College, cognitive ability, and socioeconomic disadvantage: policy lessons from the UK in 1960-2004

ONLINE APPENDIX

Andrea Ichino Aldo Rustichini Giulio Zanella

November 7, 2024

Appendix to Section 3.1

The 1970 British Cohort Study

The 1970 British Cohort Study (BCS70) is a longitudinal study following the lives of 17,198 individuals born in England, Scotland, Wales, and Northern Ireland in a single week of April 1970. The BCS70 is representative of the British population at each respective age at which participants are observed. After the initial birth survey, conducted in 1970, data were collected in 10 additional surveys of the same participants. The last data collection was carried out in 2021 when the cohort members were 51 years old. Each survey monitored the cohort members' health, education, and economic outcomes. The data can be downloaded using the UK Data Service at this URL https://ukdataservice.ac.uk. For additional details see (Dodgeon et al., 2020).

We use the second wave (1975), third wave (1980), and seventh wave (2004), where cohort members were 5, 10, and 34 years old, respectively. In these waves, cohort members completed tests assessing their verbal and logical/mathematical skills. Excluding observations with missing test scores, the usable sample sizes for these waves are 7,479, 6,927, and 9,448. Excluding high school dropouts, the sample sizes are 5,748, 5,308, and 7,270. After imputing the missing scores as described in the Online Appendix to Section 3.3, the usable sample size of subjects with tertiary education or certification that allows them to apply for tertiary education is 7,369 for all waves.

University Statistical Record

The University Statistical Record (USR) contains administrative information on the universe of students enrolled at UK universities between 1972 and 1993. USR was initiated following the Robbins Report, which stressed the need for better data for the proper design of higher education policies. It was subsequently discontinued and replaced by the Higher Education Statistics Agency (HESA) in 1993. Unfortunately, pre-1993 USR information was not merged into HESA. Out of the initial 8,103,977 person/year records of students enrolled in a UK higher education institution in this period, we keep the 6,889,425 records of white individuals born in the UK, so as to match the final USoc sample. These records correspond, after some minor data cleaning, to 1,523,192 students born during 1948-1976.

Appendix to Section 3.3

USoc

Table A-1: Eigenvector of the PCA of cognitive ability measures in USoc

Immediate word recall	0.457	Help in immediate word recall	-0.011
Delayed word recall	0.449	Help in delayed word recall	0.004
Correct subtractions	0.318	Help in substractions test	-0.050
Number series	0.413	Help in number series test	-0.034
Verbal ability	0.365	Help in verbal ability test	-0.015
Numeric ability	0.423	Help in numeric ability test	-0.004
		Material aid in recall test	0.011
		Material aid in subtraction test	-0.040

Notes: The table reports the eigenvector of the Principal Components Analysis of the 14 cognitive ability measures contained in the USoc sample with cognitive information described in the central panel of Table 1, which is representative of the UK population. The First Principal Components (FPC) is the measure of ability that we use in our analysis. It has an eigenvalue of 2.55 and explains 18.2% of the data variability. The left panel of the table displays the positive values of the eigenvector terms for the fractions of correct answers in the 6 cognitive questions. The right panel, shows instead that the eigenvector values are negative for 6 out of 8 help dummies. For the the two remaining help dummies the values are positive but close to zero.





Notes: The figure illustrates the empirical distribution of the FPC of the 14 cognitive ability variables in the USoc sample with cognitive information described in the central panel of Table 1



Figure A-2: Evolution of the cognitive ability score with different standardizations

Notes: The left panel displays the mean, the 10th and the 90th percentiles of the average ability score standardized within each birth year in the USoc sample of 22,175 white respondents born in the UK between 1940 and 1984, with non-missing education and ability score (see the central panel in Table 1). The right panel displays the mean, the 10th and the 90th percentiles of the average ability score standardized over all birth years in the same USoc sample.

The 1970 British Cohort Study

Imputation in the BCS70

Not all BCS70 subjects have all cognitive scores at all ages and there are several codes for missing information. We explain here in greater detail the different imputation strategies that we follow in the panels of Figure 2 of the main text. In panel (A) or (B) no imputation of such missing information is attempted. In panel (C), the missing answers were coded as follows. First, "Not stated", "no answer", and "more than one answer" were treated as incorrect answers, and "Not applicable" and "not scorable" were coded as missing. Then, the missing answers were imputed as follows:

- If an observation had some missing answers in a specific test, but not all answers were missing, the missing answers were coded to be incorrect.
- If all answers for a test were missing, but the score for at least one test of the same age was available, the score of the missing test was set to be equal to the child's average score from the other tests.
- If all answers to all the tests were missing, the child's score in each test was set to the average score of the other children on the test, conditional on educational attainment.

When creating a general cognitive score using the PCA, a dummy for having missing values in all tests was used.

PCA in the BCS70

	Including imputed scores	Excluding imputed scores
Schonell Reading Test	0.222	0.224
Human Figure Drawing 1	0.515	0.518
Human Figure Drawing 2	0.522	0.525
English Picture Vocabulary Test	0.354	0.340
Complete a Profile	0.319	0.323
Copying Designs	0.432	0.430
Dummy for all tests missing	0.005	_

Table A-2: Eigenvector of the PCA of cognitive ability measures in BCS70 age 5

Notes: The table reports the eigenvector of the Principal Components Analysis of the 6 cognitive ability measures contained in BCS70 sweep 2 (age 5). Two samples are presented: one including imputed scores and one excluding them. For both samples, the First Principal Components (FPC) is the measure of ability that we use in our analysis. It has eigenvalues of 2.56 and 2.51, explaining 36.5% and 41.8% of data variability in the imputed and non-imputed samples, respectively.

	Including imputed scores	Excluding imputed scores
Pictorial Language Comprehension	0.345	0.412
Spelling Dictation Task	0.410	0.395
Friendly Maths Test	0.486	0.469
Edinburgh Reading Test	0.497	0.487
British Ability Scales (BAS)	0.480	0.466
Dummy for all tests missing	-0.002	_

Table A-3: Eigenvector of the PCA of cognitive ability measures in BCS70 age 10

Notes: The table reports the eigenvector of the Principal Components Analysis of the 5 cognitive ability measures contained in BCS70 sweep 3 (age 10). Two samples are presented: one including imputed scores and one excluding them. For both samples, the First Principal Components (FPC) is the measure of ability that we use in our analysis. It has eigenvalues of 2.82 and 3.35, explaining 47% and 67% of data variability in the imputed and non-imputed samples, respectively.

Table A-4:	Eigenvector	of the PCA	of cognitive	ability	measures	in	BCS70	age	34
------------	-------------	------------	--------------	---------	----------	----	-------	-----	----

	Including imputed scores	Excluding imputed scores
Literacy Skills	0.705	0.707
Numeracy Skills	0.705	0.707
Dummy for all tests missing	-0.069	—

Notes: The table reports the eigenvector of the Principal Components Analysis of the 2 cognitive ability measures contained in BCS70 sweep 7 (age 34). Two samples are presented: one including imputed scores and one excluding them. For both samples, the First Principal Components (FPC) is the measure of ability that we use in our analysis. It has eigenvalues of 1.67 and 1.65, explaining 55.9% and 82.6% of data variability in the imputed and non-imputed samples, respectively.

Variations in cognitive ability between college and high school graduates at different ages

With reference to Figure 2 in the main text, in order to test formally that the ability of college graduates does not increase (neither relative to that of high school graduates nor in absolute terms) between age 10 and 34, we estimate the following equation:

$$\theta_{it} = \sum_{t=1}^{3} b_{1t} + \sum_{t=1}^{3} b_{2t} k_i \tag{A-1}$$

where θ_{it} is the cognitive ability of individual *i* at age *t*, k_i is a dummy that takes value 1 if the individual eventually obtains a college degree, and $t \in \{1, 2, 3\}$, where 1 indexes age 5, 2 is for age 10 and 3 is for age 34. Estimation is based on 4,386 subjects from the 1970 British Cohort Study, as used in Panel A of Figure 2 in the main text.

Table A-5: Cognitive ability of college and high school graduates at different ages

Param.	Estimate	(s.e.)	
b_{11}	99.337	(0.255)	
b_{12}	98.634	(0.253)	
b_{13}	99.487	(0.253)	
b_{21}	6.778	(0.487)	
b_{22}	11.959	(0.419)	
b_{23}	9.375	(0.352)	
Total observations	13,1	58	
Individuals	4,386		

Notes: The table reports OLS estimates of the parameters in equation (A-1). Robust standard errors in parentheses. Sample: 4,386 subjects from the 1970 British Cohort Study. See the note to Panel A of Figure 2 in the main text for additional details.

Using these estimates we test the null hypothesis $H_0: b_{22} \leq b_{23}$ against the alternative $H_A: b_{22} > b_{23}$ and we reject the null (*p*-value < 0.0001). We also test the null hypothesis $H_0: b_{12} + b_{22} \leq b_{13} + b_{23}$ against the alternative $H_A: b_{12} + b_{22} > b_{13} + b_{23}$ and we reject the null again (*p*-value < 0.0001). Therefore, we can conclude at conventional confidence levels that between ages 10 and 34 the ability of college graduates does not increase relative to that of high school graduates or in absolute terms. Note that, instead the ability gap between the two groups increases until age 10. These results are consistent with the literature cited in the text, claiming that general cognitive ability is possibly malleable in the first years of life but unlikely to change beyond a very young age (for a survey, see Figure 2 in the main text also shows that the cognitive ability of individuals who attain a college degree is higher and grows during these sensitive years relative to those who don't. This may be one of the reasons why they eventually self-select into college, in line with our model.

Appendix to Section 3.4

Table A-6:	Eigenvector of	f the PCA	of sc	ocioeconomic	factors	generating	advantage	in	college
enrollment	and graduatio	n in USoc	,						

Father education	0.277
Mother education	0.290
Mother work	0.191
Mother dead	-0.125
Mother absent	-0.224
Father work	0.617
Father dead	-0.416
Father absent	-0.428

Notes: The table reports the eigenvector of the Principal Components Analysis of the eight socioeconomic background variables (referring retrospectively to when the respondent was 14 years of age) on which we base our measure of socioeconomic disadvantage. This analysis is conducted using the USoc sample with cognitive information described in the central panel of Table 1, which is representative of the UK population. The First Principal Component (FPC) has an eigenvalue of 1.76 and explains 22% of the data variability. The table displays negative values for the variables that, as expected, reduce the FPC and increase disadvantage: whether either parent was dead or absent when the respondent was 14 years of age.

Figure A-3: Distribution of socioeconomic disadvantage



Notes: The figure illustrates the empirical distribution of our measure of disadvantage The measure is the FPC of 8 socioeconomic variables at age 14, rescaled so that the minimum is zero.

Appendix to Section 5.1

Bootstrap procedure

As mentioned in the main text, standard errors for the estimated structural parameters are obtained from 1,000 bootstrap replications, in samples obtained from random draws with replacement. We redraw with replacement separately from each college cohort sample. However, the ability and disadvantaged measures are *not* recomputed, so that for each individual in the bootstrap samples these measures are those computed from the PCA in the original sample (as if they were data when bootstrapping). Standard errors are given by the standard deviation of each parameter's estimate across the 1,000 replications.

Starting values

It is plausible that our MD estimation criterion function, $C(\gamma, \tau, \delta, \beta, \alpha_1, \alpha_2, \alpha_3; \rho)$, has local minima, and given that the grid is finite, the "wrong" starting point for the search process may yield estimates that correspond to one of them. This is particularly worrisome because there is no reference scale for policy parameters γ , τ , δ , and β , and no reference value in the literature for technology parameters α_j estimated in models where productivity depends on both educational attainment and cognitive ability. Thus, one does not know where the grid should be centered in \mathbb{R}^7 in order not to get stuck into a local minimum. We solve this problem by noting that: (i) a researcher not interested in disentangling the impact of higher education policy $G = (\gamma, \tau, \delta, \beta)$ from changing technology and socioeconomic characteristics or not interested in using the model for equilibrium policy analysis, can obtain a partial set of estimates by Nonlinear Least Squares (NLS) from the supply-side equation (18), after replacing $\Delta \ln w_i(G)$ with its empirical analog, $\ln \widehat{w}(1,j) - \ln \widehat{w}(0,j)$; and that (ii) the demand-side equation (13) can be used in isolation to calibrate the technological productivity ratios $(\alpha_1, \alpha_2, \alpha_3)$. Specifically, for each ability group j and cohort, we solve equation (13) for α_j and then plug into the resulting equation the empirical values of the odds of college graduation, ξ_j , and the wage ratio, r_j to obtain a numerical value for α_j .

These NLS estimates of $(\gamma, \tau, \delta, \beta)$ and calibrated values for $(\alpha_1, \alpha_2, \alpha_3)$ are reported in Table A-7, and provide reasonable starting values for our grid-search procedure, even if they ignore the equilibrium effects of higher education policy or of technological change.

Grid search

Our MD estimates and standard errors are then obtained as follows. Starting from the initial values in Table A-7, we set up a grid to locate the global minimum of criterion function $C(\gamma, \tau, \delta, \beta, \alpha_1, \alpha_2, \alpha_3; \rho)$, calibrating parameter ρ to the value of 0.584 estimated for the UK by Card and Lemieux (2001). Anchoring the grid search process to these initial values – which are essentially partial-equilibrium guesses of the parameters that we want to estimate – increases our confidence that the MD algorithm – which instead takes into account general equilibrium effects – does not end up at a local minimum.

In order to mitigate the curse of dimensionality, we design an algorithm that starts from a small grid composed by $3^7 = 2,187$ points, 3 for each of the seven parameters (the initial guess and two neighboring points, at distance 0.01 for γ , τ , and the three α_j 's, and

[A] NLS on supply-side eq. (18)									
College cohort									
	1960 - 1974	1975-1989	1990-2004						
γ	9.902	3.870	3.982						
	(0.800)	(0.879)	(0.622)						
au	-5.874	-1.841	-2.096						
	(0.670)	(0.623)	(0.439)						
δ	0.000	0.047	0.031						
	(0.000)	(0.015)	(0.013)						
β	0.000	0.006	0.015						
	(0.000)	(0.001)	(0.001)						
N	4,716	6,922	6,252						
$[\mathbf{B}]$	Calibration	of demand-si	de eq. (13)						
	C	College cohor	rt						
	1960-1974	1975-1989	1990-2004						
α_1	0.446	0.606	0.646						
α_2	0.740	0.816	0.949						
α_3	1.190	1.230	1.382						

Table A-7: Initial estimates of policy parameters

Notes: Panel [A] reports Nonlinear Least Squares (NLS) estimates of parameters in equation (18), after replacing $\Delta \ln w_j(G)$ with its empirical counterparts, i.e., $\ln w(1, j) - \ln w(0, j)$, for each ability group j. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school certification that allows for college enrollment (see the right panel of Table 1 in the paper). Panel [B] reports our calibration of the technological productivity ratios across ability groups, α_j , obtained by solving equation (13) for α_j and then plugging into the resulting equation the empirical values of the odds of college graduation, ξ_j , and the wage ratio, r_j . Parameter ρ is set to the value of 0.584 estimated for the UK by Card and Lemieux (2001).

distance 0.001 for δ and β). We then solve numerically for the model's equilibrium at each point of this grid by finding the unique fixed point of equation (20) for that particular combination of $(\gamma, \tau, \delta, \beta, \alpha_1, \alpha_2, \alpha_3)$, and we obtain a MD estimate by locating the minimum of $C(\gamma, \tau, \delta, \beta, \alpha_1, \alpha_2, \alpha_3; \rho)$ over the grid. If this MD estimate hits a grid boundary (for example, if the estimate for γ is the minimum or the maximum in the vector of values for γ that is used to build the grid), then a point is added to enlarge that boundary and estimation is repeated over the expanded grid. This process is iterated until the MD estimates are at an interior point of the grid. The grid is expanded considerably before convergence is achieved, which indicates that partial and general equilibrium estimates may differ in non-negligible ways. In the initial, actual sample that produces the point estimates in Table 4 of the main text, the size of the grids are: 709,800 points for college cohort 1960-1974 (the minimum value of the criterion function is $\min = 0.0241$); 173,502,800 points for cohort 1975-1989 (min = 0.0102); and 33,546,240 points for cohort 1990-2004 (min = 0.0787). Then, to produce bootstrap estimates from which standard errors are obtained, we build a new initial grid centered around these point estimates and we iterate according to the "no boundary estimates" rule described above in the 1,000 bootstrap samples. The distribution of the resulting 1,000 bootstrap estimates of each parameter is illustrated in Figure A-4 for three three college cohorts. The vertical lines mark the averages, which are virtually identical to our point estimates obtained in the actual sample. The standard deviations are our bootstrap standard errors in Table 4 of the main text.

Identification

A crucial question in our econometric analysis is whether the MD algorithm that we employ produces estimates that correspond to a global minimum or not. To increase our confidence that it does, we inspect two- and three-dimensional sections of the criterion function over a much wider grid than the one used by the computational algorithm. The two-dimensional sections are shown in Figure A-5 for each college cohort. Each panel in this figure plots the value of the log of the MD criterion, i.e., $\ln C(\gamma, \tau, \delta, \beta, \alpha_1, \alpha_2, \alpha_3; \rho)$, letting only one parameter at the time vary while keeping the remaining six parameters (in addition to ρ) fixed, at the point estimates in Table 4 of the main text. Both the global minimum and local minima are clearly visible in each panel. The vertical line corresponds to the free parameter. Note that despite the appearance of a cusp, the function is smooth around the minimum – such appearance is due to the log scale, which is convenient but produces a large negative value at the minimum, where it is very close to zero. Local minima are clearly visible, but in all cases, our MD estimate corresponds to the global minimum.

This exercise can be extended to allowing two parameters to vary while keeping the remaining five (in addition to ρ) fixed at the point estimates in Table 4 of the main text, thus producing three-dimensional sections of the log criterion function that can be represented graphically. Such sections are shown in Figure A-6 for college cohort 1960-1974 and in Figure A-7 for college cohort 1990-2004. The possible $\binom{7}{2} = 21$ combinations are represented for either cohort, and each panel illustrates the contour lines of the log MD criterion. The intersections of the straight continuous and dashed lines represent our MD estimates and the global minimum, respectively. Again, local minima are visible and we avoid them with the only relevant exception of the first panel in Figure A-6. In this case, our MD algorithm selects a local minimum for parameters γ and τ . Yet this local minimum is sufficiently close to the global one to leave our conclusion that from the 1960-1974 cohort to the 1990-2004 cohort γ declined and τ increased substantially (which results in a non-meritocratic expansion policy) unchanged.



Figure A-4: Distribution of MD estimates across 1,000 bootstrap samples

Notes: The figure illustrates the distribution of MD estimates of the seven structural parameters of interest across 1,000 bootstrap samples, by college cohort. The vertical line is the mean of the distribution. The standard errors reported in Table 4 of the main text are the standard deviations of these distributions.



Figure A-5: 2D sections of criterion function, log scale

Notes: Each panel plots the value of the log of the MD criterion as a function of one parameter, keeping the remaining six parameters fixed at the MD estimates obtained in the actual (as opposed to bootstrap) sample. The dashed line marks the global minimum, which corresponds to our MD estimate in the original sample. Local minima are clearly visible, and anchoring the grid to our initial estimates (see Table A-7) helps avoid them. Despite the appearance of cusps, the function is smooth around the global minimum, which takes a large negative value on the log scale because it is close to zero.



Figure A-6: 3D sections of criterion function for cohort 1960-1974, log scale

Notes: see Appendix text



Figure A-7: 3D sections of criterion function for cohort 1990-2004, log scale

Notes: see Appendix text

Appendix to Section 5.2

	[A] Colle	ege graduat	ion odds		[E	B] Wage rat	io	
	1960-1974	1975-1989	1990-2004		1960-1974	1975-1989	1990-2004	
		ξ_1		Ability group	1	r_1		
model	0.136	0.161	0.176	model	1.122	1.241	1.133	
	(0.013)	(0.032)	(0.013)		(0.036)	(0.033)	(0.007)	
data	0.139	0.161	0.186	data	1.012	1.297	1.301	
	(0.014)	(0.011)	(0.014)		(0.099)	(0.057)	(0.040)	
		ξ_2		Ability group	2	r_2		
model	0.294	0.351	0.558	model	1.214	1.237	1.211	
	(0.012)	(0.009)	(0.012)		(0.023)	(0.009)	(0.011)	
data	0.274	0.346	0.469	data	1.267	1.269	1.300	
	(0.018)	(0.018)	(0.018)		(0.086)	(0.040)	(0.033)	
		ξ_3		Ability group	3	r_3		
model	0.656	0.796	1.251	model	1.406	1.275	1.221	
	(0.023)	(0.015)	(0.023)		(0.049)	(0.010)	(0.052)	
data	0.695	0.870	1.264	data	1.384	1.303	1.254	
	(0.033)	(0.036)	(0.033)		(0.080)	(0.036)	(0.030)	
	[C] C	Cognitive A	bility		[D] SES disadvantage			
	1960-1974	1975-1989	1990-2004		1960-1974	1975-1989	1990-2004	
		$\mathbb{E}(\Theta K=1)$	C	 College graduat	tes	$\mathbb{E}(\Lambda K=1)$		
model	110.8	109.4	108.1	model	1.92	1.57	1.37	
	(0.4)	(0.4)	(0.3)		(0.03)	(0.03)	(0.03)	
data	110.3	109.0	108.2	data	1.85	1.58	1.21	
	(0.4)	(0.3)	(0.3)		(0.03)	(0.03)	(0.02)	
		$\mathbb{E}(\Theta K=0)$	Hig	h school gradi	iates	$\mathbb{E}(\Lambda K=0)$		
model	101.6	99.4	97.1	model	2.19	1.88	1.71	
	(0.3)	(0.3)	(0.4)		(0.02)	(0.02)	(0.03)	
data	101.8	99.5	97.3	data	2.22	1.88	1.79	
	(0.2)	(0.2)	(0.2)		(0.02)	(0.02)	(0.02)	
	. ,	. ,	. ,		. ,	. ,	. ,	

Table A-8: Empirical vs model-predicted targeted moments

Notes: The table reports the model-predicted and the empirical values (in the actual sample) of the ten targeted moments in our minimum distance (MD) estimation of the policy and technology parameters reported in Table 4 in the paper: the college graduation odds and wage ratio in the ability groups defined by the terciles of the ability distribution in the UK population, ξ_j and r_j , respectively, for i = 1, 2, 3); and the average ability and disadvantage of college and high school graduates. Bootstrap standard errors (1,000 replications) in parentheses. The MD criterion function is given by equation (21) in the paper, and the weighting matrix is the identity matrix. The college cohorts to which columns refer are defined by the period of actual or potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school certification that allows for college enrollment (see the right panel of Table 1 in the paper.)

	Ability distribution				Disadvantage distribution					
	(College cohor	·t		College cohort					
	1960-1974	1975-1989	1990-2004		1960-1974	1975-1989	1990-2004			
	Colle	ge, 25^{th} perc	entile		Colle	ge, 25^{th} perc	entile			
model	$103.1 \\ (0.5)$	$101.6 \\ (0.5)$	$99.9 \\ (0.4)$		$1.49 \\ (0.03)$	$1.02 \\ (0.01)$	$0.80 \\ (0.05)$			
data	103.0 (0.5)	101.6 (0.3)	$100.3 \\ (0.4)$		$1.32 \\ (0.03)$	$0.99 \\ (0.01)$	$0.64 \\ (0.01)$			
	Colle	ge, 75^{th} perc	entile		$College, 75^{th} percentile$					
model	$118.7 \\ (0.5)$	$117.9 \\ (0.3)$	$117.0 \\ (0.3)$		$2.16 \\ (0.06)$	$1.83 \\ (0.02)$	$1.52 \\ (0.02)$			
data	$117.7 \\ (0.4)$	$117.6 \\ (0.3)$	117.1 (0.3)		2.03 (0.04)	1.78 (0.02)	$1.46 \\ (0.01)$			
	High sc.	$hool$, 25^{th} pc	ercentile		High school , 25^{th} percentile					
model	$93.6 \\ (0.3)$	$91.0 \\ (0.3)$	88.4 (0.4)		$1.61 \\ (0.02)$	$1.24 \\ (0.02)$	$0.98 \\ (0.00)$			
data	$93.9 \\ (0.3)$	$91.3 \\ (0.3)$	88.9 (0.3)		1.64 (0.00)	1.34 (0.02)	1.01 (0.02)			
	High school, 75^{th} percentile			High school , 75^{th} percentile						
model	$111.1 \\ (0.3)$	109.3 (0.2)	$106.7 \\ (0.3)$		$2.31 \\ (0.01)$	$1.95 \\ (0.01)$	$1.74 \\ (0.03)$			
data	$111.2 \\ (0.3)$	109.4 (0.2)	106.9 (0.3)		2.31 (0.01)	$1.95 \\ (0.01)$	1.74 (0.01)			

Table A-9: Empirical vs model-predicted untargeted moments

.

Notes: The table reports the model-predicted and the empirical values of eight untargeted moments (in the actual sample). Bootstrap standard errors (1,000 replications) in parentheses. A college cohort is defined by the period of actual or potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1) in the paper).

Robustness to the inclusion of the "Big Five" personality traits

Using the "Big Five" (B5) personality traits to produce alternative measures of ability or disadvantage does not alter our conclusions. This is shown in Table A-10 and Figure A-8.

In the table, row 1 reproduces (for convenience) the correlations between ability Θ and disadvantage Λ reported in the main text. The remaining rows report this correlation when: the B5 traits are added to the socioeconomic indicators in the PCA for disadvantage (row 2), only the B5 variables contribute to the PCA for disadvantage (row 3), the B5 variables are added to the cognitive ability indicators in the PCA to produce a broader ability measure (row 4).

Figure A-8 follows the same exposition pattern to describe how the cost shifts $\Omega(\Lambda)$ and $\Gamma(\Theta)$ changed between the first and the final college cohorts, depending on whether and where the B5 variables are included. In all cases, our central conclusion stands: the cost of study effort decreased more for low-ability students than for high-ability ones and increased more for high-disadvantage students than for low-disadvantage ones.

	College cohort				
	1960-1974	1975-1989	1990-2004		
$\Theta = FPC(cognitive variables); \Lambda = FPC(SES variables)$	-0.129	-0.143	-0.139		
	(0.016)	(0.014)	(0.014)		
$\Theta = FPC(cognitive variables); \Lambda = FPC(SES \& B5 variables)$	-0.131	-0.151	-0.150		
	(0.016)	(0.014)	(0.014)		
$\Theta = FPC(cognitive variables); \Lambda = FPC(B5 variables)$	-0.029	-0.047	-0.041		
	(0.016)	(0.014)	(0.015)		
$\Theta = \mathrm{FPC}(\mathrm{cognitive}\ \&\ \mathrm{B5}\ \mathrm{variables}); \ \Lambda = \mathrm{FPC}(\mathrm{SES}\ \mathrm{variables})$	-0.127	-0.150	-0.141		
	(0.016)	(0.014)	(0.013)		
N	4,716	6,922	$6,\!252$		

Table A-10: Correlation η between alternative measures of cognitive ability Θ and socioeconomic disadvantage Λ in USoc

Notes: The table reports the correlation between alternative measures of ability (Θ) and disadvantage (Λ) that combine in different ways cognitive ability variables, socioeconomic status (SES) variables at the time the respondent was 14, and the "Big Five" (B5) personality variables available in USoc: in the first row, Θ is the FPC of cognitive variables and Λ is the FPC of SES variables (this first row is identical to the correlation reported in the paper); in the second row, Θ is the FPC of cognitive variables and Λ is the FPC of SES and B5 variables; in the third row, Θ is the FPC of cognitive variables and Λ is the FPC of B5 variables; in the fourth row, Θ is the FPC of cognitive and B5 variables, and Λ is the FPC of SES variables. Standard errors are produced via the delta method. Cross-sectional response weights are applied. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1) in the paper).

Figure A-8: Estimated study effort cost shifts under alternative definitions of ability and disadvantage that include the "Big Five" (B5) personality traits



Notes: The figure shows the study effort cost shifts $\Gamma(\cdot)$ and $\Omega(\cdot)$ implied by NLS estimates of the policy parameters, for the 1960-1974 or 1990-2004 college cohorts, as a function of ability (left panel) or disadvantage (right panel), using alternative definitions of ability and disadvantage that combine in different ways cognitive ability variables, socioeconomic status (SES) variables at the time the respondent was 14, and the "Big Five" (B5) personality variables available in USoc, following the same order of Table A-10. To facilitate a comparison, the first row corresponds to the definition employed in the paper, i.e., ability is the FPC of cognitive variables and disadvantage is the FPC of socioeconomic variables. This first row is identical to Figure 10 in the paper. In the second row, ability is the FPC of cognitive variables and disadvantage is the FPC of B5 variables; in the fourth row, ability is the FPC of cognitive and B5 variables, and disadvantage is the FPC of SES variables.

Appendix References

- CARD, D. AND T. LEMIEUX (2001): "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *The Quarterly Journal of Economics*, 116, 705–746.
- DODGEON, B., J. ELLIOTT, M. HANCOCK, J. JOHNSON, S. PARSONS, A. PETERS, AND P. SHEPHERD (2020): "1970 British Cohort Study Age 34 Sweep User Guide," UCL Centre for Longitudinal Studies.