The Path to College Education: Are Verbal Skills More Important than Math Skills?*

Esteban M. Aucejo†

London School of Economics and Political Science

Jonathan James‡

California Polytechnic State University

November 2, 2015

Abstract

Skills are multiple in nature and used with different intensity across activities. The aim of this paper is to study the differential roles of math and verbal skills for educational outcomes. By estimating a multi-period factor model of skills, using a rich panel database that follows all students in England from elementary school to university, we find that verbal skills play a greater role in explaining university enrollment than math skills. In addition, we use our framework to study the process of skill development during compulsory schooling. Results show that 40% of the variance of skills acquired by the end of compulsory education is determined after the second grade, which indicates some scope for overcoming initial skill disadvantages. Finally, we study the gender gaps in college enrollment and STEM fields, showing that verbal skills and comparative advantage in skills are key determinants of these gaps.

*We thank Ghazala Azmat, Monica Langella, and Tyler Ransom for valuable comments. We also thank seminar participants at Royal Holloway, Cal Poly, and the Federal Reserve Bank of St. Louis.

†Email: E.M.Aucejo@lse.ac.uk

‡Email: jjames04@calpoly.edu. Homepage: `http://www.calpoly.edu/~jjames04`

1 Introduction

The earnings and employment prospects of less educated workers have experienced a sharp decline since the early 1980's (Acemoglu and Autor, 2011). As formal education becomes an increasingly important determinant of lifetime income (Castex and Dechter, 2014), understanding the factors that influence schooling decisions is essential from a policy perspective. A large literature has established that cognitive skills are an important determinant of educational attainment (Heckman et al., 2006; Cawley et al., 2001; Cameron and Heckman, 2001). However, skills are multiple in nature (Cunha and Heckman, 2007) and, therefore more attention should be given to understanding exactly which skills have the greatest influence on post-secondary educational outcomes, and how these specific skills evolve over the schooling career.

The aim of this paper is threefold. First, we study whether math and verbal skills, developed during compulsory education, differentially impact university enrollment and other university outcomes. Understanding how math and verbal skills affect schooling decisions is important for predicting which types of students will be most influenced by public policies aimed at increasing university enrollment.[1] Given that the returns to formal education have increased in recent years (Castex and Dechter, 2014), acknowledging the channel in which skills influence educational attainment has become increasingly important for policy decisions.[2] Moreover, bringing attention to the math-verbal distinction complements previous work that has largely focused on socio-emotional skills (Almlund et al., 2011). Second, we study the rate at which these skills are shaped over the formative years of compulsory education. In this regard, we investigate the malleability of skills developed between the first and last years of compulsory schooling in order to determine how shocks to the skill formation process (e.g. having a better teacher in later schooling years) can help to overcome initial skill disadvantages. Finally, we analyze whether differences in math and verbal skills between males and females explain the well established gender gaps in college enrollment

[1] For example, if attending university for the marginal student relies more heavily on the level of a given type of skill, then school curricula could be adjusted to deliver those levels of skills.

[2] In fact, Castex and Dechter (2014) show that the direct labor market return to skills have experienced a sharp decline since the 1980's, where the returns to math and verbal skills have fallen by more than 50%. Over this same period, they show that the returns to formal education have increased, which suggests that the channel through which skills influence educational attainment may be the dominant role through which skills impact labor market outcomes.

and in STEM (Science, Technology, Engineering, and Math) fields of study. Evidence from many developed countries has shown that boys are less likely to attend university than girls but they are more likely to enroll in STEM fields. However, little is known regarding the role of skills in explaining these gaps.

We study these questions using a large administrative dataset that covers nearly an entire cohort of students in England.[3] This dataset contains the complete history of educational outcomes of more than 500,000 students from age 5 to 22. For each student we observe a rich set of demographic characteristics, the elementary school attended, as well as university records that show whether the student progresses to university, their institution and field of study, and whether they graduated. Importantly, for each student we observe the results from more than 30 subject specific exams taken over compulsory education.[4] We use these 30 performance measures to estimate a multi-period factor model which precisely recovers the latent skills (i.e. math and verbal) for each individual at different points in the schooling career. After characterizing these skills, we study how these skills influence educational outcomes.

Our results indicate that verbal skills have a substantially stronger influence on university enrollment and graduation than math skills. Specifically, while a one standard deviation increase in math skills improves the probability of university enrollment by 9.6 percentage points, a similar increase in verbal skills leads to an improvement of 18.7 percentage points. This relative importance of verbal skills is robust to different model specifications, databases, and additional checks.[5] We also find that about 40% of the variance in skills at age 16 is determined after age 7, suggesting that there is scope during compulsory education for overcoming initial skill disadvantages. Results also indicate that females have a large advantage in verbal skills relative to males (i.e. around 38% of a standard deviation) while differences in math skills are negligible by the last year of compulsory education. These gender differences in verbal skills constitute the main source of the gender gap in

[3] Our data contains all students in public schools in England. Data from private schools is not available, however this sector is very small in the United Kingdom (approximately 6.5% of all students).

[4] Our data contain student performance on 70 different tests, of which only a subset are required.

[5] First, we assess whether our verbal skills could be capturing other determinants of college enrollment, for example behavioral characteristics of the student like socio-emotional skills or family background characteristics like parental education. Second, we estimate our model using only a sample of white students. Finally, we perform a regression analysis using high school transcript data from the National Survey of Youth of 1997 (NLSY97) to explore cross-country evidence of our main results. Our findings are robust to all of these robustness specifications.

college enrollment. After controlling for math and verbal skills, we show that girls are slightly *less* likely to attend college than boys (i.e. the gap is reversed), suggesting that females "preferences" for educational attainment contribute in closing the gap rather than increasing it. Finally, we present evidence indicating that comparative advantage in skills plays a more relevant role for males in their decision to enroll in STEM fields than for females.

We believe that our findings further contribute to the literature that studies the impact of (high) school curriculum on long-term outcomes. While previous work (Altonji, 1995; Levine and Zimmerman, 1995; Rose and Betts, 2004; Joensen and Nielsen, 2009; Dougherty et al., 2015) and policymakers have (mainly) focused their attention on the role of high school math on various outcomes, this paper brings to the center of the analysis the importance of verbal skills (acquired during compulsory education) to predict university outcomes.[6] In this regard, our results are somewhat consistent with the findings of Cortes et al. (2015), which shows that the overall effect on college enrollment of a policy that increases rigor in high school algebra is mainly due to its impact on below-average readers. This unexpected pattern may reflect, according to the authors, the intervention's focus on reading and writing skills in the context of learning algebra.

This paper also contributes from a methodological perspective by offering a tractable method to estimate a high-dimensional, correlated factor model (i.e. nine factors in total) that contains more than 140,000 parameters. We demonstrate how the model can be easily estimated with a simple expectation and maximization algorithm (Dempster et al., 1977; Ruud, 1991). For our context, this is implemented by iterative single equation least squares estimation. This approach is ideal given our big data setting, which contains close to 15 million data points (30 measures for each 500,000 individuals).

The rest of the document is organized as follows. Section 2 describes the educational institutional setting in England and the data. Section 3 shows preliminary evidence on the main data patterns. Section 4 describes the empirical strategy. Section 5 discusses the estimation techniques. Section 6 shows the main estimation results. Section 7 studies possible alternative explanations to our findings. Section 8 analyzes the development of skills during the schooling years. Section 9

[6] See Dougherty et al. (2015) for a review of policies that aimed to increase math coursework in high school.

studies the gender gaps in college enrollment and field of study. Section 10 concludes.

2 Institutional Setting and Data

2.1 The English School System

Compulsory education in England is organized in four Key Stages (KS). Each stage ends with nationally assessed standardized tests, in addition to teacher assessments on different subjects.[7] Table 1 summarizes the English compulsory education system. Students enter school at age 4, the Foundation Stage, then proceed to Key Stage 1 (KS1), spanning ages 5 and 6, and Key Stage 2 (KS2, involving ages 7 to 11).[8] At the end of KS2 children move to secondary school, where they progress to Key Stage 3 (KS3, ages 12-14) and Key Stage 4 (KS4, ages 15-16). At KS4, students begin tailoring their curriculum by specializing in around 6 subjects. At age 16, compulsory education ends. Next, students decide whether to continue their studies for two more years, called Key Stage 5 (ages 17-18), where they choose either a vocational or academic curriculum, which typically concludes with qualifying exams (the A-levels). Most students study three or four A-level subjects simultaneously during Year 12 and Year 13, either in a secondary education institution or in a Sixth Form College, as part of their further education. Finally, higher education usually begins at age 19 with a three-year bachelor's degree, where admissions to university are mainly determined based on A-levels performance.

2.2 Data

Our analysis uses individual-level administrative panel data for one cohort of students who completed their compulsory education in the academic year 2006/07. The final dataset contains information on approximately 500,000 students, which is all students in England at that grade level except those in independent schools for whom the census is not available.[9] Our database links information from the census of all state school children in England with information from the Higher

[7]Recently, a series of reforms regarding the assessment of students have been implemented. However, these reforms were not in place for the years that we are analyzing.

[8]KS1 is equivalent to grades 1 and 2 in the US school system, KS2 to grades 3, 4 and 5.

[9]The independent sector educates around 6.5% of the total number of school children in the UK.

Table 1: Key Stages in English Education System.

	Age	Years	Test
Key Stage 1	5-7	1 and 2	National Programme of Assessment at the end of year 2 in Maths, English and Science, carried out by the teacher, and teacher assessment.
Key Stage 2	8-11	3-6	National Programme of Assessment at the end of year 6 in Maths, English and Science Additionally teacher assessment is provided.
Key Stage 3	12-14	7-9	National Programme of Assessment at the end of year 9 in Maths, English and Science. Additionally teacher assessment is provided.
Key Stage 4	15-16	10 and 11	General Certificate of Secondary Education (GCSE) are generally taken at the end of year 11. End of compulsory education
Key Stage 5 (A-level)	17-18	12 and 13	General Certificate of Education (GCE) Advanced Level, or A Level, is a secondary school leaving qualification in the United Kingdom, offered as a main qualification.

Education Statistics Agency (HESA) and from the Individualised Learner Records (ILR).[10] HESA and ILR collect information on all students in public founded universities and all learners in public founded Further Education (FE) programmes, respectively. The dataset allows us to track pupils throughout their entire education path. It contains detailed information on student performance in all tests, exams and teacher assessments during both compulsory education and later levels of education, and school absences (i.e. excused and unexcused). Furthermore, it includes data on several pupil characteristics, e.g. special education needs (SEN), free school meal eligibility (FSM), mother tongue, ethnic group, and schools attended. Overall, our database allows us to observe around 30 test score outcomes for each student during compulsory education, though not all students take the same subjects in KS4 (we have test score outcomes on 70 different subjects in total).[11] Finally, it also provides information on A-levels performance and on the specific major and university

[10]The final census entails data from the National Pupil Database (NPD) and the Pupil Level Annual School Census (PLASC), that has been replaced in 2007 by the School Census.

[11]We observe approximately 8 test scores per student at each key stage. More specifically, 4 scores in math and 4 in verbal at each key stage between KS1 and KS3, and around 8 GCSE exam outcomes in KS4.

attended, and graduation outcomes.

Table 2 presents summary statistics of selected key variables in our data. The top panel shows information on student background characteristics. The Income Deprivation Affecting Children Index (IDACI) is an index of poverty used in the United Kingdom. It is calculated by the Office of the Deputy Prime Minister and measures in a local area the proportion of children under the age of 16 that live in low income households. The higher the score the more impoverished is the area. The data also identify students who meet certain eligibility requirements for free school meals (FSM). According to Hobbs and Vignoles (2007), FSM status proxies children in households with family incomes below £200 per week. Finally, the special education needs (SEN) variable indicates whether a child has learning difficulties or disabilities that make it harder for him or her to learn than most other children and young people of about the same age. Overall, the data show that 15% of the students in our sample are eligible for FSM, where differences between genders are small. On the contrary, the indicator for special education needs shows that only 16.9% of female students are included is this category while 27.9% of boys require special education. Finally, this panel also shows that 94.7% of the students in our sample have a mother who speaks English, and more than 89% of the students are white. The second panel of the table shows overall performance in national assessment exams in math and english at each stage of the schooling career.[12] The data show that females largely outperform boys in language at each KS, while in math males seem to perform better until KS3. Finally, the last row of this panel indicates that students take 8.17 GCSE exams on average (notice that two of them, math and verbal, are compulsory), where females tend to take more exams than males. The third panel shows average absence rates (i.e. average proportion of sessions absent) in KS4 by authorized and unauthorized type. Students are absent in 9.4% of the sessions, where most of these absences are authorized, and differences between genders are small. Finally, the bottom panel provides an overview of post-secondary education outcomes. Around 36.3% (25.3%) of the students in our sample enrolled (graduated on time) in university, where approximately 54% (57%) of those enrolled (graduated on time) are females. Looking at enrollment in STEM fields, the gender gap is reversed, with females only accounting for 40% of

[12]Test scores have been standardized to have mean zero and standard deviation 1. We have information on more test scores in KS4, Table 2 only presents performance in math and verbal.

the total number of students enrolled in these fields. Finally, the variable "University Enrollment Top 24" denotes the proportion of students attending the most selective institutions in the United Kingdom (i.e. Russell group).[13] Approximately 7% of students continue their studies in these institutions, with females being overrepresented.

3 Preliminary Evidence

Before moving forward, this section characterizes the key features of the data that we wish to understand more deeply with our analysis. First, Table 3 shows the results of a linear probability model of university enrollment, controlling for each of the Key Stage test scores. These results offer three main insights. First, test scores are highly predictive of university enrollment. The regression including the Key Stage 4 test explains almost 36% of the college enrollment decision, as measured by the R-squared. Second, while tests are highly predictive of university enrollment, verbal scores appear to have a larger effect relative to math scores. The magnitude of the difference varies across the Key Stages. The difference is most pronounced in Key Stage 2, where the coefficient on the verbal score is 39% larger than the coefficient on math. Finally, Table 3 shows that including test scores in these regressions has a large impact on the coefficient for the female indicator. This suggests that skill differences between genders may play an important role in explaining the gender gap in many educational outcomes.

Table 3 provides tangential evidence on the relationship between skills and education decisions. However, there is a number of shortcoming with this analysis. First, the use of test scores in these regressions only proxies for skills, which does not directly address our research question. It is not clear *a priori* which of the observed 70 measurements should be used to proxy for which skills.[14] Given the high correlation among these scores, a regression that includes all of them would be

[13] The Russell Group represents 24 leading UK universities: University of Birmingham, University of Bristol, University of Cambridge, Cardiff University, Durham University, University of Edinburgh, University of Exeter, University of Glasgow, Imperial College London, King's College London, University of Leeds, University of Liverpool, London School of Economics & Political Science, University of Manchester, Newcastle University, University of Nottingham, University of Oxford, Queen Mary University of London, Queen's University Belfast, University of Sheffield, University of Southampton, University College London, University of Warwick, University of York.

[14] We observe around 30 measurements per student, but not all students take the same subject. The total number of scores is around 70.

Table 2: Summary Statistics. Overall and by Gender

	All		Males		Females		All	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Min.	Max.
Background Characteristics								
IDACI Index	0.210	(0.174)	0.207	(0.173)	0.213	(0.176)	0.003	0.993
Free School Meal	0.150	(0.357)	0.146	(0.353)	0.154	(0.361)	0	1
Special Education Needs	0.224	(0.417)	0.279	(0.448)	0.169	(0.374)	0	1
Mother Tongue English	0.947	(0.224)	0.947	(0.223)	0.947	(0.226)	0	1
Race								
White	89.46%		89.61%		89.30%			
Asian	5.74%		5.69%		5.78%			
Black	2.41%		2.35%		2.47%			
Other	2.40%		2.34%		2.46%			
Standardized Key Stage (KS) Test Scores								
KS1 Math	0	(1.000)	0.014	(1.036)	-0.014	(0.962)	-3.234	3.089
KS1 Verbal	0	(1.000)	-0.148	(1.020)	0.148	(0.956)	-3.017	3.604
KS2 Math	0	(1.000)	0.057	(1.014)	-0.057	(0.982)	-3.194	1.802
KS2 Verbal	0	(1.000)	-0.108	(1.002)	0.108	(0.986)	-4.000	2.692
KS3 Math	0	(1.000)	0.035	(0.999)	-0.036	(0.999)	-3.204	2.148
KS3 Verbal	0	(1.000)	-0.176	(1.001)	0.177	(0.967)	-2.104	2.706
KS4 Math	0	(1.000)	-0.003	(1.003)	0.003	(0.997)	-2.582	1.900
KS4 Verbal	0	(1.000)	-0.175	(1.017)	0.176	(0.950)	-3.057	1.998
Total GCSE Exams	8.170	(2.070)	8.009	(2.157)	8.332	(1.963)	2	18
Average Proportion of Sessions Absent in KS4								
Authorized Absences	0.074	(0.073)	0.069	(0.069)	0.079	(0.077)	0	1
Unauthorized Absences	0.020	(0.059)	0.019	(0.056)	0.021	(0.072)	0	1
University Outcomes								
University Enrollment	0.363	(0.481)	0.330	(0.470)	0.400	(0.488)	0	1
University Graduation	0.253	(0.435)	0.219	(0.413)	0.287	(0.452)	0	1
University Enrollment STEM	0.118	(0.323)	0.140	(0.347)	0.096	(0.294)	0	1
University Graduation STEM	0.075	(0.263)	0.083	(0.275)	0.067	(0.250)	0	1
University Enrollment Top 24	0.071	(0.257)	0.066	(0.248)	0.077	(0.266)	0	1

Table 3: Linear Probability Model: University Enrollment

	Baseline	KS1	KS2	KS3	KS4
Constant	0.330	0.343	0.346	0.360	0.359
Female	0.066	0.039	0.046	0.020	0.007
Math Test Score	-	0.094	0.101	0.127	0.141
Verbal Test Score	-	0.098	0.140	0.159	0.165
R-Squared	0.005	0.136	0.214	0.295	0.357
Obs.	498,736	495,197	480,157	471,095	498,736

Notes- Math and verbal test scores have been standardized to have mean zero and standard deviation 1. Standard errors not shown. All coefficients are statistically significant at the 99% level.

difficult to understand, with many of the coefficients having the wrong sign due to multicollinearity. Second, any method to weight the scores to form aggregates would be arbitrary, with no method to guide which weighting scheme best captures skills. Third, any averaging of test scores will not fully correct for measurement error, which will lead to bias. Finally, this analysis provides no insight into the sources of variation in skills and how skills evolve over the schooling career. Simply studying correlations in test scores across time periods suffers from each of the issues described above. The next section outlines the factor model that we use to address these issues. Factor analysis is a statistical method that condenses the covariance among available measures using low dimensional latent variables, which produces output that is easy to interpret and directly addresses the problem of measurement error. This approach is ideal for our investigation because it allows us to directly measure skills and enables a comprehensive use of our rich data source.

4 Methodology

In our data we observe over 70 measures of student performance from age 5 to 17. Our strategy aims to estimate a multi-period factor model that reduces these measures to a low-dimensional vector of skills that is easy to interpret.[15] After recovering the latent skills, we study the three main questions mentioned earlier. First, what influence do math and verbal skills have on decisions in higher education? Second, how do these skills evolve over the course of compulsory education? Third, what is the role of these skills in explaining the gender gaps in university enrollment and STEM fields?

Factor Model of Skills Time is indexed by t where $t \in \{1, 2, 3, 4\}$, represents each of the four key stages of compulsory education. At time period t, skills for student i are denoted by θ_{it}. We model a total of nine skills, which includes math and verbal skills for each of the four Key Stages and an additional factor in Key Stage 4 that captures the motivation of the student to attend higher education. Let $\theta_i = \begin{bmatrix}\theta'_{i1} & \theta'_{i2} & \theta'_{i3} & \theta'_{i4}\end{bmatrix}'$ be the complete vector of these nine factors, which are modeled jointly as:

$$\theta_i = female_i * \psi + \Phi x_i + \xi_i \tag{1}$$

The factors are determined by characteristics observed by the econometrician (gender and x_i), and an unobserved component ξ_i. We assume that the unobserved component of skills is drawn from a mixture of C multivariate normal distributions.[16] That is, ξ_i is drawn from $\mathcal{N}(\delta_c, \Sigma)$ with probability π_c for $c = [1, 2, \ldots, C]$.[17] Each component of the mixture has a different mean but shares a common covariance. Skills are allowed to be correlated contemporaneously and across time, so Σ is a full covariance matrix. We model the nine factors in Eq. (1) as a joint distribution because this offers more concise matrix notation.[18]

[15] Our approach follows the spirit of Heckman et al. (2013). However, we use different econometric methods than Heckman et al. (2013) due to the high dimensionality of our data (i.e. around half million observations).

[16] The mixture of normals allows to approximate many distributions, for example, the distributions can be skewed.

[17] Note that δ_1 is normalized to zero for identification purposes.

[18] The joint distribution can be re-written in terms of conditional distributions, which maps our model directly into a dynamic factor model (e.g. Cunha and Heckman (2007)) [see Appendix A]. However, given that we are not aiming to estimate the technology of skills formation, but rather providing a characterization of how skills evolve over time during compulsory education (among other things), it is unnecessary to re-express our model as a dynamic factor model.

Measurement Equations The factors produced from Eq. (1) are not observed. However, our data contain frequent and extensive measurements for each student that we use to recover these latent skills (i.e. θ_i). We observe a total of M measures for each student. Each of the measures, denoted w_{im}, is characterized by:

$$w_{im} = \mu_m + \theta_i' \lambda_m + \eta_{im} \tag{2}$$

μ_m is the mean of the m^{th} measure and λ_m are the loadings on the factors for measurement m. Our factors span multiple time periods, so only the factors associated with the time period of measurement will have non-zero loadings. Finally, η is the remaining portion of the measurement that is not explained by the factors and is assumed to be independent and normally distributed with mean zero and variance $\text{Var}(\eta_m)$.

The measures used in the analysis combine multiple nationally assessed tests as well as teacher evaluations for a broad range of subjects. A detailed description of the measurements as well as the normalizations made to identify the factor model are given in Appendix B.

Outcome Equations Key Stage 4 is the final period of compulsory schooling. At this stage, each individual is characterized by a vector of skills θ_{i4}, which contains three elements: a math skill, verbal skill, and a measure of motivation for pursuing further education. After this period, students may conclude their formal education or make additional investments. We are interested in understanding how these measures of skill influence these decisions. There are K outcomes for each individual, with realization y_{ik}^* where $k = 1, \ldots, K$ that follows:

$$y_{ik}^* = female_i * \gamma_k + \theta_{i4}' \alpha_k + x_i' \beta_k + \varepsilon_{ik} \tag{3}$$

γ_k is the influence of gender on the observed outcome (x_i includes a constant). θ_{i4} is the vector of skills at the end of compulsory schooling, x_i includes other control variables, and ε_{ik} includes the remaining determinants of the outcome variable that cannot be explained by the other parts of the model and is assumed to be independent. The outcomes we wish to study in this paper are university enrollment, university quality, major field of study, and university graduation. All of these outcomes are discrete, so y^*_{ik} represents the underlying latent variable process. We assume that ε is distributed type-I extreme value, and we only observe the outcome $y_{ik} = 1$ if $y^*_{ik} > 0$ and zero otherwise.

Endogeneity To identify the causal impact of skills on the outcome equations we need to address certain aspects of endogeneity between the measures used to extract the skills and the outcome equations themselves. Specifically, we need to be certain that ξ_i and ε_{ik} are not correlated. The first potential source of endogeneity is parental inputs. Parental inputs are crucial for skill development and influential in the college decision. Without shutting down this channel we risk confounding our estimates. We follow two strategies to deal with this issue. The first is to include observed variables in the explained portion of the factors that can account for heterogeneity in parental inputs. To this extent, we include the IDACI score of the neighborhood in which the student resides, an indicator if the student qualifies for a free or reduced price lunch, race, the native language of the student's mother, and a full set of elementary school fixed effects. Since we use a single cohort of students, each elementary school has on average 30 students, which gives us more than 15,000 fixed effects for each of the factors (a total of 135,000 fixed effects). The second strategy we use to address the endogeneity of parental inputs and college enrollment exploits the multi-period modeling of skills. Assuming that performance on earlier tests will proxy for parental input, we are able to use marginal changes in skill development in later periods to identify the effect of specific skills on the outcomes.

The second concern of endogeneity occurs later in the academic career. Many students may have made the decision not to go to college prior to the completion of compulsory schooling. These students in turn may put in low effort and have low performance on some of the measures. To address this concern, we include a third factor in the final year of school to capture motivation. While motivation is a difficult characteristic to capture, we identify this factor from three sources. The first two sources are the total number of excused and unexcused absences during Key Stage 4. The third source is the number of subject specific tests taken during Key Stage 4. Since students are required to take at least six subject specific test during Key Stage 4 before advancing to Key

Stage 5, we use this information to identify their intent for higher education. By including this third factor in Key Stage 4, our goal is to identify the causal impact of the other skills on college outcomes, holding motivation constant.

5 Estimation

The factor structure condenses the information from multiple measures into a more manageable and interpretable set of variables. This section discusses the estimation of the model. Given the distributional assumptions on all of the unobserved data, the parameters can be estimated with maximum likelihood. We use a two stage estimation approach similar to Heckman et al. (2013). In the first stage we use all of the observed measurements to jointly estimate the parameters of the measurement system and the factor structure. In the second stage, we use the parameter estimates to construct distributions for latent skills for each individual and estimate the coefficients of the outcome equations, integrating over these distributions. This process is formalized below.

Given data for N individuals, let $w_i = [w_{i1}, \ldots, w_{iM}]$ be the observed measurements and $y_i = [y_{i1}, \ldots, y_{iK}]$ be the observed outcomes for individual i. To form the likelihood, we write $L(y_i|x_i, \theta)$, the likelihood of observing the outcome y_i for a given value of the unobserved skills, θ. Conditional on θ, the outcome equations are independent, so $L(y_i|x_i, \theta)$ is a product of logit probabilities. In addition, we compute $L(w_i|\theta)$, the likelihood of observing the measurement variables conditional on θ, which is a product of univariate normal probability density functions. Conditional on θ, the remaining unobserved components of the measurements and the outcomes are independent, so the joint probability is the product of these two likelihoods. Therefore, the parameters are estimated by maximizing the integrated likelihood function:

$$LL = \sum_{i=1}^{N} \ln \left[\int_{\theta} L(y_i|x_i, \theta) L(w_i|\theta) f(\theta|x_i) d\theta \right] \tag{4}$$

Where $f(\theta|x_i)$ is the probability density function for a mixture of normals specified in Eq. (1).

Next we apply Bayes' rule to the likelihood to estimate the parameters in the measurement system separately from the parameters in the outcome equation. Let $h(\theta|w_i, x_i)$ be the probability density function of θ conditional on both the measurements and the covariates. From Bayes rule we have the identity

$$L(w_i|\theta)f(\theta|x_i) = h(\theta|w_i, x_i)\left[\int_{\theta'} L(w_i|\theta')f(\theta'|x_i)d\theta'\right] \tag{5}$$

Plugging this into the likelihood, Eq. (4) can be written as,

$$LL = \sum_{i=1}^{N} \ln\left[\int_{\theta} L(y_i|x_i, \theta)h(\theta|w_i, x_i)d\theta\right] + \sum_{i=1}^{N} \ln\left[\int_{\theta'} L(w_i|\theta')f(\theta'|x_i)d\theta'\right] \tag{6}$$

The additive separability of this likelihood allows us to estimate the parameters in two stages. First we maximize the last component of the likelihood, which contains the parameters in the measurement equations and the factor distribution. Second, using these estimates we construct the conditional distributions $h(\theta|w_i, x_i)$ and maximize the portion of the likelihood containing the outcome equations. As stated by Heckman et al. (2013), the two step approach is less efficient than joint estimation of the entire model because the information from the outcome equations is not used to help identify the factor distribution parameters. However, in addition to being more tractable, this approach is beneficial because it provides transparency in how these parameters are identified.

In the first step, we search for the parameters that maximize the likelihood function

$$LL_{step-one} = \sum_{i=1}^{N} \ln\left[\int_{\theta} L(w_i|\theta)f(\theta|x_i)d\theta\right] \tag{7}$$

One potential concern with this estimation approach in our setting is that we do not observe all measures for all students. Since we do not directly model the selection process, we need to make additional assumptions to proceed. We observe nearly complete coverage for the measures used in Key Stages 1 to 3. For Key Stage 4, all students take a mandatory math test and english test. In addition, at Key Stage 4, they choose up to four additional subject tests. We use the score on these subject tests to recover the factors. This creates a potential selection problem since those with high math ability will likely choose a different set of subjects than those with lower math ability.

Our assumption is that, once we condition on the observed scores for all of the mandatory tests, including the math and english test at Key Stage 4, the choice of subject test occurs approximately at random.

Given the dimensionality and complexity of this problem, finding the parameters that maximize the likelihood in Eq. (7) poses a number of challenges. First, we use data on more than 15,000 elementary schools. Since we include elementary school fixed effects for each of the nine factors, the model contains more than 135,000 fixed effects. In total, our baseline specification includes around 140,000 parameters. Conventional numerical optimizers based on Newton's method are not possible because the hessian matrix cannot be stored in read/write memory. Second, even if a suitable large scale algorithm is found, the factor model contains many constraints that are difficult to impose during estimation. For example we require the nine dimensional covariance matrix to be positive semi-definite.

We overcome these computational challenges by maximizing Eq. (7) with the expectation and maximization (EM) algorithm (Dempster et al., 1977). The EM algorithm is an iterative procedure for maximizing complicated integrated likelihood functions. In the expectation step, the current iteration parameters are used to construct individual densities of the unobserved data. In the maximization step, new parameters are found by maximizing an augmented data likelihood that treats the unobserved data as observed, integrating over these densities. The appeal of this method is that it is easy to implement. For example, the parameters in Eq. (1) when θ_i is observed can be found using equation by equation OLS. This only requires the inversion of a 15,000 by 15,000 matrix of fixed effects, which is the same for each regression and can be calculated and inverted outside of the algorithm. Second, many of the constraints, like positive semi-definiteness of the covariance matrix, are naturally imposed by the algorithm. One shortcoming of the EM algorithm is that it has, in some cases, a slow rate of convergence. To speed up the convergence of the algorithm we use the SQUAREM accelerator in Varadhan and Roland (2008), which can deliver super-linear rates of convergence.

Once the first stage parameters are estimated, we calculate the conditional distribution of the latent skills for each individual, $h(\theta|w_i, x_i)$ and maximize the likelihood.

$$LL_{step-two} = \sum_{i=1}^{N} \ln \left[\int_{\theta} L(y_i|x_i, \theta) h(\theta|w_i, x_i) d\theta \right] \tag{8}$$

Since the outcome equations are all discrete, this integrated likelihood function represents a mixed logit model. The maximization of this likelihood is only over the parameters in $L(y_i|\cdot)$, denoted $\Psi = [\gamma, \alpha, \beta]$. The parameters that maximize this likelihood are the ones that are a root to the score function.

$$\begin{aligned}
\frac{\partial LL_{step-two}}{\partial \Psi} &= \sum_{i=1}^{N} \frac{1}{\int_{\theta'} L(y_i|x_i, \theta') h(\theta'|w_i, x_i) d\theta'} \int_{\theta} \frac{\partial L(y_i|x_i, \theta)}{\partial \Psi} h(\theta|w_i, x_i) d\theta && (9) \\
&= \sum_{i=1}^{N} \int_{\theta} \frac{\partial \ln [L(y_i|x_i, \theta)]}{\partial \Psi} \frac{L(y_i|x_i, \theta) h(\theta|w_i, x_i)}{\int_{\theta'} L(y_i|x_i, \theta') h(\theta'|w_i, x_i) d\theta'} d\theta && (10) \\
&= \sum_{i=1}^{N} \int_{\theta} \frac{\partial \ln [L(y_i|x_i, \theta)]}{\partial \Psi} h(\theta|y_i, w_i, x_i) d\theta && (11)
\end{aligned}$$

The density function $h(\theta|y_i, w_i, x_i)$ is the conditional density of the unobserved factor conditional on all of the data. In this second stage, we impose the restriction on the data that the outcome equations provide no more additional information on the factors once we condition on the observed measurements, i.e. $h(\theta|y_i, w_i, x_i) = h(\theta|w_i, x_i)$. The primary purpose of this assumption is that it facilitates validation of the model. Our goal is to understand how much variation of the outcome variables can be explained by these skills. If we were to relax this assumption, maximum likelihood would choose parameters that explain as much of the variation as possible. By imposing this restriction we risk having less explanatory power. However, as described previously, this approach provides more transparency over the identification of the parameters. A second benefit of this assumption is that the likelihood reduces to an integrated standard logit, which requires maximizing,

$$LL_{step-two} = \sum_{i=1}^{N} \int_{\theta} \ln [L(y_{ik}|x_i, \theta)] \, h(\theta|w_i, x_i) d\theta \tag{12}$$

Because the integral is outside of the log function, maximizing this likelihood can be done equation by equation. There is no known closed form for the integral in this likelihood, so we approximate it with 10 simulated draws.

6 Main Results

The estimates from the factor model facilitate a deeper investigation of the patterns observed in the simple analysis in Section 3. The main benefit of the factor structure is that it allows us to extract measurements of unobserved skills from a large set of data, which can then be used to study the outcome equations. This section summarizes our main results. First, we characterize the distribution of the recovered skills in the population. Second, we report the factor loadings on the measurement equations and the residual variance, in order to exemplify how the factor structure allows us to extract and condense the information contained in the 70 measures used in the analysis. Finally, we analyze the role that these skills have on affecting educational decisions.

6.1 Factor Distributions

Our empirical strategy recovers nine factors (i.e. math and verbal factors for each of the four key stages, and one motivation factor just for KS4), that are mutually correlated.[19] Table 4 shows the full correlation matrix and the standard deviation of each factor. As expected, we observe that the factors are highly correlated over time and across skills. For example, the correlation between contemporaneous math and verbal skills is around 0.8 at each key stage. However, the correlation among factors that are more distant in time show a declining trend. For example, the coefficient of correlation between KS1 and KS2 math (verbal) factors is 0.742 (0.818) while between KS1 and KS4 math is around 0.558 (0.641). In addition, notice that the correlation between KS3 and KS4 math (verbal) factors is 0.927 (0.907), suggesting that skills in KS4 are mostly established in previous stages. To conclude the description of factor correlations, the motivation factor (that aims to capture aspirations for pursuing further education), measured in the last stage of the compulsory schooling career, shows a similar correlation with the verbal and math factors. Finally, the last row of Table 4 displays the standard deviation of the factors, where conditional on key stage math and

[19] The motivation factor aims to capture schooling aspirations of the students.

verbal skills show similar variance.

In order to explore how skill distributions in KS4 differ across demographic characteristics, Figure 1 shows kernel densities of each factor by gender. While both groups show similar distributions in math skills, differences in verbal skills are substantial in favor of women. More specifically, Table 5 shows that the difference in verbal skills in KS4 between males and females represents 38% of a standard deviation. Similarly, girls seem to outperform boys in motivation, though the difference is much smaller (8% of standard deviation). Table 5 also shows the evolution of the gender differences in skills across key stages. Even though boys perform better than females in math at early stages, differences follow an (inverse) u-shape with larger male advantage in KS2 and KS3 than in KS1 and KS4, where disparities are negligible. On the contrary, differences in verbal skills are always larger in favor of females. Finally, the estimation of the factor distributions points out that males show (in most periods) larger dispersion in skills than females, particularly in verbal skills. Overall, these results highlight large differences in verbal skills between males and females, suggesting that the gender gap in college enrollment that has been described in subsection 2.2 could be explained by the differences in verbal skills. Section 9 investigates this issue in more detail.

6.2 Factor Loadings

The factor model described in Section 4 makes use of the multiple KS test scores to characterize the latent skills. Table 6 displays the loadings on the math and verbal factors and the residual variance (i.e. variance of the "uniqueness") of a subset of KS4 measurements.[20] By looking at these estimates, it is possible to assess how skills load on each of the different test scores, and to analyze the importance of measurement error. For example, Table 6 shows (as expected) that statistics mainly loads on math skills while social science relies heavily on verbal skills. However, courses such as geography and design and technology load similarly on both skills. Finally, the last column of Table 6 provides the variance of the uniqueness. Given that each measure has variance 1, then the reported residual variance denotes the proportion of the total variance that can be interpreted as "noise". Our estimates indicate that measurement error is pervasive on some of our measures.

[20]Table 15 shows the complete list of loading estimates for each of the 70 measurements.

Table 4: Factor Correlation Matrix

	KS1 Math	KS1 Verbal	KS2 Math	KS2 Verbal	KS3 Math	KS3 Verbal	KS4 Math	KS4 Verbal	KS4 Motivation
KS1 Math	1.0	-	-	-	-	-	-	-	-
KS1 Verbal	0.826	1.0	-	-	-	-	-	-	-
KS2 Math	0.742	0.673	1.0	-	-	-	-	-	-
KS2 Verbal	0.712	0.818	0.826	1.0	-	-	-	-	-
KS3 Math	0.736	0.682	0.939	0.826	1.0	-	-	-	-
KS3 Verbal	0.633	0.733	0.725	0.907	0.822	1.0	-	-	-
KS4 Math	0.623	0.603	0.795	0.750	0.927	0.810	1.0	-	-
KS4 Verbal	0.558	0.641	0.659	0.803	0.787	0.907	0.890	1.0	-
KS4 Motive	0.455	0.510	0.555	0.626	0.693	0.722	0.820	0.844	1.0
Std. Dev.	0.872	0.902	0.939	0.888	0.762	0.832	0.922	0.922	1.410

Note- All correlations are statistically significant at the 1% level

Figure 1: Distribution of Key Stage 4 Factors by Gender

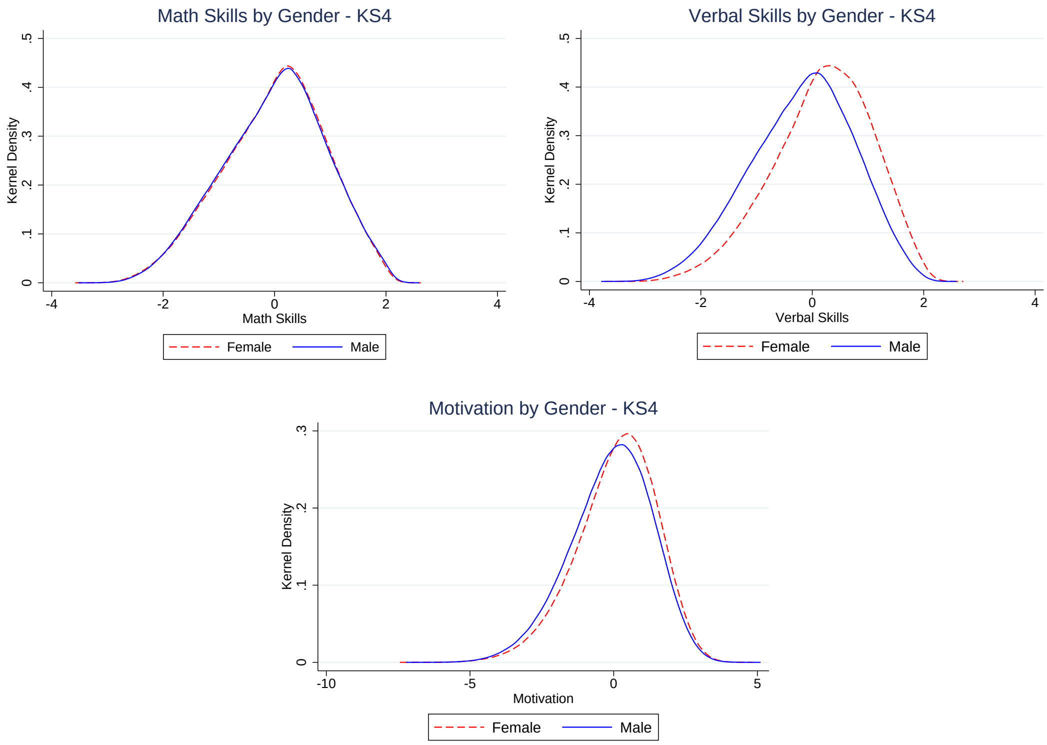


Table 5: Factor Means and Standard Deviations by Gender KS1-KS4

	Male		Female		Difference
	Mean	S.D.	Mean	S.D.	
Math					
KS1	0.008	0.916	-0.008	0.847	0.016
KS2	0.051	0.979	-0.051	0.947	0.102
KS3	0.017	0.785	-0.017	0.770	0.034
KS4	-0.002	0.937	0.002	0.937	-0.004
Verbal					
KS1	-0.137	0.930	0.137	0.873	-0.274
KS2	-0.120	0.916	0.121	0.875	-0.241
KS3	-0.163	0.844	0.164	0.808	-0.327
KS4	-0.179	0.932	0.179	0.896	-0.358
Motivation					
KS4	-0.100	1.431	0.100	1.400	-0.200

Note-Differences in means are statistically significant at the 1% level with the only exception being KS4 math.

For example, noise is substantially large in design and technology (51.5%), while this is not the case in geography (21.9%). To summarize, Table 6 highlights the advantages of using a factor model in the context of the English educational system because it avoids creating arbitrary indexes of math and verbal skills, and tackles problems of measurement error.[21]

Table 6: Factor Loadings: Selected Measurements only loading in Key Stage 4

Measurement	KS4 Math	KS4 Verbal	Residual Variance
Math	$1^{\dagger}$	0	0.127 (0.001)
English	0	$1^{\dagger}$	0.132 (0.001)
Design and Technology: Resistant Materials Technology	0.396 (0.007)	0.429 (0.007)	0.515 (0.004)
Geography	0.508 (0.006)	0.593 (0.006)	0.219 (0.001)
Social Science	0.088 (0.012)	0.864 (0.010)	0.369 (0.003)
Statistics	1.047 (0.010)	0.067 (0.008)	0.246 (0.004)
Physics	1.568 (0.007)	0	0.194 (0.003)
Chemistry	1.548 (0.007)	0	0.198 (0.004)
English Literature	0	1.045 (0.001)	0.220 (0.001)
Home Economics: Child Development	0.240 (0.021)	0.799 (0.021)	0.371 (0.005)
Double Science	1.056 (0.002)	0	0.191 (0.001)

Note- Residual variance denotes the variance of the noise of each measurement (i.e. $var(\eta_{im})$). Given that each measurement has been standardized to have mean 0 and standard deviation to 1, the residual variance can be interpreted as the proportion of the total variance that constitutes "measurement error". Table 15 shows the complete list of loading estimates for each of the 70 measurements.
† denotes normalized to 1.

[21] Table 16 reports the coefficients corresponding to Eq. (1) (i.e. skill coeficients).

Table 7: Logistic Regression: University Enrollment and Key Stage 4 Skills

	(1)	(2)	(3)	(4)
Average Marginal Effects				
KS4 Math	0.273 (0.001)	–	–	0.096 (0.001)
KS4 Verbal	–	0.298 (0.001)	–	0.187 (0.001)
KS4 Motive	–	–	0.301 (0.001)	0.029 (0.001)

Note- Results are from a logistic regression with controls for gender, IDACI Index, free school lunch, special education needs, race, mother tongue, and school fixed effects. Skills have been standardized to have mean 0 and standard deviation 1. Bootstrapped standard errors at school level.

6.3 The Role of Skills in Education Decisions

This section studies how skills affect the probability of attending and graduating from university. Table 7 shows the marginal effects of a logistic model on university enrollment using only the skills recovered in Key Stage 4. We focus on KS4 skills because this is the period when compulsory school finishes, and students have the ability to make their own educational decisions. The regression analysis also includes additional controls for background characteristics and school fixed effects. The first three columns of Table 7 show the effect of each skill separately, while the last column considers all of the them jointly. Columns (1) to (3) point out that each skill by itself has a strong effect on university enrollment. For example, the computed average marginal effect of a one standard deviation increase in (just) one of these skills will improve college attainment by more than 27 percentage points. However, column (4) shows that when the skills are jointly considered, verbal stands out, and the impact of the others falls drastically.[22] For example, the effect of motivation becomes substantially less important once conditioning on verbal and math. Similarly, the coefficient on math skills falls by nearly a factor of three. This result is a consequence of the large correlation among the factors shown in subsection 6.1. In summary, Table 7 shows that verbal skills have an effect on college enrollment that is almost twice as large as math skills.

[22] Appendix E shows that this result holds if we estimate a factor model with general, math, and verbal skills, where the math and verbal factor are orthogonal to the general skill factor.

One possible concern with these findings is that verbal skills may simply be proxying for another aspect that we are not accounting for in the model. Section 4 discusses two possible sources of endogeniety. First, despite the fact that we control for background characteristics, unobserved family inputs might still be driving our results. To further investigate this issue, we exploit the multi-period nature of our empirical strategy. Given that our model estimates skills at every time period, we can recover the contribution of each factor to the level of skills that we observed in Key Stage 4. To take advantage of this, we will let $\hat{\theta}_{i4|3,2,1,0} = \mathrm{E}(\theta_{i4}|\theta_{i3}, \theta_{i2}, \theta_{i1}, x_i)$ denote the expected value of skills in Key Stage 4, conditional on everything that has happened to the student as of Key Stage 3 (note that the value 0 only refers to background characteristics, i.e. IDACI, gender, race, among others). We can re-write θ_{i4} as

$$\theta_{i4} = \underbrace{\left(\theta_{i4} - \hat{\theta}_{i4|3,2,1,0}\right)}_{\text{residual change in skill occurring in KS4}} + \hat{\theta}_{i4|3,2,1,0} \tag{13}$$

Writing the variable in this way is useful because it informs us about the portion of these skills that was determined in previous periods and the portion that was determined in the current period. In fact, we can further decompose this variable period by period.

$$\theta_{i4} = \underbrace{\left(\theta_{i4} - \hat{\theta}_{i4|3,2,1,0}\right)}_{\text{KS4 Shock}} + \underbrace{\left(\hat{\theta}_{i4|3,2,1,0} - \hat{\theta}_{i4|2,1,0}\right)}_{\text{KS3 Shock}} + \underbrace{\left(\hat{\theta}_{i4|2,1,0} - \hat{\theta}_{i4|1,0}\right)}_{\text{KS2 Shock}} + \underbrace{\left(\hat{\theta}_{i4|1,0} - \hat{\theta}_{i4|0}\right)}_{\text{KS1 Shock}} + \hat{\theta}_{i4|0} \tag{14}$$

Where $\hat{\theta}_{i4|2,1,0} = \mathrm{E}(\theta_{i4}|\theta_{i2}, \theta_{i1}, x_i)$, $\hat{\theta}_{i4|1,0} = \mathrm{E}(\theta_{i4}|\theta_{i1}, x_i)$, and $\hat{\theta}_{i4|0} = \mathrm{E}(\theta_{i4}|x_i)$

While the results in Table 7 show the total effect of the skills when each of the five components in Eq. (14) are combined, an alternative approach would be to regress the outcome on all of these differences, which will allow us to recover the marginal effect of the new information received at each key stage. The benefit of this approach is that it is possible to study the relative contribution to university enrollment of math and verbal skills, after conditioning on, for example, all the information available through KS1 (i.e. covariates for background characteristics, and performance in math and verbal national assessments). If we assume earlier years' performance in national tests serve as a sufficient statistic for background characteristics, in particular parental inputs, (Todd

Table 8: Logistic Regression: College Enrollment and the Role of Shocks

Average Marginal Effects	
KS1 Math Shock	0.075 (0.004)
KS2 Math Shock	0.099 (0.001)
KS3 Math Shock	0.113 (0.001)
KS4 Math Shock	0.122 (0.010)
KS1 Verbal Shock	0.194 (0.004)
KS2 Verbal Shock	0.197 (0.001)
KS3 Verbal Shock	0.200 (0.001)
KS4 Verbal Shock	0.180 (0.006)
Motivation (not shown)	

Note- Results are from a logistic regression with controls for gender, IDACI Index, free school lunch, and special education needs. Shocks are defined in Equation 14. Note that the shocks have been normalize into KS4 standard deviation units. Bootstrapped standard errors at the school level.

and Wolpin, 2003), then conditioning on these tests in this way will correct for the endogeneity.

Table 8 shows the results of a logistic regression that includes the decomposition of Key Stage 4 math and verbal skills as described in Eq. (14). This regression allows us to consider the following counterfactual: does performing better than expected in math, once we condition on earlier test performance, have a differential impact from performing better than expected in verbal? The results in this table show that at each key stage, performing better than expected in verbal has a much larger impact on the college enrollment decision than performing better than expected in math. For example, a one standard deviation increase in the predicted value of KS4 verbal skills occurring from an outcome in KS2, after conditioning on family background characteristics and KS1 information, leads to an increase in college enrollment of 19.7 percentage points, while a similar shock in math would have lead to an increase of 9.9 percentage points. If we assume parental inputs are fixed, these results strongly suggest that the conclusions from our earlier analysis were not driven by endogeneity in parental inputs.

The second source of endogeneity that we need to address is motivation. While we model this

Table 9: Logistic Regression: Institution Type and Key Stage 4 Skills

	Enrolled in Non-Selective Institution (conditional on not attending selective institution)	Enrolled in Selective Institution (conditional on attending university)
Average Marginal Effects		
KS4 Math	0.090 (0.001)	0.147 (0.001)
KS4 Verbal	0.185 (0.001)	0.151 (0.002)
KS4 Motive	0.032 (0.001)	-0.008 (0.001)
Obs.	463141	180830

Note- Results are from a logistic regression with controls for gender, IDACI Index, free school lunch, special education needs, race, mother tongue, and school fixed effects. Bootstrapped standard errors at the school level.

directly, it is possible that it is not completely captured and might be driving our main result. The issue is that students at KS4 (age 16) may have already made their educational decisions (i.e. whether attending university, and field of study if attending), and therefore the effort that they may exert is a function of these decisions. For example, students who have already planned to obtain a degree in history may not spend enough time studying math. One way to address this concern is to look at the relative importance of skills using earlier measures (e.g. KS1, age 7), when effort and motivation at school is less likely to be determined by decisions that have to be made 11 years later. Table 7 and Appendix Table 17 show that verbal skills continue to have a larger effect on university attainment than math skills at each stage stage of the schooling career. Moreover, the magnitude of the differential effect is sizable across the board, which further substantiates our main findings.

So far, we have focused on university enrollment. However, we also have access to other educational outcomes, i.e. graduation, field of study, and selectiveness of the university attended. In order to further analyze the role of skills on these alternative outcomes, Tables 9 and 10 show the average marginal effects from logistic regressions of enrollment outcomes conditional on institution

Table 10: Logistic Regression: Graduation and Key Stage 4 Skills

	All Majors (controls for major)	Conditional on Enrolled in STEM	Conditional on Enrolled in Non-STEM
Average Marginal Effects			
KS4 Math	0.011 (0.001)	0.038 (0.002)	0.004 (0.002)
KS4 Verbal	0.091 (0.001)	0.101 (0.003)	0.085 (0.001)
KS4 Motive	0.029 (0.001)	0.032 (0.005)	0.028 (0.002)
Obs.	180830	54206	126624

Note- Results are from a logistic regression with controls for gender, IDACI Index, free school lunch, special education needs, race, mother tongue, and school fixed effects. Bootstrapped standard errors at school level.

selectiveness (i.e. being a Russell Group member or not), and graduation outcomes conditional on field of study. Consistent with our previous results, the first column of Table 9 shows that verbal skills have a larger impact than math skills on enrolling in university conditional on not attending a selective institution. However, the second column of this table shows that conditional on enrollment, math and verbal skills have a similar effect on the probability of attending a selective institution. This result is expected, as most selective institutions are able to enroll students from the very top of both skills distributions. Overall, the results in both columns suggest that the larger effect of verbal skills on math skills are likely driven by those students who are closer in skills to the marginal enrolled student. Finally, Table 10 shows different graduation outcomes conditional on enrollment. Column (1) shows that, once conditioning on enrollment, verbal skills have a substantially larger effect than math skills in explaining overall university graduation rates. Similarly, once we condition on enrolling in STEM or non-STEM fields, verbal plays a more relevant role than math. While this result regarding non-STEM fields does not seem to be surprising, this is not the case for STEM graduation.[23] The STEM result reinforces our main finding that points towards a key role of verbal skills in educational attainment.

[23]Note, that if we analyze enrollment in STEM fields conditional on college enrollment, math skills are substantially more important that verbal skills. Section 9 discusses this point in detail.

7 Addressing Alternative Explanations

We now investigate three possible avenues that may be conflating our main findings. First, we analyze whether verbal skills are proxying for externalizing behavior[24] more than math skills, and to what extent more detailed family covariates (not included in our previous analysis) could differentially affect math and verbal tests scores. Second, we study the role played by features of university supply on the disproportionate effect of verbal skills in college enrollment. Finally, we investigate whether the relative importance of verbal skills over math skills is due to some idiosyncrasy of the United Kingdom. We find no evidence that any of these alternative explanations are driving our main result.

7.1 Interrelation between Test Scores and Externalizing Behavior, Family Background Characteristics, and IQ

It is possible that our verbal factor is disproportionately capturing other types of skills that affect schooling outcomes, while this may not be true for the math factor. For example, if externalizing behavior/socio-emotional skills are more related to outcomes in verbal test scores rather than math test scores, then we could be confounding the larger effect of verbal skills with the role played by externalizing behavior. While the multi-period factor model addresses this concern by controlling for special education needs, we further investigate this issue using a database that contains richer measures of externalizing behavior. More specifically, we make use of the ALSPAC database (The Avon Longitudinal Study of Parents and Children) which is a large scale longitudinal study of children born in Avon (United Kingdom) during the early 1990s. This data cannot be linked to our main database (i.e. NPD-HESA). However, it is useful for further analysis given that it has very rich information on student background characteristics, and individuals in this sample not only have a similar age to the students from our main database, but also belong to the same educational system. Our proxies for externalizing behavior are obtained from the Strengths and Difficulties Questionnaire (SDQ) that was completed by the teachers of the students at the age of 7

[24]Childhood behaviors characterized by impulsivity, disruptiveness, aggression, antisocial features, and overactivity are called externalizing behavior.

years.[25] We have measures for emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problem. Higher scores (scale of 0 to 10) indicate greater levels of severity. In addition, we have a measure for pro-social behavior that takes values from 0 to 10, where a higher value denotes a more prosocial behavior.

This database also provides information on schooling records. For example we have access to overall students performance in each key stage assessment in math and verbal. In order to study whether these proxies for externalizing behavior/socio-emotional skills show a higher correlation with verbal skills than math skills, we perform a regression analysis where the dependent variables are performance in KS2 math or verbal and the independent variables are the SDQ measures.[26] Table 11 shows regression outcomes where each coefficient corresponds to a separate regression (in each of them we control for gender).[27] Panel A of Table 11 shows that, while these proxies are highly correlated with math and verbal skills, they do not seem to show a larger correlation with verbal than math test scores (i.e. all components of the SDQ questionnaire show similar effect in math and verbal). For example, a one-point increase in hyperactivity problems decreases verbal test scores by 0.169 of a standard deviation, which is very similar to the effect in math (i.e. 0.165). Therefore, these results suggest that the larger effect attributed to verbal skills in explaining college enrollment is not likely to be given by a larger correlation between verbal skills and externalizing behavior.

Family background characteristics might have a larger impact on verbal than math skills. Therefore, failing to control for this may misattribute the effect of parental characteristics on university enrollment to verbal skills. In our empirical model, we address this issue by following two strategies. First, we include the following controls for family background characteristics: free school meal eligibility, race, mother tongue, special education needs, IDACI index (a poverty index), and school attended. Second, we further examine our findings by conditioning on early key stage skills.[28] In

[25]The SDQ is a behavioral screening questionnaire for children and adolescents ages 2 through 17 years old, developed by the child psychiatrist Robert N. Goodman.

[26]The test scores on KS2 math and verbal have been standardized to have mean 0 and standard deviation 1

[27]We did not include all the measures in one regression because they are highly correlated, making the interpretation of the coefficients difficult due to multicolinearity problems.

[28]Table 8 shows evidence on this point.

other words, we use early life skills as sufficient statistics for family background charateristics.[29] In order to perform a final check on this assumption, we make further use of the ALSPAC database that provides more detailed information on parental background characteristics. Panel B of Table 11 shows OLS regressions of family background covariates such as parental education (i.e. parents holding a college degree), and proxies for family composition (i.e. father living at home) and income (i.e. home ownership status) on KS2 math and verbal performance (note that each coefficient corresponds to a separate regression). Overall, the results seem to indicate that there is no differential effect of family background characteristics on math and verbal performance. For example, having a mother (father) with a college degree increases KS2 english and math performance by 0.789 (0.753) and 0.756 (0.727) of a standard deviation, respectively. In summary, these findings further suggest that our main results are not likely to be driven by differential effects of family characteristics on math and verbal test scores.

Finally, we also explore, using ALSPAC database, the correlation between IQ tests and KS2 test scores. Panel C of Table 11 shows the interrelation between the IQ test Wechsler Intelligence Scale for Children (WISC) and KS2 exams.[30] Results show that both math and verbal scores are highly and similarly correlated with the WISC score, suggesting that verbal tests are not necessary better proxying students, IQ.

7.2 Can University Supply Explain the Higher Importance of Verbal Skills Relative to Math Skills?

An additional possible explanation to the large effect of verbal skills on post-secondary enrollment could be the natural consequence of universities in the United Kingdom mainly offering programs that do not require an intensive use of math skills. To study this possibility, we create a variable indicating whether a specific A-level exam (KS5) is required in order to be admitted into a given

[29]Similar assumptions have been done in the literature of teacher value-added (Todd and Wolpin, 2003).

[30]The Wechsler Intelligence Scale for Children (WISC) is an intelligence test for children between the ages of 6 and 16. The total IQ score represents a child's general intellectual ability. It also provides five primary index scores (i.e., verbal comprehension index, visual spatial index, fluid reasoning index, working memory index, and processing speed index). In this sample the mean of the IQ score is 104 points, and the standard deviation is 16.1. The raw correlations between the verbal and math test scores with the WISC index are 0.69 and 0.61 respectively.

Table 11: Linear Regression Model: Key Stage 2 Test Scores and ALSPAC Data

	Verbal	Math
Panel A:Socio-Emotional Skills		
Hyperactivity Problems (obs. = 5,434)	-0.169 (0.005)	-0.165 (0.005)
Emotional Problems (obs. = 5,464)	-0.092 (0.007)	-0.112 (0.006)
Conduct Problems (obs. = 5,460)	-0.163 (0.009)	-0.149 (0.009)
Peer Problems (obs. = 5,464)	-0.094 (0.007)	-0.103 (0.007)
Pro-social (obs. = 5,461)	0.084 (0.006)	0.081 (0.006)
Panel B: Family Background Characteristics		
Parents Own House (obs. = 9,356)	0.549 (0.023)	0.534 (0.024)
Father Live at Home (obs. = 7,985)	0.327 (0.032)	0.322 (0.032)
Mother College Degree (obs. = 10,232)	0.789 (0.028)	0.756 (0.029)
Father College Degree (obs. = 9,845)	0.753 (0.025)	0.727 (0.025)
Panel C: IQ Test		
WISC IQ Test (obs. = 6,427)	0.035 (0.023)	0.038 (0.024)

Notes- Each coefficient corresponds to a separate regression. The dependent variables, overall Key Stage 2 math and verbal test scores have been standardized to have mean 0 and standard deviation 1. All specifications include controls for gender.

major when applying to university.[31] Then, we combine this information with the field of study of the enrolled students. We find that 44.6% of students enrolled in university were required to obtain a qualification in the sciences (e.g. math, physics, etc), 30.92% a qualification in non-sciences, and 25.54% did not face any requirement. In addition, we expand our previous definition of "math requirements" by also considering what type of A-level exam is (generally) recommended to be taken even though they may not be a pre-requisite for a given major.[32] By jointly taking into account requirements and recommendations, the proportion of students attending university that were (at least) suggested to take a science related A-level is 62.4%. Overall, these simple statistics suggest that math skills are in fact required by universities and, therefore, our results are not likely to be driven by the type of majors that are offered in the higher education system.

A second piece of evidence consistent with these findings can be found in the last column of Table 10, which shows that, after conditioning on STEM enrollment, verbal skills are more important than math skills in explaining STEM graduation. While surprising, this fact suggest that verbal skills are also crucial to graduate in STEM fields, highlighting their importance in educational outcomes.

7.3 Is the Larger Effect of Verbal Skills on University Enrollment a Specific Phenomenon of the UK Educational System?

It is still possible that our finding indicating that verbal skills have a larger effect on college enrollment than math skills could be a consequence of particular institutional features of the UK. In order to assess this possibility, we make use of the high school transcript data from the National Longitudinal Survey of Youth of 1997 (NLSY97). The NLSY97 is a nationally representative sample of youths from the United States who were 12 to 17 years old when they were first surveyed in 1997. It collects extensive information on family background characteristics, educational experiences and labor market outcomes through time. Table 12 shows OLS results from simple econometric models where the dependent variable is college enrollment, and the independent variables are gender, race,

[31] This information was extracted from the document "Informed Choices" created by the Institute of Career Guidance and the Russell Group universities. This publication aims to provide information to all students considering A-level and equivalent options.

[32] This information was also extracted from the document "Informed Choices."

and performance on math and english courses in high school. We measure performance in these subjects using two different definitions. First, we include high school GPA separately for math and english courses. Second, we focus on performance in Algebra I and english in grade 9 which are generally taken at the beginning of secondary education. We expect the specifications using the later set of variables to be less driven by schooling aspirations. Results show a similar pattern as in the UK data, where performance in subjects that are related to verbal skills have a larger effect on college enrollment. Specifically, while a one standard deviation increase in high school math GPA increases the probability of attending university by 3.8 percentage points, a one standard deviation increase in english GPA leads to a gain of 16.3 percentage points. Similarly, the last column of Table 12 shows that the performance in english in grade 9 has a larger impact on the probability of attending and graduating from university than performance in Algebra I. In summary, these findings using US data suggest that the key role of verbal skills in explaining college enrollment and graduation does not appear to be unique to the UK educational system.

8 Development of Skills

We now study how math and verbal skills evolve over the course of a student's compulsory education career. Table 4 presents the full correlation of these nine factors, showing that they are highly correlated. Most likely a large portion of the Key Stage 4 math and verbal factors are determined from outcomes in Key Stage 1 to Key Stage 3. To formalize the analysis, we return to the tools used in Eq. (14) and Table 8. In this table we reported the effect of changes in expectations of skills, once we condition on past skills. From our model we can compute empirically how likely deviations from the expectations occur in the data. For example, we can compute $\mathrm{Var}(\hat{\theta}_{i4|1,0} - \hat{\theta}_{i4|0})$, that describes the extent to which the events of Key Stage 1 affect the expected values of skills in Key Stage 4, once we condition on background characteristics.[33] This is an empirical measure in the population of the Key Stage 4 skill shocks attributed to outcomes in Key Stage 1.

Table 13 computes the contribution of each Key Stage shock in explaining the total variance

[33] Here we are quantifying to what extent the new information that arrives in each key stage contributes in explaining the total variability of skills by the end of compulsory schooling.

Table 12: Linear Probability Model: University Enrollment (NLSY97)

	(1)	(2)	(3)
Constant	0.494 (0.016)	0.647 (0.016)	0.656 (0.018)
Female	0.100 (0.013)	0.012 (0.013)	0.035 (0.016)
High School GPA Math	–	0.038 (0.009)	–
High School GPA English	–	0.163 (0.009)	–
GPA Algebra I	–	–	0.074 (0.009)
GPA English Grade 9	–	–	0.122 (0.010)
Race Controls	Yes	Yes	Yes
R-Squared	0.024	0.161	0.130
Obs.	5,768	5,262	3,763

Notes- High school regressors are standardized to have mean zero and standard deviation 1.

of KS4 skills, measured in terms of R-squared. In addition, this table also lists the computed probability of receiving a skill shock that would move a student up 0.5 standard deviation units (or more) in the aggregate skill distribution. The first row of the table shows the contribution of background characteristics (i.e. IDACI index, free school meal eligibility, special education needs, elementary school attended, mother tongue, race and gender) in explaining the total variation (in terms of R-squared) in each of the KS4 factors. As expected, these elements have large explanatory power, on the order of 46% in math and verbal skills and 50% in motivation. Moreover, since these family background variables show similar predictive power in math and verbal skills, it suggests that these skills are similarly shaped at home. Moving forward, the outcome in Key Stage 1 adds an additional 14% and 13% to the explained variance of KS4 math and verbal skills, respectively. This implies that, by age 7, approximately 60% of the variance in both skills at age 16 can already been explained. This result highlights a key point. While a large portion of skills variation in KS4 is already determined at an early age, there remains a large amount (40%) to be explained at later stages, suggesting some scope for overcoming initial disadvantages. However, by the end of KS3 (age 14), around 96% (88%) of the variation in KS4 math (verbal) skills can be explained by KS1-KS3 stages information and family background characteristics. Overall, the first three columns of Table 13 show that each of the key stages plays a role in explaining the variance of skills at the end of compulsory education. Finally, the last three columns of Table 13 indicate how likely it is that a student could move 0.5 standard deviation units in the aggregate skill distribution for each skill given the new information obtained in the given key stage. For example, the probability that a student moves up 0.5 standard deviations in their KS4 math skills given a positive shock in KS3 is 0.13 (i.e. this implies getting a positive shock that is more than 1 standard deviation larger than the mean shock in that period). Since Table 7 reports the average marginal effect for a one standard deviation increase in KS4 skills, the impact of this shock would be a 4.5 percentage point increase in the probability of college enrollment.

In summary, this section highlights two main points. First, around half of the variation in skills at the end of compulsory education is explained by "new information" that is arriving between KS1-KS4, suggesting that a sizable portion of skills is determined at later stages. However, the

Table 13: Development of Key Stage 4 Skills

	Contribution to R-squared			Probability of Receiving Shock to Move Up 0.5 S.D. in the Aggregate Distribution		
	Math	Verbal	Motive	Math	Verbal	Motive
Background Characteristics	0.465 (0.002)	0.466 (0.002)	0.501 (0.003)	0.232 (0.000)	0.232 (0.000)	0.240 (0.001)
Key Stage 1 Shocks	0.140 (0.001)	0.130 (0.001)	0.067 (0.002)	0.091 (0.001)	0.082 (0.001)	0.027 (0.001)
Key Stage 2 Shocks	0.173 (0.002)	0.160 (0.001)	0.081 (0.002)	0.114 (0.001)	0.106 (0.001)	0.039 (0.002)
Key Stage 3 Shocks	0.209 (0.001)	0.149 (0.002)	0.154 (0.004)	0.137 (0.001)	0.098 (0.001)	0.101 (0.003)
Key Stage 4 Shocks	0.040 (0.001)	0.117 (0.001)	0.211 (0.003)	0.006 (0.000)	0.072 (0.001)	0.138 (0.002)
Total	1.00	1.00	1.00	–	–	–

Note: Key Stage 4 shocks are defined as $\theta_{i4} - \mathrm{E}(\theta_{i4}|\theta_{i3}, \theta_{i2}, \theta_{i1}, x_i)$. Key Stage 3 shocks are defined as $\mathrm{E}(\theta_{i4}|\theta_{i3}, \theta_{i2}, \theta_{i1}, x_i) - \mathrm{E}(\theta_{i4}|\theta_{i2}, \theta_{i1}, x_i)$. The other Key Stage shocks are defined similarly.

probability of receiving a sufficiently positive shock at any given key stage is relatively small.

While the inclusion of the the motivation factor into our analysis was largely to correct for endogeneity, the results on motivation in Table 13 are interesting in their own right. The entirety of the motivation factor is either explained by background characteristics or events occurring towards the end of compulsory education. One of the measurements used to extract motivation was the number of subject specific exams taken in Key Stage 4. The results in this table suggest that the events that occur during Key Stage 1, after conditioning on background characteristics, have virtually no predictive power on the number of subject tests taken in Key Stage 4. Therefore, if the KS4 skills are thought to be endogenous with University enrollment, the results from motivation in this table validates the use of these early skills as exogenous regressors.

9 Understanding the Gender Gap in College Enrollment and STEM Major

Subsection 6.1 shows the presence of small differences in math skills between genders contrasted with a large female advantage in verbal skills. This empirical regularity combined with our findings on the importance of verbal skills suggests a possible explanation for the gender gap in college in college enrollment and major. Table 14 studies whether the gender gap in college enrollment and STEM fields can be explained by differences in skills. Column (1) shows that females are 4.4% more likely to attend university than males. This gap is smaller than the 6.6% reported earlier because this specification already controls for special education needs.[34] Column (2) shows that, after controlling for skills in KS4, females are less likely to attend university than males, suggesting that gender differences in verbal skill levels play a key role in explaining the gender gap in college enrollment.[35] This also implies that if males had similar preferences for college enrollment as females, then the observed gap would be even larger. Columns (3) and (4) analyze whether skills have differential effect on college enrollment once conditioning on gender. These columns show results that are based on the male and female sample separately. Differences in the average marginal effects between columns (3) and (4) are small, suggesting that math and verbal skills do not play a differential role on males and females in the decision to enroll in college.

We also investigate the gender gap in STEM fields, which (as is well known) favors males. Columns (5) to (8) analyze enrollment in STEM fields conditional on enrolling in college. Column (5) shows that after controlling for family background characteristics (i.e. IDACI Index, free school lunch, race, mother tongue, and special education needs), females are 17.1 percentage points less likely to enroll in STEM fields. After controlling for skills in column (6), the gap drops by about 40%. The negative coefficient on verbal skills suggests that comparative advantage in math skills is a key factor in the STEM enrollment decision. As shown in Figure 2, males are more likely to perform better in math than in verbal (when considering their percentile in the respective skill

[34] Boys are substantially more likely to be assigned into this category. Aucejo (2015) shows that socio-emotional skills also play a role in explaining the gender gap in college enrollment.

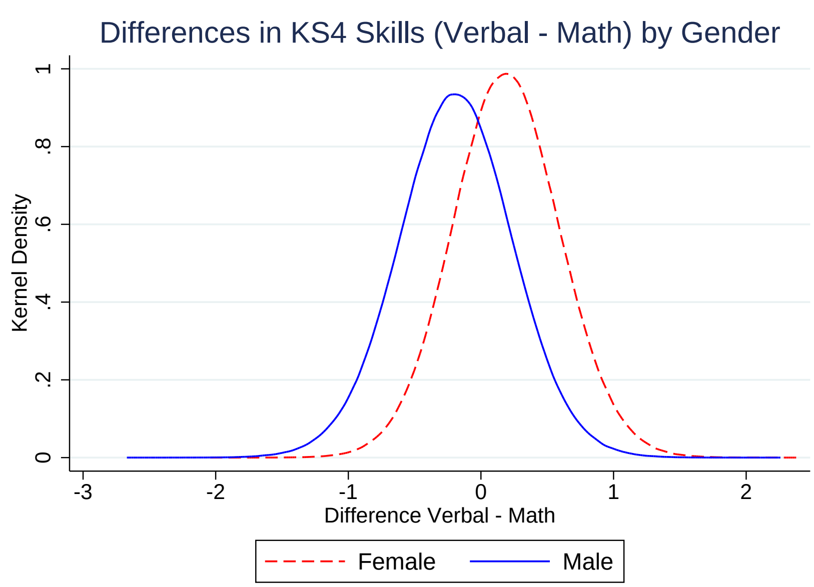
[35] Table D shows that similar results are obtained if instead we control for skills in KS3.

Table 14: Logistic Regression: Gender, University Outcomes and Key Stage 4 Skills

	University Enrollment				Enrollment in STEM (conditional on enrollment)			
	All		Male	Female	All		Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average Marginal Effects								
Female	0.044 (0.001)	-0.008 (0.001)	–	–	-0.171 (0.003)	-0.105 (0.002)	–	–
KS4 Math	–	0.096 (0.001)	0.099 (0.002)	0.095 (0.002)	–	0.218 (0.002)	0.273 (0.003)	0.167 (0.004)
KS4 Verbal	–	0.187 (0.001)	0.179 (0.000)	0.200 (0.002)	–	-0.159 (0.003)	-0.225 (0.002)	-0.097 (0.004)
KS4 Motive	–	0.029 (0.001)	0.024 (0.001)	0.031 (0.001)	–	0.003 (0.001)	-0.002 (0.003)	0.005 (0.001)

Note- Results are from a logistic regression with controls for IDACI Index, free school lunch, race, mother tongue, and special education needs. Bootstrapped standard errors at school level.

Figure 2: Differences in Skills



distributions), while for females the opposite is true. More specifically, 65% of males perform better in math than in verbal while this figure is 35% for females. Therefore, this fact, combined with the large negative coefficient on verbal skills, suggest that males' comparative advantage in math leads them to be more likely to choose STEM fields. Finally, the last columns of Table 14 analyze STEM enrollment by gender. Results indicate that skills have a differential effect between genders. For example, while a one standard deviation increase in verbal (math) skills for males reduces (increases) the probability of enrolling in STEM by 22.5 (27.3) percentage points, females show a decrease (increase) of 9.7 (16.7) percentage points. To conclude, these results suggest that males are more sensitive to the level of skills than females when deciding to enroll in STEM fields, where skill comparative advantage seems to be substantially more important for males in their decision to enroll in scientific fields.

10 Conclusion

This paper makes use of a rich panel database that follows all students in England from elementary school to university to assess the effect of math and verbal skills on schooling outcomes (i.e. university enrollment, field of study and graduation). By estimating a multi-period factor model of skills, we find that the effect of verbal skills on university enrollment is almost two times larger than math skills. Our findings also indicate that, while 60% of the variance in skills at age 16 can be explained by school outcomes and background characteristics by age 7, a large portion of skill variation at the last year of compulsory school is explained between the ages of 7 and 16. This suggests that there is some scope for overcoming initial skills disadvantages. Finally, we show that, after controlling for math and verbal skills, females are slightly less likely to attend college than males (i.e. the gap is reversed), suggesting that female "preferences" for educational attainment contribute to close the gap rather than increasing it. Moreover, we present evidence indicating that females' outperformance in verbal contribute to explain the gender gap in STEM fields.

References

D. Acemoglu and D. Autor. Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics*, 4:1043–1171, 2011.

M. Almlund, A. L. Duckworth, J. Heckman, and T. Kautz. Chapter 1 - personality psychology and economics. In S. M. Eric A. Hanushek and L. Woessmann, editors, *Handbook of The Economics of Education*, volume 4 of *Handbook of the Economics of Education*, pages 1 – 181. Elsevier, 2011. doi: http://dx.doi.org/10.1016/B978-0-444-53444-6.00001-8.

J. Altonji. The effects of high school curriculum on education and labor market outcomes. *The Journal of Human Resources*, 30(3):409–438, 1995.

E. Aucejo. Explaining cross-racial differences in the educational gender gap. *Working Paper, London School of Economics and Political Sciences*, 2015.

S. Cameron and J. Heckman. The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy*, 109(3):pp. 455–499, 2001.

G. Castex and E. K. Dechter. The changing roles of education and ability in wage determination. *Journal of Labor Economics*, 32(4):pp. 685–710, 2014.

J. Cawley, J. Heckman, and E. Vytlacil. Three observations on wages and measured cognitive ability. *Labour Economics*, 8(4):pp. 419–42, 2001.

K. Cortes, J. Goodman, and T. Nomi. Intensive math instruction and educational attainment long-run impacts of double-dose algebra. *Journal of human Resources*, 50(1):108–158, 2015.

F. Cunha and J. Heckman. The technology of skill formation. Technical report, National Bureau of Economic Research, 2007.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

S. Dougherty, J. Goodman, D. Hill, E. Litke, and L. Page. Early math coursework and college readiness: Evidence from targeted middle school math acceleration. *NBER working Paper Series*, pages 1–50, 2015.

J. Heckman, R. Pinto, and P. Savelyev. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *The American Economic Review*, 103(6):1–35, 2013.

J. J. Heckman, J. Stixrud, and S. Urzua. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482, 2006.

G. Hobbs and A. Vignoles. *Is free school meal status a valid proxy for socio-economic status (in schools research)?* Centre for the Economics of Education, London School of Economics and Political Science, 2007.

J. S. Joensen and H. S. Nielsen. Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources*, 44(1):171–198, 2009.

P. B. Levine and D. J. Zimmerman. The benefit of additional high-school math and science classes for young men and women. *Journal of Business & Economic Statistics*, 13(2):137–149, 1995.

H. Rose and J. R. Betts. The effect of high school courses on earnings. *Review of Economics and Statistics*, 86(2):497–513, 2004.

P. A. Ruud. Extensions of estimation methods using the em algorithm. *Journal of Econometrics*, 49(3):305–341, 1991.

P. E. Todd and K. I. Wolpin. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113:3–33, 2003.

R. Varadhan and C. Roland. Simple and globally convergent methods for accelerating the convergence of any em algorithm. *Scandinavian Journal of Statistics*, 35(2):335–353, 2008.

Appendix

A Mapping the joint factor model to a dynamic factor model

In our factor specification, we allow the factor error to be correlated across time period. An alternative specification would allow the factors to be a function of the earlier factors and model the factor error as independent (once we condition on the earlier factor outcomes). This alternative approach is commonly used in dynamic factor models. This section demonstrates that for any value of the parameters, the factor structure we use in Eq. (1) maps directly to a dynamic factor model.

In a dynamic factor model, the current period factors are written as a linear function of previous factor realizations and an additive, independent error term. For example, the factors in Key Stage 3 would be written as:

$$\theta_{i3} = female_i * \widetilde{\psi}_3 + \widetilde{\Phi}_3 x_i + \Omega_3 \begin{bmatrix} \theta_{i1}' & \theta_{i2}' \end{bmatrix}' + \widetilde{\xi}_i \tag{15}$$

The tilde's represent that the parameters in this specification perform a similar role to the ones in Eq. (1) but are not identical. The additional variable Ω_3 is a weighting matrix that describes how Key Stage 1 and 2 factors influence Key Stage 3 factors. Finally, the error term, $\widetilde{\xi}_i \sim \mathcal{N}(0, \widetilde{\Sigma}_3)$. This gives:

$$\theta_{i3} \sim \mathcal{N}\left(female_i * \widetilde{\psi}_3 + \widetilde{\Phi}_3 x_i + \Omega_3 \begin{bmatrix} \theta_{i1}' & \theta_{i2}' \end{bmatrix}', \widetilde{\Sigma}_3\right) \tag{16}$$

To map this model to the parameters in Eq. (1), we can write the distribution of θ_{i3} conditional on θ_{i1} and θ_{i2} as $\theta_{i3} \sim \mathcal{N}\left(\mathrm{E}(\theta_{i3}|\theta_{i1}, \theta_{i2}), \mathrm{Var}(\xi_{i[5:6]}|\theta_{i1}, \theta_{i2})\right)$, where $\xi_{i[5:6]}$ is the 5^{th} and 6^{th} components of the Key Stage 3 factor residual. This gives:

$$\begin{aligned}
\mathrm{E}(\theta_{i3}|\theta_{i1}, \theta_{i2}) =& female_i * \psi_3 + \Phi_3 x_i + \dots \\
& + \Omega_3 \begin{bmatrix} (\theta_{i1} - female_i * \psi_1 - \Phi_1 x_i)' & (\theta_{i2} - female_i * \psi_2 - \Phi_2 x_i) \end{bmatrix}' \\
=& female_i * \underbrace{(\psi_3 - \Omega_3 (\psi_1 + \psi_2))}_{\widetilde{\psi}_3} + \underbrace{(\Phi_1 - \Omega_3 (\Phi_2 + \Phi_3))}_{\widetilde{\Phi}_3} x_i + \Omega_3 \begin{bmatrix} \theta_{i1}' & \theta_{i2}' \end{bmatrix}'
\end{aligned}$$

$$\Omega_3 = \Sigma_{[5:6,1:4]} \Sigma^{-1}_{[1:4,1:4]}$$

$$\begin{aligned}
\mathrm{Var}(\widetilde{\xi}_i) =& \mathrm{Var}(\xi_{i[5:6]}|\theta_{i1}, \theta_{i2}) \\
=& \Sigma_{[5:6,5:6]} - \Sigma_{[5:6,1:4]} \Sigma^{-1}_{[1:4,1:4]} \Sigma_{[1:4,5:6]}
\end{aligned}$$

Where $\Sigma_{[5:6,5:6]}$ represents the fifth to six row and fifth to six column partition of the matrix Σ.

We can apply this technique similarly to redefine all of the skills in Key Stage 1, 2 and 4.

B Measurement Equations and Normalizations

To identify the model we need to make two normalizations. First the factor specification in Eq. (1) does not contain a constant since we include a constant, μ_m, in the measurement equation. Second, in order to identify the full covariance matrix Σ we need to make some normalizations to λ_m.

Table 15: Factor Loadings

No.	Key Stage	Description	KS1 Math	KS1 Verbal	KS2 Math	KS2 Verbal	KS3 Math	KS3 Verbal	KS4 Math	KS4 Verbal	KS4 Motive	Residual Variance
1	1	Math Test	1†	0	0	0	0	0	0	0	0	0.224 (0.002)
2	1	Math Using and Applying TA	0.985 (0.002)	0	0	0	0	0	0	0	0	0.246 (0.002)
3	1	Math Number and Algebra TA	1.013 (0.002)	0	0	0	0	0	0	0	0	0.202 (0.001)
4	1	Math Shapes and Measure TA	0.992 (0.004)	0	0	0	0	0	0	0	0	0.235 (0.002)
5	1	Writing Test	0	1†	0	0	0	0	0	0	0	0.172 (0.001)
6	1	Writing TA	0	0.962 (0.002)	0	0	0	0	0	0	0	0.231 (0.001)
7	1	Reading TA	0	0.937 (0.002)	0	0	0	0	0	0	0	0.271 (0.001)
8	1	Listening TA	0	0.827 (0.003)	0	0	0	0	0	0	0	0.431 (0.001)
9	2	Math Test Paper A	0	0	1†	0	0	0	0	0	0	0.126 (0.000)
10	2	Math Test Paper B	0	0	0.981 (0.001)	0	0	0	0	0	0	0.158 (0.000)
11	2	Math Arithmetic Test	0	0	0.960 (0.001)	0	0	0	0	0	0	0.194 (0.001)
12	2	Math TA	0	0	0.908 (0.001)	0	0	0	0	0	0	0.237 (0.002)
13	2	Reading Test	0	0	0	1†	0	0	0	0	0	0.250 (0.001)
14	2	Writing Test	0	0	0	0.878 (0.002)	0	0	0	0	0	0.421 (0.001)
15	2	Spelling Test	0	0	0	0.876 (0.001)	0	0	0	0	0	0.424 (0.002)
16	2	English TA	0	0	0	0.972 (0.002)	0	0	0	0	0	0.233 (0.001)
17	3	Math Test Paper 1	0	0	0	0	0.984 (0.002)	0	0	0	0	0.423 (0.001)

(*Continued on next page*)

Note-TA denotes teacher assessment.

Table 15: Factor Loadings

No.	Key Stage	Description	KS1 Math	KS1 Verbal	KS2 Math	KS2 Verbal	KS3 Math	KS3 Verbal	KS4 Math	KS4 Verbal	KS4 Motive	Residual Variance
18	3	Math Test Paper 2	0	0	0	0	0.996 (0.003)	0	0	0	0	0.409 (0.001)
19	3	Math Arithmetic Test	0	0	0	0	$1^{\dagger}$	0	0	0	0	0.406 (0.001)
20	3	Math TA	0	0	0	0	1.189 (0.001)	0	0	0	0	0.147 (0.001)
21	3	Writing Test (Longer)	0	0	0	0	0	$1^{\dagger}$	0	0	0	0.343 (0.002)
22	3	Reading Test	0	0	0	0	0	1.061 (0.002)	0	0	0	0.259 (0.001)
23	3	Writing Test (Shorter)	0	0	0	0	0	1.016 (0.001)	0	0	0	0.322 (0.002)
24	3	Reading Test (Shakespeare)	0	0	0	0	0	0.932 (0.003)	0	0	0	0.429 (0.002)
25	3	English TA	0	0	0	0	0	0.999 (0.002)	0	0	0	0.295 (0.001)
26	4	Math	0	0	0	0	0	0	$1^{\dagger}$	0	0	0.127 (0.001)
27	4	English	0	0	0	0	0	0	0	$1^{\dagger}$	0	0.132 (0.001)
28	4	Design and Technology: Graphic Products	0	0	0	0	0	0	0.301 (0.006)	0.579 (0.007)	0	0.477 (0.006)
29	4	Design and Technology: Resistant Materials Technology	0	0	0	0	0	0	0.396 (0.007)	0.429 (0.007)	0	0.515 (0.004)
30	4	Design and Technology: Textiles Technology	0	0	0	0	0	0	0.379 (0.008)	0.557 (0.009)	0	0.415 (0.002)
31	4	Art and Design	0	0	0	0	0	0	0.064 (0.006)	0.711 (0.005)	0	0.530 (0.003)
32	4	History	0	0	0	0	0	0	0	1.078 (0.002)	0	0.245 (0.001)
33	4	Geography	0	0	0	0	0	0	0.508 (0.006)	0.593 (0.006)	0	0.219 (0.001)
34	4	French	0	0	0	0	0	0	0.320 (0.006)	0.766 (0.006)	0	0.322 (0.002)
35	4	German	0	0	0	0	0	0	0.392 (0.007)	0.720 (0.009)	0	0.358 (0.003)

(*Continued on next page*)

Note-TA denotes teacher assessment.

Table 15: Factor Loadings

No.	Key Stage	Description	KS1 Math	KS1 Verbal	KS2 Math	KS2 Verbal	KS3 Math	KS3 Verbal	KS4 Math	KS4 Verbal	KS4 Motive	Residual Variance
36	4	Business Studies	0	0	0	0	0	0	0.405 (0.009)	0.700 (0.006)	0	0.322 (0.001)
37	4	Religious Studies	0	0	0	0	0	0	0	1.052 (0.004)	0	0.281 (0.002)
38	4	Short Religious Studies	0	0	0	0	0	0	0	0.970 (0.002)	0	0.337 (0.002)
39	4	Physical Education	0	0	0	0	0	0	0.593 (0.008)	0.342 (0.008)	0	0.469 (0.002)
40	4	Physics	0	0	0	0	0	0	1.568 (0.007)	0	0	0.194 (0.003)
41	4	Chemistry	0	0	0	0	0	0	1.548 (0.007)	0	0	0.198 (0.004)
42	4	Biology	0	0	0	0	0	0	1.102 (0.011)	0.241 (0.008)	0	0.187 (0.003)
43	4	Drama	0	0	0	0	0	0	-0.005 (0.007)	0.860 (0.007)	0	0.478 (0.003)
44	4	Information Technology	0	0	0	0	0	0	0.423 (0.013)	0.503 (0.010)	0	0.411 (0.004)
45	4	Short Information Technology	0	0	0	0	0	0	0.423 (0.008)	0.421 (0.006)	0	0.461 (0.003)
46	4	Spanish	0	0	0	0	0	0	0.295 (0.012)	0.781 (0.015)	0	0.363 (0.004)
47	4	Music	0	0	0	0	0	0	0.298 (0.008)	0.584 (0.008)	0	0.465 (0.004)
48	4	Social Science	0	0	0	0	0	0	0.088 (0.012)	0.864 (0.010)	0	0.369 (0.003)
49	4	Design and Technology: Electronic Products	0	0	0	0	0	0	0.528 (0.022)	0.330 (0.020)	0	0.492 (0.009)
50	4	Design and Technology: System and Control	0	0	0	0	0	0	0.540 (0.021)	0.334 (0.026)	0	0.493 (0.012)
51	4	English Literature	0	0	0	0	0	0	0	1.045 (0.001)	0	0.220 (0.001)
52	4	Design and Technology: Food Technology	0	0	0	0	0	0	0.202 (0.009)	0.742 (0.006)	0	0.372 (0.003)
53	4	Science	0	0	0	0	0	0	0.670 (0.006)	0.342 (0.006)	0	0.282 (0.003)

(*Continued on next page*)

Note-TA denotes teacher assessment.

Table 15: Factor Loadings

No.	Key Stage	Description	KS1 Math	KS1 Verbal	KS2 Math	KS2 Verbal	KS3 Math	KS3 Verbal	KS4 Math	KS4 Verbal	KS4 Motive	Residual Variance
54	4	Statistics	0	0	0	0	0	0	1.047 (0.010)	0.067 (0.008)	0	0.246 (0.004)
55	4	Medial, Film and Television Studies	0	0	0	0	0	0	-0.027 (0.011)	1.026 (0.009)	0	0.333 (0.003)
56	4	Fine Art	0	0	0	0	0	0	0.011 (0.015)	0.792 (0.014)	0	0.487 (0.006)
57	4	Office Technology	0	0	0	0	0	0	0.484 (0.016)	0.489 (0.015)	0	0.354 (0.003)
58	4	Home Economics: Child Development	0	0	0	0	0	0	0.240 (0.021)	0.799 (0.021)	0	0.371 (0.005)
59	4	Italian	0	0	0	0	0	0	0.224 (0.053)	0.710 (0.043)	0	0.521 (0.023)
60	4	Urdu	0	0	0	0	0	0	-0.033 (0.074)	0.763 (0.078)	0	0.637 (0.019)
61	4	Additional Applied Science	0	0	0	0	0	0	0.678 (0.050)	0.481 (0.043)	0	0.311 (0.008)
62	4	Leisure and Tourism	0	0	0	0	0	0	0.112 (0.011)	0.876 (0.007)	0	0.446 (0.007)
63	4	Applied ICT	0	0	0	0	0	0	0.302 (0.016)	0.572 (0.013)	0	0.506 (0.004)
64	4	Applied Science	0	0	0	0	0	0	0.478 (0.016)	0.559 (0.018)	0	0.457 (0.005)
65	4	Health and Social Care	0	0	0	0	0	0	0.164 (0.010)	0.827 (0.012)	0	0.450 (0.007)
66	4	Applied Business	0	0	0	0	0	0	0.416 (0.012)	0.602 (0.012)	0	0.407 (0.004)
67	4	Double Science	0	0	0	0	0	0	1.056 (0.002)	0	0	0.191 (0.001)
68	4	Total GCSE Exams Taken	0	0	0	0	0	0	0	0	$1^{\dagger}$	1.710 (0.017)
69	4	Authorize Absences	0	0	0	0	0	0	0	0	-0.016 (0.000)	0.005 (0.000)
70	4	Unauthorized Absences	0	0	0	0	0	0	0	0	-0.017 (0.000)	0.003 (0.000)

Note-TA denotes teacher assessment.

C Skill Coefficients

Table 16: Skill Coefficients

	KS1 Math	KS1 Verbal	KS2 Math	KS2 Verbal	KS3 Math	KS3 Verbal	KS4 Math	KS4 Verbal	KS4 Motive
Demographics									
female	-0.118	0.159	-0.222	0.116	-0.124	0.232	-0.083	0.274	0.096
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)
race: Asian	-0.101	-0.046	0.023	-0.025	0.082	0.082	0.236	0.209	0.289
	(0.015)	(0.011)	(0.012)	(0.009)	(0.010)	(0.008)	(0.014)	(0.011)	(0.018)
race: Black	-0.098	0.058	-0.095	-0.022	-0.076	0.019	0.034	0.110	0.311
	(0.005)	(0.007)	(0.010)	(0.011)	(0.010)	(0.008)	(0.014)	(0.008)	(0.023)
race: Other	0.015	0.068	0.050	0.083	0.037	0.098	0.057	0.107	0.100
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.014)
mother's tougue	0.139	0.191	-0.006	0.060	-0.042	-0.025	-0.183	-0.164	-0.497
	(0.016)	(0.015)	(0.012)	(0.011)	(0.011)	(0.010)	(0.015)	(0.012)	(0.023)
SEN	-0.939	-1.104	-1.069	-1.190	-0.838	-0.916	-0.837	-0.849	-1.075
	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.009)
FSM	-0.185	-0.233	-0.205	-0.224	-0.210	-0.251	-0.336	-0.347	-0.700
	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)
IDACI	-0.399	-0.468	-0.470	-0.517	-0.514	-0.635	-0.784	-0.818	-1.371
	(0.012)	(0.014)	(0.010)	(0.008)	(0.009)	(0.010)	(0.009)	(0.013)	(0.017)
Elementary School Fixed Effects									
Std. of 15,353 FE	0.323	0.314	0.281	0.274	0.230	0.276	0.303	0.303	0.661
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Distribution of Unobservables									
Type 1 (prob type 1 = 0.24)	0	0	0	0	0	0	0	0	0
Type 2 (prob type 2 = 0.76)	0.531	0.449	1.177	0.633	0.758	0.475	0.659	0.501	0.527
	(0.004)	(0.006)	(0.004)	(0.004)	(0.003)	(0.005)	(0.003)	(0.006)	(0.017)
Std(ξ)	0.663	0.649	0.572	0.595	0.517	0.597	0.692	0.687	1.009
	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.006)

Note-These estimates correspond to the coefficients of Eq. (1) (i.e. skill coefficients)

D University Enrollment with Early Factors

Table 17: Logistic Regression: University Enrollment and Early Factors

	Key Stage 1	Key Stage 2	Key Stage 3
Average Marginal Effects			
Female	0.031 (0.001)	0.032 (0.001)	0.010 (0.001)
Math	0.067 (0.001)	0.062 (0.001)	0.107 (0.001)
Verbal	0.116 (0.001)	0.184 (0.001)	0.173 (0.001)

Note- Results are from a logistic regression with controls for gender, IDACI Index, free school lunch, special education needs, race, mother tongue, and school fixed effects. Skills are normalized to mean zero, standard deviation 1. Bootstrapped standard errors at school level.

E Robustness Check. College Enrollment

Table 18: Logistic Regression, University Enrollment, Alternative Specifications

	Full Sample – No Mixture	Full Sample – No Mixture, KS4 Factors only (controlling for earlier test scores)	White, British Sample Only – No Mixture, KS4 Factors only (controlling for earlier test scores)‡
Average Marginal Effects			
KS4 Math	0.097	0.096	0.095
KS4 Verbal	0.189	0.190	0.196
KS4 Motive	0.030	0.030	0.025

†

‡ Results are from a logistic regression on the white, British subsample whose mother's native tongue is english. Model controls for gender, IDACI Index, free school lunch, special education needs, and school fixed effects. Skills are normalized to mean zero, standard deviation 1.