

Chapter 1 Online Appendix: Changes in SES Skill Differentials over Childhood

Here we provide a framework in which to interpret changes in gaps between SES groups in skills or achievements (say test scores) over childhood and into adulthood. To do so, we posit a simple value-added model of the evolution of skills as the child ages (Todd and Wolpin 2003, 2007). It relates skills at given age or life cycle stage, denoted as t , to *investments*, or actions undertaken by families and schools to improve skills at a particular age or life stage and to the level of skills achieved previously:

$$y_{tj} = I_{tj} + b_t y_{t-1j}, \quad t \geq 1, \quad 0 < b_t \leq 1$$

(1)

where y_{tj} denotes the skill or ability level of child at age t for SES group j , and I_{tj} denotes the outcomes of the investments. If, for example, y_{tj} referred to a child's height, we would expect $b_t = 1$ and at some age t^* , $I_{tj} = 0$ for $t > t^*$, because at some point further investments in nutrition have no impact on height.

Depreciation in skills over time, making it necessary to continue to invest enough at each age to maintain skills, would be represented by $b_t < 1$.

We want to use this simple model to express SES differentials and how these change as the child ages. Suppose there are two SES groups, high and low ($j=H,L$) and three stages of childhood. Then the SES skill difference at stage t is defined as

$$\Delta y_t = y_{tH} - y_{tL}, \text{ and}$$

$$\Delta y_1 = I_{1H} - I_{1L} + b_1(y_{0H} - y_{0L})$$

$$\Delta y_2 = I_{2H} - I_{2L} + b_2(I_{1H} - I_{1L}) + b_2 b_1(y_{0H} - y_{0L})$$

(2)

$$\Delta y_3 = I_{3H} - I_{3L} + b_3(I_{2H} - I_{2L}) + b_3 b_2(I_{1H} - I_{1L}) + b_3 b_2 b_1(y_{0H} - y_{0L})$$

From this representation, we see that SES differences in the current skill level depend on SES differences in all past investments plus the SES difference initial conditions, such as health and development at birth (as in the model of Cunha and Heckman 2007). We need to take the differences

between these expressions to see what drives changes in SES differentials as the child ages. For example, the change in the SES differential between childhood stages two and three is given by the following equation:

$$\Delta y_3 - \Delta y_2 = I_{3H} - I_{3L} + (b_3 - 1)[(I_{2H} - I_{2L}) + b_2(I_{1H} - I_{1L}) + b_2 b_1(y_{0H} - y_{0L})]$$

(3)

It clearly depends in part on differences in current stage investment, $\{EQ\}_{I_{3H} - I_{3L}}\{EQ\}$; it also depends on b_3 . If, for example, $b_3 = 1$ (as for $y = \text{height}$), then whether SES differentials widen or narrow depends entirely on whether $I_{3H} - I_{3L} > 0$. If, however, $b_3 < 1$, as estimated by Petra Todd and Kenneth Wolpin for math and reading percentile score measures of y_{ij} (2007), then SES differentials could narrow even in the case in which high SES parents invest more in children at stage 3 than low SES parents (that is, $I_{3H} - I_{3L} > 0$) if b_3 is low enough.

As children pass into and through school, the SES investment differential ($I_{tH} - I_{tL}$) is likely to alter. For instance, equalizing schools may cause the SES investment differential to narrow as schools substitute more for families in skill acquisition, but it could also be that better-off parents may locate near or otherwise direct their children to better schools, causing the investment differential to widen. Although not assumed in our simple model, it is possible that children from more affluent families experience less depreciation (have higher b_3) because of environmental differences or dinner conversation or school quality.

Falvio Cunha and James Heckman (2007) make the case that the productivity of investments in producing skills depends on past investments, which in turn encourages more investment. In our simple framework, this would mean that current investments have bigger impacts on skills when the level of acquired skills is higher (producing a positive correlation between I_{ij} and y_{t-1j})—so called dynamic complementarity.

Chapter 3 Online Appendix

Carina Mood
Jan O. Jonsson
Erik Bihagen

Swedish Institute for Social Research (SOFI)
Stockholm University

Table 3A.1. Descriptive statistics for variables used in the analyses

	All				For the income analyses				For the education analyses			
	<i>Mean</i>	<i>Std dev</i>	<i>Min/Max</i>	<i>N</i>	<i>Mean</i>	<i>Std dev</i>	<i>Min/Max</i>	<i>N</i>	<i>Mean</i>	<i>Std dev</i>	<i>Min/Max</i>	<i>N</i>
Child's income	0.30	0.73	-1.24/4.00	191413	0.32	0.73	-1.24/4.00	156837	---	---	---	---
Father's income	0.45	0.87	-1.55/4.00	200383	0.48	0.86	-1.53/4.00	156837	0.46	0.87	-1.55/4.00	179696
Social maturity	3.11	0.71	1.00/5.00	211083	3.15	0.68	1.00/5.00	156837	3.12	0.70	1.00/5.00	179696
Intensity	2.95	0.88	1.00/5.00	211096	2.98	0.87	1.00/5.00	156837	2.96	0.88	1.00/5.00	179696
Psych energy	3.15	0.71	1.00/5.00	211094	3.18	0.69	1.00/5.00	156837	3.16	0.70	1.00/5.00	179696
Emotional stability	3.04	0.67	1.00/5.00	211090	3.07	0.64	1.00/5.00	156837	3.06	0.66	1.00/5.00	179696
Psychological funct	5.11	1.66	1.00/9.00	217593	5.21	1.60	1.00/9.00	156837	5.16	1.63	1.00/9.00	179696
BMI dev	2.04	1.88	0.00/37.12	220642	2.02	1.83	0.00/37.12	156837	2.03	1.85	0.00/37.12	179696
Cogn: Log-inductive	23.93	8.36	0.00/40.00	222390	24.87	7.59	0.00/40.00	156837	24.65	7.68	0.00/40.00	179696
Cogn: Verbal	24.13	7.45	0.00/40.00	222390	24.91	6.57	0.00/40.00	156837	24.77	6.64	0.00/40.00	179696
Cogn: Spatial	21.03	8.67	0.00/40.00	222390	21.77	8.18	0.00/40.00	156837	21.63	8.22	0.00/40.00	179696
Cogn: Technical	24.12	6.63	0.00/40.00	222509	24.85	5.68	0.00/40.00	156837	24.74	5.74	0.00/40.00	179696
Cognitive ability total	93.21	27.15	0.00/159.00	222390	96.40	23.61	0.00/159.00	156837	95.79	23.90	0.00/159.00	179696
Height	179.25	6.52	155.00/205.00	220802	179.38	6.48	155.00/205.00	156837	179.34	6.50	155.00/205.00	179696
Physical capacity	5.96	1.86	1.00/9.00	221194	6.04	1.86	1.00/9.00	156837	6.01	1.86	1.00/9.00	179696
Education. Years	12.67	2.49	8.20/21.7	225039	12.85	2.51	8.20/21.7	156837	12.76	2.50	8.20/21.70	179696
Fathers educ. years	9.97	3.25	7.20/20.3	221282	10.05	3.26	7.20/20.3	156837	10.05	3.27	7.20/20.30	179696
Fath manual work	0.45	0.50	0.00/1.00	227835	0.45	0.50	0.00/1.00	156837	0.45	0.50	0.00/1.00	179696
Fath low non-man	0.12	0.32	0.00/1.00	227835	0.13	0.33	0.00/1.00	156837	0.13	0.33	0.00/1.00	179696
Fath low salariat	0.20	0.40	0.00/1.00	227835	0.21	0.41	0.00/1.00	156837	0.21	0.41	0.00/1.00	179696
Fath high salariat	0.10	0.31	0.00/1.00	227835	0.11	0.32	0.00/1.00	156837	0.11	0.32	0.00/1.00	179696
Fath self-emp	0.13	0.33	0.00/1.00	227835	0.10	0.30	0.00/1.00	156837	0.11	0.31	0.00/1.00	179696
Fath occ prestige	40.89	11.92	6.00/78.00	224888	41.39	12.02	6.00/78.00	156837	41.36	12.03	6.00/78.00	179696
Leadership capacity					5.31	1.53	1/9	105031				
GPA (1972)									3,16	0,69	0.00/5.00	35371

Note: Incomes are averages of yearly incomes that are z-standardized in relation to the entire income distribution among men for the respective years. The average incomes here are above 0, which means that the fathers and the sons in our cohorts had on average higher incomes than other men in the labour market during these income years.

Table 3A.2. Correlation matrix, all included variables

	Ch inc	Fath. inc	Ch educ	Fa educ	Social	Intense	Energy	Emotion	Pers tot	Logic	Verbal	Spatial	Technic	Cogn tot	Height	Bmi dev	Phys
Child inc	1.00																
Fath inc	0.31	1.00															
Ch educ	0.40	0.33	1.00														
Fath educ	0.26	0.56	0.38	1.00													
Social	0.30	0.20	0.29	0.21	1.00												
Intensity	0.21	0.11	0.16	0.09	0.42	1.00											
Energy	0.28	0.17	0.28	0.18	0.59	0.52	1.00										
Emotion	0.29	0.18	0.26	0.18	0.59	0.45	0.53	1.00									
Ps tot	0.34	0.21	0.31	0.20	0.75	0.78	0.76	0.77	1.00								
Logic	0.35	0.26	0.49	0.29	0.35	0.20	0.31	0.30	0.37	1.00							
Verbal	0.32	0.27	0.50	0.32	0.34	0.16	0.29	0.28	0.34	0.72	1.00						
Spatial	0.27	0.21	0.40	0.25	0.27	0.11	0.23	0.23	0.27	0.61	0.51	1.00					
Technic	0.29	0.22	0.40	0.26	0.29	0.17	0.25	0.26	0.31	0.62	0.59	0.63	1.00			-	
Cogn tot	0.36	0.28	0.53	0.33	0.37	0.19	0.32	0.31	0.38	0.88	0.83	0.83	0.82	1.00			
Height	0.12	0.10	0.12	0.09	0.11	0.09	0.10	0.11	0.13	0.13	0.14	0.11	0.11	0.14	1.00		
Bmi dev	-0.09	-0.06	-0.09	-0.07	-0.10	-0.15	-0.11	-0.10	-0.15	-0.09	-0.08	-0.07	0.07	-0.09	-0.01	1.00	
Phys	0.21	0.14	0.23	0.14	0.29	0.47	0.36	0.31	0.47	0.22	0.19	0.14	0.16	0.21	0.27	-0.14	1.00
Fath siops	0.23	0.54	0.31	0.59	0.18	0.08	0.16	0.15	0.18	0.25	0.27	0.21	0.21	0.28	0.07	-0.06	0.11
<i>Leadersh</i>	<i>0.34</i>	<i>0.21</i>	<i>0.30</i>	<i>0.21</i>	<i>0.72</i>	<i>0.59</i>	<i>0.66</i>	<i>0.74</i>	<i>0.87</i>	<i>0.24</i>	<i>0.22</i>	<i>0.13</i>	<i>0.18</i>	<i>0.26</i>	<i>0.11</i>	<i>-0.13</i>	<i>0.38</i>
<i>GPA</i>	<i>--</i>	<i>0.29</i>	<i>0.60</i>	<i>0.34</i>	<i>0.43</i>	<i>0.27</i>	<i>0.40</i>	<i>0.37</i>	<i>0.46</i>	<i>0.66</i>	<i>0.62</i>	<i>0.50</i>	<i>0.50</i>	<i>0.68</i>	<i>0.11</i>	<i>-0.15</i>	<i>0.32</i>

Chapter 4 Online Appendix

by Bruce Bradbury, Miles Corak, Jane Waldfogel and Elizabeth Washbrook

1. Survey details

Longitudinal Study of Australian Children (LSAC). LSAC contains data from two separate nationally representative samples of children and their families. The ‘K’ cohort (not used in our study) comprises 4,983 children born between March 1999 and February 2000, and aged 4-5 during the first round of interviews. The data used in this study are from the more recent ‘B’ cohort – 5,107 children born between March 2003 and February 2004, and aged 0-1 during the first round of interviews.

Random stratified samples were drawn from the Health Insurance Commission database on which practically all children (and their date of birth) are registered; since it is through this that they have access to free public health care. The reweighted sample is designed to be representative of all Australian children in the ‘B’ cohort age range.

Full waves of face to face interviews have been collected every two years for the LSAC ‘B’ cohort – when the children were age 0 to 1, 2 to 3 and 4 to 5. In addition, postal questionnaires were sent to parents at Waves 1.5 and Waves 2.5 (ie., half way through the two year intervals between the main interviews).

National Longitudinal Study of Children and Youth (NLSCY). Unlike the other datasets used in this study, the NLSCY is a long-running longitudinal study of Canadian children, with different cohorts brought into the sample depending on the year. The so-called “original cohort” of 0 to 11 year olds were first surveyed in 1994/95 (Cycle 1), with information collected every two years since that time. The latest collection of data occurred in 2006/7 when the original cohort was 12

to 23 years of age (Cycle 7). In addition, each biennial survey adds an “Early Childhood Development” (ECD) cohort, a representative cohort of 0 to 1 year olds born in 1996, 1998, 2000, 2002, 2004, and 2006. In cycle 7 there are a total of 20,470 children in all available ECD cohorts.

Our analysis draws on two cohorts of children: children born in 2000, age 0-1 in Cycle 4, age 2-3 in Cycle 5, and age 4-5 in Cycle 6; and children born in 2002, age 0-1 in Cycle 5, age 2-3 in Cycle 6, and age 4-5 in Cycle 7.

The NLSCY sample is meant to be representative of children and is based upon a sampling frame of all Canadian Households. Typically children are selected from households sampled by Statistics Canada’s Labour Force Survey (LFS) that collects labour market data from a national sample of 54000 households on a monthly basis. The LFS is the Canadian equivalent of the US Current Population Survey (CPS). This is based upon a stratified, multistage design using probability sampling at all stages.

The NLSCY produces three sets of weights at each cycle: two longitudinal, one cross-sectional. The NLSCY weighting strategy is based on a series of adjustments applied to the NLSCY design weight. Each child’s NLSCY design weight is equal to the inverse of his/her probability of selection. The Labour Force Survey (LFS) subweight accounts for all of the sample design information for the LFS sample. For the NLSCY, children are selected from the households selected from the LFS. To reflect these additional sample design steps, the LFS subweight is multiplied by several other factors. The final weight is obtained by applying nonresponse and post-stratification adjustments to the NLSCY design weight. The ‘post-stratification’ weight adjustment procedure was carried out to ensure consistency between the

estimates produced by NLSCY and Statistics Canada's population estimates by age, sex and province.

The NLSCY collects has three broad components: Child Component, Youth Component and an Adult Component. The Person Most Knowledgeable (PMK) about the child and his/her spouse answered the questions in the Adult and Child Component (for children aged 0-17). The Youth component was used for selected respondents aged 16 and above. The survey has been designed to interview up to two respondents in the Youth or Child category. The NLSCY therefore makes an important distinction between 'households' and 'children/youth'. The interviews were carried out using computer-assisted interviewing (CAI) methods and the use of paper questionnaires. The former consisted of computer-assisted personal interviewing (CAPI) with the help of a laptop as well as computer-assisted telephone interviewing (CATI).

Interviews were conducted in English or French. The interviewers were given freedom to switch between both languages if the respondent encountered any difficulty in understanding a question or a phrase. Also, a small number of interviews were conducted in languages other than French or English with the interviewer translating the questions into the respondent's preferred language.

The NLSCY is jointly conducted by Statistics Canada and sponsored by Human Resources and Social Development Canada (HRSDC). The data is available in Research Data Centres that were established through a joint effort by Statistics Canada and university consortia. The access to these Data Centres is restricted to researchers working on projects approved by Social Sciences and Humanities Research Council (SSHRC) and Statistics Canada.

Millennium Cohort Study (MCS). Children eligible for inclusion in the MCS were those born between 1 September 2000 and 31 August 2001 (for England and Wales), and between 23 November 2000 and 11 January 2002 (for Scotland and Northern Ireland), alive and living in the UK at age nine months.

The geography of electoral wards was used as a sampling frame. There were 11,090 electoral wards in the UK at the time of sampling, which were combined into 9,548 ‘superwards’ in order to eliminate very small units with less than 24 expected births in a year. 398 of these wards and ‘superwards’ were selected for sampling. The sample is clustered geographically and disproportionately stratified to over-represent: (1) the three smaller countries of the UK (Wales, Scotland and Northern Ireland); (2) areas in England with higher minority ethnic populations in 1991 (where at least 30 per cent of the population were Black or Asian); and (3) disadvantaged areas (drawn from the poorest 25 per cent of wards based on the Child Poverty Index). Of the 398 wards sampled, 50% were in England, 18% in Wales, 16% in Scotland and 16% in Northern Ireland; 5% were high ethnic minority wards and 48% were disadvantaged wards.

A list of all nine month old children living in the sampled wards was derived from Child Benefit records provided by the Department of Social Security (subsequently Department for Work and Pensions and then HM Revenue and Customs). Child Benefit claims cover virtually all of the child population except those ineligible due to recent or temporary immigrant status. 27,201 families were identified as eligible from the Child Benefit records, of which 24,180 (90%) were issued to the field.

Two types of weights are provided with the MCS. The sample design or probability weights are used to correct for MCS cases having unequal probabilities of selection that result from the stratified cluster sample design. These are fixed, and do not change over time and vary

only by the nine stratum (advantaged and disadvantaged wards in each of the four countries, plus ethnic wards in England only). Non-response weights adjust for possible biases generated by systematic unit non-response and vary by sweep and cohort family. The non-response weight at the current sweep is the inverse of the predicted probability of responding based on a logistic regression model using data from previous sweep(s).

The MCS surveyed cohort families three times, when the cohort members were roughly 9 months, 3 years and 5 years of age. Future sweeps are planned but not used in this study. At each sweep there were separate questionnaires for the Main Carer and the Main Carer's partner (if present in the household). Interviews were carried out using computer-assisted personal interview (CAPI) software on a laptop, and using a confidential computer-assisted self-completion interview (CASI) for sensitive subjects. Direct child assessments of cognitive ability and anthropomorphic measurements were carried out at sweeps 2 and 3. Older siblings aged 10 to 15 (in English families only) completed a paper questionnaire at Sweeps 2 and 3.

Roughly 4% of interviews with the main carer were conducted wholly or partly in a language other than English in Wave 1, falling to around 2% in Wave 3. Common languages used were: Welsh, Arabic, Bengali, Gujarati, Hindi, Kurdish, Punjabi, Somali, Tamil, Turkish and Urdu. The child cognitive assessments were conducted in English only (25 Welsh children completed a Welsh version of the Naming Vocabulary test, the only assessment for which this option was offered).

The Millennium Cohort Study is funded by the Economic and Social Research Council and a consortium of Government Departments headed by the Office for National Statistics (ONS). Data are publicly available from the UK Data Archive.

The Early Childhood Longitudinal Study – Birth Cohort (ECLS-B). The ECLS-B selected a nationally representative probability sample of children born in the United States in 2001. Births were sampled within a set of primary sampling units (PSUs) and in some cases secondary sampling units (SSUs) in order to control data collection costs. (A PSU is an individual county or group of contiguous counties. An SSU is a subdivision of a PSU.) The core ECLS-B sample consists of births sampled within 96 PSUs. An additional sample of 18 PSUs was selected from a supplemental frame consisting of areas where the population has a higher proportion of American Indian births (see below). The ECLS-B samples included births occurring in counties within 46 states and in Washington, D.C.

The ECLS-B was designed to support statistical analyses in the following analytic domains: (1) child's race/ethnicity, as derived from birth certificate data (American Indian including Alaska Natives; Chinese; Other Asian or Pacific Islander, i.e. excluding Chinese; Hispanic; Black, non-Hispanic; White, non-Hispanic); (2) birth weight (very low, moderately low and normal) and (3) plurality (twin and non-twin). When combined, these analytic domains result in 36 distinct birth certificate sampling strata. Six strata required oversampling: the American Indian, Chinese, and Other Asian/Pacific Islander groups; those with very low birth weight (under 1,500 grams) and those with moderately low birth weight (between 1,500 and 2,500 grams); and twins.

Children were sampled via registered births from the National Center for Health Statistics (NCHS) vital statistics system. Individual birth certificates within PSUs or SSUs were sampled from data files provided by state registrars. Within each case stratum, each sampled birth had an equal overall probability of selection. Membership of an ECLS-B PSU was determined using state and county of occurrence, although birth certificates also provide information on state of residence. Approximately 14,000 births were sampled and fielded and yielded 10,688 9-month completed cases.

The ECLS-B weights were developed in three steps: (1) base weights were calculated using the overall selection probabilities; (2) weights were adjusted for survey nonresponse; and (3) raking (adjusting sums of weights to known or estimated totals) was used to adjust for undercoverage and to improve precision of survey estimates. The development of the ECLS-B weights was a sequential process. The 9-month weights were developed from the base weight. Weights for subsequent sweeps were then developed as adjustments to the weights from the previous wave.

The ECLS-B surveyed cohort families three times, when the cohort members were roughly 9 months and 2 years of age, and during the preschool year when most cohort members were 4 years of age. A future kindergarten wave is planned but not used in this study. A Parent Computer-Assisted Personal Interview (CAPI) was administered at each of the three sweeps, consisting of an in-person interview in which the interviewer asked questions aloud and entered the responses in a computer. In addition, for items that could be considered sensitive, parent self-completion instruments were used.

The full scope of the ECLS-B includes a number of data collection instruments not used in this study. These are Resident and Non-resident Father Questionnaires; selected variables taken directly from the birth certificate; a Child Care Provider telephone interview; and in some cases a Child Care Observation.

Computer-assisted personal interviews (CAPIs) and self-administered paper questionnaires were available in English and Spanish versions, and bilingual interviewers were trained to conduct interviews in either English or Spanish. An interpreter, either a community or household member, was used for families who spoke languages other than English or Spanish. Around 6-7% of parent interviews were conducted in Spanish at each sweep, and further 1.5-2% were conducted in a language other than English and Spanish. A Spanish version of the ECLS-B Preschool Direct Cognitive Assessment (Wave 3) was used in the field when necessary, but

assessment scores were not computed/provided because sample sizes were too small (n=120). Hence all of the preschool wave assessment data that is usable was administered in English.

The primary sponsor of the ECLS-B is the National Center for Education Statistics (NCES), Institute of Education Sciences (IES), U.S. Department of Education (ED). Due to NCES' confidentiality legislation, ECLS-B data in micro-data form (i.e., CD-ROM) are available *only* through a restricted-use data license agreement.

2. ISCED categorizations

Table 4A.1. National qualifications in the ISCED framework

	Examples of national qualifications	% at Wave 3
Australia		
ISCED 2	Did not complete final year of high school	8.2%
ISCED 3 to 4	Final year of school completed or trade certificate	42.6%
ISCED 5B	Diploma	10.8%
ISCED 5A/6	University degree	38.4%
Parental education missing		0.0%
Total		100% (N = 4,386)
Canada		
ISCED 2	Did not graduate high school	6.2%
ISCED 3 to 4	High school, Some community college, Trade/vocational school or university	16.4%
ISCED 5B	Diploma from trade, technical or vocational school, Community college, CEGEP or nursing school	23.2%
ISCED 5A/6	Bachelor's degree, Master's degree	54.2%
Parental education missing		0.4%
Total		100% (N = 6812)
UK		
ISCED 2	GCSE D-F	12.2%
ISCED 3 to 4	GCSE A-C; A-level	42.2%
ISCED 5B	Diploma in HE; Nursing qualification	10.7%
ISCED 5A/6	First degree; Professional qualification at degree level; Bachelor's degree	31.1%
Parental education missing		0.8%
Total		100% (N = 15,460)
US		
ISCED 2	Less than high school	10.4%
ISCED 3 to 4	High school; GED	25.0%
ISCED 5B	Associate's degree	31.6%
ISCED 5A/6	Bachelor's degree, Master's degree	33.0%
Parental education missing		-
Total		100% (N = 8950*)

3. Derivation of average gross household income variables

As described in the main chapter, the US measure of income is the least detailed of the four countries, and we “downgrade” income measures from the other three surveys to minimize the differences in measurement error. The US measure asked respondents to place their total gross household income in one of 13 bands, ranging from “Less than \$5,000” to “\$200,001 or more”.

First, we calculated the proportion of families falling into each of the 13 bands at each of three waves, separately for two-parent and single-parent households. To illustrate 1.7% of two-parent households reported incomes in the lowest band at Wave 1 and 3.0% reported incomes in the second lowest band. Equivalent percentages of the sample of single-parent households were 17.8% and 15.6%. The basis of our income harmonization was then to define categorical income variables -- for each country, wave and family structure group -- with the same distributions as the US variables. The categorical variables were then converted back to a “lumpy” continuous variables by assigning a representative dollar value to each band. This procedure differed slightly across surveys.

Australia and Canada. The income questions in both the LSAC and the NLSCY asked for gross household income in a continuous format. After dividing families into the proportions described above on the basis of their reported income, each family in a given band was assigned the median value of income in that band, overwriting the family-specific reported value.

UK. Income questions in the MCS related to net rather than gross household income, and also required responses in bands, although in this case 19 bands were offered with the boundaries differing for single- and two-parent families. To deal with this, we used external nationally representative data from the Family Resources Survey (FRS) 2001 to 2007 on the gross and net incomes of families with children under 6. The median value of *gross* income for FRS families with *net* incomes within each pair boundary values was calculated and assigned to the relevant MCS families. Hence this step both converted net to gross values and created continuous variables from the MCS banded data in the most detailed way possible. This intermediary variable was then transformed in the same way as the Australian and Canadian continuous gross

income variables, by dividing it up according to the US distribution and over-writing with the median band value.

US. In order to assign a dollar value to each of the US income bands, and in particular to “close” to open top band, we again use external nationally representative data, in this case from the June Current Population Survey (CPS). The value of income assigned to each band is the median gross income of families with children under 6 with income inside the boundary values.

For all countries these nominal derived gross household income variables were deflated to 2006 values using national price indices, converted to US dollars using OECD purchasing power parity (PPP) indices, and normalized by the square root of household size. The three observations of real gross equivalized household income for each family were then averaged and the survey weights were used to define nationally-representative quintile boundaries.

More specifically, the Australian price indices for the three waves of the data were 2004 = 0.941, 2006 = 1, and 2008 = 1.068 (Source: ABS 6401.0, Annual CPI), the Canadian indices were 2000 = 0.874, 2002 = 0.917, 2004 = 0.958 and 2006 = 1 (Source: Statistics Canada, CPI for Canada V41690973 series), the UK 2001 = 0.875, 2002 = 0.889, 2003 = 0.915, 2004 = 0.942, 2005 = 0.969, 2006 = 1, 2007 = 1.043 (Source: ONS, All Items RPI) and the US 2001 = 0.878, 2002 = 0.892, 2003 = 0.913, 2004 = 0.937, 2005 = 0.969, 2006 = 1, 2007 = 1.028 (Source: BLS CPI-U). The PPP indices used were AU = 1.375, CN = 1.210, UK = 0.617 and US = 1 (Source: PPP index for 'actual individual consumption' for 2006, Source OECD National Accounts Statistics, Volume I - Exchange Rate, PPPs, and Population Vol 2009 release 01, downloaded 15 Dec 09).

4. Additional information on outcome measures

As described in the main text, the picture vocabulary tests used in the Australian, Canadian and US surveys all used items from the Peabody Picture Vocabulary Test (PPVT). The instruments administered, however, were not identical. The LSAC used a short version of the Third Edition of the PPVT (PPVT-III), adapted specifically for the survey and containing 40 items – 20 core, 10 basal (to which children performing poorly on the core items were routed), and 10 ceiling items (for children scoring highly on the core items). The NLSCY administered the PPVT Revised Version in full (PPVT-R; the prior version to the PPVT-III), with a French adaptation (EVIP) available for all French-speakers. The ECLS-B, like the LSAC, used only selected items from the PPVT-III, but in this case only 15 items were selected in total.

Early math and number skills were assessed in three of the countries, Australia being the exception. We provide some results relating to these outcomes but note they were assessed in quite different ways in different surveys and results may not be as comparable across countries as for some other outcomes. The Canadian Number Knowledge assessment is a mental arithmetic test designed to measure children's comprehension of the system of whole numbers. Children were not permitted to use a pencil and paper but instead were required to rely on a "mental counting line" in order to answer the orally-given questions. The ECLS-B Mathematics assessment was designed specifically for the US survey and included questions in following content areas: number sense, geometry, counting, operations, and patterns. As with the vocabulary assessments, the items administered were routed according to the child's responses and scored consistently using IRT techniques. The UK dataset contains no assessments of math skills in Wave 3. However, in Wave 2 (when the children were 3 years of age) the Bracken School Readiness Assessment (BRSA) was administered, which is composed of six sub-scales. We derive a measure of math skills as the sum of the Numbers, Sizes, Comparisons and Shapes

sub-scales, a broad definition that is perhaps most comparable with the US Mathematics assessment. Assessments of copying skills are available for children in all countries except the UK. The tasks are designed to assess the child's ability to conceptualize and reconstruct a geometrical shape and provide a non-verbal test of cognitive ability that provides a contrast to the verbal vocabulary measure. The use of the ability to copy geometrical figures to assess the level of development in children has been long established, indeed dating back to the original Simon Binet test. Research indicates that copying skills are strongly associated with subsequent school achievement, are valid across different cultural groups, and provide a reliable measure of development at the time of assessment (de Lemos, 2002). In all three countries the child recorded his or her response in a booklet, which was then scored centrally by trained researchers. The Australian and Canadian children were assessed using the same instrument: the Copying scale of the Who Am I (WAI) assessment, which requires children to copy five shapes (circle, triangle, cross, square, and diamond) with each response assessed on a four-point scale. In the US task the child was asked to copy seven shapes (vertical line, horizontal line, circle, square, cross, triangle, and asterisk) with each items scored either one (a pass) or zero (fail).

Several other cognitive assessment measures are available for single countries or pairs of countries only. For Canada and Australia, the Symbols sub-scale of the WAI assessment, composed of a set of five writing tasks (printing their name, printing some letters, numbers, words and a sentence), assesses the ability of the child to understand and use symbolic representations such as numbers, letters and words. The WAI Total score is the sum of the WAI Copying and Symbols sub-scales. The US Literacy Assessment, again designed specifically for the ECLS-B and using IRT scoring techniques, assesses a range of content areas related to emergent literacy: letter recognition, in both receptive and expressive modes; letter sounds; early reading (recognition of simple words); phonological awareness; knowledge of print conventions; and matching words. Finally, the total BRSA score for the UK children

(the sum of the four math sub-scales plus the Letters and Colors sub-scales) provides a summary measure of cognitive ability that combines a number of literacy and math skills.

5. Characteristics by Income and Education

Table 4A.2. Average characteristics of families with 4 to 5 year old children, by parental education and country

	AU (N = 4265)	CN (N = 6812)	UK (N = 15,460)	US (N = 8,500)*
A. LOW EDUCATION GROUP				
Mean household income (SD)	16,090 (12,248)	14,712 (7,641)	11,412 (7,034)	9,680 (6,524)
Single parent household at W3	48.0%	36.9%	47.7%	43.8%
Mother <20 at birth	15.0%	12.4%	21.1%	24.6%
Mother >30 at birth	43.0%	30.5%	25.5%	17.1%
Under 18s in household at W3	2.76 (1.41)	2.57 (1.48)	2.80 (1.38)	2.80 (1.40)
Foreign-born parent	29.0%	29.8%	17.0%	49.3%
White (non-Hispanic for US)	-	76.3%	77.0%	17.8%
Black (non-Hispanic for US)	-	5.0%	4.6%	18.5%
Hispanic	-	-	-	58.1%
Asian	-	-	-	2.0%
South Asian	-	1.8%	-	-
Pakistani/Bangladeshi	-	-	10.2%	-
Indian	-	-	1.9%	-
Chinese	-	1.7%	-	-
Indigenous (AU)/ Aboriginals (CN)	15.5%	8.2%	-	-
Mixed	-	-	4.4%	-
Race/ethnicity NEC	-	7.1%	1.9%	3.6%
B. MIDDLE EDUCATION GROUP				
Mean household income (SD)	21,416 (10,703)	23,738 (12,586)	21,902 (13,980)	19,699 (15,187)
Single parent household at W3	17.0%	19.5%	21.7%	26.7%
Mother <20 at birth	5.0%	5.3%	8.7%	14.5%
Mother >30 at birth	43.0%	31.7%	33.7%	21.5%
Under 18s in household at W3	2.55 (1.11)	2.34 (1.02)	2.37 (1.05)	2.44 (1.17)
Foreign-born parent	29.0%	30.3%	10.2%	20.7%
White (non-Hispanic for US)	-	81.0%	88.3%	49.0%
Black (non-Hispanic for US)	-	4.2%	2.4%	17.5%
Hispanic	-	-	-	27.1%
Asian	-	-	-	1.3%
South Asian	-	4.5%	-	-
Pakistani/Bangladeshi	-	-	3.9%	-
Indian	-	-	1.5%	-

	AU (N = 4265)	CN (N = 6812)	UK (N = 15,460)	US (N = 8,500)*
Chinese	-	1.4%	-	-
Indigenous (AU)/ Aboriginals (CN)	6.1%	2.2%	-	-
Mixed	-	-	2.9%	-
Race/ethnicity NEC	-	6.7%	1.0%	5.1%

C. HIGH EDUCATION GROUP

Mean household income (SD)	33,362 (17,908)	36,002 (19,610)	41,149 (21,197)	49,613 (34,779)
Single parent household at W3	5.0%	8.1%	6.3%	6.5%
Mother <20 at birth	1.0%	1.1%	1.0%	0.8%
Mother >30 at birth	63.0%	50.9%	57.2%	53.7%
Under 18s in household at W3	2.41 (0.86)	2.23 (0.87)	2.29 (0.85)	2.33 (0.98)
Foreign-born parent	39.0%	32.1%	15.9%	19.8%
White (non-Hispanic for US)	-	81.4%	88.1%	74.0%
Black (non-Hispanic for US)	-	2.5%	2.7%	6.1%
Hispanic	-	-	-	11.3%
Asian	-	-	-	4.9%
South Asian	-	5.5%	-	-
Pakistani/Bangladeshi	-	-	2.1%	-
Indian	-	-	2.3%	-
Chinese	-	3.1%	-	-
Indigenous (AU)/ Aboriginals (CN)	1.1%	0.8%	-	-
Mixed	-	-	3.4%	-
Race/ethnicity NEC	-	6.6%	1.3%	3.8%

* ECLS-B frequencies rounded to the nearest 50 in accordance with NCES reporting rules.

Table 4A.3. Average characteristics of families with 4 to 5 year old children, by parental income group and country

	AU (N = 4239)	CN (N = 6848)	UK (N = 15,460)	US (N = 8,500)*
A. LOW INCOME GROUP				
Low education (ISCED 2)	21.7%	19.9%	35.9%	28.9%
Middle education (ISCED 3/5B)	63.3%	53.1%	56.4%	68.0%
High education (ISCED 5A/6)	15.0%	27.0%	6.0%	3.2%
Mean household income (SD)	9,784 (2,839)	11,026 (2,954)	7,648 (1,678)	6,003 (2,536)
Single parent household at W3	41.0%	38.5%	55.7%	47.0%
Mother <20 at birth	10.0%	10.6%	22.6%	22.3%
Mother >30 at birth	40.0%	30.0%	21.1%	17.6%
Under 18s in household at W3	2.87 (1.44)	2.53 (1.25)	2.61 (1.29)	2.84 (1.39)
Foreign-born parent	37.0%	46.7%	16.2%	29.7%
White (non-Hispanic for US)	-	68.5%	74.2%	26.6%
Black (non-Hispanic for US)	-	6.6%	5.9%	31.3%
Hispanic	-	-	-	35.8%
Asian	-	-	-	1.2%
South Asian	-	8.8%	-	-
Pakistani/Bangladeshi	-	-	10.6%	-
Indian	-	-	1.7%	-
Chinese	-	2.6%	-	-
Indigenous (AU)/ Aboriginals (CN)	12.6%	4.8%	-	-
Mixed	-	-	6.0%	-
Race/ethnicity NEC	-	8.7%	1.6%	5.1%
B. MIDDLE INCOME GROUP				
Low education (ISCED 2)	5.8%	3.7%	7.8%	7.5%
Middle education (ISCED 3/5B)	59.2%	42.4%	61.2%	65.1%
High education (ISCED 5A/6)	35.0%	53.9%	30.5%	27.4%
Mean household income (SD)	23,029 (5,633)	26,835 (7,202)	23,257 (8,447)	22,367 (9,218)
Single parent household at W3	11.0%	9.7%	13.2%	18.4%
Mother <20 at birth	3.0%	2.1%	5.3%	10.5%
Mother >30 at birth	48.0%	39.3%	38.5%	26.7%
Under 18s in household at W3	2.50 (0.94)	2.23 (0.91)	2.40 (1.02)	2.39 (1.09)
Foreign-born parent	29.0%	28.2%	10.8%	23.2%
White (non-Hispanic for US)	-	83.1%	89.8%	54.9%

	AU (N = 4239)	CN (N = 6848)	UK (N = 15,460)	US (N = 8,500)*
Black (non-Hispanic for US)	-	2.8%	2.1%	11.2%
Hispanic	-	-	-	26.9%
Asian	-	-	-	2.4%
South Asian	-	4.5%	-	-
Pakistani/Bangladeshi	-	-	2.9%	-
Indian	-	-	1.8%	-
Chinese	-	1.9%	-	-
Indigenous (AU)/ Aboriginals (CN)	3.8%	1.4%	-	-
Mixed	-	-	2.4%	-
Race/ethnicity NEC	-	6.3%	1.1%	4.6%
C. HIGH INCOME GROUP				
Low education (ISCED 2)	1.7%	0.5%	0.6%	0.3%
Middle education (ISCED 3/5B)	26.0%	16.9%	23.4%	19.0%
High education (ISCED 5A/6)	72.3%	82.6%	75.5%	80.6%
Mean household income (SD)	49,081 (16,130)	57,889 (18,,239)	59,395 (15,383)	70,489 (34,766)
Single parent household at W3	3.0%	4.7%	2.9%	6.1%
Mother <20 at birth	0.0%	0.2%	0.3%	0.4%
Mother >30 at birth	67.0%	61.3%	66.3%	61.4%
Under 18s in household at W3	2.21 (0.79)	2.04 (0.79)	2.16 (0.77)	2.19 (0.88)
Foreign-born parent	40.0%	29.0%	13.5%	16.8%
White (non-Hispanic for US)	-	87.1%	92.4%	79.8%
Black (non-Hispanic for US)	-	1.7%	1.4%	3.7%
Hispanic	-	-	-	8.2%
Asian	-	-	-	4.4%
South Asian	-	1.9%	-	-
Pakistani/Bangladeshi	-	-	0.6%	-
Indian	-	-	1.6%	-
Chinese	-	3.4%	-	-
Indigenous (AU)/ Aboriginals (CN)	0.9%	0.4%	-	-
Mixed	-	-	3.1%	-
Race/ethnicity NEC	-	5.5%	0.7%	3.9%

* ECLS-B frequencies rounded to the nearest 50 in accordance with NCES reporting rules.

6. Regression estimates and pairwise t-tests of country differences

Table 4A.4. SES gradients in vocabulary outcomes

	By parental education			By parental income			
	Low	High	High-Low	Low	High	High-Low	Corr.
	<u>Unadjusted</u>						
AU PPVT	-0.496 (0.085)	0.352 (0.032)	0.848 (0.090)	-0.541 (0.053)	0.331 (0.035)	0.872 (0.058)	0.282
CN PPVT	-0.333 (0.094)	0.327 (0.039)	0.661 (0.094)	-0.491 (0.059)	0.252 (0.045)	0.743 (0.062)	0.246
UK BAS-NV	-0.537 (0.038)	0.432 (0.026)	0.969 (0.045)	-0.520 (0.032)	0.428 (0.030)	0.947 (0.044)	0.339
US PPVT	-0.610 (0.052)	0.601 (0.034)	1.212 (0.063)	-0.462 (0.043)	0.622 (0.040)	1.083 (0.052)	0.373
	<u>Adjusted for race/ethnicity/immigration</u>						
AU PPVT	-0.461 (0.080)	0.365 (0.032)	0.826 (0.086)	-0.484 (0.051)	0.354 (0.036)	0.838 (0.056)	
CN PPVT	-0.380 (0.115)	0.331 (0.038)	0.712 (0.114)	-0.493 (0.070)	0.215 (0.044)	0.708 (0.070)	
UK BAS-NV	-0.431 (0.028)	0.434 (0.024)	0.865 (0.032)	-0.403 (0.023)	0.406 (0.028)	0.809 (0.032)	
US PPVT	-0.367 (0.044)	0.496 (0.034)	0.863 (0.057)	-0.353 (0.038)	0.489 (0.041)	0.841 (0.045)	
	<u>Adjusted for all controls</u>						
AU PPVT	-0.388 (0.083)	0.291 (0.031)	0.679 (0.088)	-0.404 (0.053)	0.278 (0.035)	0.681 (0.059)	
CN PPVT	-0.327 (0.116)	0.312 (0.039)	0.639 (0.118)	-0.462 (0.073)	0.177 (0.046)	0.639 (0.078)	
UK BAS-NV	-0.331 (0.029)	0.378 (0.025)	0.709 (0.035)	-0.308 (0.025)	0.341 (0.030)	0.649 (0.036)	
US PPVT	-0.319 (0.043)	0.437 (0.034)	0.756 (0.057)	-0.291 (0.039)	0.420 (0.042)	0.711 (0.047)	
	<u>Majority native ethnic group</u>						
AU PPVT	-0.223 (0.074)	0.374 (0.034)	0.598 (0.074)	-0.326 (0.057)	0.301 (0.039)	0.627 (0.059)	
CN PPVT	-0.420 (0.109)	0.328 (0.043)	0.748 (0.109)	-0.431 (0.074)	0.201 (0.046)	0.632 (0.076)	
UK BAS-NV	-0.413 (0.031)	0.413 (0.027)	0.826 (0.035)	-0.401 (0.027)	0.370 (0.030)	0.771 (0.035)	
US PPVT	-0.361 (0.083)	0.455 (0.031)	0.816 (0.083)	-0.381 (0.052)	0.397 (0.034)	0.778 (0.055)	

Standard errors (adjusted for complex survey design) in parenthesis. All numbers (other than correlations) are standard deviation units. Low and High denote the deviations in the mean outcome from the (omitted) middle SES group among the low and high SES groups respectively.

Race/ethnicity/immigration controls: Dummy variable for either parent foreign-born; set of country-specific dummies for race/ethnic category (see Table 5).

All controls: Race/ethnicity/immigration controls described above, plus: dummies for mother<20; mother>30; single parent household at Wave 3; number of under-18s in the household at Wave 3 (continuous variable).

Majority native ethnic group defined as neither parent foreign-born and race/ethnicity as follows (% total sample): AU – Non-indigenous (67%); CN – White (70%); UK – White (79%); US – White non-Hispanic (42%).

Table 4A.5. Pairwise t-tests of country differences in SES vocabulary gradients

	By parental education			By parental income		
	Low	High	High-Low	Low	High	High-Low
<u>Unadjusted</u>						
AU vs CN	-1.29	0.51	1.44	-0.62	1.40	1.52
AU vs UK	0.44	-1.94	-1.20	-0.33	-2.08	-1.04
AU vs US	1.15	-5.33	-3.30	-1.15	-5.46	-2.72
CN vs UK	2.02	-2.26	-2.95	0.43	-3.26	-2.68
CN vs US	2.58	-5.33	-4.85	-0.40	-6.18	-4.20
UK vs US	1.14	-3.96	-3.14	-1.09	-3.87	-2.00
<u>Adjusted for race/ethnicity/immigration</u>						
AU vs CN	-0.58	0.68	0.80	0.10	2.41	1.45
AU vs UK	-0.36	-1.74	-0.43	-1.46	-1.14	0.45
AU vs US	-1.03	-2.83	-0.36	-2.07	-2.49	-0.04
CN vs UK	0.43	-2.30	-1.29	-1.23	-3.64	-1.31
CN vs US	-0.11	-3.25	-1.19	-1.76	-4.56	-1.60
UK vs US	-1.23	-1.50	0.03	-1.12	-1.69	-0.58
<u>Adjusted for all controls</u>						
AU vs CN	-0.43	-0.43	0.27	0.64	1.73	0.44
AU vs UK	-0.65	-2.20	-0.32	-1.62	-1.37	0.47
AU vs US	-0.74	-3.15	-0.74	-1.70	-2.60	-0.39
CN vs UK	0.03	-1.42	-0.57	-1.99	-2.98	-0.12
CN vs US	-0.06	-2.39	-0.89	-2.06	-3.89	-0.79
UK vs US	-0.23	-1.40	-0.70	-0.36	-1.54	-1.05
<u>Majority native ethnic group</u>						
AU vs CN	1.49	0.84	-1.14	1.12	1.66	-0.05
AU vs UK	2.36	-0.89	-2.78	1.19	-1.41	-2.11
AU vs US	1.23	-1.75	-1.96	0.72	-1.86	-1.87
CN vs UK	-0.06	-1.67	-0.68	-0.38	-3.06	-1.66
CN vs US	-0.43	-2.39	-0.50	-0.55	-3.40	-1.55
UK vs US	-0.59	-1.03	0.11	-0.34	-0.59	-0.11

T-statistics > |1.96| marked in bold. T-tests relate to the OLS coefficients in Table A4.

Table 4A.6. SES gradients in externalizing behaviour problems

	By parental education			By parental income			Corr.
	Low	High	High-Low	Low	High	High-Low	
	<u>Unadjusted</u>						
AU SDQ	-0.298 (0.083)	0.345 (0.032)	0.643 (0.083)	-0.301 (0.053)	0.243 (0.037)	0.543 (0.056)	0.157
CN	-0.047 (0.085)	0.188 (0.040)	0.235 (0.082)	-0.300 (0.060)	0.061 (0.045)	0.361 (0.067)	0.125
UK SDQ	-0.431 (0.037)	0.368 (0.021)	0.798 (0.037)	-0.404 (0.028)	0.312 (0.026)	0.716 (0.032)	0.245
US	-0.280 (0.049)	0.365 (0.029)	0.645 (0.052)	-0.310 (0.038)	0.316 (0.028)	0.627 (0.044)	0.224
	<u>Adjusted for race/ethnicity/immigration</u>						
AU SDQ	-0.278 (0.085)	0.338 (0.032)	0.616 (0.084)	-0.282 (0.054)	0.236 (0.038)	0.518 (0.056)	
CN	-0.012 (0.104)	0.175 (0.042)	0.187 (0.101)	-0.318 (0.074)	0.042 (0.046)	0.360 (0.080)	
UK SDQ	-0.426 (0.037)	0.365 (0.020)	0.791 (0.037)	-0.406 (0.028)	0.312 (0.026)	0.718 (0.033)	
US	-0.266 (0.049)	0.341 (0.031)	0.607 (0.054)	-0.293 (0.041)	0.293 (0.027)	0.586 (0.047)	
	<u>Adjusted for all controls</u>						
AU SDQ	-0.212 (0.082)	0.278 (0.034)	0.491 (0.083)	-0.210 (0.053)	0.176 (0.038)	0.386 (0.057)	
CN	0.032 (0.105)	0.153 (0.041)	0.121 (0.102)	-0.280 (0.072)	0.015 (0.048)	0.295 (0.080)	
UK SDQ	-0.334 (0.038)	0.292 (0.020)	0.625 (0.039)	-0.267 (0.032)	0.242 (0.027)	0.509 (0.039)	
US	-0.219 (0.050)	0.274 (0.037)	0.493 (0.061)	-0.235 (0.044)	0.224 (0.031)	0.459 (0.056)	
	<u>Majority native ethnic group</u>						
AU SDQ	-0.266 (0.113)	0.387 (0.040)	0.652 (0.110)	-0.270 (0.072)	0.319 (0.044)	0.589 (0.075)	
CN	-0.099 (0.123)	0.133 (0.046)	0.232 (0.119)	-0.388 (0.081)	0.050 (0.052)	0.438 (0.091)	
UK SDQ	-0.433 (0.044)	0.382 (0.022)	0.815 (0.044)	-0.453 (0.031)	0.326 (0.027)	0.778 (0.036)	
US	-0.337 (0.088)	0.378 (0.033)	0.716 (0.089)	-0.534 (0.055)	0.265 (0.037)	0.799 (0.060)	

. Standard errors (adjusted for complex survey design) in parenthesis. See notes to Table A4.

Table 4A.7. Pairwise t-tests of country differences in SES externalizing behavior gradients

	By parental education			By parental income		
	Low	High	High-Low	Low	High	High-Low
<u>Unadjusted</u>						
AU vs CN	-2.10	3.07	3.51	-0.01	3.12	2.10
AU vs UK	1.46	-0.60	-1.70	1.73	-1.52	-2.69
AU vs US	-0.19	-0.46	-0.02	0.15	-1.57	-1.18
CN vs UK	4.13	-4.02	-6.26	1.56	-4.85	-4.80
CN vs US	2.37	-3.59	-4.23	0.14	-4.86	-3.34
UK vs US	-2.44	0.08	2.40	-1.99	-0.10	1.63
<u>Adjusted for race/ethnicity/immigration</u>						
AU vs CN	-1.98	3.07	3.26	0.39	3.25	1.62
AU vs UK	1.60	-0.70	-1.90	2.04	-1.66	-3.07
AU vs US	-0.12	-0.06	0.09	0.16	-1.24	-0.92
CN vs UK	3.75	-4.07	-5.61	1.12	-5.06	-4.16
CN vs US	2.21	-3.17	-3.66	-0.30	-4.70	-2.44
UK vs US	-2.60	0.65	2.79	-2.28	0.51	2.29
<u>Adjusted for all controls</u>						
AU vs CN	-1.83	2.36	2.81	0.78	2.66	0.93
AU vs UK	1.35	-0.35	-1.47	0.92	-1.41	-1.77
AU vs US	0.07	0.09	-0.02	0.36	-0.98	-0.91
CN vs UK	3.28	-3.03	-4.62	-0.16	-4.15	-2.41
CN vs US	2.16	-2.19	-3.13	-0.53	-3.70	-1.68
UK vs US	-1.83	0.43	1.82	-0.59	0.44	0.73
<u>Majority native ethnic group</u>						
AU vs CN	-1.00	4.15	2.60	1.09	3.96	1.28
AU vs UK	1.38	0.10	-1.37	2.33	-0.14	-2.28
AU vs US	0.50	0.17	-0.45	2.90	0.94	-2.20
CN vs UK	2.56	-4.85	-4.59	0.75	-4.73	-3.47
CN vs US	1.57	-4.30	-3.26	1.48	-3.40	-3.32
UK vs US	-0.97	0.10	1.00	1.27	1.34	-0.30

T-statistics > |1.96| marked in bold. T-tests relate to the OLS coefficients in Table A6.

Table 4A.8. SES gradients in other cognitive outcomes (no controls)

	By parental education			By parental income			
	Low	High	High-Low	Low	High	High-Low	Corr.
<u>Math/number skills</u>							
CN Number	-0.276	0.315	0.591	-0.288	0.306	0.594	0.200
Knowledge	(0.098)	(0.040)	(0.099)	(0.066)	(0.047)	(0.073)	
UK Bracken	-0.503	0.530	1.033	-0.494	0.480	0.974	0.347
Math (age 3)	(0.026)	(0.035)	(0.039)	(0.027)	(0.043)	(0.045)	
US Math	-0.472	0.745	1.217	-0.573	0.670	1.243	0.430
	(0.052)	(0.035)	(0.058)	(0.033)	(0.036)	(0.044)	
<u>Copying</u>							
AU (WAI)	-0.179	0.257	0.436	-0.188	0.187	0.376	0.109
	(0.072)	(0.031)	(0.074)	(0.045)	(0.035)	(0.052)	
CN (WAI)	-0.277	0.209	0.486	-0.221	0.180	0.401	0.131
	(0.108)	(0.041)	(0.106)	(0.061)	(0.052)	(0.071)	
US	-0.156	0.347	0.503	-0.296	0.282	0.578	0.211
	(0.051)	(0.034)	(0.052)	(0.039)	(0.036)	(0.043)	
<u>Who Am I (WAI) Symbols</u>							
AU	-0.285	0.358	0.643	-0.355	0.248	0.603	0.183
	(0.072)	(0.034)	(0.076)	(0.045)	(0.035)	(0.051)	
CN	-0.311	0.248	0.559	-0.215	0.204	0.419	0.154
	(0.085)	(0.041)	(0.084)	(0.066)	(0.049)	(0.075)	
<u>Who Am I (WAI) Total</u>							
AU	-0.262	0.345	0.607	-0.310	0.243	0.553	0.166
	(0.072)	(0.033)	(0.074)	(0.045)	(0.034)	(0.051)	
CN	-0.329	0.256	0.585	-0.243	0.215	0.458	0.160
	(0.084)	(0.041)	(0.082)	(0.063)	(0.050)	(0.073)	
<u>Other</u>							
UK Bracken	-0.551	0.551	1.102	-0.544	0.502	1.046	0.375
Total (age 3)	(0.028)	(0.034)	(0.039)	(0.028)	(0.041)	(0.043)	
UK BAS-NV	-0.537	0.357	0.895	-0.523	0.307	0.831	0.305
(age 3)	(0.040)	(0.025)	(0.048)	(0.035)	(0.028)	(0.044)	
US Literacy	-0.489	0.737	1.225	-0.497	0.688	1.186	0.398
	(0.049)	(0.031)	(0.051)	(0.040)	(0.037)	(0.046)	

. Standard errors (adjusted for complex survey design) in parenthesis. See notes to Table A4.

Table 4A.9. Pairwise t-tests of country differences in SES gradients in other cognitive outcomes

	By parental education			By parental income		
	Low	High	High-Low	Low	High	High-Low
	<u>Math/number skills</u>					
CN vs UK	2.24	-4.03	-4.15	2.87	-2.72	-4.46
CN vs US	1.77	-8.05	-5.45	3.84	-6.13	-7.63
UK vs US	-0.53	-4.36	-2.62	1.84	-3.37	-4.27
	<u>Copying</u>					
AU vs CN	0.75	0.94	-0.39	0.43	0.12	-0.29
AU vs US	-0.26	-1.94	-0.74	1.81	-1.87	-2.99
CN vs US	-1.01	-2.58	-0.14	1.04	-1.60	-2.13
	<u>Who Am I (WAI) Symbols</u>					
AU vs CN	0.24	2.06	0.74	-1.77	0.73	2.03
	<u>Who Am I (WAI) Total</u>					
AU vs CN	0.60	1.69	0.20	-0.86	0.47	1.06

T-statistics > |1.96| marked in bold. T-tests relate to the OLS coefficients in Table A8.

Table 4A.10. SES gradients in other behavior outcomes

	By parental education			By parental income			Corr.
	Low	High	High-Low	Low	High	High-Low	
<u>Hyperactivity/inattention</u>							
AU	-0.275 (0.088)	0.322 (0.032)	0.597 (0.086)	-0.281 (0.051)	0.206 (0.039)	0.487 (0.056)	0.134
CN	-0.109 (0.082)	0.177 (0.041)	0.286 (0.079)	-0.254 (0.060)	0.044 (0.045)	0.298 (0.066)	0.109
UK	-0.332 (0.035)	0.350 (0.020)	0.681 (0.035)	-0.325 (0.026)	0.283 (0.027)	0.609 (0.032)	0.210
US	-0.175 (0.044)	0.368 (0.029)	0.542 (0.047)	-0.215 (0.038)	0.308 (0.034)	0.523 (0.045)	0.187
<u>Conduct problems</u>							
AU	-0.229 (0.086)	0.260 (0.031)	0.488 (0.086)	-0.225 (0.056)	0.206 (0.035)	0.431 (0.056)	0.131
CN	0.046 (0.092)	0.131 (0.039)	0.085 (0.088)	-0.204 (0.061)	0.068 (0.043)	0.272 (0.066)	0.095
UK	-0.468 (0.037)	0.299 (0.020)	0.767 (0.037)	-0.418 (0.029)	0.278 (0.025)	0.696 (0.033)	0.239
US	-0.306 (0.057)	0.263 (0.031)	0.569 (0.060)	-0.329 (0.039)	0.242 (0.029)	0.571 (0.046)	0.203
<u>SDQ Total Difficulties Score</u>							
AU	-0.394 (0.090)	0.300 (0.032)	0.694 (0.089)	-0.409 (0.058)	0.236 (0.035)	0.645 (0.060)	0.190
UK	-0.449 (0.037)	0.367 (0.023)	0.816 (0.038)	-0.471 (0.032)	0.348 (0.025)	0.820 (0.034)	0.280
<u>Other</u>							
UK External (age 3)	-0.406 (0.050)	0.391 (0.025)	0.798 (0.053)	-0.428 (0.033)	0.352 (0.028)	0.780 (0.041)	0.279
UK SDQ TDS (age 3)	-0.439 (0.058)	0.432 (0.025)	0.871 (0.060)	-0.531 (0.039)	0.364 (0.028)	0.895 (0.043)	0.321

. Standard errors (adjusted for complex survey design) in parenthesis. See notes to Table A4.

Table 4A.11. Pairwise t-tests of country differences in SES gradients in other behaviour outcomes

	By parental education			By parental income		
	Low	High	High-Low	Low	High	High-Low
<u>Hyperactivity/inattention</u>						
AU vs CN	-1.38	2.81	2.66	-0.34	2.71	2.20
AU vs UK	0.60	-0.74	-0.90	0.77	-1.62	-1.90
AU vs US	-1.01	-1.06	0.56	-1.02	-1.97	-0.50
CN vs UK	2.73	-4.21	-5.02	1.19	-5.26	-4.74
CN vs US	0.71	-3.78	-2.79	-0.55	-4.66	-2.83
UK vs US	-2.81	-0.50	2.37	-2.36	-0.58	1.55
<u>Conduct problems</u>						
AU vs CN	-2.18	2.56	3.28	-0.25	2.47	1.82
AU vs UK	2.56	-1.06	-2.99	3.07	-1.66	-4.07
AU vs US	0.75	-0.08	-0.78	1.53	-0.79	-1.92
CN vs UK	5.56	-4.29	-7.76	3.49	-4.84	-6.39
CN vs US	3.24	-2.64	-4.56	1.72	-3.34	-3.69
UK vs US	-2.38	0.98	2.83	-1.84	0.94	2.21
<u>SDQ Total Difficulties Score</u>						
AU vs UK	0.56	-1.72	-1.26	0.93	-2.58	-2.53

T-statistics > |1.96| marked in bold. T-tests relate to the OLS coefficients in Table A10.

7. Sample sizes

Table 4A.12. Observations in OLS regressions

	AU (Eligible N=4386)		CN (Eligible N=7147)			UK (Eligible N=15460)		US (Eligible N=8941)	
	Ed	Inc	Ed	Inc		Ed	Inc	Ed	Inc
				Reg	Corr				
Vocabulary	4265	4239	6284	6234	4843	15036	14971	8450	8350
Externalizing behavior	3822	3798	6722	6758	5367	13383	13333	8900	8800
Math/number skills			6194	6228	4837	12489	12397	8300	8250
Copying	4227	4201	5947	5980	4589			8000	7900
WAI Symbols	4227	4201	5947	5980	4589				
WAI Total	4227	4201	5947	5980	4589				
Bracken SRA Total						12489	12397		
Vocabulary (Wave 2)						13104	12999		
Literacy								8250	8200
Hyperactivity/ inattention	3822	3798	6729	6765	5374	13751	13702	8900	8850
Conduct problems	3824	3800	6748	6784	5393	14238	14186	8900	8850
SDQ Total Behavioral Difficulties	3820	3796				11782	11732		
Externalizing behavior (Wave 2)						10837	10775		
SDQ Total Behavioral Difficulties (Wave 2)						9190	9132		
Vocabulary majority sample	2837	2825	4282	4391		11849	11836	3500	3500
Externalizing behavior majority sample	2599	2589	4600	4716		10948	10920	3600	3600
Vocabulary (+controls sample)			5356	5373					
Externalizing behavior (+controls sample)			5796	5815					

Chapter 9 Online Appendix

Greg J. Duncan, Lars Bergman, Kathryn Duckworth, Katja Kokko, Anna-Liisa Lyyra, Molly Metzger,
Lea Pulkkinen, Sharon Simonton

Data Appendix

The U.S. Beginning School Study (BSS)

The Beginning School Study (BSS) has followed a group of 838 individuals from their first grade year in 1982. Sampling began with a stratified random sample of 20 Baltimore, Maryland (U.S.) public schools. From there, roughly 12 first graders were randomly sampled from each first grade classroom, with a participation rate of 97% among those selected.

Interviews were conducted recurrently between first grade and ages 28/29. For adult outcomes, the BSS's "Mature Adult" survey consists of 660 (79%) of the original participants at the age of 28/29.

Many children attending Baltimore public schools in the early 1980s came from disadvantaged families, although these children were not as uniformly disadvantaged as the children in many urban school districts today. Of the respondents in the age 28/29 interviews, 56% are African-American, with virtually all of the remainder Caucasian. Only about a third of the analytic sample lived with a single parent at the baseline year, but over two-thirds were eligible for a free or reduced price lunch at some point during their elementary school years.

Interviewing rules limited the sample size in some years. In the Grade 2 and 3 follow-ups, the study followed only children attending the originally sampled schools. In the Grade 4, 5, & 6 follow-ups, the researchers attempted to follow all children still attending any Baltimore public school. Beginning with the Grade 7 follow-up, they attempted to contact the entire original sample.

For more on sampling methods and sample description, see: Entwisle, D. R., Alexander, K. L., and Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education*, 80, 114–138.

The Swedish Individual Development and Adaptation (IDA)

The longitudinal research program Individual Development and Adaptation (IDA) was initiated by David Magnusson in the early 1960s; and he directed it until 1996, when Lars Bergman became the principal investigator. General descriptions of the IDA data base are provided in Magnusson (1998). The data base consists of three whole school grade cohorts, but the present study uses only data from the cohort born in 1955. The sample characteristics of this cohort are described below.

In the present study, data were used from the first data collection in 1965 for the complete school grade cohort of children in grade 3 from the town of Örebro, who were then about 10 years of age. This cohort constituted our target sample and included 517 boys and 510 girls. Basic data from grade 3 were available for 958 of these children or 93 percent of the target sample. It is fairly representative of a Swedish urban population, except that the socioeconomic level of the children's families was slightly above average (Bergman, 1973). Two extensive data collections were performed when the individuals in question were middle-aged, one for females in 1998 when they were 43 and one for males in 2002 when they were 47. Four hundred and thirty females and 390 males took part (84 percent and 75 percent of the target sample, respectively). With regard to school achievement and the parents' education in grade 3, there were no significant differences between those who took part in the data collections in middle age and those who did not.

The Finnish Jyväskylä Longitudinal Study of Personality and Social Development (JYLS)

The Jyväskylä Longitudinal Study of Personality and Social Development (JYLS) was begun by Lea Pulkkinen in 1968 when she randomly selected 12 second-grade school classes in the town of Jyväskylä, Finland to become part of the study sample. All the participants in the 12 classes participated in the study; the original sample included 173 girls and 196 boys. Ninety-five percent of the participants were born in 1959 (the rest either in 1958 or 1960); the participants were about 8 years old. At age 8, children's social behavior (the main focus was on emotional and behavioral regulation) was assessed

using teacher ratings and peer nominations, and information about school success was collected from teachers. The next main data collection phase took place in 1974 when the participants were 14 years old.

All of the participants from the original sample were again contacted in 1986, at the age of 27. Data were then gathered by means of a mailed Life Situation Questionnaire (LSQ1) and semi-structured psychological interview, which yielded information about such factors as family relationships, housing, financial situation, education, occupation, work, employment, leisure activities, personality characteristics, satisfaction with life, use of alcohol, smoking, self-rated health, life events, and antisocial acts. The LSQ1 was completed by 155 women (90% of the original sample) and 166 men (85%), and 142 women (82%) and 150 men (77%) were interviewed.

The original participants were again contacted in 1995 when they were 36 years old. The LSQ2 and semi-structured psychological interview yielded information about topics similar to those covered at age 27. The LSQ2 was completed by 150 women (87% of the original sample) and 161 men (82%); 137 women (79%) and 146 men (76% of the available sample – two men had died) were interviewed.

The next main data collection phase took place in 2001 when the participants were 42 years old. The mailed LSQ3 and semi-structured psychological interview covered topics similar to those addressed at ages 36 and 42; some new topics and methods were added. For the first time, the data collection included a medical examination with laboratory tests. By age 42, the available sample was reduced to 158 women and 186 men: one woman and five men had died and 14 women (8% of the original sample) and five men (3%) had refused to take part in the study. Of this available sample, 134 women (85%) and 151 men (81%) participated in a follow-up study. The LSQ3 was completed by 132 women (84%) and 147 men (79%); 120 women (76%) and 123 men (66%) were interviewed; and 120 women (76%) and 121 men (65%) participated in the medical examination. The most recent data collection was completed in 2009 with 50-years-old participants. The data collection methods were similar to those at age 42. By age 50, the available sample was reduced to 149 women and 174 men: Four women and eight men had died and 20 women (12% of the original sample) and 14 men (7%) had refused to take part in the study. Of this available sample, 127 women (85%) and 141 men (81%) participated in a follow-up study. The LSQ4

was completed by 127 women (85%) and 140 men (80%); 111 women (74%) and 113 men (65%) were interviewed; and 114 women (77%) and 110 men (63%) participated in the medical examination.

At ages 42 and 50, the participants provide a good representation of the original random sample, indicating that there has been no selective attrition. When compared on characteristics such as marital status, number of children and employment gathered by Statistics Finland, the age-42 sample also appears representative of the Finnish age-cohort group born in 1959.

The Finnish measure of completed schooling represents highest level of education rather than actual years spent in school. Since university schooling is state-supported (but the support is time-limited), there is less pressure to graduate quickly than in those countries where there are tuition fees. Individuals can spend seven to eight years at the university or three years in vocational school and then another three years in vocational college. Consequently, all the university graduates were assigned the same number of study years. The same logic applies to all the other educational institutions. The ISCED years refer to the years assumed to be used to obtain the highest level of completed education. This may have underestimated the schooling years of those who have first graduated from, let's say vocational school and then from the vocational college.

For more information about the JYLS data collection phases (only the main data collection phases are described above), methods, and theoretical background, see

<http://www.jyu.fi/ytk/laitokset/psykologia/en/research/jyls>

The National Child Development Study 1958 Birth Cohort (NCDS)

The 1958 National Child Development Study is a longitudinal study of British children who were born during the week of March 3 through 9, 1958. A total of 17,414 mothers, representing 98% of all births that week, were interviewed (Shepherd, 1985). Follow-up interviews were conducted when the children were age 7 (1965; n = 15,468), 11 (1969; n = 15,503), and 16 years (1974; n = 14,761). These three ages were selected since they were important transition points in the children's educational progress through the British school system. Adult follow-up survey interviews were conducted when the

participants were 23 (1981; n = 12,537), 33 (1991; n = 11, 469), 42 (2000; n = 11, 419) and 46 (2004; n = 9,534) years of age.

For more on sampling methods and sample description, see: Shepherd, Peter M. (1985). *The National Child Development Study: An introduction to the background of the study and the methods of data collection* (Working Paper No. 1, Social Statistics Research Unit, City University, London).

The 1970 British Cohort Study (UK)

The UK 1970 British Birth Cohort (BCS) is a nationally representative longitudinal study which has followed into adulthood a cohort of children born in England, Scotland, Wales and Northern Ireland during one week in April 1970. The birth sample of 17,287 infants was approximately 97% of the target birth population. Since the birth survey there have been six other major data collection sweeps aimed at monitoring these children's health, education, social and economic circumstances. These were carried out in 1975 (age 5), 1980 (age 10), 1986 (age 16), 1996 (age 26), 2000 (age 30) and 2004 (age 34).

Data for the BCS70 have been collected from a variety of sources including questionnaires completed by mothers, midwives present at birth and other healthcare professionals, head and class teachers, school health service personnel and the cohort member themselves. These data have also been collected through a number of means including paper and electronic questionnaires, clinical records, medical examinations, tests of ability, educational assessments and diaries.

A teacher strike in 1986 in England and Wales meant that much of the educational data for age 16 is totally missing for a big swathe of the cohort. Moreover, the age 16 maths data was only recently (and only partially) made available. There are only N=3677 maths test scores available to analyse while there are N=6003 spelling and vocabulary assessments. Add to this the fact that those schools who did administer the tests (i.e. those not on strike) are likely to be the more advantaged schools, the data available are unlikely to be as representative of the full cohort as in the NCDS sample.

Attrition reduced the achieved sample to 10,833 in the age-30 (2000) sweep of the survey (Ferri, Bynner and Wadsworth, 2003) and to 9,665 in the age-34 survey (Dodgeon *et al.*, 2006).

Representativeness of the original birth cohort has been maintained with only slight biases in the currently participating sample towards women and towards the more educated (Ferri *et al.*, 2003, op cit). However, missing data at the item response level (again maintaining broadly the representativeness of the original cohort) reduces the effective data set for most analyses to between 9,000 and 10,000 cases for the age-30 adult outcomes and further still if using age-34 adult outcomes.

Chapter 10 Online Appendix
Appendix 10A. Descriptive statistics

Table 10A.1. U.S. Descriptive statistics by parent education

	Parent Education									
	Full sample KF(5.7)		Level 2		Level 3		Level 5B		Level 5A/6	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Parent Education										
Level 2	.07	.25								
Level 3-4	.26	.44								
Level 5B	.34	.47								
Level 5A-6	.34	.47								
Family Income	55782	58968	18822	16449	33602	31935	46392	40414	88860	77122
Preschool	.79	.41	.66	.47	.73	.44	.78	.41	.86	.35
Race/Ethnicity										
White	.62	.49	.30	.46	.54	.50	.61	.49	.75	.43
Black	.15	.36	.30	.46	.23	.42	.16	.37	.06	.25
Hispanic	.12	.33	.28	.45	.14	.35	.13	.34	.07	.26
Others	.11	.31	.12	.32	.10	.29	.10	.30	.12	.32
Family Structure										
Bio-two parents	.67	.47	.38	.48	.53	.50	.65	.48	.85	.36
Blend families-step parents	.08	.27	.08	.28	.11	.31	.09	.29	.04	.19
Single parent-mother/father	.22	.41	.46	.50	.31	.46	.23	.42	.08	.27
Any adopted or others	.04	.20	.08	.27	.04	.20	.04	.19	.03	.18
School Type-Private	.23	.42	.03	.18	.11	.32	.20	.40	.38	.48
Disability Status	.14	.35	.15	.36	.15	.36	.15	.36	.13	.34
Gender-female	.49	.50	.47	.50	.49	.50	.49	.50	.49	.50
Sample size	15,654		1,083		4,017		5,296		5,258	

Table 10A.2. U.S. Descriptive statistics by income quintile

	Income Quintile									
	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Parent Education										
Level 2	.23	.42	.11	.31	.04	.21	.01	.11	.00	.07
Level 3-4	.43	.49	.40	.49	.30	.46	.18	.38	.07	.25
Level 5B	.27	.45	.40	.49	.43	.50	.40	.49	.20	.40
Level 5A-6	.06	.25	.10	.29	.22	.42	.41	.49	.73	.45
Family Income	9538	4662	22760	3474	37899	5619	59364	7271	122674	88053
Preschool	.76	.43	.74	.44	.73	.44	.80	.40	.87	.33
Race/Ethnicity										
White	.32	.47	.48	.50	.65	.48	.74	.44	.78	.42
Black	.39	.49	.22	.42	.13	.33	.07	.26	.05	.21
Hispanic	.16	.37	.18	.38	.13	.34	.10	.30	.08	.27
Others	.13	.34	.12	.32	.10	.30	.09	.29	.10	.30
Family Structure										
Bio-two parents	.27	.45	.50	.50	.69	.46	.83	.38	.88	.32
Blend families-step parents	.08	.27	.10	.30	.11	.31	.07	.26	.04	.20
Single parent-mother/father	.58	.49	.35	.48	.16	.37	.08	.27	.05	.22
Any adopted or others	.07	.26	.05	.21	.04	.19	.03	.16	.03	.17
School Type-Private	.06	.23	.11	.31	.19	.39	.30	.46	.40	.49
Disability Status	.17	.37	.17	.38	.14	.35	.13	.34	.12	.33
Gender-female	.49	.50	.48	.50	.49	.50	.50	.50	.48	.50
Sample size	2,630		2,393		3,549		3,424		3,658	

Table 10A.3. England descriptive statistics by parental education

	Parent education									
	Full sample		Level 2		Level 3-4		Level 5B		Level 5A-6	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Parent Education										
Level 2	.11	.31								
Level 3	.48	.50								
Level 5b	.22	.41								
Level 5a6	.19	.39								
Family Income	15,124	7,199	9,406	4,908	13,712	6,511	15,711	6,437	21,372	6,388
Nursery (age 3 to 4)	.45	.50	.41	.49	.42	.49	.42	.49	.57	.49
Nonwhite	.05	.22	.06	.24	.05	.21	.04	.21	.06	.23
Family Structure (age 3.9)										
Bio-two parents	.83	.38	.65	.48	.81	.40	.86	.34	.94	.24
Blend families-step parents	.05	.21	.07	.26	.05	.23	.04	.19	.01	.11
Single parent-mother/father	.12	.32	.26	.44	.13	.33	.09	.28	.05	.21
Any adopted or others	.01	.10	.02	.14	.01	.11	.01	.10	.00	.06
Special Educational Needs (age 11)	.19	.39	.31	.46	.20	.40	.17	.38	.10	.31
Female	.48	.50	.49	.50	.48	.50	.49	.50	.48	.50
Sample size	12,986		1,407		6,292		2,845		2,442	

Notes. Sample is all ALSPAC core cohort children with at least one valid Key Stage Math or Reading record. Multiple imputation used for missing values. Income is average annual net household income at ages 2 and 3 in 1995 British pounds.

Table 10A.4. England descriptive statistics by income quintile

	Parent education									
	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Parent Education										
Level 2	.26	.44	.13	.33	.08	.27	.03	.18	.01	.11
Level 3	.56	.50	.56	.50	.54	.50	.44	.50	.29	.45
Level 5b	.14	.35	.23	.42	.27	.45	.27	.45	.18	.38
Level 5a6	.04	.20	.08	.27	.10	.31	.25	.44	.52	.50
Family Income	6,164	1,325	11,286	1,392	14,506	1,061	19,703	2,327	26,859	797
Nursery (age 3 to 4)	.41	.49	.40	.49	.40	.49	.46	.50	.57	.49
Nonwhite	.09	.29	.05	.21	.03	.17	.04	.19	.03	.18
Family Structure (age 3.9)										
Bio-two parents	.49	.50	.85	.35	.95	.22	.94	.23	.98	.15
Blend families-step parents	.09	.29	.06	.24	.02	.15	.03	.16	.01	.11
Single parent-mother/father	.40	.49	.08	.26	.02	.15	.02	.15	.01	.08
Any adopted or others	.02	.15	.01	.11	.01	.08	.01	.08	.00	.06
Special Educational Needs (age 11)	.28	.45	.21	.40	.17	.38	.14	.35	.12	.32
Female	.48	.50	.48	.50	.49	.50	.50	.50	.49	.50
Sample size	3,063		2,485		2,449		2,801		2,187	

Notes. Sample is all ALSPAC core cohort children with at least one valid Key Stage Math or Reading record. Multiple imputation used for missing values. Income is average annual net household income at ages 2 and 3 in 1995 British pounds.

Appendix 10B: Gradients in socio-emotional development

Outcomes

Our analysis of gradients in socio-emotional development mirrors that conducted for achievement outcomes as far as possible. However, a number of differences in the samples used and the underlying data sources were unavoidable.

The ECLS-K provides teacher reports of children's externalizing and internalizing behavior problems at five assessment points between age 5 and age 11.¹ Externalizing problem behavior or aggressive behavior is measured by a five-item scale that asks teachers how frequently the child fights, argues, gets angry, acts impulsively, or disturbs ongoing activities. The assessment of internalizing problems asks teachers about children's problems such as depression and anxiety. The response scale ranges from 1 (child never displays this behavior) to 4 (child often displays this behavior). Thus, a lower score on these scales represents better behavior or fewer behavior problems.

Behavior outcomes in the ECLS-K were collected directly along with achievement outcomes by researchers at each survey wave. In contrast, while achievement data for the England sample were taken from the matched administrative NPD, behavior outcomes were collected with ALSPAC postal questionnaires, and for this reason differ from the achievement outcomes in both the timing of measurement and the proportion of the sample with valid information. After imputing missing data, the England behaviour sample size is 9,465 as compared with 12,986 for achievement outcomes, reflecting attrition (which occurred disproportionately among less advantaged groups) in the later waves of the ALSPAC survey.

Behavior problems are measured in ALSPAC through parental reports about their child in response to the Strengths and Difficulties Questionnaire (SDQ), a very widely used measure of behavior problems. We use a subset of the SDQ sub-scales to derive a measure of externalizing problems (the sum of the Hyperactivity and Conduct problems sub-scales) and a measure of internalizing problems (the Emotional symptoms sub-scale), corresponding to the externalizing and internalizing scales in the ECLS-K. The timing of the SDQ assessments in ALSPAC was determined by the child's exact age in months, rather than their school grade, with scores collected at ages 6, 8, 9,

11 and 13.² For this reason, there is no one-to-one correspondence between SDQ assessment and the child's grade or year of schooling.

Development of the gradients in socio-emotional development

Figure 10B.1 displays trends in raw (unstandardized) externalizing and internalizing behavior scores by parental education and income group in the US sample, with regression estimates provided in Table 10B.1. As mentioned earlier, these scores measure the teachers' rating of the frequency with which the child displays a set of behavior problems, with lower scores indicating better functioning. Already at school entry at age 5, children of the most highly educated parents have much lower externalizing behavior scores, and this advantage widens over time. While levels of externalizing problems continue to be low for the children of the most educated, reported problems rise in frequency with age for the other groups, with a particularly noteworthy increase for the children of the least-educated parents. Externalizing problems decline for all groups between age 9 and 11, but nevertheless sizable gaps by parental education are still apparent at that time point. Trends by family income quintile are similar (again, the graphs by family income display more linear gradients than is evident when comparing parental education groups).

Although patterns are similar for internalizing problems, with children of the most highly educated (and, to a lesser extent, children of the highest income group) displaying an increased advantage over the other groups over time, this is not because their rating is flat over time. In fact, internalizing problems for the children of the most educated increase slightly throughout the period including between age 9 and 11.

Results for standardized behavior problem scores for children in the US sample, shown in Figure B2 (with regression results in Table B2), are similar. Gradients in both externalizing and internalizing problems fan out between school entry and age 9, before narrowing a bit between age 9 and 11. Measured on a comparable standardized scale, gaps in internalizing problems tend to be smaller than those for externalizing problems. At age 11, the gap in externalizing problems between the children of the most highly educated and least educated is .48 standard deviations, up from .26 at age 5, in contrast to the corresponding gaps for internalizing problems, which are .32 and .30. Comparing children from the top income quintile to those from the bottom income quintile, the gap in

externalizing problems grows from .28 at age 5 to .47 at age 11, while the gap in internalizing problems grows from .25 to .34. A comparison of Figure 10B.2 with the equivalent charts for achievement outcomes reveals clearly that SES gradients in socio-emotional outcomes are much smaller than in academic achievement outcomes at the same ages.

Figures 10B.3 and 10B.4 provide results for England. As mentioned earlier, we must exercise caution in comparing these results to the US ones because they rely on parent-reported measures (as opposed to teacher-reported measures in the US) and cover slightly older children (age 6 through 13, as opposed to age 5 through 11 in the US). In the raw scores (Figure 10B.3 and Table 10B.3), there are sizable gaps in both externalizing and emotional symptom problems (by both education and income group). There is some evidence that the gaps are largest at the bottom of the SES distribution, with children of the lowest educated and lowest income parents exhibiting substantially more behavior problems than middle-SES children. In the standardized scores (Figure 10B.4 and Table 10B.4), gaps tend to be larger for externalizing problems than for emotional symptoms but there is little systematic variation in their development over time. Between age 6 and 13, the gap in emotional symptom problems between children of the most and least highly educated parents grows from .085 (not significant) to .21 standard deviations, while the gap in externalizing problems grows only slightly from .44 to .48 standard deviations. However, if we compare children of the highest income families to the lowest income families, again between age 6 and 13, the gap in emotional symptom problems is roughly constant at .25 and .24, while the gap in externalizing problems grows from .38 to .46. Again, it is notable that gradients in socio-emotional outcomes are smaller than those for achievement outcomes.

Figure 10B.1. Mean US Behavior problems raw scores, by SES group

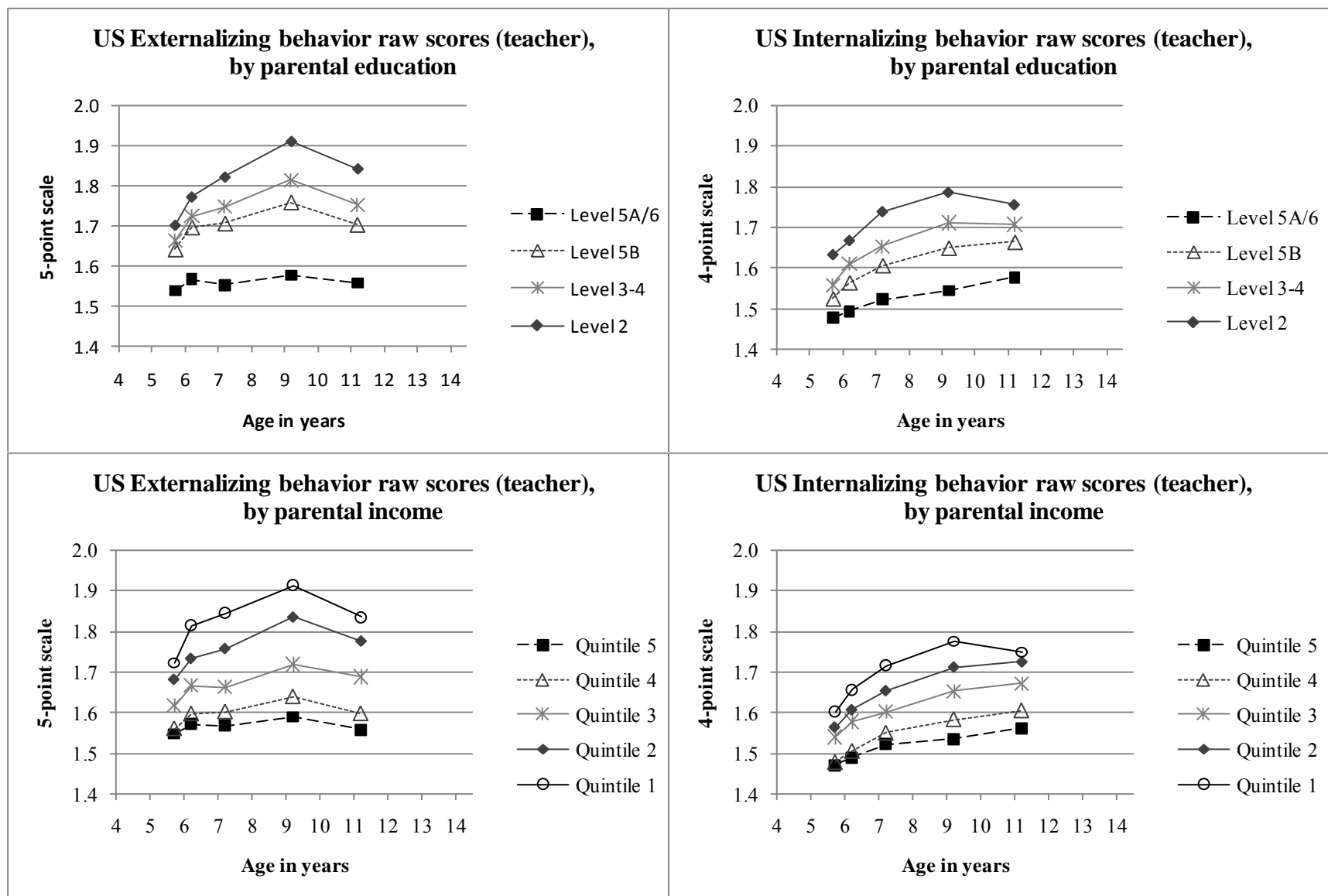


Figure 10B.2. Mean US Behavior problems standardized scores, by SES group

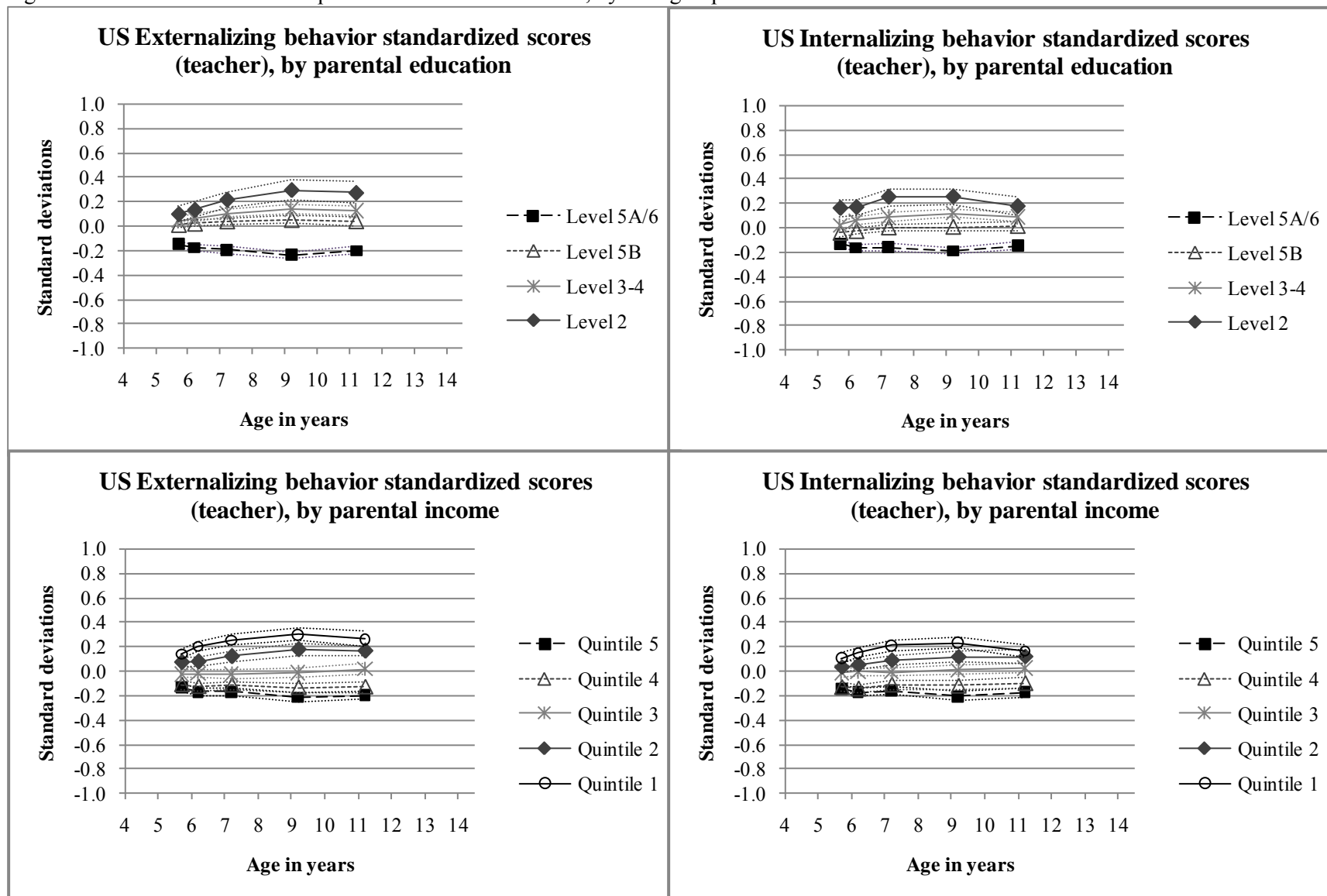


Figure 10B.3. Mean England behavior problems raw scores, by SES group

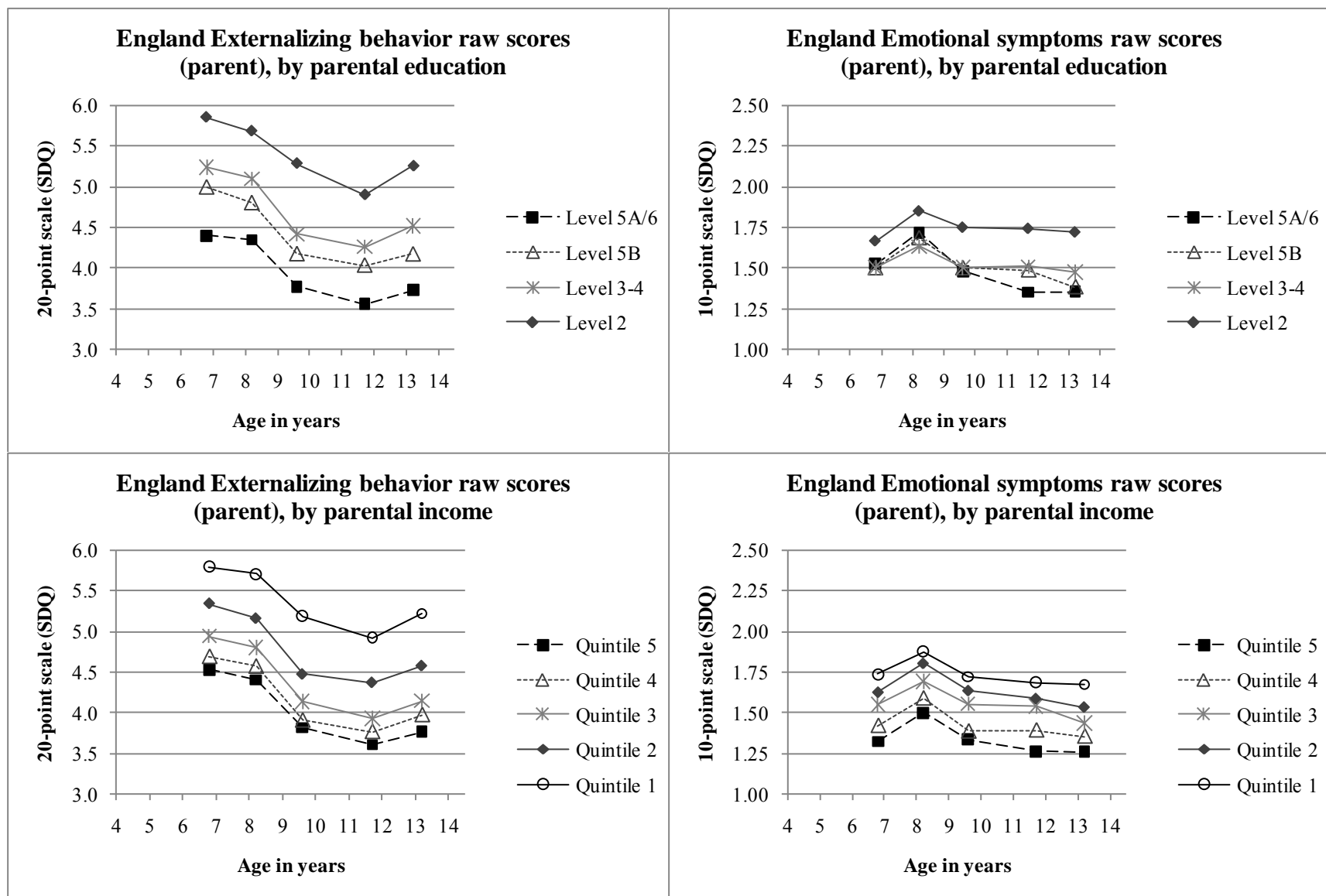


Figure 10B.4. Mean England behavior problems standardized scores, by SES group

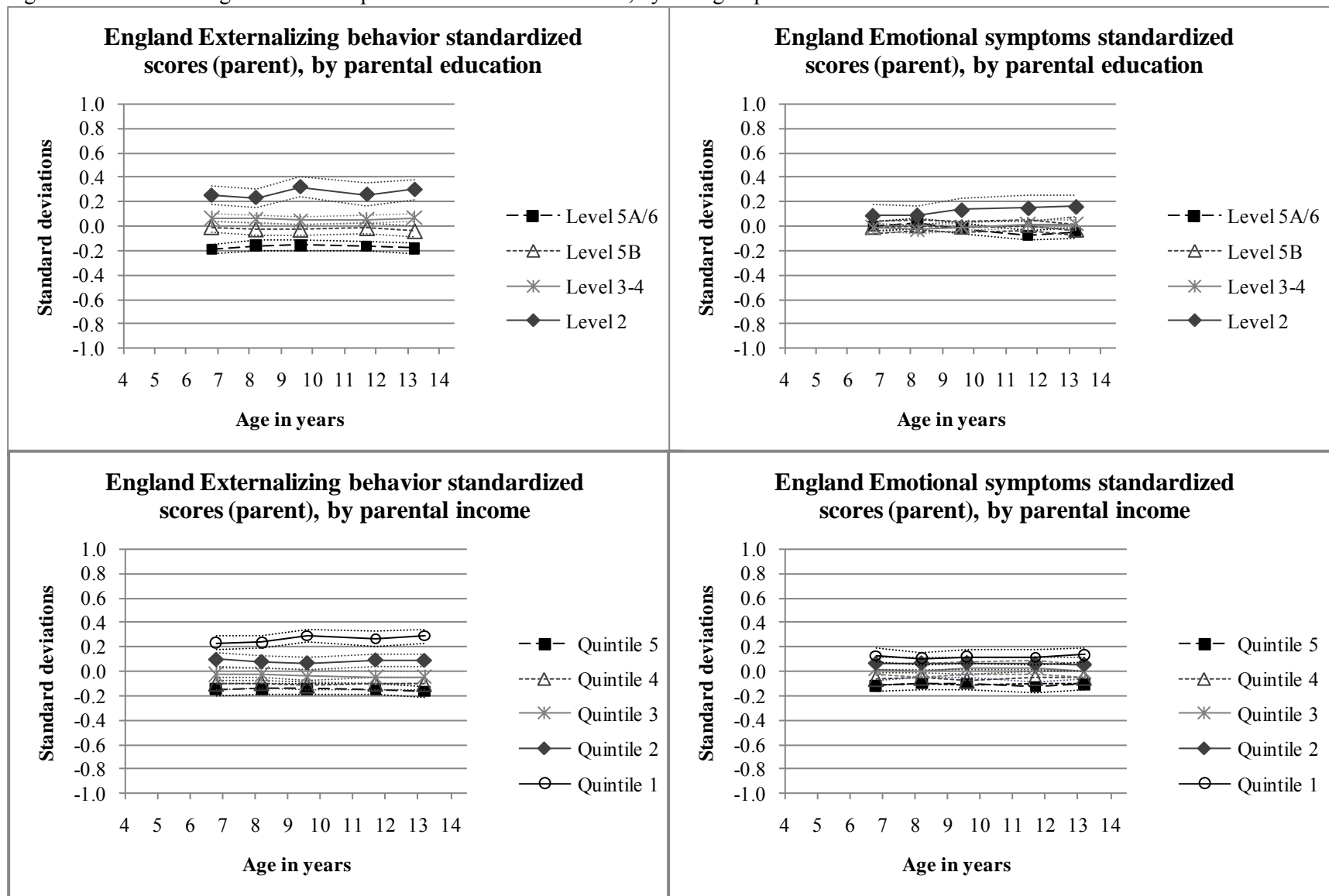


Table 10B.1. US Behavior outcomes – raw outcome variables (unstandardized)

10B.1a. Summary of Regressions of External Behavior (Unstandardized) on Parental Education

Education	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Level 3	-.04 (.02)	-.05* (.02)	-.07** (.02)	-.10** (.03)	-.09** (.03)	
Level 5B	-.06* (.02)	-.07** (.02)	-.12** (.02)	-.15** (.03)	-.14** (.03)	
Level 5A6	-.16** (.02)	-.21** (.02)	-.27** (.02)	-.33** (.03)	-.28** (.03)	
Constant	1.70** (.02)	1.77** (.02)	1.82** (.02)	1.91** (.03)	1.84** (.03)	
Observations	15486	15486	15486	15486	15486	
R-squared	.01	.01	.02	.03	.02	

** p<.01, * p<.05

Note. Level 2 is reference group. Behavior measures are unstandardized and scales from 1 to 5.

10B.1b. Summary of Regressions of External Behavior (Unstandardized) on Income Quintile

Family Income	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Quintile 2	-.04* (.02)	-.08** (.02)	-.09** (.02)	-.08** (.02)	-.06** (.02)	
Quintile 3	-.10** (.02)	-.15** (.02)	-.18** (.02)	-.20** (.02)	-.15** (.02)	
Quintile 4	-.16** (.02)	-.22** (.02)	-.24** (.02)	-.27** (.02)	-.24** (.02)	
Quintile 5	-.17** (.02)	-.24** (.02)	-.28** (.02)	-.32** (.02)	-.28** (.02)	
Constant	1.72** (.01)	1.81** (.01)	1.85** (.02)	1.91** (.02)	1.84** (.02)	
Observations	15486	15486	15486	15486	15486	
R-squared	.01	.02	.02	.03	.03	

** p<.01, * p<.05

Note. Income quintile 1 (lowest) is reference group. Behavior measures are unstandardized and scales from 1 to 5.

Continues

Table 10B.1 *continued*. US behavior outcomes – raw outcome variables (unstandardized)

10B.1c. Summary of Regressions of Internal Behavior (Unstandardized) on Parental Education						
Education	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Level 3	-.07** (.02)	-.06** (.02)	-.09** (.02)	-.07** (.02)	-.05 (.03)	
Level 5B	-.11** (.02)	-.10** (.02)	-.13** (.02)	-.14** (.02)	-.09** (.02)	
Level 5A6	-.15** (.02)	-.17** (.02)	-.21** (.02)	-.24** (.02)	-.18** (.02)	
Constant	1.63** (.02)	1.67** (.02)	1.74** (.02)	1.79** (.02)	1.76** (.02)	
Observations	15456	15456	15456	15456	15456	
R-squared	.01	.01	.01	.02	.01	

** p<.01, * p<.05

Note. Level 2 is reference group. Behavior measures are unstandardized and scaled from 1 to 4.

10B.1d. Summary of Regressions of Internal Behavior (Unstandardized) on Income Quintile						
Family Income	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Quintile 2	-.04* (.02)	-.05** (.02)	-.06** (.02)	-.06** (.02)	-.02 (.02)	
Quintile 3	-.06** (.01)	-.08** (.01)	-.11** (.02)	-.12** (.02)	-.08** (.02)	
Quintile 4	-.12** (.02)	-.15** (.01)	-.17** (.02)	-.19** (.02)	-.14** (.02)	
Quintile 5	-.13** (.02)	-.17** (.01)	-.19** (.02)	-.24** (.02)	-.19** (.02)	
Constant	1.60** (.01)	1.66** (.01)	1.72** (.01)	1.78** (.01)	1.75** (.02)	
Observations	15456	15456	15456	15456	15456	
R-squared	.01	.01	.02	.02	.02	

** p<.01, * p<.05

Note. Income quintile 1 (lowest) is reference group. Behavior measures are unstandardized and scaled from 1 to 4.

Table 10B.2. US Behavior outcomes – standardized outcome variables

10B.2a. Summary of Regressions of External Behavior (Standardized) on Parental Education						
Education	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Level 3	-.06 (.04)	-.07* (.03)	-.11** (.04)	-.16** (.04)	-.15** (.05)	
Level 5B	-.09* (.03)	-.11** (.03)	-.18** (.04)	-.24** (.04)	-.23** (.05)	
Level 5A6	-.25** (.03)	-.31** (.03)	-.41** (.04)	-.53** (.04)	-.47** (.04)	
Constant	.10** (.03)	.14** (.03)	.22** (.03)	.30** (.04)	.27** (.05)	
Observations	15486	15486	15486	15486	15486	
R-squared	.01	.01	.02	.03	.02	

** p<.01, * p<.05

Note. Level 2 is reference group. Behavior measures are standardized.

10B.2b. Summary of Regressions of External Behavior (Standardized) on Income Quintile						
Family Income	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Quintile 2	-.06* (.03)	-.12** (.03)	-.13** (.03)	-.13** (.03)	-.10** (.04)	
Quintile 3	-.16** (.03)	-.23** (.03)	-.28** (.03)	-.31** (.03)	-.25** (.04)	
Quintile 4	-.25** (.03)	-.33** (.03)	-.37** (.03)	-.44** (.03)	-.40** (.03)	
Quintile 5	-.27** (.03)	-.37** (.03)	-.42** (.03)	-.52** (.03)	-.46** (.03)	
Constant	.14** (.02)	.21** (.02)	.26** (.02)	.30** (.03)	.26** (.03)	
Observations	15486	15486	15486	15486	15486	
R-squared	.01	.02	.02	.03	.03	

** p<.01, * p<.05

Note. Income quintile 1 (lowest) is reference group. Behavior measures are standardized.

Continues

Table 10B.2 *continued*. US behavior outcomes – standardized outcome variables

10B.2c. Summary of Regressions of Internal Behavior (Standardized) on Parental Education						
Education	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Level 3	-.14** (.04)	-.11** (.04)	-.16** (.04)	-.14** (.04)	-.09 (.05)	
Level 5B	-.20** (.04)	-.20** (.03)	-.25** (.04)	-.25** (.04)	-.16** (.04)	
Level 5A6	-.29** (.04)	-.33** (.03)	-.41** (.04)	-.44** (.04)	-.32** (.04)	
Constant	.16** (.03)	.17** (.03)	.25** (.04)	.25** (.03)	.18** (.04)	
Observations	15456	15456	15456	15456	15456	
R-squared	.01	.01	.01	.02	.01	

** p<.01, * p<.05

Note. Level 2 is reference group. Behavior measures are standardized.

10B.2d. Summary of Regressions of Internal Behavior (Standardized) on Income Quintile						
Family Income	KF (5.7)	KS (6.2)	1S (7.2)	3S (9.2)	5S (11.2)	8S (14.2)
Quintile 2	-.07* (.03)	-.09** (.03)	-.12** (.03)	-.12** (.03)	-.04 (.04)	
Quintile 3	-.12** (.03)	-.15** (.03)	-.22** (.03)	-.22** (.03)	-.14** (.03)	
Quintile 4	-.23** (.03)	-.29** (.03)	-.32** (.03)	-.35** (.03)	-.26** (.04)	
Quintile 5	-.25** (.03)	-.32** (.03)	-.37** (.03)	-.44** (.04)	-.34** (.04)	
Constant	.11** (.02)	.15** (.02)	.21** (.02)	.23** (.03)	.16** (.03)	
Observations	15456	15456	15456	15456	15456	
R-squared	.01	.01	.02	.02	.02	

** p<.01, * p<.05

Note. Income quintile 1 (lowest) is reference group. Behavior measures are standardized.

Table 10B.3. England behavior outcomes – raw outcome variables (unstandardized)

10B.3a. Summary of Regressions of Externalizing Behavior (raw) on Parental Education					
	SDQ6	SDQ8	SDQ9	SDQ11	SDQ13
Education	(6.8)	(8.2)	(9.6)	(11.7)	(13.2)
Level 3	-.61** (.14)	-.58** (.15)	-.87** (.15)	-.64** (.17)	-.74** (.16)
Level 5B	-.86** (.15)	-.88** (.16)	-1.11** (.16)	-.87** (.19)	-1.09** (.15)
Level 5A6	-1.45** (.15)	-1.34** (.15)	-1.51** (.15)	-1.34** (.18)	-1.53** (.15)
Constant	5.85** (.13)	5.68** (.14)	5.29** (.14)	4.91** (.16)	5.26** (.14)
Observations	9435	9435	9435	9435	9435
R-squared	.01	.01	.02	.01	.02

** p<.01, * p<.05

Note. Level 2 is reference group. Behavior measures are unstandardized and scaled from 0 to 20.

10B.3b. Summary of Regressions of Externalizing Behavior (raw) on Income Quintile					
	SDQ6	SDQ8	SDQ9	SDQ11	SDQ13
Family Income	(6.8)	(8.2)	(9.6)	(11.7)	(13.2)
Quintile 2	-.46** (.13)	-.55** (.12)	-.71** (.13)	-.55** (.12)	-.64** (.12)
Quintile 3	-.85** (.12)	-.90** (.13)	-1.05** (.12)	-1.00** (.14)	-1.08** (.13)
Quintile 4	-1.11** (.13)	-1.14** (.13)	-1.28** (.11)	-1.16** (.13)	-1.25** (.13)
Quintile 5	-1.26** (.13)	-1.30** (.13)	-1.37** (.11)	-1.32** (.12)	-1.45** (.11)
Constant	5.80** (.09)	5.71** (.09)	5.19** (.09)	4.93** (.10)	5.23** (.10)
Observations	9435	9435	9435	9435	9435
R-squared	.02	.02	.02	.02	.02

* p<.05, † p<.1

Note. Income quintile 1 (lowest) is reference group. Behavior measures are unstandardized and scaled from 0 to 20.

Continues

Table 10B.3 *continued*. England behavior outcomes – raw outcome variables (unstandardized)

10B.3c. Summary of Regressions of Emotional Symptoms (raw) on Parental Education					
	SDQ6	SDQ8	SDQ9	SDQ11	SDQ13
Education	(6.8)	(8.2)	(9.6)	(11.7)	(13.2)
Level 3	-.16 (.09)	-.22** (.08)	-.25** (.09)	-.23* (.11)	-.25** (.09)
Level 5B	-.17 (.09)	-.16* (.08)	-.25** (.09)	-.26* (.11)	-.34** (.08)
Level 5A6	-.14 (.09)	-.13 (.08)	-.27** (.10)	-.39** (.10)	-.37** (.09)
Constant	1.67** (.08)	1.85** (.07)	1.75** (.09)	1.75** (.10)	1.72** (.08)
Observations	9435	9435	9435	9435	9435
R-squared	.00	.00	.00	.00	.00

** p<.01, * p<.05

Note. Level 2 is reference group. Behavior measures are unstandardized and scaled from 0 to 1.

10B.3d. Summary of Regressions of Emotional Symptoms (raw) on Income Quintile					
	SDQ6	SDQ8	SDQ9	SDQ11	SDQ13
Family Income	(6.8)	(8.2)	(9.6)	(11.7)	(13.2)
Quintile 2	-.11 (.07)	-.07 (.07)	-.08 (.07)	-.10 (.07)	-.14* (.07)
Quintile 3	-.18** (.07)	-.19* (.08)	-.17* (.07)	-.15* (.07)	-.24** (.07)
Quintile 4	-.31** (.07)	-.29** (.06)	-.33** (.07)	-.29** (.08)	-.32** (.07)
Quintile 5	-.41** (.06)	-.38** (.06)	-.38** (.07)	-.42** (.07)	-.42** (.07)
Constant	1.74** (.05)	1.88** (.05)	1.72** (.05)	1.69** (.06)	1.68** (.05)
Observations	9435	9435	9435	9435	9435
R-squared	.01	.01	.01	.01	.01

* p<.05, † p<.1

Note. Income quintile 1 (lowest) is reference group. Behavior measures are unstandardized and scaled from 0 to 1.

Table 10B.4. England behavior outcomes – standardized outcome variables

10B.4a. Summary of Regressions of Externalizing Behavior (std) on Parental Education

	SDQ6	SDQ8	SDQ9	SDQ11	SDQ13
Education	(6.8)	(8.2)	(9.6)	(11.7)	(13.2)
Level 3	-.19** (.03)	-.17** (.04)	-.28** (.05)	-.20** (.05)	-.23** (.05)
Level 5B	-.26** (.05)	-.26** (.05)	-.35** (.05)	-.28** (.06)	-.34** (.05)
Level 5A6	-.44** (.05)	-.39** (.04)	-.48** (.05)	-.42** (.06)	-.48** (.05)
Constant	.25** (.04)	.23** (.04)	.32** (.04)	.26** (.05)	.30** (.04)
Observations	9435	9435	9435	9435	9435
R-squared	.01	.01	.02	.01	.02

** p<.01, * p<.05

Note. Level 2 is reference group.

10B.4b. Summary of Regressions of Externalizing Behavior (std) on Income Quintile

	SDQ6	SDQ8	SDQ9	SDQ11	SDQ13
Family Income	(6.8)	(8.2)	(9.6)	(11.7)	(13.2)
Quintile 2	-.14** (.04)	-.16** (.04)	-.22** (.04)	-.17** (.04)	-.20** (.04)
Quintile 3	-.26** (.04)	-.26** (.04)	-.33** (.04)	-.31** (.04)	-.34** (.04)
Quintile 4	-.33** (.04)	-.33** (.04)	-.41** (.04)	-.37** (.04)	-.39** (.04)
Quintile 5	-.38** (.04)	-.38** (.04)	-.43** (.04)	-.42** (.04)	-.46** (.04)
Constant	.23** (.03)	.24** (.03)	.29** (.03)	.27** (.03)	.29** (.03)
Observations	9435	9435	9435	9435	9435
R-squared	.02	.02	.02	.02	.02

* p<.05, † p<.1

Note. Income quintile 1 (lowest) is reference group.

Continues

Table 10B.4 *continued*. England behavior outcomes – standardized outcome variables

10B.4c. Summary of Regressions of Emotional Symptoms (std) on Parental Education

	SDQ6 (6.8)	SDQ8 (8.2)	SDQ9 (9.6)	SDQ11 (11.7)	SDQ13 (13.2)
Level 3	-.10 (.05)	-.12** (.04)	-.14** (.05)	-.13* (.06)	-.14** (.05)
Level 5B	-.10 (.06)	-.09* (.05)	-.14** (.05)	-.15* (.06)	-.20** (.05)
Level 5A6	-.09 (.05)	-.07 (.05)	-.15** (.06)	-.23** (.06)	-.21** (.05)
Constant	.09 (.05)	.09* (.04)	.14** (.05)	.15** (.06)	.16** (.05)
Observations	9435	9435	9435	9435	9435
R-squared	.00	.00	.00	.00	.00

** p<.01, * p<.05

Note. Level 2 is reference group.

10B.4d. Summary of Regressions of Emotional Symptoms (std) on Income Quintile

	SDQ6 (6.8)	SDQ8 (8.2)	SDQ9 (9.6)	SDQ11 (11.7)	SDQ13 (13.2)
Family Income					
Quintile 2	-.07 (.04)	-.04 (.04)	-.05 (.04)	-.06 (.04)	-.08* (.04)
Quintile 3	-.11** (.04)	-.10* (.04)	-.10* (.04)	-.08* (.04)	-.14** (.04)
Quintile 4	-.19** (.04)	-.16** (.03)	-.19** (.04)	-.17** (.04)	-.19** (.04)
Quintile 5	-.25** (.04)	-.21** (.03)	-.22** (.04)	-.24** (.04)	-.24** (.04)
Constant	.13** (.03)	.11** (.03)	.12** (.03)	.12** (.03)	.14** (.03)
Observations	9435	9435	9435	9435	9435
R-squared	.01	.01	.01	.01	.01

* p<.05, † p<.1

Note. Income quintile 1 (lowest) is reference group.

¹ Teacher reports on children's behavior are not collected after age 11 (fifth grade). The ECLS-K asks children to report on their own behavior starting at age 9 (third grade). In results not shown, we examined trends in SES gaps in these child-reported data, but we prefer to focus on the teacher reports because they are a more widely used measure for school-age children and are available over more assessment points.

² We limit the SDQ age range to the school years for comparability with other outcomes. As a result we do not make use of an available measure at age 3.

Chapter 13 Online Appendix

Appendix Table for Table 13.3

a) Italy (base = Vocational education)			
	Technical education	Academic oriented education	Teachers' school
ISCED3*Fair	-0.105* (0.0638)	-0.161 (0.110)	0.209* (0.124)
ISCED3*Good	-0.0753 (0.0991)	-0.193 (0.128)	0.115 (0.154)
ISCED3*Excellent	-0.210 (0.140)	-0.251 (0.156)	0.467** (0.190)
ISCED4-6*Fair	-0.133 (0.165)	-0.338* (0.176)	0.275 (0.235)
ISCED4-6*Good	-0.00644 (0.213)	-0.383* (0.216)	0.510* (0.292)
ISCED4-6*Excellent	0.242 (0.314)	-0.179 (0.309)	0.705* (0.363)
b) Germany (base = low)			
	intermediate	high	other
ISCED 4-6*intermediate	0.397 (0.948)	-2.097 (1.659)	-0.711 (1.162)
ISCED 4-6*high	-0.454 (0.891)	-2.503 (1.553)	-0.268 (1.055)
ISCED 4-6*other	-1.015 (0.946)	-1.82 (1.648)	-0.555 (1.283)

Appendix Table for Table 13.4

	Repeated grades	Changed to higher track	Changed to lower track
a) Italy			
ISCED3* Technical	-0.0561 (0.0687)	0.114 (0.226)	
ISCED3* Academic oriented	-0.172 (0.0884)		-0.312*** (0.110)
ISCED3* Teachers' school	-0.188 (0.137)	0.627 (0.525)	-0.158 (0.205)
ISCED4-6* Technical	-0.378* (0.174)	-0.144 (0.364)	
ISCED4-6* Academic oriented	-0.669*** (0.169)		-0.420** (0.213)
ISCED4-6* Teachers' school	-0.384 (0.259)	0.807 (0.631)	0.0504 (0.338)
b) Germany			
ISCED 4-6*intermediate	-2.009*** (0.759)	-0.026 (0.739)	0.229 (0.733)
ISCED 4-6*high	-0.985 (0.762)		
ISCED 4-6*other	-1.492* (0.890)		

Appendix Table for Table 13.5

Italy		Germany	
ISCED3* Technical	0.0752 (0.0492)	ISCED 4-6*intermediate	-0.196 (0.545)
ISCED3* Academic oriented	0.0237 (0.0612)	ISCED 4-6*high	-0.160 (0.546)
ISCED3* Teachers' school	0.231** (0.0902)	ISCED 4-6*other	1.090 (0.982)
ISCED4-6* Technical	0.102 (0.125)		
ISCED4-6* Academic oriented	0.352*** (0.118)		
ISCED4-6* Teachers' school	0.365** (0.166)		

Appendix Table for Table 13.6 (Italy only)

	University enrollment	University drop out
ISCED3* Technical	0.194*** (0.0621)	-0.0690 (0.117)
ISCED3* Academic oriented	0.123 (0.108)	-0.156 (0.144)
ISCED3* Teachers' school	0.366*** (0.107)	-0.401** (0.195)
ISCED4-6* Technical	0.348** (0.139)	0.163 (0.217)
ISCED4-6* Academic oriented	0.335* (0.178)	-0.293 (0.225)
ISCED4-6* Teachers' school	0.282 (0.218)	-0.137 (0.340)

Chapter 14 Online Appendix

Table 14A.1 LONGITUDINAL DATA FOR THE U. S. AND CANADA

	U.S. (PSID)	Canada (NLSCY)	Canada (YITS)
Sampling frame and representativeness	<p>The PSID is a nationally representative sample of families in the U.S. with an oversample of low-income families. The PSID contains annually collected data for 1968-1997 and bi-annual data for 1999-2009 and sample members are followed as they split off into new households; our sample uses data through the 2007 wave. The CDS sampled all PSID families with children aged 0-12 years during the calendar year of 1997. Up to two children per family were selected. These children were interviewed in 1997, 2002-03, 2005, and 2007. For those members of the CDS who reached age 18, a new study called “Transition into Adulthood” collected additional data in 2005 and 2007.</p>	<p>The target population comprises the non-institutionalized civilian population (aged 0 to 11 at the time of their selection) in Canada's 10 provinces. The survey excludes children living on Indian reserves or Crown lands.</p> <p>The NLSCY is a probability survey designed to collect detailed information every two years. The longitudinal samples are representative of the original longitudinal populations (i.e., the populations at the time of sample selection). Typically, children in the NLSCY are drawn from the Labour Force Survey's (LFS) sample of respondent households. The LFS is a monthly survey that collects labour market data from a national sample of more than 52,000 dwellings. The LFS</p>	<p>The target population comprises persons who were born in 1984 and in the 1999/2000 school year were attending any form of schooling in the provinces of Canada. Students living in the territories or on Indian reserves were excluded. The survey design consisted of a two-stage approach. In the first stage, a stratified sample of schools was selected to ensure adequate coverage in the ten provinces included in the national desired target population (including adequate coverage of minority school systems in certain provinces). The stratification was based on the enrolment of 15 year olds in the school in the previous academic year. In the second stage, a simple random sample</p>

		sample design is based on a stratified, multistage design using probability sampling at all stages of the design. The design principles of the LFS are the same for each province.	of 15 year-old students within the school was selected
Maturity of the panel (cohorts and time periods included)	Information on the parents (and grandparents) of the children is available in the PSID since 1968. More detailed information on the children is available in the CDS for 1997, 2002-03, 2005, and 2007	Collection for the first cycle of the NLSCY began in 1994 with one large cohort of 0- to 11-year-olds who lived in any province. This sample is referred to as the original cohort. At Cycle 7, the NLSCY original cohort sample consists of children aged 12 to 23 years (original cohort).	The YITS is strictly a longitudinal survey. The initial sample of 15 year-olds selected at cycle 1 (2000) is surveyed every two years for as many as five cycles. No attempts were made to top-up the sample for any YITS cycles to ensure a cross-section representation of the population. Non-respondents at a specific cycle are not followed-up for subsequent cycles of the survey.
Panel attrition: prevalence and corrections	In 1997, out of 2,705 families selected for the first wave of the CDS, 2,394 families participated (88%) for a total sample of 3,563 children. In 2002-03, 91 percent of these families were successfully re-interviewed. Sample weights adjust for attrition.	The proportion of respondents remaining in the survey relative to the number of respondents at Cycle 1 is 67 percent. Statistics Canada provides longitudinal weights for point estimation: each child's final survey weight has been adjusted for nonresponse, and post-	In 2006, out of 26,063 students selected for the first wave of the YITS (with parental questionnaire), 17,321 were re-interviewed (66%). Sample weights partially adjust for attrition.

		stratified by province, age and sex to match known population totals at the time of sample selection.	
Parental resources (P)	Educational attainment of head and spouse. Head income (including sources), “wife” income (including sources), total family income. Wealth.	Educational attainment of the person most knowledgeable and spouse (highest grade completed). Household income and personal income of the person most knowledgeable and spouse (Canadian dollars).	Educational attainment of head and spouse. Head income (including sources), “wife” income (including sources), total family income.

Chapter 15 Online Appendix

by M. Bratti and L. Cappellari

Variables description

DEGREE FINAL GRADE. It is the final grade obtained in the undergraduate degree.

PROBABILITY TO CONTINUE IN POSTGRADUATE EDUCATION. It is a dummy that takes on value one if an individual enrolled in postgraduate education and zero otherwise.

LOG HOURLY WAGE. It is the natural logarithm of hourly wage.

JOB SATISFACTION ABOUT WAGES. It is a discrete ordered variable that takes on the values: 4= very satisfied, 3 = quite satisfied, 2 = not very satisfied, 1 = not satisfied at all.

SHORT DEGREE. It is a dummy that takes on value one if an individual graduated in a short degree and zero otherwise.

HISCED. It is the highest between mother's and father's educations according to the International Standard Classification of Education (ISCED). We group HISCED into HISCED 1-2 (primary and lower secondary), HISCED 3-4 (upper secondary and post-secondary non tertiary education), HISCED 5-6 (first stage and second stage of tertiary education). HISCED 1-2 is used as the reference group in the econometric models.

AGE 25-29. A dummy for individuals 25 or older (the estimation sample only includes individuals younger than 30). Younger graduates (less than 24) are the reference group.

FEMALE. A dummy which equals one if the individual is female and zero otherwise.

UPPER SECONDARY SCHOOL TYPE. Dummies for the secondary school track. We consider the following seven school tracks provided by ISTAT, which can be grouped in three main tracks. Academic track: scientific lyceum, classical lyceum, language lyceum; Technical track: technical school; Vocational track: art school, pedagogic school, vocational school. Scientific lyceum is the reference group.

UPPER SECONDARY SCHOOL FINAL GRADE. Final grade obtained at the upper secondary schooling central exit examination. It ranges between 36 and 60.

WORKING WHILE STUDYING. It is a dummy which takes value one if the individual worked while studying and zero otherwise.

COLLEGE MAJOR. Dummies for the college major of the first degree. We consider the following groups provided by ISTAT: scientific; chemical-pharmaceutical; geo-biological; medical; engineering; architecture; agriculture; economics-statistics; politics-sociology; law; literature; languages; psychology; physical education; defense and police. Economics-statistics is the reference group.

HIGHER EDUCATION INSTITUTIONS (HEIs). Dummies for HEIs. The reference is set to the University of Padua (located in North-Eastern Italy).

SWITCHER. It is a dummy that takes on value one if an individual who enrolled in the pre-reform system switched to the new regime (short degree) and zero otherwise. Non-switchers are the reference group.

REGION. Dummies for 20 Italian administrative regions (NUTS-2). Piedmont is the reference group.

PART-TIME JOB. It is a dummy that takes on value one if an individual has a part-time job and zero otherwise. Full-time jobs are the reference group.

TEMPORARY JOB. A dummy that takes on value one if an individual is working with a fixed-term contract and zero otherwise. Permanent jobs are the reference group.

Supplementary tables

Table 15A.1. Degree final grade

	Control Group: Long Degrees in Time						Control Group: Long Degrees with One Year of Delay at Maximum					
	all				men	women	all				men	women
	1	2	3	4	5	6	7	8	9	10	11	12
short degree (sd)	-1.109*** (0.177)	-1.185*** (0.284)	-1.375*** (0.279)	-1.393*** (0.347)	-1.481** (0.645)	-1.292*** (0.346)	-0.716*** (0.162)	-0.750*** (0.259)	-1.117*** (0.254)	-1.080*** (0.312)	-1.207** (0.564)	-0.971*** (0.314)
HISCED 3-4	-0.518*** (0.172)	-0.623** (0.283)	-0.729*** (0.269)	-0.696** (0.274)	-0.805 (0.489)	-0.634** (0.293)	-0.534*** (0.159)	-0.549** (0.225)	-0.590*** (0.210)	-0.577*** (0.210)	-0.836* (0.425)	-0.456** (0.228)
HISCED 5-6	-0.307 (0.214)	-0.320 (0.342)	-0.436 (0.306)	-0.526* (0.298)	-1.340** (0.520)	-0.138 (0.348)	-0.443** (0.198)	-0.468* (0.272)	-0.471* (0.241)	-0.473* (0.249)	-0.890*** (0.313)	-0.219 (0.328)
HISCED 3-4 * sd		0.177 (0.349)	0.446 (0.333)	0.360 (0.402)	0.036 (0.732)	0.496 (0.397)		0.034 (0.304)	0.275 (0.288)	0.239 (0.380)	0.037 (0.726)	0.318 (0.346)
HISCED 5-6 * sd		0.007 (0.399)	0.597 (0.374)	0.712* (0.370)	0.998 (0.670)	0.565 (0.470)		0.062 (0.341)	0.595* (0.320)	0.690* (0.347)	0.594 (0.522)	0.631 (0.430)
majors fixed effects			yes	yes	yes	yes			yes	yes	yes	yes
institutions fixed effects				yes	yes	yes				yes	yes	yes
R ²	0.180	0.180	0.264	0.317	0.339	0.279	0.174	0.174	0.270	0.328	0.333	0.297
No. observations	15,824	15,824	15,824	15,809	6,589	9,220	20,105	20,105	20,105	20,086	8,533	11,553

Source: Authors' calculations

Note. The dependent variable is the degree final grade, which varies in the range 66-111 (110 *cum laude*). All models are estimated with OLS. Estimates use probability weights. Heteroskedasticity robust standard errors in parentheses. Errors are clustered by HEIs in the model using HEIs fixed effects. The models also include controls for age, gender (except the gender specific regressions), secondary school track, upper secondary school final grade, grade by track interactions, and dummies for working while studying and being a switcher (to the new system).

* p < .10, ** p < .05, *** p < .01

Table 15A.2. Job satisfaction regarding wages

	Control Group: Long Degrees in Time							Control Group: Long Degrees with One Year of Delay at Maximum						
	all					men	women	all					men	women
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
short degree (sd)	0.178** (0.074)	0.077 (0.131)	0.249* (0.133)	0.227 (0.143)	0.259* (0.145)	0.182 (0.280)	0.289* (0.165)	0.119* (0.066)	0.027 (0.105)	0.223** (0.109)	0.226** (0.104)	0.256** (0.110)	0.307* (0.163)	0.198 (0.129)
HISCED 3-4	0.004 (0.076)	-0.074 (0.132)	-0.072 (0.133)	-0.062 (0.152)	-0.083 (0.153)	-0.388 (0.254)	0.109 (0.182)	0.024 (0.065)	-0.015 (0.091)	-0.012 (0.093)	-0.003 (0.097)	-0.015 (0.098)	-0.003 (0.151)	0.014 (0.156)
HISCED 5-6	0.170* (0.088)	0.092 (0.140)	0.064 (0.144)	0.066 (0.199)	0.049 (0.192)	-0.197 (0.247)	0.228 (0.232)	0.155** (0.079)	0.088 (0.105)	0.067 (0.107)	0.063 (0.150)	0.051 (0.147)	0.129 (0.124)	0.037 (0.212)
degree final grades	-0.007 (0.005)	-0.007 (0.005)	0.004 (0.005)	0.007 (0.006)	0.008 (0.006)	0.008 (0.006)	0.009 (0.010)	-0.009** (0.004)	-0.009** (0.004)	0.001 (0.004)	0.003 (0.004)	0.004 (0.004)	0.005 (0.006)	0.004 (0.007)
HISCED 3-4 * sd		0.137 (0.155)	0.107 (0.157)	0.112 (0.162)	0.150 (0.169)	0.379 (0.282)	-0.056 (0.199)		0.096 (0.123)	0.066 (0.124)	0.067 (0.113)	0.096 (0.118)	0.043 (0.183)	0.052 (0.167)
HISCED 5-6 * sd		0.140 (0.167)	0.067 (0.171)	0.076 (0.212)	0.107 (0.204)	0.233 (0.229)	-0.061 (0.279)		0.179 (0.140)	0.106 (0.142)	0.120 (0.175)	0.149 (0.175)	0.001 (0.132)	0.164 (0.267)
majors fixed effects			yes	yes	yes	yes	yes			yes	yes	yes	yes	yes
Institutions fixed effects				yes	yes	yes	yes				yes	yes	yes	yes
job characteristics					yes							yes		
Pseudo R2	0.072	0.072	0.086	0.093	0.098	0.130	0.089	0.068	0.068	0.078	0.083	0.090	0.110	0.079
No. observations	7,718	7,718	7,718	7,713	7,713	3,174	4,539	10,135	10,135	10,135	10,128	10,128	4,400	5,728

*significant at 10%; **significant at 5%; ***significant at 1%.

Source: Authors' calculations

Note. The dependent variable is a discrete ordered variable taking on the values: 4 = very satisfied, 3 = quite satisfied, 2 = not very satisfied, 1 = not satisfied at all. All models are estimated with an ordered logit, and the table reports the model coefficients. Estimates use probability weights. Heteroskedasticity robust standard errors in parentheses. Errors are clustered by HEIs in the model using HEIs fixed effects. The models also include controls for age, gender (except the gender specific regressions), secondary school track, upper secondary school final grade, grade by track interactions, dummies for working while studying and being a switcher (to the new system), region where the individual works and log hourly wage. Job characteristics are two dummies for part-time (vs. full-time) and temporary (vs. permanent) jobs, respectively. The estimation sample includes only individuals who found their current job after university graduation.

* $p < .10$, ** $p < .05$, *** $p < .01$

Chapter 16 Online Appendix

Silke Anger (DIW Berlin)

Data

This chapter's analysis is based on the German Socio-Economic Panel Study (SOEP), which is a representative household panel survey that started in 1984 (Wagner et al., 2007). The SOEP conducts annual personal interviews with all household members aged 18 and above, and provides rich information on socio-demographic characteristics, family background, and childhood environment. In more recent years, a Youth Questionnaire was implemented for adolescents at age 17. The SOEP data used in this project come from the samples of adult respondents, where parents and their adult children can be identified. In addition, data from the Youth Questionnaire is used to match adolescent children to their parents from the adult samples. Thus, the intergenerational transmission of skills will be analyzed separately for adolescent children aged around 17 and for young adult children aged 18 to 29. Parents and children who were not of German nationality were excluded from the study, since individuals with a migration background may be disadvantaged as compared to native speakers due to inadequate language skills when taking the tests or when rating their personality.

Measures of Cognitive and Noncognitive Skills of Adult Respondents

Since information on cognitive skills was only collected in 2006 and on noncognitive skills only in 2005 from adult respondents, this study uses these two cross-sections for young adult children and all parents. In 2006, about one third of all respondents (only those with a CAPI interview) participated in two ultra-short IQ tests lasting 90 seconds each (Lang et al., 2007): a word fluency test and a symbol correspondence test. Both tests correspond to different modules of the Wechsler Adult Intelligence Scale (WAIS). The symbol correspondence test is conceptually related to the mechanics of cognition or fluid intelligence and comprises general abilities. The test involved asking respondents to match as many

numbers and symbols as possible within 90 seconds according to a given correspondence list which is permanently visible to the respondents on a screen. The word fluency test is conceptually related to the pragmatics of cognition or crystallized intelligence. It involves the fulfillment of specific tasks that improve with knowledge and skills acquired in the past. The word fluency test implemented in the SOEP was based on the animal-naming task (Lindenberger and Baltes, 1995): respondents name as many different animals as possible within 90 seconds. While verbal fluency is based on learning, speed of cognition is related to individuals' innate abilities (Cattell, 1987). The scores are added together across the 90 seconds per test to generate an index which ranges from 0 to 60 (symbol correspondence test), respectively from 0 to 99 (word fluency test). In addition, a measure of general intelligence is generated by averaging the two ability test scores.

One year previously, in 2005, detailed measures of personality were part of the SOEP questionnaire for all respondents in the adult sample (Dehne and Schupp, 2007). These included self-rated measures that were related to the Five Factor Model (McCrae and Costa, 1999) and comprise the five basic psychological dimensions – openness to experience, conscientiousness, extraversion, agreeableness, neuroticism (Big Five) – as well as measures of locus of control. All items related to the personality traits had to be answered on 7-point Likert-type scales (1 – “disagree completely” to 7 – “agree completely”). The scores are summed up to create an index ranging from 1 to 7.

Measures of Cognitive and Noncognitive Skills of Adolescent Respondents

Since 2006, all adolescents entering the SOEP at age 17 have participated in somewhat more complex intelligence tests which cover the following domains: verbal skills, numerical skills, and abstract reasoning. The tests are modified versions of the I-S-T 2000-Test (Solga et al., 2005) and allow for a total time of 27 minutes for completion of all 60 tasks. Each of the three domains contains 20 individual tasks. In the first part (analogies), the respondent is asked to correctly assign expressions to a sequence of words according to a particular rule. These tasks test the ability to combine based on the vocabulary of the respondent, and, thus, measure verbal potential. In the second part (numerical series) the respondent is

asked to insert the correct arithmetic operator into an incomplete equation. These tasks measure numerical potential by testing the adolescent's abstract ability to recombine and logical reasoning. The third part (matrices) measures abstract reasoning. The respondent is asked to select the correct piece out of five possible figures according to a particular logical rule as provided by a displayed sequence of figures. The allotted times for completing each of the task groups are: 7 minutes for analogies, 10 minutes for numerical series, and 10 minutes for matrices. The scores are added together across the 20 individual tasks per domain to generate an index ranging from 0 to 20. An integrated additive index of verbal and numerical skills provides an adequate assessment of the adolescent's crystallized intelligence, i.e., skills that improve with knowledge acquired in the past, whereas abstract reasoning is related to fluid intelligence and, thus, comprises largely innate abilities. Since 2006, the SOEP questionnaire for adolescents has included items that relate to the Five Factor Model comprising the five basic psychological dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Big Five). Furthermore, measures of the locus of control were collected from adolescents every year. Again, 7-point Likert type scales (1 – “disagree completely” to 7 – “agree completely”) have been used for the items related to the personality traits. As for the sample of adults, the scores can be added together to create an index ranging from 1 to 7.

Table 16A.1: Summary Statistics: IQ Test Scores, Personality Traits, Family Background, and Childhood Environment

Variable	Adolescent Children				Young Adult Children			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Children's Characteristics								
Cognitive Skills								
Verbal skills	7.66	3.71	1	19	-	-	-	-
Numerical skills	12.36	4.92	1	20	-	-	-	-
Abstract reasoning	8.98	3.51	0	18	-	-	-	-
Word fluency	-	-	-	-	25.62	10.67	1	82
Coding speed	-	-	-	-	32.71	10.52	5	60
Noncognitive Skills								
B5: Openness	4.74	1.06	1	7	4.62	1.20	1	7
B5: Conscientiousness	4.94	1.17	1.3	7	5.44	1.06	1.3	7
B5: Extraversion	4.94	1.17	1.3	7	4.95	1.20	1	7
B5: Agreeableness	5.37	0.95	1	7	5.36	0.96	1.3	7
B5: Neuroticism	3.84	1.16	1	7	3.90	1.20	1	7
Locus of control: internal	4.90	0.73	2.25	7	4.83	0.75	1.8	7
Locus of control: external	3.66	0.93	1	6.5	3.64	0.94	1.2	7
Age	17.54	0.80	17	19	22.18	3.22	18	29
Single parent	0.24	0.43	0	1	0.19	0.40	0	1
First born	0.51	0.50	0	1	0.44	0.50	0	1
Number of brothers	0.88	0.88	0	4	0.86	1.03	0	7
Number of sisters	0.79	0.96	0	6	0.82	0.99	0	6
Height (in cm)	174.36	9.47	154	202	175.73	9.05	150	200
Good health	0.83	0.37	0	1	0.79	0.41	0	1
Childhood area: rural	0.31	0.47	0	1	0.30	0.46	0	1
Childhood area: town	0.26	0.44	0	1	0.17	0.38	0	1
Childhood area: city	0.20	0.40	0	1	0.17	0.38	0	1
Childhood area: urban	0.23	0.42	0	1	0.27	0.45	0	1
Childhood area: missing	-	-	-	-	0.08	0.27	0	1

Parents' Characteristics

Cognitive Skills

Word fluency	25.90	10.56	1	62	25.66	10.32	1	59.5
Coding speed	27.80	8.22	7	56	25.87	8.62	4	49
Noncognitive Skills ^a								
B5: Openness	6.33	1.67	1	10.5	6.08	1.69	1	10.5
B5: Conscientiousness	8.53	1.56	2.7	10.5	8.23	1.68	3.3	10.5
B5: Extraversion	6.91	1.58	2.3	10.5	6.62	1.60	1.7	10.5
B5: Agreeableness	7.77	1.55	2.7	10.5	7.52	1.63	2.7	10.5
B5: Neuroticism	5.68	1.61	1	10.5	5.55	1.62	1.3	10.5
Locus of control: internal	6.81	1.41	3	10.5	6.60	1.51	2.5	10.5
Locus of control: external	5.33	1.46	1.5	10.5	5.09	1.37	1.5	9.6
Low education	0.09	0.29	0	1	0.06	0.24	0	1
Medium education	0.63	0.48	0	1	0.64	0.48	0	1
High education	0.28	0.45	0	1	0.30	0.46	0	1
Number of Individuals (cognitive skills) ^a	280				446			
Number of Individuals (noncognitive skills)	1184				2228			

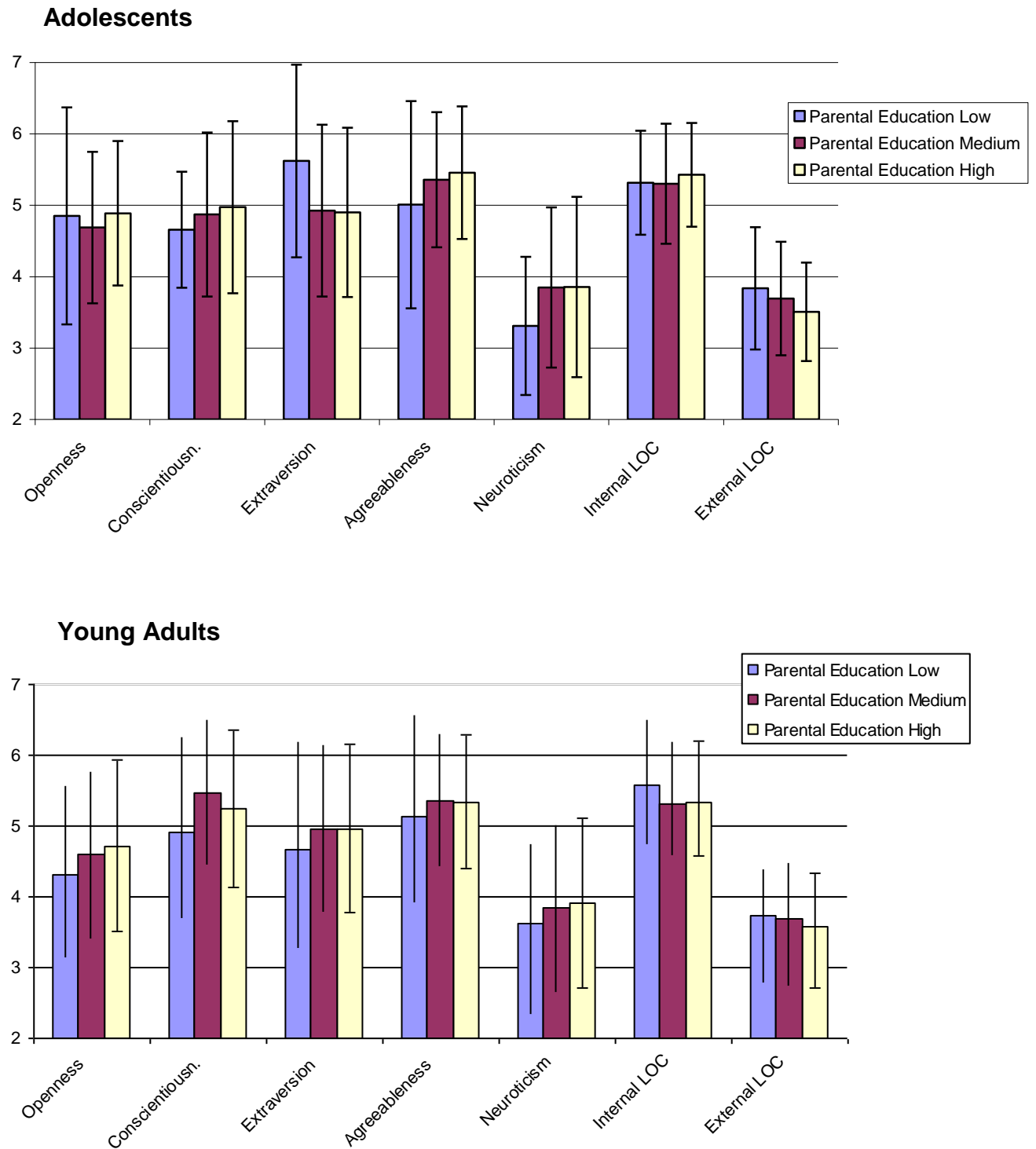
Source: Author's calculations based on SOEP v26, years 2005-2008. Weighted averages.

Adolescent children: verbal and numerical skills (word analogies, arithmetic operations) are added together to generate an index for crystallized intelligence, whereas abstract reasoning (matrix test) relates to fluid intelligence.

Young adult children and parents: word fluency (animal-naming task) relates to crystallized intelligence, whereas coding speed (symbol correspondence test) refers to fluid intelligence.

With the exception of the means for the personality traits, all summary statistics are taken from this smaller sample. However, the summary statistics of the bigger sample (noncognitive skills) are virtually the same.

Figure 16A.1: Children's Personality Scores According to Parental Education



Source: Author's calculations based on SOEP v26, years 2005-2008.

Chapter 17 Online Appendix

Table 17A.1a: Regression of log earnings on parental education and age and all intervening variables

A. With height

		Men-Sweden	Men-UK	Women-UK
(Intercept)		4.770 (0.204)	6.198 (0.618)	3.584 (0.858)
avggrade		0.503 (0.015)	0.570 (0.136)	1.026 (0.191)
avgparage		0.024 (0.006)	-0.020 (0.029)	0.067 (0.039)
birthweight		0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
avgparage ² /10		-0.003 (0.001)	0.002 (0.003)	-0.008 (0.005)
ISCEDkid(omitted: 1)	2	0.681 (0.166)	0.224 (0.087)	0.262 (0.172)
	3	0.785 (0.166)	0.291 (0.086)	0.428 (0.172)
	4	0.761 (0.166)	0.378 (0.096)	0.548 (0.180)
	5	0.793 (0.166)	0.515 (0.087)	0.853 (0.172)
ISCEDpar(omitted: 1)	2	-0.002 (0.012)	0.177 (0.071)	0.117 (0.099)
	3	0.007 (0.009)	0.011 (0.042)	-0.030 (0.059)
	4	0.028 (0.011)	0.033 (0.056)	-0.072 (0.079)
	5	0.032 (0.011)	0.109 (0.047)	0.000 (0.065)
lowbw		-0.031 (0.017)	-0.027 (0.079)	0.021 (0.115)
zheight		0.024 (0.003)	0.035 (0.017)	0.065 (0.022)
n		43620	1255	1371
k		15	15	15
σ		0.524	0.536	0.782
Adj R ²		0.0558	0.135	0.161

Table 17A.1b: Regression of log earnings on parental education and age and all intervening variables

B. Without height

		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Intercept)		4.770 (0.204)	6.198 (0.618)	89.160 (NaN)	3.584 (0.858)
avggrade		0.503 (0.015)	0.570 (0.136)	1.334 (NaN)	1.026 (0.191)
avgparage		0.024 (0.006)	-0.020 (0.029)	-4.199 (NaN)	0.067 (0.039)
birthweight		0.000 (0.000)	0.000 (0.000)	0.001 (NaN)	-0.000 (0.000)
avgparage ² /10		-0.003 (0.001)	0.002 (0.003)	0.500 (NaN)	-0.008 (0.005)
ISCEDkid(omitted: 1)	2	0.681 (0.166)	0.224 (0.087)		0.262 (0.172)
	3	0.785 (0.166)	0.291 (0.086)		0.428 (0.172)
	4	0.761 (0.166)	0.378 (0.096)		0.548 (0.180)
	5	0.793 (0.166)	0.515 (0.087)		0.853 (0.172)
ISCEDpar(omitted: 1)	2	-0.002 (0.012)	0.177 (0.071)	2.119 (NaN)	0.117 (0.099)
	3	0.007 (0.009)	0.011 (0.042)	0.019 (NaN)	-0.030 (0.059)
	4	0.028 (0.011)	0.033 (0.056)		-0.072 (0.079)
	5	0.032 (0.011)	0.109 (0.047)		0.000 (0.065)
lowbw		-0.031 (0.017)	-0.027 (0.079)		0.021 (0.115)
zheight		0.024 (0.003)	0.035 (0.017)	0.271 (NaN)	0.065 (0.022)
n		43620	1255	12	1371
k		15	15	8	15
σ		0.524	0.536	NaN	0.782
Adj R ²		0.0558	0.135	NaN	0.161

Table 17A.2: Descriptives of log earnings, parental education and age and all intervening variables

A. With height

		Men-Sweden	Men-UK	Women-UK
(Dependent var)		6.314 (0.540)	6.628 (0.576)	6.035 (0.854)
avggrade		0.476 (0.208)	0.543 (0.134)	0.564 (0.129)
avgparage		40.757 (4.973)	40.176 (5.592)	40.228 (5.622)
birthweight		3560.060 (539.641)	3397.878 (540.543)	3274.508 (481.386)
avgparage ² /10		168.586 (42.354)	164.539 (47.660)	164.994 (47.552)
ISCEDkid(omitted: 1)	2	0.074 (0.261)	0.191 (0.393)	0.191 (0.393)
	3	0.515 (0.500)	0.250 (0.433)	0.220 (0.414)
	4	0.147 (0.354)	0.096 (0.295)	0.100 (0.300)
	5	0.264 (0.441)	0.426 (0.495)	0.473 (0.499)
ISCEDpar(omitted: 1)	2	0.083 (0.276)	0.056 (0.229)	0.056 (0.230)
	3	0.497 (0.500)	0.298 (0.457)	0.302 (0.459)
	4	0.152 (0.359)	0.111 (0.314)	0.110 (0.313)
	5	0.180 (0.384)	0.278 (0.448)	0.282 (0.450)
lowbw		0.029 (0.168)	0.056 (0.229)	0.049 (0.216)
zheight		0.004 (0.999)	0.077 (0.946)	0.037 (1.016)
n		43620	1255	1371

Table 17A.2: Descriptives of log earnings, parental education and age and all intervening variables

B. Without height

		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Dependent var)		6.314 (0.540)	6.628 (0.576)	5.993 (0.502)	6.035 (0.854)
avggrade		0.476 (0.208)	0.543 (0.134)	0.434 (0.193)	0.564 (0.129)
avgparage		40.757 (4.973)	40.176 (5.592)	41.375 (3.830)	40.228 (5.622)
birthweight		3560.060 (539.641)	3397.878 (540.543)	3612.500 (627.968)	3274.508 (481.386)
avgparage ² /10		168.586 (42.354)	164.539 (47.660)	172.656 (31.993)	164.994 (47.552)
ISCEDkid(omitted: 1)	2	0.074 (0.261)	0.191 (0.393)		0.191 (0.393)
	3	0.515 (0.500)	0.250 (0.433)	0.375 (0.484)	0.220 (0.414)
	4	0.147 (0.354)	0.096 (0.295)	0.375 (0.484)	0.100 (0.300)
	5	0.264 (0.441)	0.426 (0.495)	0.125 (0.331)	0.473 (0.499)
ISCEDpar(omitted: 1)	2	0.083 (0.276)	0.056 (0.229)	0.125 (0.331)	0.056 (0.230)
	3	0.497 (0.500)	0.298 (0.457)	0.750 (0.433)	0.302 (0.459)
	4	0.152 (0.359)	0.111 (0.314)		0.110 (0.313)
	5	0.180 (0.384)	0.278 (0.448)		0.282 (0.450)
lowbw		0.029 (0.168)	0.056 (0.229)	0.000 (0.000)	0.049 (0.216)
zheight		0.004 (0.999)	0.077 (0.946)	-0.023 (0.966)	0.037 (1.016)
n		43620	1255	12	1371

Table 17A.3: Regression of log earnings on parental education and age

A. Regression results						
		Men-Sweden	Men-UK	Women-Sweden	Women-UK	
(Intercept)		5.111 (0.120)	6.758 (0.385)	4.701 (0.117)	4.465 (0.541)	
avgparage		0.051 (0.006)	-0.015 (0.019)	0.054 (0.006)	0.057 (0.026)	
avgparage ² /10		-0.006 (0.001)	0.002 (0.002)	-0.006 (0.001)	-0.006 (0.003)	
ISCEDpar(omitted: 1)	2	0.018 (0.012)	0.117 (0.041)	0.057 (0.012)	0.042 (0.061)	
		0.059 (0.009)	0.133 (0.027)	0.075 (0.009)	0.218 (0.039)	
	4	0.123 (0.011)	0.212 (0.038)	0.121 (0.011)	0.308 (0.054)	
		0.165 (0.010)	0.330 (0.028)	0.213 (0.010)	0.476 (0.041)	
	n		47949	3299	43864	3146
	k		7	7	7	7
	σ		0.555	0.593	0.515	0.832
	Adj R ²		0.0117	0.0393	0.0184	0.0511

B. Means and standard deviations					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Dependent var)		6.293 (0.558)	6.570 (0.605)	5.983 (0.520)	5.986 (0.854)
avgparage		40.800 (4.987)	39.444 (5.758)	40.720 (5.003)	39.493 (5.749)
avgparage ² /10		168.952 (42.518)	158.900 (47.861)	168.315 (42.549)	159.274 (47.909)
ISCEDpar(omitted: 1)	2	0.083 (0.276)	0.077 (0.267)	0.084 (0.278)	0.074 (0.262)
		3	0.496 (0.500)	0.274 (0.446)	0.494 (0.500)
	4		0.149 (0.356)	0.093 (0.291)	0.149 (0.356)
			0.181 (0.385)	0.230 (0.421)	0.182 (0.385)
	n		47949	3299	43864

Table 17A.4: Regression of birthweight on parental education and age

A. Regression results					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Intercept)		2192.308 (114.767)	1926.717 (261.506)	2565.187 (115.879)	2425.721 (241.780)
avgparage		58.072 (5.407)	65.319 (12.885)	34.094 (5.485)	32.613 (11.796)
avgparage ² /10		-5.952 (0.633)	-7.410 (1.560)	-3.208 (0.644)	-3.159 (1.415)
ISCEDpar(omitted: 1)	2	-18.623 (11.930)	49.266 (27.969)	-8.484 (11.941)	10.480 (27.410)
	3	-8.578 (8.962)	62.140 (19.342)	6.055 (9.028)	68.890 (18.301)
	4	-6.695 (10.374)	73.569 (28.718)	20.063 (10.396)	72.426 (26.234)
		18.867 (9.957)	96.154 (20.891)	29.229 (9.972)	102.651 (19.906)
n		49722	5345	45642	5000
k		7	7	7	7
σ		541	547	519	502
Adj R ²		0.00761	0.0115	0.00572	0.0135

B. Means and standard deviations					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Dependent var)		3552.823 (543.069)	3367.970 (549.883)	3424.202 (520.290)	3255.526 (505.319)
avgparage		40.835 (5.021)	39.087 (5.877)	40.725 (4.995)	39.291 (5.868)
avgparage ² /10		169.270 (42.846)	156.236 (48.513)	168.345 (42.479)	157.821 (48.882)
ISCEDpar(omitted: 1)	2	0.083 (0.275)	0.089 (0.284)	0.083 (0.276)	0.083 (0.276)
		3	0.496 (0.500)	0.254 (0.435)	0.491 (0.500)
	4		0.148 (0.356)	0.083 (0.276)	0.150 (0.357)
		0.182 (0.386)	0.203 (0.402)	0.185 (0.388)	0.207 (0.405)
	n	49722	5345	45642	5000

Table 17A.5: Regression of low birthweight on parental education and age

A. Regression results					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Intercept)		0.298 (0.036)	0.285 (0.113)	0.188 (0.043)	0.337 (0.117)
avgparage		-0.012 (0.002)	-0.011 (0.006)	-0.007 (0.002)	-0.012 (0.006)
avgparage ² /10		0.001 (0.000)	0.001 (0.001)	0.001 (0.000)	0.001 (0.001)
ISCEDpar(omitted: 1)	2	-0.007 (0.004)	-0.018 (0.012)	0.006 (0.004)	-0.021 (0.013)
	3	-0.003 (0.003)	-0.015 (0.008)	0.002 (0.003)	-0.032 (0.009)
	4	-0.005 (0.003)	-0.023 (0.012)	-0.004 (0.004)	-0.033 (0.013)
		-0.008 (0.003)	-0.018 (0.009)	-0.005 (0.004)	-0.034 (0.010)
n		49722	5345	45642	5000
k		7	7	7	7
σ		0.172	0.236	0.194	0.242
Adj R ²		0.00137	0.00113	0.0005	0.0042

B. Means and standard deviations					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Dependent var)		0.031 (0.172)	0.059 (0.236)	0.039 (0.194)	0.063 (0.243)
avgparage		40.835 (5.021)	39.087 (5.877)	40.725 (4.995)	39.291 (5.868)
avgparage ² /10		169.270 (42.846)	156.236 (48.513)	168.345 (42.479)	157.821 (48.882)
ISCEDpar(omitted: 1)	2	0.083 (0.275)	0.089 (0.284)	0.083 (0.276)	0.083 (0.276)
	3	0.496 (0.500)	0.254 (0.435)	0.491 (0.500)	0.264 (0.441)
	4	0.148 (0.356)	0.083 (0.276)	0.150 (0.357)	0.093 (0.290)
		0.182 (0.386)	0.203 (0.402)	0.185 (0.388)	0.207 (0.405)
n		49722	5345	45642	5000

Table 17A.6: Regression of average grade on parental education and age

A. Regression results					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Intercept)		-0.320 (0.042)	0.137 (0.066)	-0.266 (0.044)	0.142 (0.064)
avgparage		0.031 (0.002)	0.012 (0.003)	0.031 (0.002)	0.014 (0.003)
avgparage ² /10		-0.003 (0.000)	-0.001 (0.000)	-0.003 (0.000)	-0.001 (0.000)
ISCEDpar(omitted: 1)	2	0.014 (0.004)	-0.010 (0.007)	0.022 (0.005)	-0.015 (0.008)
	3	0.054 (0.003)	0.078 (0.005)	0.057 (0.003)	0.067 (0.005)
	4	0.131 (0.004)	0.089 (0.007)	0.114 (0.004)	0.093 (0.007)
		0.211 (0.004)	0.149 (0.005)	0.180 (0.004)	0.138 (0.005)
		n	48768	44804	4169
		k	7	7	7
		σ	0.196	0.195	0.118
		Adj R ²	0.123	0.0942	0.195

B. Means and standard deviations					
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
(Dependent var)		0.472 (0.210)	0.497 (0.132)	0.529 (0.205)	0.519 (0.131)
avgparage		40.819 (5.009)	39.695 (5.814)	40.708 (4.980)	39.696 (5.822)
avgparage ² /10		169.128 (42.703)	160.952 (48.654)	168.193 (42.326)	160.967 (48.867)
ISCEDpar(omitted: 1)	2	0.083 (0.275)	0.074 (0.261)	0.083 (0.276)	0.068 (0.252)
	3	0.497 (0.500)	0.272 (0.445)	0.493 (0.500)	0.278 (0.448)
	4	0.149 (0.356)	0.091 (0.287)	0.150 (0.357)	0.101 (0.301)
		0.181 (0.385)	0.226 (0.418)	0.184 (0.387)	0.225 (0.418)
		n	48768	44804	4169

Table 17A.7: Regression of height on parental education and age

A. Regression results

	Men-Sweden	Men-UK
(Intercept)	-1.790 (0.224)	-1.296 (0.833)
avgparage	0.073 (0.011)	0.048 (0.040)
avgparage ² /10	-0.008 (0.001)	-0.005 (0.005)
ISCEDpar(omitted: 1) 2	0.002 (0.023)	0.155 (0.094)
3	0.077 (0.018)	0.234 (0.056)
4	0.173 (0.020)	0.176 (0.077)
	0.225 (0.019)	0.402 (0.058)
n	44890	2150
k	7	7
σ	0.995	0.982
Adj R ²	0.00893	0.0234

B. Means and standard deviations

	Men-Sweden	Men-UK
(Dependent var)	0.003 (0.999)	0.024 (0.993)
avgparage	40.766 (4.974)	40.101 (5.672)
avgparage ² /10	168.659 (42.369)	164.029 (48.020)
ISCEDpar(omitted: 1) 2	0.083 (0.276)	0.061 (0.239)
3	0.494 (0.500)	0.287 (0.453)
4	0.152 (0.359)	0.101 (0.302)
	0.183 (0.387)	0.260 (0.439)
n	44890	2150

Table 17A.8: Regression of offspring education level on parental education and age

		A. Regression results			
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
1 2		0.543 (0.001)	-0.143 (1.302)	-0.648 (0.001)	0.961 (1.361)
2 3		5.098 (0.142)	1.749 (1.301)	3.850 (0.185)	3.238 (1.358)
3 4		8.075 (0.143)	3.008 (1.302)	6.932 (0.187)	4.431 (1.359)
4 5		8.835 (0.144)	3.418 (1.302)	7.571 (0.187)	4.860 (1.360)
avgparage		0.283 (0.007)	0.059 (0.063)	0.258 (0.009)	0.143 (0.067)
avgparage ² /10		-0.029 (0.001)	-0.003 (0.008)	-0.027 (0.001)	-0.014 (0.008)
ISCEDpar(omitted: 1)	2	0.213 (0.045)	0.405 (0.137)	0.268 (0.046)	0.411 (0.143)
	3	0.754 (0.034)	1.028 (0.093)	0.772 (0.035)	0.908 (0.091)
	4	1.558 (0.039)	1.375 (0.132)	1.469 (0.040)	1.398 (0.126)
		2.411 (0.038)	2.027 (0.104)	2.244 (0.040)	2.090 (0.107)
n		48662	3272	44359	3570
k		10	10	10	10
AIC		1.08e+05	8.93e+03	9.61e+04	9.42e+03
		B. Means and standard deviations			
		Men-Sweden	Men-UK	Women-Sweden	Women-UK
Dependent	2				
	3				
	4				
	5				
avgparage		40.837 (5.024)	39.481 (5.816)	40.720 (5.004)	39.487 (5.686)

avgparage ² /10		169.288	159.254	168.316	159.157
		(42.868)	(48.560)	(42.561)	(47.297)
ISCEDpar(omitted: 1)	2	0.083	0.078	0.084	0.070
		(0.276)	(0.268)	(0.277)	(0.256)
	3	0.497	0.266	0.493	0.287
		(0.500)	(0.442)	(0.500)	(0.453)
	4	0.149	0.094	0.149	0.100
		(0.356)	(0.291)	(0.356)	(0.300)
		0.180	0.232	0.183	0.223
		(0.385)	(0.422)	(0.386)	(0.416)
n		48662	3272	44359	3570