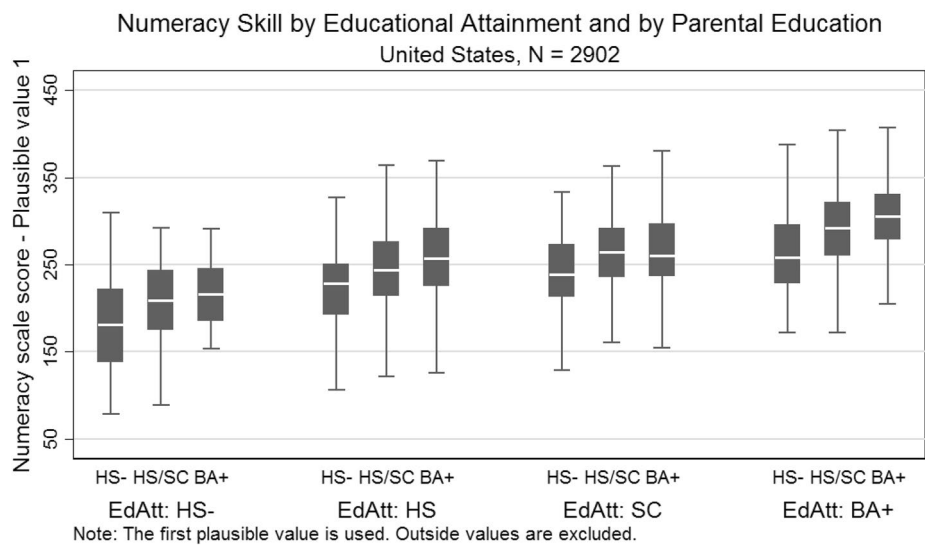
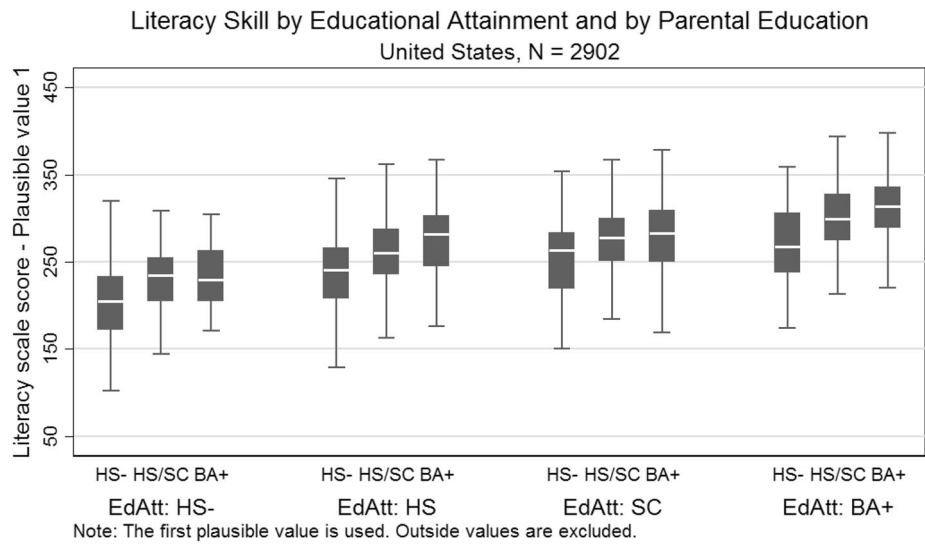
Country	Percentage excluded (%)
AUT	19.0
BEL	28.4
CAN	23.0
CZE	33.0
DEU	21.9
DNK	19.1
ESP	42.0
EST	31.2
FIN	17.8
FRA	24.9
GBR	29.6
IRL	37.1
ITA	35.7
JPN	22.5
KOR	25.5
NLD	21.1
NOR	16.9
POL	37.1
SVK	34.5
SWE	15.0
USA	28.2

Appendix 3

Boxplots of Literacy and Numeracy by Educational Attainment and Parental Education for Germany, Sweden and United States.







Appendix 4

See Table 6.

Table 6 D-statistics for goodness of fit: Q4 analyses

	Model 1 D	Model 2 D	Model 3 D
AUT	0.10	0.16	0.21
BEL	0.11	0.18	0.25
CZE	0.10	0.12	0.15
DEN	0.10	0.18	0.26
DEU	0.12	0.23	0.30
ESP	0.08	0.15	0.23
EST	0.15	0.18	0.21
FIN	0.11	0.17	0.22
FRA	0.10	0.17	0.24
GBR	0.08	0.15	0.21
IRL	0.05	0.13	0.18
ITA	0.09	0.11	0.14
JPN	0.21	0.25	0.26
KOR	0.12	0.16	0.19
NLD	0.14	0.23	0.29
NOR	0.12	0.18	0.21
POL	0.06	0.09	0.11
SVK	0.02	0.03	0.06
SWE	0.10	0.14	0.17
USA	0.10	0.16	0.21

Model 1 = Gender, age, ParEd, # of books at home

Model 2 = Model 1 + Cognitive Skill (standardized Literacy and Numeracy score)

Model 3 = Model 2 + Education Attainment

Appendix 5

See Table 7.

Table 7 Logistic regressions for Q4 analyses

Variable	Model 1		Model 2		Model 3	
	B	Sig.	B	Sig.	B	Sig.
AUT						
Basic demo						
Female	− 0.76	0.00	− 0.66	0.00	− 0.70	0.00
Age 25–34	− 1.00	0.00	− 1.18	0.00	− 1.27	0.00
Age 35–44	− 0.07	0.62	− 0.16	0.26	− 0.14	0.32
ParEd = below HS	− 0.53	0.00	− 0.25	0.18	− 0.06	0.78
ParEd = HS	− 0.17	0.23	− 0.06	0.66	0.08	0.57
≤ 25 books at home	− 0.97	0.00	− 0.52	0.01	− 0.31	0.10
26–100 books at home	− 0.54	0.00	− 0.32	0.02	− 0.22	0.11
Cognitive Skills						
Std. (Lit + Num)			0.74	0.00	0.57	0.00
Education Attainment						
EDATT = below HS					− 1.93	0.00
EDATT = HS					− 1.13	0.00
EDATT = some college					− 0.49	0.00
Reference group	1.35	0.00	1.06	0.00	1.79	0.00
Observation	1851		1851		1851	
BEL						
Basic demo						
Female	− 0.82	0.00	− 0.73	0.00	− 0.95	0.00
Age 25–34	− 1.55	0.00	− 1.81	0.00	− 1.97	0.00
Age 35–44	− 0.51	0.00	− 0.63	0.00	− 0.68	0.00
ParEd = below HS	− 0.77	0.00	− 0.39	0.02	0.03	0.88
ParEd = HS	− 0.44	0.00	− 0.26	0.09	− 0.01	0.96
≤ 25 books at home	− 0.63	0.00	− 0.33	0.05	− 0.18	0.33
26–100 books at home	− 0.01	0.93	0.10	0.50	0.14	0.39
Cognitive Skills						
Std. (Lit + Num)			0.84	0.00	0.48	0.00
Education Attainment						
EDATT = below HS					− 2.44	0.00
EDATT = HS					− 1.96	0.00
EDATT = some college					− 1.26	0.00
Reference group	0.39	0.01	− 0.16	0.28	1.03	0.00
Observation	1739		1739		1739	
CZE						
Basic demo						
Female	− 1.19	0.00	− 1.16	0.00	− 1.27	0.00
Age 25–34	0.27	0.29	0.12	0.64	0.13	0.61
Age 35–44	0.36	0.15	0.30	0.23	0.30	0.26
ParEd = below HS	− 1.45	0.00	− 1.22	0.00	− 0.82	0.01
ParEd = HS	− 0.86	0.00	− 0.74	0.00	− 0.46	0.02
≤ 25 books at home	− 1.14	0.00	− 0.80	0.05	− 0.70	0.08
26–100 books at home	− 0.42	0.04	− 0.30	0.13	− 0.17	0.40
Cognitive Skills						
Std. (Lit + Num)			0.39	0.00	0.21	0.09
Education Attainment						
EDATT = below HS					− 1.35	0.02

Table 7 (continued)

CZE						
EDATT = HS					− 0.99	0.00
EDATT = some college					− 0.30	0.38
Reference group	− 0.06	0.81	− 0.22	0.41	0.28	0.32
Observation	1808		1808		1807	
DEN						
Basic demo						
Female	− 0.93	0.00	− 0.90	0.00	− 1.13	0.00
Age 25–34	− 0.93	0.00	− 1.16	0.00	− 1.43	0.00
Age 35–44	− 0.09	0.50	− 0.24	0.09	− 0.30	0.04
ParEd = below HS	− 0.47	0.00	− 0.25	0.12	0.01	0.95
ParEd = HS	− 0.31	0.02	− 0.09	0.51	0.18	0.27
≤ 25 books at home	− 0.77	0.00	− 0.39	0.02	− 0.21	0.23
26–100 books at home	− 0.73	0.00	− 0.55	0.00	− 0.40	0.00
Cognitive Skills						
Std. (Lit + Num)			0.93	0.00	0.66	0.00
Education Attainment						
EDATT = below HS					− 2.13	0.00
EDATT = HS					− 1.68	0.00
EDATT = some college					− 1.19	0.00
Reference group	0.34	0.01	− 0.12	0.44	1.04	0.00
Observation	2478		2478		2476	
DEU						
Basic demo						
Female	− 0.94	0.00	− 0.88	0.00	− 1.09	0.00
Age 25–34	− 1.14	0.00	− 1.38	0.00	− 1.44	0.00
Age 35–44	− 0.09	0.48	− 0.29	0.04	− 0.20	0.17
ParEd = below HS	− 1.46	0.00	− 0.96	0.01	− 0.57	0.12
ParEd = HS	− 0.39	0.00	− 0.25	0.04	0.07	0.60
≤ 25 books at home	− 1.05	0.00	− 0.37	0.06	− 0.14	0.52
26–100 books at home	− 0.50	0.00	− 0.27	0.07	− 0.10	0.53
Cognitive Skills						
Std. (Lit + Num)			1.08	0.00	0.76	0.00
Education Attainment						
EDATT = below HS					− 2.51	0.00
EDATT = HS					− 1.80	0.00
EDATT = some college					− 1.37	0.00
Reference group	0.63	0.00	0.16	0.30	1.19	0.00
Observation	1696		1696		1696	
ESP						
Basic demo						
Female	− 0.59	0.00	− 0.47	0.00	− 0.90	0.00
Age 25–34	− 1.08	0.00	− 1.09	0.00	− 1.07	0.00
Age 35–44	− 0.34	0.01	− 0.42	0.00	− 0.43	0.00
ParEd = below HS	− 0.50	0.01	− 0.30	0.12	0.13	0.53
ParEd = HS	0.01	0.95	0.15	0.45	0.41	0.07
≤ 25 books at home	− 1.02	0.00	− 0.53	0.01	− 0.31	0.14
26–100 books at home	− 0.37	0.02	− 0.23	0.17	− 0.24	0.17

Table 7 (continued)

ESP						
Cognitive Skills						
Std. (Lit + Num)			0.81	0.00	0.42	0.00
Education Attainment						
EDATT = below HS					− 2.06	0.00
EDATT = HS					− 1.48	0.00
EDATT = some college					− 0.96	0.00
Reference group	0.62	0.00	− 0.05	0.80	0.84	0.00
Observation	1790		1790		1789	
EST						
Basic demo						
Female	− 1.36	0.00	− 1.34	0.00	− 1.56	0.00
Age 25–34	0.33	0.00	0.26	0.03	0.41	0.00
Age 35–44	0.14	0.26	0.13	0.27	0.22	0.07
ParEd = below HS	− 0.87	0.00	− 0.74	0.00	− 0.48	0.00
ParEd = HS	− 0.60	0.00	− 0.50	0.00	− 0.37	0.00
≤ 25 books at home	− 1.00	0.00	− 0.68	0.01	− 0.64	0.01
26–100 books at home	− 0.52	0.00	− 0.34	0.01	− 0.30	0.02
Cognitive Skills						
Std. (Lit + Num)			0.45	0.00	0.28	0.00
Education Attainment						
EDATT = Below HS					− 1.09	0.00
EDATT = HS					− 1.13	0.00
EDATT = some college					− 0.82	0.00
Reference group	0.54	0.00	0.33	0.02	0.96	0.00
Observation	2562		2562		2561	
FIN						
Basic demo						
Female	− 1.22	0.00	− 1.22	0.00	− 1.55	0.00
Age 25–34	− 1.38	0.00	− 1.58	0.00	− 1.72	0.00
Age 35–44	− 0.30	0.02	− 0.36	0.00	− 0.45	0.00
ParEd = below HS	− 0.31	0.04	− 0.14	0.39	0.17	0.29
ParEd = HS	− 0.19	0.15	− 0.06	0.67	0.18	0.21
≤ 25 books at home	− 0.82	0.00	− 0.51	0.01	− 0.32	0.10
26–100 books at home	− 0.35	0.00	− 0.23	0.05	− 0.13	0.27
Cognitive Skills						
Std. (Lit + Num)			0.61	0.00	0.36	0.00
Education Attainment						
EDATT = below HS					− 1.77	0.00
EDATT = HS					− 1.59	0.00
EDATT = some college					− 0.99	0.00
Reference group	0.39	0.01	0.17	0.26	0.94	0.00
Observation	2154		2154		2154	
FRA						
Basic demo						
Female	− 0.62	0.00	− 0.59	0.00	− 0.87	0.00
Age 25–34	− 1.20	0.00	− 1.53	0.00	− 1.78	0.00
Age 35–44	− 0.48	0.00	− 0.66	0.00	− 0.86	0.00

Table 7 (continued)

FRA						
ParEd = below HS	− 0.89	0.00	− 0.49	0.00	− 0.04	0.79
ParEd = HS	− 0.73	0.00	− 0.48	0.00	− 0.13	0.37
≤ 25 books at home	− 1.01	0.00	− 0.52	0.00	− 0.37	0.06
26–100 books at home	− 0.40	0.00	− 0.27	0.04	− 0.18	0.16
Cognitive Skills						
Std. (Lit + Num)			0.93	0.00	0.56	0.00
Education Attainment						
EDATT = below HS					− 2.15	0.00
EDATT = HS					− 1.80	0.00
EDATT = some college					− 0.63	0.00
Reference group	0.54	0.00	− 0.05	0.77	0.97	0.00
Observation	2180		2180		2179	
GBR						
Basic demo						
Female	− 0.61	0.00	− 0.52	0.00	− 0.65	0.00
Age 25–34	− 0.84	0.00	− 0.96	0.00	− 1.10	0.00
Age 35–44	0.11	0.48	0.05	0.78	0.03	0.85
ParEd = below HS	− 1.25	0.00	− 0.87	0.00	− 0.64	0.01
ParEd = HS	− 0.75	0.00	− 0.67	0.00	− 0.44	0.02
≤ 25 books at home	− 0.06	0.76	0.38	0.07	0.40	0.06
26–100 books at home	− 0.04	0.79	0.13	0.33	0.19	0.17
Cognitive Skills						
Std. (Lit + Num)			0.74	0.00	0.60	0.00
Education Attainment						
EDATT = below HS					− 1.18	0.00
EDATT = HS					− 1.14	0.00
EDATT = some college					− 0.93	0.00
Reference group	0.68	0.00	0.14	0.48	0.74	0.00
Observation	2581		2581		2574	
IRL						
Basic demo						
Female	− 0.19	0.07	− 0.07	0.55	− 0.31	0.01
Age 25–34	− 1.00	0.00	− 1.16	0.00	− 1.35	0.00
Age 35–44	− 0.02	0.89	− 0.13	0.39	− 0.19	0.24
ParEd = below HS	− 0.62	0.00	− 0.45	0.00	− 0.17	0.17
ParEd = HS	− 0.32	0.01	− 0.25	0.07	− 0.04	0.79
≤ 25 books at home	− 0.26	0.14	0.22	0.27	0.39	0.05
26–100 books at home	− 0.13	0.35	0.01	0.97	0.09	0.55
Cognitive Skills						
Std. (Lit + Num)	0.94	0.00	0.52	0.00	1.37	0.00
Education Attainment						
EDATT = below HS					− 1.51	0.00
EDATT = HS					− 1.43	0.00
EDATT = some college					− 1.21	0.00
Reference group	0.94	0.00	0.52	0.00	1.37	0.00
Observation	1892		1892		1892	

Table 7 (continued)

ITA						
Basic demo						
Female	− 0.74	0.00	− 0.72	0.00	− 0.94	0.00
Age 25–34	− 1.34	0.00	− 1.44	0.00	− 1.60	0.00
Age 35–44	− 0.48	0.00	− 0.50	0.00	− 0.51	0.00
ParEd = below HS	− 0.77	0.01	− 0.59	0.05	− 0.15	0.65
ParEd = HS	− 0.41	0.15	− 0.36	0.22	− 0.05	0.88
≤ 25 books at home	− 0.87	0.00	− 0.58	0.00	− 0.33	0.14
26–100 books at home	− 0.57	0.00	− 0.47	0.01	− 0.34	0.07
Cognitive Skills						
Std. (Lit + Num)			0.46	0.00	0.37	0.00
Education Attainment						
EDATT = below HS					− 1.28	0.00
EDATT = HS					− 1.08	0.00
EDATT = some college					− 0.88	0.10
Reference group	0.78	0.00	0.43	0.15	0.95	0.00
Observation	1577		1577		1577	
JPN						
Basic demo						
Female	− 1.64	0.00	− 1.62	0.00	− 1.53	0.00
Age 25–34	− 1.88	0.00	− 2.05	0.00	− 2.05	0.00
Age 35–44	− 0.53	0.00	− 0.64	0.00	− 0.61	0.00
ParEd = below HS	− 0.45	0.03	− 0.34	0.10	− 0.12	0.54
ParEd = HS	− 0.41	0.00	− 0.35	0.01	− 0.20	0.19
≤ 25 books at home	− 0.97	0.00	− 0.73	0.00	− 0.62	0.00
26–100 books at home	− 0.64	0.00	− 0.51	0.00	− 0.47	0.00
Cognitive Skills						
Std. (Lit + Num)			0.51	0.00	0.39	0.00
Education Attainment						
EDATT = below HS					− 0.61	0.10
EDATT = HS					− 0.84	0.00
EDATT = some college					− 0.77	0.00
Reference group	1.63	0.00	1.46	0.00	1.79	0.00
Observation	1812		1812		1811	
KOR						
Basic demo						
Female	− 1.27	0.00	− 1.18	0.00	− 1.09	0.00
Age 25–34	− 1.72	0.00	− 2.01	0.00	− 2.13	0.00
Age 35–44	− 0.19	0.12	− 0.39	0.00	− 0.48	0.00
ParEd = Below HS	− 0.82	0.00	− 0.64	0.00	− 0.46	0.00
ParEd = HS	− 0.19	0.22	− 0.08	0.62	0.00	0.99
≤ 25 books at home	− 0.46	0.01	− 0.18	0.30	0.07	0.70
26–100 books at home	− 0.14	0.34	− 0.01	0.96	0.06	0.71
Cognitive Skills						
Std. (Lit + Num)			0.56	0.00	0.37	0.00
Education Attainment						
EDATT = below HS					− 1.97	0.00
EDATT = HS					− 0.90	0.00
EDATT = some college					− 0.61	0.00

Table 7 (continued)

KOR						
Reference group	0.33	0.08	0.07	0.71	0.51	0.02
Observation	2707		2707		2707	
NLD						
Basic demo						
Female	− 1.02	0.00	− 0.95	0.00	− 1.19	0.00
Age 25–34	− 1.49	0.00	− 1.80	0.00	− 1.84	0.00
Age 35–44	− 0.14	0.33	− 0.34	0.03	− 0.32	0.05
ParEd = below HS	− 0.60	0.00	− 0.45	0.01	− 0.10	0.62
ParEd = HS	− 0.19	0.28	− 0.18	0.31	0.03	0.88
≤ 25 Books at Home	− 0.76	0.00	− 0.25	0.16	− 0.10	0.57
26–100 books at home	− 0.47	0.00	− 0.26	0.09	− 0.20	0.20
Cognitive Skills						
Std. (Lit + Num)			0.90	0.00	0.64	0.00
Education Attainment						
EDATT = below HS					− 1.62	0.00
EDATT = HS					− 1.35	0.00
EDATT = some college					− 0.09	0.79
Reference group	1.21	0.00	0.89	0.00	1.53	0.00
Observation	1384		1384		1384	
NOR						
Basic demo						
Female	− 1.16	0.00	− 1.10	0.00	− 1.30	0.00
Age 25–34	− 1.04	0.00	− 1.14	0.00	− 1.10	0.00
Age 35–44	− 0.08	0.49	− 0.19	0.11	− 0.20	0.09
ParEd = below HS	− 0.65	0.00	− 0.43	0.00	− 0.20	0.17
ParEd = HS	− 0.39	0.00	− 0.28	0.01	− 0.16	0.16
≤ 25 books at home	− 0.86	0.00	− 0.31	0.11	− 0.23	0.23
26–100 books at home	− 0.34	0.00	− 0.08	0.50	− 0.03	0.83
Cognitive Skills						
Std. (Lit + Num)			0.86	0.00	0.68	0.00
Education Attainment						
EDATT = below HS					− 1.12	0.00
EDATT = HS					− 0.94	0.00
EDATT = some college					− 0.17	0.20
Reference group	0.41	0.00	0.04	0.74	0.42	0.00
Observation	2154		2154		2153	
POL						
Basic demo						
Female	− 0.92	0.00	− 0.92	0.00	− 1.10	0.00
Age 25–34	− 0.24	0.27	− 0.35	0.13	− 0.46	0.06
Age 35–44	− 0.21	0.34	− 0.26	0.24	− 0.30	0.19
ParEd = below HS	− 1.18	0.00	− 0.96	0.01	− 0.73	0.06
ParEd = HS	− 0.35	0.10	− 0.19	0.36	0.01	0.97
≤ 25 books at home	− 1.47	0.00	− 1.05	0.00	− 0.81	0.00
26–100 books at home	− 0.62	0.00	− 0.43	0.03	− 0.33	0.10
Cognitive Skills						
Std. (Lit + Num)			0.59	0.00	0.45	0.00

Table 7 (continued)

POL						
Education Attainment						
EDATT = below HS					0.04	0.94
EDATT = HS					− 1.06	0.00
EDATT = some college					− 0.51	0.11
Reference group	− 0.77	0.00	− 0.77	0.00	− 0.41	0.16
Observation	2061		2061		2061	
SVK						
Basic demo						
Female	− 1.41	0.00	− 1.36	0.00	− 1.48	0.00
Age 25–34	− 0.21	0.54	− 0.25	0.47	− 0.24	0.49
Age 35–44	0.08	0.80	0.02	0.94	0.08	0.80
ParEd = below HS	− 0.70	0.24	− 0.34	0.57	0.20	0.74
ParEd = HS	− 0.37	0.29	− 0.20	0.57	0.19	0.58
≤ 25 books at home	− 1.10	0.03	− 0.68	0.18	− 0.26	0.61
26–100 books at home	− 0.55	0.05	− 0.37	0.16	− 0.12	0.63
Cognitive Skills						
Std. (Lit + Num)			0.73	0.00	0.56	0.01
Education Attainment						
EDATT = Below HS					− 2.61	0.00
EDATT = HS (or some college)					− 1.34	0.00
Reference group	− 1.90	0.00	− 2.55	0.00	− 2.15	0.00
Observation	2066		2066		2066	
SWE						
Basic demo						
Female	− 0.84	0.00	− 0.85	0.00	− 1.02	0.00
Age 25–34	− 0.86	0.00	− 1.01	0.00	− 1.01	0.00
Age 35–44	− 0.20	0.11	− 0.27	0.05	− 0.26	0.07
ParEd = below HS	− 0.54	0.00	− 0.37	0.01	− 0.18	0.24
ParEd = HS	− 0.22	0.10	− 0.15	0.27	− 0.04	0.75
≤ 25 books at Home	− 0.97	0.00	− 0.40	0.07	− 0.37	0.09
26–100 books at home	− 0.70	0.00	− 0.36	0.01	− 0.26	0.05
Cognitive Skills						
Std. (Lit + Num)			0.78	0.00	0.57	0.00
Education Attainment						
EDATT = below HS					− 1.10	0.00
EDATT = HS					− 0.88	0.00
EDATT = some college					− 0.11	0.50
Reference group	0.83	0.00	0.53	0.00	0.98	0.00
Observation	1690		1690		1690	
USA						
Basic demo						
Female	− 0.62	0.00	− 0.53	0.00	− 0.78	0.00
Age 25–34	− 0.96	0.00	− 1.03	0.00	− 1.07	0.00
Age 35–44	− 0.26	0.06	− 0.30	0.04	− 0.35	0.02
ParEd = below HS	− 1.44	0.00	− 0.91	0.00	− 0.63	0.01
ParEd = HS	− 0.30	0.01	− 0.13	0.23	0.02	0.83
≤ 25 books at home	− 0.61	0.00	− 0.17	0.30	0.09	0.57
26–100 books at home	− 0.30	0.04	− 0.15	0.30	− 0.07	0.63

Table 7 (continued)

USA						
Cognitive Skills						
Std. (Lit + Num)			0.79	0.00	0.48	0.00
Education Attainment						
EDATT = below HS					− 3.01	0.00
EDATT = HS					− 1.39	0.00
EDATT = some college					− 1.02	0.00
Reference group	0.95	0.00	0.44	0.00	1.25	0.00
Observation	1700		1700		1700	

Model 1 = Gender, Age, ParEd, # of books at home

Model 2 = M1 + Cognitive Skill (standardized Literacy & Numeracy score)

Model 3 = M2 + Education Attainment

Reference Group = Male, Age 45–54, ParEd = At least one BA +, 101 books or more, EDATT = BA+

Appendix 6

See Table 8.

Table 8 D-statistics for goodness of fit: Q1 analyses

	Model 1 D	Model 2 D	Model 3 D
AUT	0.08	0.09	0.11
BEL	0.11	0.17	0.21
CAN	0.01	0.01	0.01
CZE	0.09	0.11	0.14
DEU	0.04	0.10	0.13
DNK	0.05	0.07	0.09
ESP	0.09	0.13	0.19
EST	0.03	0.06	0.07
FIN	0.11	0.14	0.17
FRA	0.05	0.09	0.12
GBR	0.06	0.10	0.12
IRL	0.25	0.27	0.29
ITA	0.15	0.17	0.19
JPN	0.08	0.11	0.14
KOR	0.10	0.16	0.20
NLD	0.08	0.10	0.13
NOR	0.13	0.15	0.19
POL	0.12	0.15	0.20
SVK	0.03	0.04	0.06
SWE	0.11	0.15	0.19
USA	0.02	0.04	0.06

Model 1 = Gender, Age, ParEd, # of books at home

Model 2 = M1 + Cognitive Skill (standardized Literacy & Numeracy score)

Model 3 = M2 + Education Attainment

Appendix 7

See Table 9.

Table 9 Logistic Regressions for Q1 Analyses

Variable	Model 1		Model 2		Model 3	
	B	Sig.	B	Sig.	B	Sig.
AUT						
Basic demo						
Female	1.49	0.00	1.45	0.00	1.45	0.00
Age 25–34	0.35	0.01	0.39	0.01	0.44	0.00
Age 35–44	0.21	0.15	0.23	0.11	0.24	0.09
ParEd = below HS	0.35	0.08	0.24	0.25	0.12	0.56
ParEd = HS	0.01	0.94	−0.02	0.90	−0.10	0.59
≤ 25 books at home	0.23	0.13	0.02	0.91	−0.13	0.39
26–100 books at home	0.20	0.17	0.10	0.48	0.03	0.83
Cognitive Skills						
Std. (Lit + Num)			−0.28	0.00	−0.16	0.04
Education Attainment						
EDATT = below HS					1.08	0.00
EDATT = HS					0.69	0.00
EDATT = some college					0.18	0.36
Reference group	−2.71	0.00	−2.59	0.00	−3.03	0.00
Observation	2581		2581		2581	
BEL						
Basic demo						
Female	1.21	0.00	1.16	0.00	1.39	0.00
Age 25–34	0.58	0.00	0.77	0.00	0.76	0.00
Age 35–44	0.26	0.01	0.35	0.00	0.35	0.00
ParEd = below HS	0.53	0.00	0.20	0.21	−0.10	0.53
ParEd = HS	0.37	0.00	0.24	0.05	0.07	0.57
≤ 25 books at home	0.58	0.00	0.30	0.01	0.15	0.22
26–100 books at home	0.07	0.52	−0.04	0.74	−0.09	0.42
Cognitive Skills						
Std. (Lit + Num)			−0.67	0.00	−0.44	0.00
Education Attainment						
EDATT = below HS					1.58	0.00
EDATT = HS					1.24	0.00
EDATT = some college					0.39	0.01
Reference group	−1.85	0.00	−1.55	0.00	−2.21	0.00
Observation	2288		2288		2286	
CAN						
Basic demo						
Female	0.75	0.00	0.72	0.00	0.79	0.00
Age 25–34	0.41	0.04	0.46	0.02	0.46	0.02
Age 35–44	−0.15	0.45	−0.11	0.56	−0.08	0.67
ParEd = below HS	0.09	0.67	0.00	0.99	−0.14	0.53
ParEd = HS	−0.15	0.43	−0.18	0.35	−0.29	0.14
≤ 25 books at home	−0.13	0.56	−0.25	0.25	−0.27	0.21
26–100 books at home	−0.58	0.00	−0.63	0.00	−0.66	0.00
Cognitive Skills						
Std. (Lit + Num)			−0.22	0.01	−0.12	0.19
Education Attainment						
EDATT = below HS					0.69	0.05

Table 9 (continued)

CAN						
EDATT = HS					0.73	0.00
EDATT = some college					0.45	0.01
Reference group	− 3.21	0.00	− 3.10	0.00	− 3.46	0.00
Observation	11,620		11,620		11,612	
CZE						
Basic demo						
Female	1.00	0.00	0.96	0.00	1.01	0.00
Age 25–34	− 0.30	0.18	− 0.15	0.54	− 0.12	0.63
Age 35–44	0.10	0.68	0.16	0.50	0.17	0.50
ParEd = Below HS	1.25	0.00	1.02	0.00	0.73	0.02
ParEd = HS	0.62	0.01	0.50	0.03	0.28	0.25
≤ 25 books at home	0.72	0.01	0.38	0.19	0.35	0.24
26–100 books at home	0.50	0.01	0.38	0.04	0.30	0.10
Cognitive Skills						
Std. (Lit + Num)			− 0.38	0.00	− 0.28	0.02
Education Attainment						
EDATT = below HS					0.95	0.02
EDATT = HS					0.81	0.02
EDATT = some college					0.43	0.31
Reference group	− 2.12	0.00	− 2.01	0.00	− 2.48	0.00
Observation	2080		2080		2079	
DEN						
Basic demo						
Female	0.70	0.00	0.69	0.00	0.80	0.00
Age 25–34	0.69	0.00	0.75	0.00	0.84	0.00
Age 35–44	0.13	0.31	0.20	0.11	0.25	0.05
ParEd = Below HS	0.01	0.96	− 0.18	0.26	− 0.38	0.02
ParEd = HS	− 0.04	0.80	− 0.17	0.24	− 0.28	0.05
≤ 25 books at home	0.61	0.00	0.30	0.06	0.19	0.25
26–100 books at home	0.29	0.05	0.13	0.36	0.05	0.72
Cognitive Skills						
Std. (Lit + Num)			− 0.62	0.00	− 0.49	0.00
Education Attainment						
EDATT = below HS					1.14	0.00
EDATT = HS					0.44	0.00
EDATT = some college					− 0.09	0.51
Reference group	− 2.34	0.00	− 2.14	0.00	− 2.46	0.00
Observation	3084		3084		3082	
DEU						
Basic demo						
Female	1.63	0.00	1.60	0.00	1.67	0.00
Age 25–34	0.29	0.05	0.40	0.01	0.38	0.01
Age 35–44	− 0.18	0.14	− 0.06	0.65	− 0.12	0.36
ParEd = below HS	0.61	0.01	0.28	0.23	− 0.02	0.92
ParEd = HS	0.16	0.14	0.06	0.58	− 0.15	0.18
≤ 25 books at home	0.53	0.00	0.17	0.33	0.01	0.94
26–100 books at home	0.18	0.14	0.01	0.92	− 0.10	0.45

Table 9 (continued)

DEU						
Cognitive Skills						
Std. (Lit + Num)			− 0.51	0.00	− 0.31	0.00
Education Attainment						
EDATT = below HS					1.82	0.00
EDATT = HS					1.13	0.00
EDATT = some college					0.69	0.00
Reference group	− 2.39	0.00	− 2.24	0.00	− 2.91	0.00
Observation	2487		2487		2487	
ESP						
Basic demo						
Female	1.28	0.00	1.23	0.00	1.51	0.00
Age 25–34	0.67	0.00	0.68	0.00	0.64	0.00
Age 35–44	0.27	0.04	0.33	0.01	0.34	0.01
ParEd = below HS	0.09	0.56	− 0.04	0.82	− 0.44	0.01
ParEd = HS	− 0.35	0.06	− 0.43	0.03	− 0.68	0.00
≤ 25 books at home	0.99	0.00	0.69	0.00	0.45	0.00
26–100 books at home	0.41	0.01	0.31	0.03	0.23	0.11
Cognitive Skills						
Std. (Lit + Num)			− 0.45	0.00	− 0.20	0.01
Education Attainment						
EDATT = below HS					1.62	0.00
EDATT = HS					1.31	0.00
EDATT = some college					0.97	0.00
Reference group	− 2.24	0.00	− 1.93	0.00	− 2.69	0.00
Observation	2235		2235		2234	
EST						
Basic demo						
Female	0.83	0.00	0.82	0.00	0.96	0.00
Age 25–34	− 0.47	0.00	− 0.41	0.00	− 0.53	0.00
Age 35–44	− 0.18	0.12	− 0.18	0.12	− 0.26	0.03
ParEd = below HS	0.47	0.00	0.36	0.01	0.14	0.33
ParEd = HS	0.07	0.62	− 0.01	0.94	− 0.13	0.35
≤ 25 Books at home	0.42	0.01	0.17	0.34	0.07	0.68
26–100 books at home	0.10	0.42	− 0.06	0.65	− 0.12	0.35
Cognitive Skills						
Std. (Lit + Num)			− 0.34	0.00	− 0.20	0.00
Education Attainment						
EDATT = below HS					1.19	0.00
EDATT = HS					0.93	0.00
EDATT = some college					0.58	0.00
Reference group	− 1.83	0.00	− 1.71	0.00	− 2.24	0.00
Observation	2956		2956		2955	
FIN						
Basic demo						
Female	1.10	0.00	1.13	0.00	1.49	0.00
Age 25–34	0.80	0.00	1.00	0.00	1.02	0.00
Age 35–44	0.19	0.15	0.26	0.04	0.32	0.01

Table 9 (continued)

FIN						
ParEd = below HS	0.09	0.53	− 0.09	0.57	− 0.33	0.04
ParEd = HS	0.02	0.89	− 0.12	0.39	− 0.33	0.03
≤ 25 books at home	0.83	0.00	0.53	0.00	0.37	0.02
26–100 books at home	0.29	0.00	0.17	0.06	0.06	0.56
Cognitive Skills						
Std. (Lit + Num)			− 0.57	0.00	− 0.36	0.00
Education Attainment						
EDATT = below HS					1.64	0.00
EDATT = HS					1.53	0.00
EDATT = some college					0.68	0.00
Reference group	− 1.89	0.00	− 1.79	0.00	− 2.64	0.00
Observation	2528		2528		2528	
FRA						
Basic demo						
Female	1.04	0.00	1.04	0.00	1.11	0.00
Age 25–34	− 0.17	0.33	0.00	0.99	0.06	0.73
Age 35–44	0.03	0.83	0.10	0.47	0.16	0.22
ParEd = below HS	0.32	0.10	0.06	0.76	− 0.12	0.57
ParEd = HS	0.27	0.19	0.13	0.53	0.01	0.94
≤ 25 books at home	0.46	0.00	0.12	0.47	0.01	0.96
26–100 books at home	− 0.06	0.65	− 0.16	0.28	− 0.21	0.16
Cognitive Skills						
Std. (Lit + Num)			− 0.52	0.00	− 0.39	0.00
Education Attainment						
EDATT = below HS					0.89	0.00
EDATT = HS					0.58	0.00
EDATT = some college					0.09	0.63
Reference group	− 2.98	0.00	− 2.76	0.00	− 3.10	0.00
Observation	2610		2610		2609	
GBR						
Basic demo						
Female	1.48	0.00	1.44	0.00	1.49	0.00
Age 25–34	0.57	0.00	0.59	0.00	0.69	0.00
Age 35–44	0.44	0.01	0.50	0.00	0.56	0.00
ParEd = below HS	0.70	0.00	0.39	0.05	0.16	0.46
ParEd = HS	0.21	0.22	0.09	0.63	− 0.08	0.71
≤ 25 books at home	0.48	0.02	0.16	0.44	0.12	0.55
26–100 books at home	0.27	0.13	0.16	0.38	0.11	0.56
Cognitive Skills						
Std. (Lit + Num)			− 0.52	0.00	− 0.43	0.00
Education Attainment						
EDATT = Below HS					0.97	0.00
EDATT = HS					0.56	0.01
EDATT = some college					0.40	0.13
Reference group	− 3.02	0.00	− 2.70	0.00	− 3.05	0.00
Observation	3485		3485		3476	

Table 9 (continued)

IRL						
Basic demo						
Female	0.95	0.00	0.90	0.00	1.09	0.00
Age 25–34	– 0.24	0.17	– 0.17	0.31	0.00	0.99
Age 35–44	– 0.05	0.73	0.02	0.92	0.10	0.54
ParEd = Below HS	0.52	0.02	0.34	0.14	0.03	0.90
ParEd = HS	0.14	0.53	0.05	0.80	– 0.09	0.66
≤ 25 books at home	0.56	0.01	0.27	0.22	0.06	0.79
26–100 books at home	– 0.10	0.58	– 0.17	0.36	– 0.25	0.17
Cognitive Skills						
Std. (Lit + Num)			– 0.52	0.00	– 0.33	0.00
Education Attainment						
EDATT = below HS					1.85	0.00
EDATT = HS					0.87	0.00
EDATT = some college					1.04	0.00
Reference group	– 2.49	0.00	– 2.28	0.00	– 3.11	0.00
Observation	2506		2506		2506	
ITA						
Basic demo						
Female	0.94	0.00	0.94	0.00	1.05	0.00
Age 25–34	0.52	0.00	0.59	0.00	0.68	0.00
Age 35–44	– 0.04	0.81	– 0.02	0.89	0.00	0.99
ParEd = below HS	0.36	0.26	0.16	0.63	– 0.08	0.80
ParEd = HS	– 0.24	0.48	– 0.30	0.40	– 0.40	0.25
≤ 25 books at home	0.25	0.21	– 0.10	0.60	– 0.27	0.18
26–100 books at home	– 0.02	0.91	– 0.17	0.40	– 0.23	0.27
Cognitive Skills						
Std. (Lit + Num)			– 0.46	0.00	– 0.38	0.00
Education Attainment						
EDATT = below HS					0.84	0.00
EDATT = HS					0.31	0.11
EDATT = some college					0.05	0.93
Reference group	– 2.08	0.00	– 1.78	0.00	– 2.05	0.00
Observation	1952		1952		1952	
JPN						
Basic demo						
Female	2.40	0.00	2.38	0.00	2.40	0.00
Age 25–34	0.21	0.15	0.30	0.03	0.30	0.05
Age 35–44	– 0.08	0.49	– 0.03	0.80	– 0.09	0.46
ParEd = below HS	0.44	0.01	0.33	0.07	0.03	0.84
ParEd = HS	0.35	0.01	0.28	0.03	0.08	0.53
≤ 25 Books at Home	0.56	0.00	0.30	0.03	0.13	0.34
26–100 books at home	0.26	0.03	0.14	0.25	0.04	0.74
Cognitive Skills						
Std. (Lit + Num)			– 0.44	0.00	– 0.28	0.00
Education Attainment						
EDATT = below HS					1.43	0.00
EDATT = HS					1.08	0.00

Table 9 (continued)

JPN						
EDATT = some college					0.75	0.00
Reference group	− 2.45	0.00	− 2.30	0.00	− 2.77	0.00
Observation	2358		2358		2357	
KOR						
Basic demo						
Female	1.77	0.00	1.74	0.00	1.72	0.00
Age 25–34	− 0.46	0.00	− 0.28	0.03	− 0.12	0.41
Age 35–44	− 0.43	0.00	− 0.31	0.01	− 0.23	0.06
ParEd = Below HS	0.41	0.00	0.28	0.04	0.03	0.82
ParEd = HS	0.29	0.05	0.22	0.14	0.07	0.66
≤ 25 books at home	0.41	0.00	0.21	0.13	0.00	0.97
26–100 books at home	0.05	0.69	− 0.05	0.72	− 0.10	0.42
Cognitive Skills						
Std. (Lit + Num)			− 0.36	0.00	− 0.22	0.00
Education Attainment						
EDATT = Below HS					1.29	0.00
EDATT = HS					0.97	0.00
EDATT = some college					0.42	0.01
Reference group	− 2.10	0.00	− 1.99	0.00	− 2.43	0.00
Observation	3208		3208		3208	
NLD						
Basic demo						
Female	1.88	0.00	1.82	0.00	1.88	0.00
Age 25–34	0.07	0.69	0.19	0.31	0.24	0.20
Age 35–44	0.10	0.52	0.18	0.26	0.21	0.19
ParEd = Below HS	0.17	0.40	0.02	0.91	− 0.16	0.49
ParEd = HS	0.03	0.90	− 0.06	0.80	− 0.13	0.56
≤ 25 books at home	0.56	0.00	0.25	0.17	0.17	0.36
26–100 books at home	0.41	0.01	0.30	0.06	0.28	0.09
Cognitive Skills						
Std. (Lit + Num)			− 0.47	0.00	− 0.29	0.00
Education Attainment						
EDATT = below HS					1.08	0.00
EDATT = HS					0.50	0.03
EDATT = some college					− 0.37	0.41
Reference group	− 3.59	0.00	− 3.44	0.00	− 3.82	0.00
Observation	2376		2376		2375	
NOR						
Basic demo						
Female	1.31	0.00	1.30	0.00	1.53	0.00
Age 25–34	0.76	0.00	0.74	0.00	0.85	0.00
Age 35–44	0.12	0.34	0.14	0.27	0.26	0.05
ParEd = below HS	0.33	0.03	0.09	0.59	− 0.14	0.42
ParEd = HS	0.07	0.60	− 0.04	0.78	− 0.19	0.17
≤ 25 books at home	1.04	0.00	0.56	0.00	0.45	0.00
26–100 books at home	0.38	0.00	0.18	0.13	0.05	0.67

Table 9 (continued)

NOR						
Cognitive Skills						
Std. (Lit + Num)			− 0.60	0.00	− 0.45	0.00
Education Attainment						
EDATT = below HS					1.57	0.00
EDATT = HS					1.15	0.00
EDATT = some college					1.01	0.00
Reference group	− 2.62	0.00	− 2.45	0.00	− 3.28	0.00
Observation	2617		2617		2616	
POL						
Basic demo						
Female	0.71	0.00	0.76	0.00	1.04	0.00
Age 25–34	0.26	0.11	0.35	0.04	0.55	0.00
Age 35–44	0.14	0.42	0.16	0.36	0.22	0.21
ParEd = below HS	0.73	0.00	0.62	0.02	0.13	0.64
ParEd = HS	0.42	0.04	0.34	0.11	− 0.05	0.85
≤ 25 books at home	1.10	0.00	0.79	0.00	0.39	0.01
26–100 books at home	0.51	0.00	0.33	0.04	0.14	0.40
Cognitive Skills						
Std. (Lit + Num)			− 0.47	0.00	− 0.29	0.00
Education Attainment						
EDATT = below HS					2.17	0.00
EDATT = HS					1.45	0.00
EDATT = some college					1.13	0.00
Reference group	− 2.53	0.00	− 2.38	0.00	− 3.07	0.00
Observation	2351		2351		2351	
SVK						
Basic demo						
Female	1.20	0.00	1.20	0.00	1.32	0.00
Age 25–34	0.11	0.37	0.11	0.37	0.14	0.28
Age 35–44	0.03	0.82	0.05	0.66	0.05	0.67
ParEd = below HS	1.61	0.00	1.43	0.00	1.00	0.00
ParEd = HS	0.94	0.00	0.87	0.00	0.55	0.00
≤ 25 books at home	0.74	0.00	0.50	0.00	0.26	0.15
26–100 books at home	0.39	0.00	0.29	0.01	0.15	0.21
Cognitive Skills						
Std. (Lit + Num)			− 0.39	0.00	− 0.27	0.00
Education Attainment						
EDATT = below HS					1.97	0.00
EDATT = HS (or some college)					1.03	0.00
Reference group	− 1.48	0.00	− 1.22	0.00	− 1.70	0.00
Observation	2224		2224		2224	
SWE						
Basic demo						
Female	0.76	0.00	0.74	0.00	0.85	0.00
Age 25–34	0.83	0.00	0.86	0.00	0.91	0.00
Age 35–44	0.29	0.18	0.32	0.15	0.36	0.10
ParEd = below HS	0.06	0.74	− 0.03	0.86	− 0.17	0.37

Table 9 (continued)

SWE						
ParEd = HS	− 0.31	0.15	− 0.34	0.11	− 0.48	0.03
≤ 25 books at home	0.52	0.04	0.18	0.50	0.18	0.48
26–100 books at home	− 0.21	0.26	− 0.41	0.03	− 0.44	0.02
Cognitive Skills						
Std. (Lit + Num)			− 0.39	0.00	− 0.31	0.00
Education Attainment						
EDATT = below HS					0.73	0.01
EDATT = HS					0.88	0.00
EDATT = some college					0.97	0.00
Reference group	− 2.84	0.00	− 2.70	0.00	− 3.37	0.00
Observation	2107		2107		2107	
USA						
Basic demo						
Female	0.72	0.00	0.69	0.00	0.81	0.00
Age 25–34	0.16	0.35	0.15	0.38	0.13	0.47
Age 35–44	− 0.01	0.95	0.00	0.99	0.00	0.99
ParEd = below HS	0.75	0.00	0.44	0.05	0.24	0.29
ParEd = HS	0.37	0.02	0.30	0.08	0.12	0.52
≤ 25 books at home	− 0.05	0.82	− 0.28	0.24	− 0.49	0.03
26–100 books at home	− 0.09	0.66	− 0.18	0.40	− 0.29	0.20
Cognitive Skills						
Std. (Lit + Num)			− 0.35	0.00	− 0.13	0.27
Education Attainment						
EDATT = below HS					1.71	0.00
EDATT = HS					1.42	0.00
EDATT = some college					1.35	0.00
Reference group	− 2.69	0.00	− 2.50	0.00	− 3.40	0.00
Observation	2162		2162		2162	

Model 1 = Gender, age, ParEd, # of books at home

Model 2 = M1 + Cognitive Skill (standardized Literacy & Numeracy score)

Model 3 = M2 + Education Attainment

Reference Group = Male, age 45–54, Pared = At least one BA+, 101 books or more, EDATT = BA+

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 June 2017 Accepted: 16 May 2018

Published online: 26 May 2018

References

- Alexander, K., Entwistle, D., & Olson, L. (2014). *The long shadow: family background disadvantaged urban youth, and the transition to adulthood*. New York: Russell Sage.
- Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186), 843–851. <https://doi.org/10.1126/science.1251868>.
- Bernstein, J. (2016). Wages in the United States: Trends, explanations and solutions. In I. Kirsch & H. I. Braun (Eds.), *The dynamics of opportunity in America: Evidence and perspectives*. New York: Springer.

- Bind, M.-A. C., & Rubin, D. B. (2017). Bridging observational studies and randomized experiments by embedding the former in the latter. *Statistical Methods in Medical Research*. <https://doi.org/10.1177/0962280217740609>.
- Braun, H. I. (2016). The dynamics of opportunity in America: A working framework. In I. Kirsch & H. I. Braun (Eds.), *The dynamics of opportunity in America: Evidence and perspectives*. New York: Springer.
- Braun, H. I., & von Davier, M. (2017). The use of test scores from large-scale assessment surveys: psychometric and statistical considerations. *Large-scale Assessments in Education*, 5, 17.
- Braun, H.I., Wang, A., Jenkins, F., & Weinbaum, E. (2006). The Black-White achievement gap: Do state policies matter? *Education Policy Analysis Archives*. <https://doi.org/10.14507/epaa.v14n8.2006>.
- Cappellari, L., Castelnovo, P., Checchi, D., & Leonardi, M. (2015). Skilled or educated? Educational reform, human capital and earnings. http://www.iza.org/conference_files/2015_Skill_Mismatch/cappellari_l1064.pdf. Accessed 5 Apr 2018.
- Carnoy, M. (2015). *International test score comparisons and education policy: A review of the critiques*. Boulder: National Education Policy Center.
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4), 1553–1623.
- Chmielewski, A.K. (2015). Can we close social origin gaps in Literacy over the life course? Evidence from synthetic cohorts in IALS, ALL, and PIAAC. Presentation at the PIAAC International Conference, Haarlem, Netherlands (24 November). https://www.aanmelder.nl/piaac_conference_2015/wiki/160925/presentations%20and%20papers. Accessed 5 Apr 2018.
- Chmielewski, A. K., & Dhuey, E. (2017). *The analysis of international large-scale assessments to address causal questions in education policy. Commissioned paper*. Washington, DC: National Academy of Education.
- de Broucker, P., & Underwood, K. (1998). Intergenerational education mobility: An international comparison with a focus on postsecondary education. *Education Quarterly Review*, 5(2), 30.
- Deming, D. J. (2017). The growing importance of social skills. *The Quarterly Journal of Economics*, 132(4), 1593–1640.
- Ermisch, J., Jantti, M., & Smeeding, T. (2012). *From parents to children: The intergenerational transmission of disadvantage*. New York: Russell Sage Foundation.
- Goodman, L., & Kruskal, W. (1954). Measures of association for cross-classifications. *Journal of the American Statistical Association*, 49(268), 732–764.
- Gustafsson, J.-E. (2013). Causal inference in educational effectiveness research: A comparison of three methods to investigate the effects of homework on student achievement. *School Effectiveness and School Improvement*, 24(3), 275–295.
- Hampf, F., Wiederhold, S., & Woessmann, L. (2017). Skills, earnings, and employment: exploring causality in the estimation of returns to skills. *Large-scale Assessment in Education*, 5, 12. <https://doi.org/10.1186/s40536-017-0045-7>.
- Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2013). *Returns to skills around the world: Evidence from PIAAC. (OECD Education Working Papers, No. 101)*. Paris: OECD Publishing.
- Holzer, H.J., & Lerman, R.I. (2015). Cognitive Skills in the U.S. labor market: For whom do they matter? https://static1.squarespace.com/static/51bb74b8e4b0139570dd020/t/54da7582e4b0cb4c49fc0d9d/1423603074704/Holzer_Lerman_PIAAC.pdf. Accessed 5 Apr 2018.
- IEA (2015) *Help Manual for the IDB Analyzer*. Hamburg: IEA. www.iea.nl/data. Accessed 12 Nov 2017.
- Kaplan, D. (2009). Causal inference in non-experimental educational policy research. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook on education policy research* (pp. 139–153). New York: Taylor and Francis.
- Kaplan, D., & McCarty, A.T. (2013). Data fusion with international large-scale assessments: A case study using the OECD PISA and TALIS studies. *Large-scale assessments in education*. <http://largescaleassessmentsineducation.com/content/1/1/6>. Accessed 5 Apr 2018.
- Khatiwada, I., & Sum, A. (2016). The widening socio-economic divergence in the U.S. labor market. In I. Kirsch & H. I. Braun (Eds.), *The dynamics of opportunity in America: Evidence and perspectives* (pp. 197–252). New York: Springer.
- Kirsch, I. (2001). *The International Adult Literacy Survey (IALS): Understanding what was measured*. ETS Research Report RR-01-25. Princeton: Educational Testing Service.
- Kirsch, I., Braun, H., Lennon, M.L., & Sands, A. (2016). *Choosing our future: A story of opportunity in America*. Princeton, NJ: Educational Testing Service. <http://opportunityproject.ets.org/>. Accessed 5 Apr 2018.
- Kirsch, I., & Lennon, M. L. (2017). PIAAC: A new design for a new era. *Large-scale Assessments in Education*, 5(11), 1–22.
- Levin, H. M. (2013). The utility and need for incorporating non-Cognitive Skills into large-scale educational assessments. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of international large-scale assessments: Perspectives from technology, economy and educational research* (pp. 67–86). Springer Science + Business Media: Dordrecht, Netherlands.
- Lin, D., Lutter, R., & Ruhm, C. J. (2016). *Cognitive performance and labor market outcomes. (National Bureau of Economic Research Working Paper, No. 22470)*. Boston: NBER.
- Massing, N., & Schneider, S. L. (2017). Degrees of competency: the relationship between educational qualifications and adult skills across countries. *Large Scale Assessment in Education*, 5(6), 1–34.
- Moretti, E. (2013). *The new geography of jobs*. New York: Mariner Books.
- Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the Cognitive Skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4), 547–568.
- OECD. (2013a). *Survey of adult skills technical report* (2nd ed.). Paris: OECD Publishing.
- OECD. (2013b). *OECD skills outlook 2013: First results from the survey of adult skills*. Paris: OECD Publishing.
- OECD. (2016). *Skills matter: Further results from the survey of adult skills, OECD skills studies*. Paris: OECD Publishing.
- Park, H., & Kyei, P. (2011). Literacy gaps by Educational Attainment: A cross-national analysis. *Social Forces*, 89(3), 879–904.
- Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity*. New York: Russell Sage Foundation.
- Reardon, S., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, 2(3), 1–18.
- Ritzen, J. (2013). International large-scale assessments as change agents. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of large-scale assessments: Perspectives from technology, economy and educational research*. New York: Springer.

- Rothstein, J. (2017). *Inequality of educational opportunity? Schools as mediators of the intergenerational transmission of income*. (Institute for Research on Labor and Employment. Working Paper). Berkeley: U.C. Berkeley.
- Singer, J. D., & Braun, H. I. (2018). The (mis)uses of international education assessments, with five suggestions for improvement. *Science*, 360(6384), 38–40.
- Singer, J. D., Braun, H. I., & Chudowski, N. (2018). *International education assessments: Cautions, conundrums, and common sense*. Washington, DC: National Academy of Education.
- Smeeding, T. (2016). Gates, gaps and intergenerational mobility: The importance of an even start. In I. Kirsch & H. I. Braun (Eds.), *The dynamics of opportunity in America: Evidence and perspectives*. New York: Springer.
- Tjur, T. (2008). Coefficients of determination in logistic regression models—a new proposal: The coefficient of discrimination. *The American Statistician*, 63(4), 366–372.
- von Davier, M., Sinharay, S., Oranje, A., & Beaton, A. (2007). The statistical procedures used in National Assessment of Educational Progress: recent developments and future directions. In C. R. Rao & S. Sinharay (Eds.), *Handbook of Statistics* (Vol. 26, pp. 1039–1055). New York: Elsevier.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
