Do Advanced Mathematics Skills Predict Success in Biology and Chemistry Degrees?

Michael Adkins and Andrew Noyes

School of Education, University of Nottingham

Acknowledgements

This research is part of the Rethinking the Value of Advanced Mathematics Participation project, funded by the Nuffield Foundation [EDU/41221]. The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The views expressed herein are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org.

Author Note

Professor Andrew Noyes. Address: School of Education, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK. Email:

andrew.noyes@nottingham.ac.uk

Abstract

The mathematical preparedness of science undergraduates has been a subject of debate for some time. This paper investigates the relationship between school mathematics attainment and degree outcomes in biology and chemistry across England, a much larger scale of analysis than has hitherto been reported in the literature. A unique dataset which links the National Pupil Database for England (NPD) and Higher Education Statistics Agency (HESA) data is used to track the educational trajectories of a national cohort of 16-year-olds through their school and degree programmes. Multilevel regression models indicate that students who completed advanced mathematics qualifications prior to their university study of biology and chemistry were no more likely to attain the best degree outcomes than those without advanced mathematics. The models do, however, suggest that success in advanced chemistry at school predicts outcomes in undergraduate biology and vice versa. There are important social background differences and the impact of the university attended is considerable. We discuss a range of possible explanations of these findings.

KEY WORDS: mathematics, biology, chemistry, degree outcomes, multilevel modelling

Introduction

This paper considers the extent to which school mathematics qualifications predict undergraduate success in science, in particular biology and chemistry. Many researchers point to a problematic gap between school mathematics and university applications of mathematics within the disciplines (Heck and van Gastel, 2006; Tai, Sadler & Loehr, 2005; Groen, Coupland, *et al*., 2015). Such studies support a general consensus that success in undergraduate science is built upon 'two pillars' (Sadler and Tai, 2007): the level of mathematics and discipline-specific science knowledge, though some studies demur from this position. Beyond that, the research is less clear; there is little support for the cross-over benefit of, say, school chemistry on undergraduate biology, though here too there have been some dissenting voices (e.g. Brogt, Sampson, *et al*., 2011).

Researching the relationship between school mathematics and undergraduate outcomes is made more difficult by the limited availability of high-quality, system-level data. Whilst there are several interesting single-institution studies, it is not clear that such findings are generalizable within a country let alone between different countries with their particular systems of schooling and higher education. The availability of high quality datasets in England makes it possible to combine national cohort data of school performance at age 16 and 18 and link this with sector-wide undergraduate degree outcomes data. Such national analyses are rare in the education literature and although examples do exist they are not focused on the problem discussed herein (e.g. Shulruf, Hattie & Turnen, 2008).

Despite the dearth of research relating school mathematics qualifications to science degree outcomes, some small-scale studies have been reported from several countries including Australasia (e.g. Rylands & Coady, 2009), New Zealand (Comer *et al.* 2011), the Netherlands (Heck and van Gastel, 2006) and the US (e.g Tai, Sadler & Loehr, 2005). The Australian education system shares a common problem with England; because the study of mathematics to 18 is not compulsory, many students opt out of the subject even though they might then progress to mathematically demanding undergraduate programmes. This has caused as much concern in Australia (e.g. Brown, 2009) as it has in England (Royal Society, 2011) but it does raise an interesting question: does it matter whether students have advanced mathematics when progressing to science degrees? Perhaps the most pertinent Australian study at the University of Sydney (Nicholas, Poladian *et al*, 2015) suggests that whilst higher levels of mathematical study are in general terms beneficial, higher performance at a lower level of mathematics (e.g. at age 16) is also an important predictor of success, particularly for chemistry. Nicholas *et al.* conclude that it is important to consider the underlying mathematical ability of students as well as the level of mathematics studied as failure to do so would be to overlook potentially strong science candidates. Elsewhere in Australia, at Wollongong, Armstrong *et al.* (2014) conclude that high school mathematics rather than chemistry qualifications are the best predictor of general chemistry performance. In the United States, Spencer (1996) reported a single-site study which found that performance on the mathematics SAT is a good predictor of general chemistry attainment. Tai *et al*. (2005, p 1002) commented on “the striking role of preparation in advanced mathematics on college chemistry success” and other US studies have come to the same conclusion (e.g. Donovan & Welland, 2009). It should be noted that all of these chemistry studies are a) not of degree outcomes but ‘general’, introductory and first year modules, and b) are dealing with different types of students from different qualification landscapes. This second point might help to explain the different findings. In biology, Brogt *et al.’s* (2011) single-site study in New Zealand reports no association between school mathematics preparation and first year undergraduate outcomes. They do, however, highlight the importance of school chemistry in predicting success in undergraduate biology (see also Comer, Brogt & Sampson, 2011).

In England, little attention has been paid to careful statistical analysis of the relationship between school mathematics and degree outcomes, though much has been said about the problem of transition into mathematically demanding disciplines (e.g. Hulme & de Wilde, 2014). In one recent study, Darlington and Bowyer (2016) reported findings from a cross-sector survey of undergraduate students' perception of the usefulness of their pre-university mathematics qualifications. Though interesting, their study comprises a self-selected sample with a heavy bias to elite, Russell Group, universities. It offers little insight to the actual degree performance of these students which is the focus of the research reported herein.

There seems little doubt amongst many academics in science disciplinary communities that the effective application of mathematical and statistical techniques is necessary for good science learning (e.g. Hoban, Finlayson & Nolan, 2013). That said, the extant literature does not fully support these assumptions for biology and chemistry. Indeed, Nicoll and Francisco's (2001) analysis of physical chemistry performance concludes that neither the students nor the tutors were able to identify accurately the correlates of success. What is needed is a larger-scale statistical analysis of these questions that looks beyond the particular circumstances of a university site.

The research question of interest is not about the usefulness of mathematics for science in general – few would argue against scientists needing good mathematical skills - but rather whether school qualifications (in England) predict success in these subjects. The quality of the available datasets allows us to model how the whole system works and therefore raises questions about national qualifications, their effectiveness for preparing young people for undergraduate science study and the sorts of advice giving, academic expectations and policy making that should result. The literatures discuss the role of mathematics transition modules, support centres, effective entry diagnostics and 'maths for scientists' modules on degree programmes. All of these interventions make assessing any causal relationship between school qualifications and degree outcomes more complex and our analysis is not so bold as to have identified such causes.

The context – England

The level of mathematical and scientific competence amongst England's school and university graduates has come under the spotlight in recent years (for example, CBI, 2015). In 2010, the Nuffield Foundation reported that England had one of the lowest rates of post-16 mathematics participation amongst advanced economies (Hodgen, Pepper, Sturman & Ruddock 2010). As part of ongoing qualification reforms in upper secondary schools in England, the government is aiming for the 'vast majority' of young people to be studying mathematics up until the age of 18 (Gove 2011) by the end of this decade. New 'Core Maths' qualifications have been introduced to enable an additional quarter of a million young people each year to continue with their study of mathematics to 18. This mathematics gap is a national policy problem, but it does present researchers with an opportunity to examine the differential outcomes of students progressing to particular degrees with, and without, advanced mathematics qualifications. That in turn should inform the debates about the merits of making more young people do more maths, at least in terms of their subsequent educational outcomes.

Much has been written about 'the mathematics problem' in upper secondary schools. The Royal Society's (2011) report on the transition from school and college to university study of Science, Technology, Engineering and Mathematics (STEM) subjects urged greater uptake of advanced mathematics at A-level in order to facilitate the wide range of science, engineering and technology studies available. Whilst we do not have any reason to question the logic of this argument, there is an underpinning assumption that A-level Mathematics *per se* will enable learners to acquire the sorts of knowledge and skills required for success in these undergraduate studies. Given the pattern of A-level Mathematics participation in England, many undergraduate programmes – even in science - do not require applicants to have completed A-level Mathematics as this would reduce the applicant pool.

Before proceeding we briefly set out the qualifications system in England's upper secondary school and universities. At age 16 (Year 11) students take national assessments – the General Certificate of Secondary Education (GCSE) - and can achieve grades from A*-G in each subject. Grades C and above are considered 'good' grades and students with five or more of these typically continue on an academic pathway to study advanced, A-level qualifications. A-levels are the standard university-entrance qualifications in England and most students would study three subjects over the following two years, up to the age of 18. Degrees in England's universities are not awarded with a GPA score but rather with first, second and third class outcomes. Around 20% of students get a 'first'.

Research question

In the light of the issues raised above we investigated the performance of individuals at degree level in biology and chemistry with and without A-level Mathematics. Given that degree offers are made on the basis of predicted A-levels, but also known GCSE outcomes (and that GCSE Mathematics grades predict later mathematics participation, see Author, 2009) we also consider whether GCSE Mathematics is a better predictor of science degree success than A-level Mathematics. Our research question and sub-questions are as follows:

- Does completing A-level Mathematics increase the probability of gaining a first-class degree in biology and chemistry in England's universities?
    - Does the level of achievement in A-level Mathematics matter?
    - Does the effect of A-level Mathematics vary by university attended?

Consider, for example, two 16-year-old girls from similar schools and social backgrounds. They have near identical grade profiles including an A grade in GCSE Mathematics. Jane proceeds to study A-levels in chemistry, biology and psychology whilst Alice chooses mathematics, chemistry and geography and they happen to both gain grades AAB. They meet each other at university as they start their chemistry degrees. Will the fact that Alice did A-level Mathematics make her any more likely to achieve a first class degree? Of course our statistical modelling is not concerned with any one pair of such students, but rather the averages. The multiple other factors (many of them unknowable or unmeasurable) that make Alice and Jane unique will probably have more influence on their degree outcome than just possessing A-level Mathematics.

In our analyses we are keen to understand the educational predictors of academic success alongside the social ones (e.g. gender, ethnicity and socio-economic status). We are also interested to see how outcomes vary between universities. So, if Alice and Jane went to quite different universities would their chances of attaining a first class degree be significantly different and, if so, would their mathematics qualification help to explain this difference? We might ask whether, for example, prior mathematics or science attainment matters more in an elite vs a low-ranking university. The single-site studies discussed above are blind to this problem of sector-wide differentiation.

Method

Data

A bespoke dataset was requested from the Department for Education's (DfE) National Pupil Database that included demographic and qualification outcome data for all state sectors pupils from the Year 11 (age-16) cohort of 2005/6 along with their A-level results in years 2006/7 through to 2008/9. These records were linked, using unique identifiers, to the Higher

Education Statistics Agency (HESA) database up to the academic year 2012/2013. The dataset included gender, ethnicity, the Income Deprivation Affecting Children Index (IDACI), the free school meal indicator, codes for the title and category of the degree, degree classification, level of study, institution, socio-economic status and POLAR3, a postcode-based measure of the level of participation in higher education.

Data cleaning

The NPD and HESA databases are high quality datasets providing extensive information on pupil demographics, attainment, schools and universities attended. However, they did require cleaning to generate a final working dataset and this included dealing with some duplicate cases. For the GCSE and A-level data, we stripped out unnecessary variables to reduce the overall size of the dataset, retaining qualifications and demographic data. We then processed a unique record for each student at each stage, i.e. the one including the highest grade for each of the qualifications. Subsequently, each GCSE and A-level qualification was recoded into a points score. Total and average A-level and GCSE points measures were calculated and then superfluous variables were stripped out of the data frame.

The HESA dataset was straightforward to prepare. We merged the data frames and then recoded the subject title, institutional identification, degree classification, qualification obtained, level of study, year of programme, and individual ethnicity from their numerical integer codes into factor variables. We subset the data to include only the final case in which a degree was awarded. Lastly, where an individual had had obtained an additional degree, we kept the first awarded degree only.

Having merged the data frames some additional recoding was undertaken. The dependent variable was created (1 equals ‘has a first class degree’ and 0 equals ‘all other degree classification grades’). We simplified the ethnicity variable into White, Black, Asian, Chinese, Mixed and Other categories, normalized GCSE qualification grade data to create an average GCSE points score (excluding GCSE Mathematics), created a normalised mathematics GCSE score which was mean centred, a squared version of the mathematics GCSE score, normalized A-level grades to create an average A-level points score, standardized the IDACI score, created discrete indicators for A-level Mathematics, Biology and Chemistry, and lastly, converted all variables to either numeric variables or integers as appropriate.

Sample

The final sample was 6464 with biology degrees (including genetics, microbiology, molecular biology, biophysics, biochemistry and zoology) from 96 universities, and 1980 with chemistry degrees from 61 universities. Table 1 shows some of the demographic proportions in the sample.

Table 1

Proportion of graduates with a first class degree, cohort demographics and prior mathematics attainment.

	Biology graduates (%)	Chemistry graduates (%)
Has first class degree	19.2	27.3
Female	63.4	45.7
White	87.3	90.9
Black	1.4	0.7
Asian	5.5	3.5
Chinese	0.9	0.3
Mixed	3.6	3.1
Other	0.7	0.9
A grade in A-level Mathematics	31.5	48.0
B grade in A-level Mathematics	8.0	17.0
C or below grade in A-level Mathematics	4.5	3.7
Integrated Masters	N/A	73.7

The exploratory analysis of the sample shows differences between the groups. Of the 19.2% of graduates in biology and the 27.3% of graduates in chemistry with a first, the demographic proportions are quite different. The proportion of non-White students gaining first class honours degrees are substantially lower than for White students. For example, just 7.6% of Black biology students and 8.2% of Black chemistry students achieve a first class degree. For Asian biology and chemistry students, 11.3% and 15.1% gain a first class degree respectively. Lastly, for Chinese biology and chemistry students, 11% and 6.5% gain first class degrees respectively.

Modelling

The modelling compared having obtained a first class degree with another degree classification so we used logistic regression. To be more precise, our analyses use Bayesian hierarchical logistic regression modelling estimated by Markov Chain Monte Carlo (MCMC) methods. Such hierarchical or multilevel modelling is an extension of classic regression modelling. A core assumption of regression is that cases should be independent and identically distributed (i.i.d). However, education populations are typically more complex and are often naturally clustered. For example, students within particular degree programmes or universities are likely to be more similar than students taking other degrees or studying at a different university. Multilevel modelling relaxes the i.i.d. assumption and estimates coefficients in batches (e.g. by course or by institution) making adjustments to the regression model intercept and slopes.

We followed Little's (2006) 'calibrated Bayes' approach using classical approaches to model development, particularly in assessing the model fit, before moving to Bayesian approaches for the final models. We tested a number of single level models and basic varying intercept multilevel models, gradually increasingly complexity and making adjustments to improve interpretability. Once happy with the individual level predictors, we switched to using the Bayesian inference software STAN (Stan Team, 2015) in R. Five models of increasing complexity were fitted for both science subjects: 1) a null model, 2) single level model with all predictors, 3) a multilevel null model with university attended at level 2, 4) a varying intercept multilevel model with all predictors, and finally 5) a varying intercept and varying slope multilevel model. This final model is discussed below.

For biology, the model consists of an intercept, a coefficient for female, five coefficients for ethnicity (Black, Asian, Chinese, Mixed, Other) with White as the base category, a coefficient for the IDACI score which has been mean centred and standardized by two standard deviations, a coefficient for average GCSE points which has been mean centred, two coefficients measuring GCSE Mathematics score (a mean centred variable and its square) and a coefficient for total A-level points score which, like the previous, has also been mean centred. In addition, there are three batches of A level grades (A, B and C/D/E) for Mathematics, Biology and Chemistry. For these three batches, the base category is no A-level. Furthermore, the chemistry model differs slightly with an additional coefficient representing whether someone took a three year bachelor's degree or a four year integrated master's degree. Lastly, there is an adjustment for the intercept to allow the model to vary by university attended.

While the NPD and HESA datasets are high quality, there is missing data in both the biology and chemistry subsets. The pattern of missingness is non-monotone, but fortunately, missing data was confined to a small number of the variables. We used REALCOM-IMPUTE (Carpenter, Goldstein, & Kenward, 2011), a multilevel imputation programme to deal with the missingness. We imputed 10 datasets with the default burn-in of 100 iterations and a total of 1000 (post burn-in) iterations with imputations drawn at 100 iteration intervals. The results were then post-processed in STATA, saved as individual datasets and re-imported into R. The final five models for both subject subsets were estimated via Hamiltonian MCMC. We used 2000 iterations fitting a separate imputed dataset to each of the three chains and discarded half of the samples as ‘warmup’. Leave-one-out cross-validation (LOO-CV) was used for assessing model fit.

Results

The results raise interesting questions regarding the importance of advanced mathematics qualifications - in particular A-level Mathematics in England - for undergraduate science studies. We chose to model the subjects separately and in the following section we discuss model fit, data-level estimates, group-level variation and predictions. For all of the estimates presented we include a point estimate of probability (i.e. percentage) together with a 95% ‘credible interval’ in parentheses.

Model fit

Table 2 shows the change in the intercept (with standard errors) as well as standard deviation for the university-level, and the LOO-CV statistics - p_loo (a measure of the effective number of parameters in the model) and the looic (a measure of predictive accuracy on the deviance scale, penalised for a higher number of parameters in the model). The intercept represents the average (logged) probability of gaining a first for the baseline control group. A value below 0 indicates a probability of less than 50%.

Table 2

Change in intercept and improvements in model fit (biology) on the logarithmic scale with standard errors in parentheses.

Model	Intercept	$\sigma_{university}$	p_loo	looic
1. Null Model	-1.4 (0.0)	N/A	1.0 (0.0)	6339.2 (91.0)
2. Single-level	-1.4 (0.1)	N/A	21.2 (0.7)	5788.5 (93.0)
3. ML Null	-1.6 (0.1)	0.5	48.1 (1.3)	6179.9 (91.2)
4. ML Varying Intercept	-1.1 (0.2)	0.9	84.1 (2.9)	5523.1 (94.3)
5. ML Varying Slope	-1.1 (0.2)	0.8	99.8 (3.2)	5524.9 (94.3)

The Null model equates to the descriptive analysis in Table 1. In Bayesian statistics, parameters are treated as random which is why we report the mean as well as the lower and upper bounds of the credible interval in squared brackets. For biology, this equates to a probability of 19.2% [18.3 : 20.2] of getting a first class degree and for chemistry this is 27.3% [18.1 : 39.2] (see Table 3).

Table 3

Change in intercept and improvements in model fit (chemistry) on the logarithmic scale with standard errors in parentheses.

Model	Intercept	$\sigma_{university}$	p_loo	looic
1. Null Model	-1.0 (0.1)	N/A	0.9(0.0)	2324.3 (38.8)
2. Single-level	-1.3 (0.3)	N/A	21.7(1.2)	1984.7 (46.6)
3. ML Null	-1.2 (0.1)	0.6	26.3(0.9)	2273.2 (40.1)
4. ML Varying Intercept	-1.2 (0.3)	0.9	55.4(2.8)	1889.7(48.5)
5. ML Varying Slope	-1.2 (0.3)	0.8	65.6(3.2)	1895.2(48.4)

For the single-level model which contains all the coefficients (2), these probabilities (i.e. the intercept) stayed the approximately the same with a mean of 19.9% [16.5 : 23.6] for biology and a decrease to 21.9% [15.4 : 30.0] for chemistry. The control group is male and white, with mean scores for GCSE Mathematics, overall GCSE and A-level. They are one standard deviation below the mean for IDACI score, do not have a grade for A-level Mathematics, Biology and Chemistry and finally (for the chemistry subset only) have taken a bachelor's rather than an integrated master's degree. With the multilevel null model (3), the intercept changes slightly decreasing to 15.1% for biology and increasing to 23.1% for chemistry because individuals are now processed in batches according to the university attended. The intercept here refers the average proportion of those who gain a first class degree from each university. For the varying intercept, varying slope model (5), the biology mean intercept was approximately 24.4% and the chemistry mean intercept was 22.5%.

This variance between universities for the multilevel Null model (i.e. without predictors included) is 7.1% for biology and 9.9% for chemistry (see Table 2). The LOO-CV statistic favours more parsimonious models. For both the biology and chemistry subsets, the model which fits best is the varying intercept model. However, because our research question was concerned with whether an A-level Mathematics effect varies by institution, the results below report the final varying slope model although the model fit is not quite as good.

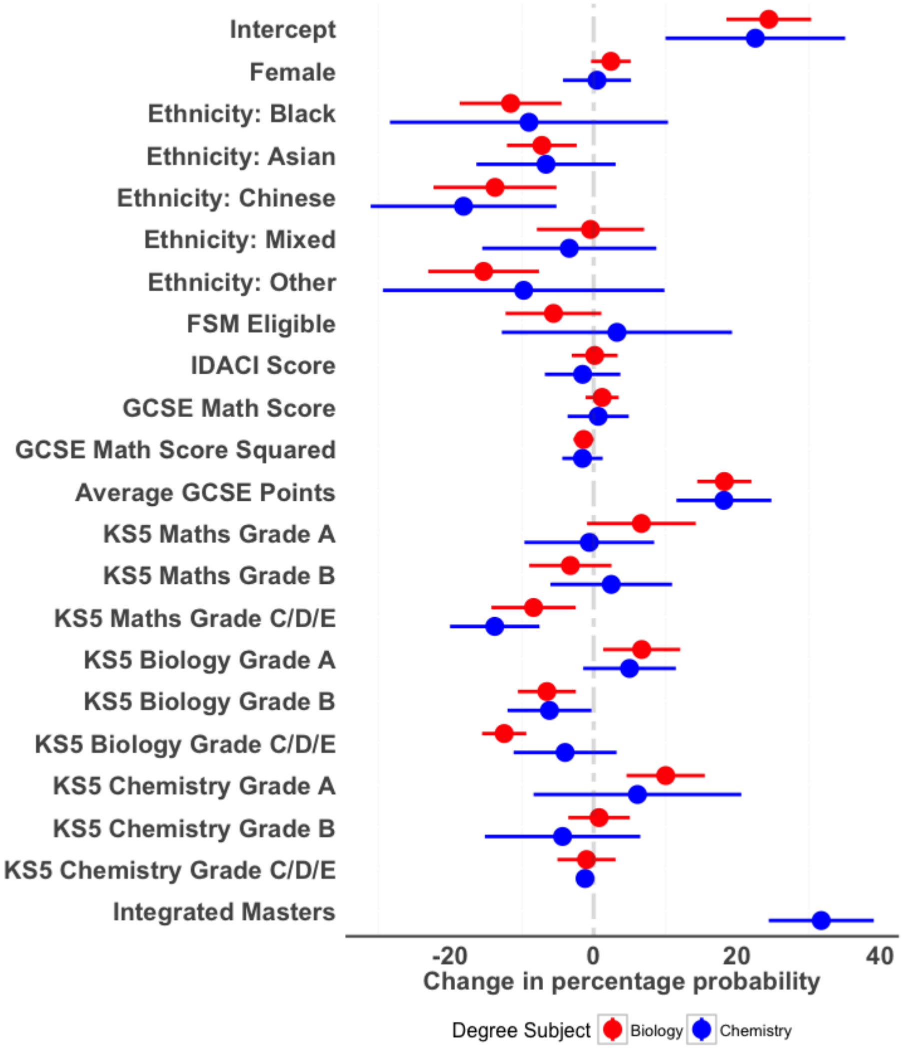
Individual level estimates

Figure 1 presents a summary of the parameter estimates for biology and chemistry graduates. The parameter values have a point estimate and uncertainty represented by the 95% credible intervals. As with classical statistical approaches, a credible interval which does not cross zero provides stronger evidence of an effect, but unlike strict interpretations of traditional confidence intervals, we can also interpret an interval where the majority of the density is either positive or negative as evidence of an effect, but with less confidence.

The estimates for biology suggest that females have on average 2.4% [-0.4 : 5.2] greater probability than the control group of gaining a first class degree. Ethnic minority students on the other hand fare less well. In particular, Black, Asian and Chinese students have much lower probabilities than the control group of gaining a first. For Black students the average effect is -11.5% [-18.7 : -4.5]. For Asian students the average effect is -7.2% [-12.1 : -2.4] and finally, for Chinese students the average effect is -13.8% [- 22.3 : -5.2]. The impact of free school meal eligibility status at age 16 (a proxy for coming from an economically deprived background) on the probabilities are reasonably strong with a mean of -5.6% [-12.2 : 1.1] and the impact for those with higher IDACI scores (another proxy for having lived in a more deprived neighbourhood) are minor, having a mean of 0.14% [-3.1 : 3.3]. *Overall, in undergraduate biology there is credible evidence of a small, positive effect for being female, a strong negative effect for most categories of ethnicity (compared to White), and some evidence of a negative effect for those students eligible for free school meals, all with regard to the overall probability of gaining a first class degree.*

Figure 1

Parameter estimates for the model variables indicating the percentage change in the probability of attaining a first class degree (A level here shortened to KS5, Key Stage 5)



For Chemistry graduates, there is little effect of being female; 0.5% [-4.3:5.2] greater probability of gaining a first-class degree, but the interval clearly crosses zero. In addition, while the estimates suggest that ethnic minority students are substantially less likely to gain a 'first', the estimates are very noisy, which is a product of the very small ethnic minority subsamples. Black students were on average 9.0% [-28.4: -10.4] less likely to gain a first compared to white students. In comparison, Asian students were on average 6.6% [-16.4:3.1] less likely and Chinese students were on average 18.1% [-31.1: -5.2] less likely to gain a first class degree compared to white students. The impact of the family background deprivation proxies are noisy with free school meal eligibility at age 16 having a mean effect of 3.2% [-12.8:19.3] but with a very wide interval. The IDACI measure had little effect and the probabilities for those one standard deviation higher than the mean for IDACI is centred on zero (-1.5), with a wide credible interval [-6.8:3.7]. *Overall, for chemistry, there is credible evidence of a negative effect for some ethnic minority students, albeit limited by the data, and little or no evidence of effects due to gender and deprivation.*

Prior attainment at age 16 (GCSE) has a major impact on the probability of gaining a first class degree. For biology, those individuals who were one unit above the average GCSE score, had an 18.2% [14.4 : 22.0] increase in the probability of attaining a first. However, the GCSE Mathematics score has little overall additional impact. For chemistry, the effect of the average GCSE is a mean increase of 18.2% [11.5 : 24.8] but here too there is little additional impact of the GCSE Mathematics score. *There is clear evidence that prior ability as measured by the average GCSE score has a strong effect on degree outcomes but there is no evidence that a higher ability in mathematics, over and above general attainment, has any additional impact on the probability of gaining a first-class degree.*

Prior attainment at age 18 (A-level) has a varied impact on the overall probabilities. In biology, there was clear effect for A-level Mathematics grade A, which increased the overall probability of obtaining a ‘first’ by 6.7% [-0.9 :14.2]. For those with a B grade for A-level Mathematics; the probability decreases by 3.3% although the tail of the distribution does intersect 0 [-9.0 : 2.5]. A-level Mathematics grades C/D/E have a clear negative effect with a mean of -8.4% [-14.3: -2.5]. Having an A-level Biology A grade had a positive effect with a mean of 6.7% [1.3: 12.1]. There were clear negative effects for A-level Biology grades B with a mean of -6.7% [-10.6: -2.5], and grades C/D/E which has a mean of -12.5% [-15.6: -

9.4]. For A-level Chemistry grades, Grade A has clear evidence of a strong positive effect of 10.0% [4.5: 15.5], but grade B and grade C/D/E have weak effects of 0.8% and -1.0%, with densities that clearly cross zero. Overall, for Biology students, while there is clear evidence that an A-level Mathematics grade A has a positive effect on the probability of gaining a first-class degree, with approximately 98% of the density of the interval being positive, students with A-level Mathematics grade B or below are less likely to achieve a first class degree than those without A-level mathematics, suggesting that this positive effect may be more of a proxy for underlying ability, rather than the impact of strong mathematics skills in general. *Overall, for biology graduates, there is clear evidence that A-level Mathematics grade A has a positive effect on the probability of gaining a first-class degree. Students with A-level Mathematics grade B or below are less likely to achieve the highest degree classification that those without A-level Mathematics. Strong performance (i.e. grade A) in A-level Biology, and even more so in A-level Chemistry, predicts contributes to success in biology degrees.*

For Chemistry students, an A grade in A-Level Mathematics had little effect on the chances of obtaining a first class degree with a mean of -0.6% [-9.7 : 8.4]; the Grade B effect has a higher mean of 2.4% [-6.0 : 11.0] but both of these intervals crossed zero. Only Grade C/D/E produced a clear effect with a mean of -13.8% [-20.0 : -7.6]. A grade A in A-level Biology had a mean effect of 5.0% [-1.5 : 11.5] indicating good evidence of a positive effect. Grades B and C/D/E show clear negative effects with means of -6.2% [-12.0 : -0.3] and -4.0% [-11.20 : 3.20] respectively. Lastly, A-level Chemistry grade A has a strong effect with a mean of 6.1% [-8.4 : 20.6] but the interval clearly shows substantial uncertainty. Grade B and Grades C/D/E show negative effects with means of -4.3% [-15.2 : 6.5] and -1.2% [-1.8 : -0.6]. In summary, for Chemistry undergraduate, grades A and B in A-Level Mathematics do not have clear positive effects, but grades C/D/E do have negative effects. *Overall, for chemistry students, grades A and B in A-Level Mathematics do not have clear positive effects* *on degree outcomes, but grades C/D/E do have negative effects compared to those with no A-level Mathematics. Strong performance in Biology A-level (i.e. grade A) appears to have a strong positive effect on chemistry degree outcomes.*

The last Chemistry parameter - Bachelors or Integrated Masters route - stands out from the rest. It has a clear positive effect with a mean of 31.7% [24.4 : 39.1] but the group is non-random. Universities generally only allow those performing well onto the Masters route so they are more likely to gain a higher degree classification.

Group-level Estimates

Figures 2 and 3 show the university-level residuals in rank order along with credible intervals. The institution attended plays an important role in predicting degree outcomes based on demographic and prior attainment. For biology graduates, based on the intercept adjustment and the mean predictions for the individual from the reference group, the average probability of gaining a first-class degree can vary considerably from 6.8% to 50.8%. For chemistry graduates, the average probability of gaining a first-class degree can vary from 4.9% to 47.6%. These are striking differences. The intervals vary in size due to the size of the cohorts.

For biology undergraduate, a move from having no mathematics at A-level to a Grade A increases the average probability of a first by 6.7% (from 24.4 to 31.1%). Adjusting for university attended this probability can vary more widely from 22.0% to 37.1%. For chemistry undergraduates, this same change decreases the average probability of a first-class degree by -0.6% (from 22.5% to 21.9%), but when accounting for the university attended, the average probability of gaining a first-class degree (for those with an A grade in Mathematics) varies from 15.5% to 28.2%.

Figure 2

University residuals and confidence intervals from multilevel chemistry model (5)

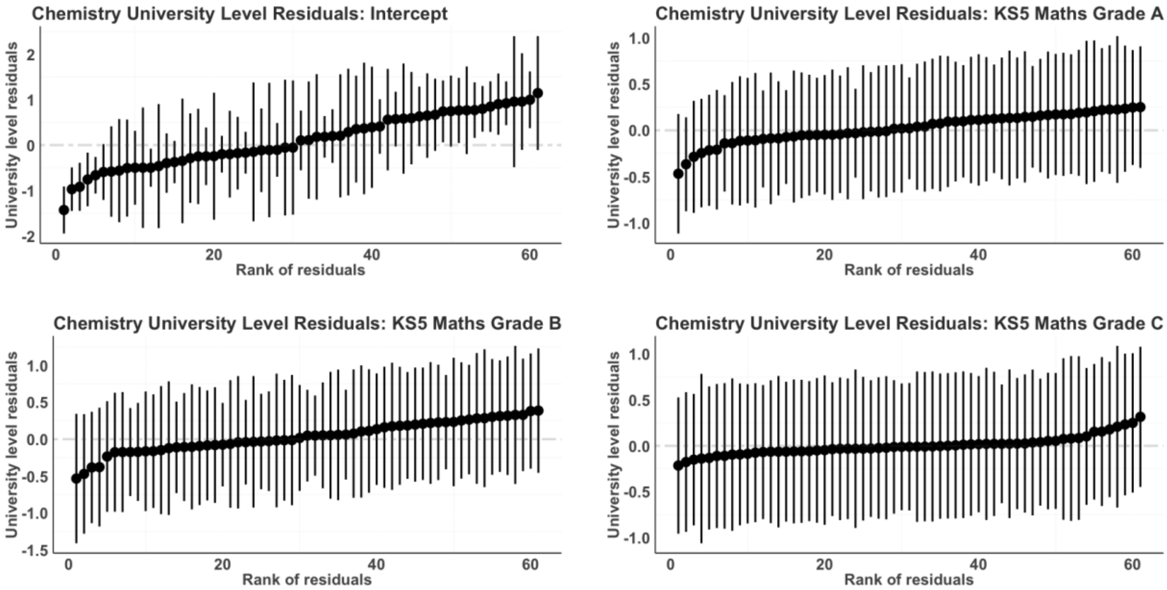
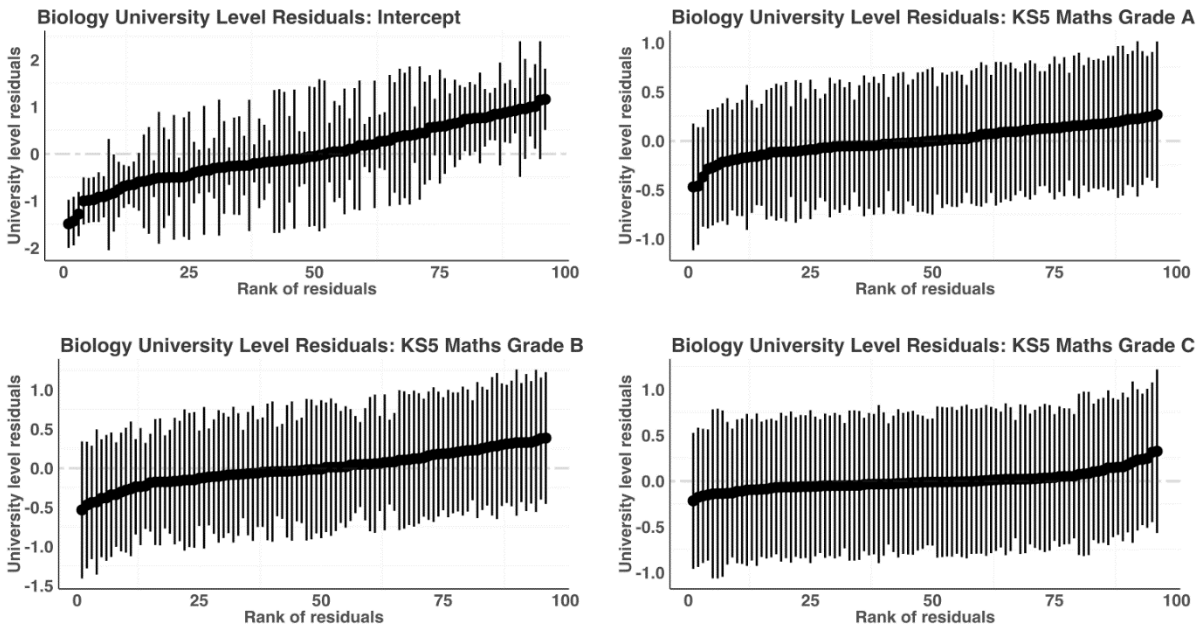


Figure 3

University residuals and confidence intervals from multilevel biology model (5)



Moving from no A-level Mathematics to a Grade B decreases the average probability of a first in biology by 3.3% (from 24.4 to 21.2%) but when also allowing for the university attended the average probability varies from 13.6 to 28.3%. For chemistry, the same change increases the average probability of a first-class degree by 2.4% (from 22.5 to 25.3%) but when adjusting for university attended the average probability varies from 18.3% to 30.1%. *Generally, there is more variation in the probability of achieving a first class degree due to the institution than whether or not one has A-level Mathematics (as shown in the wider range of logged probability changes on the left hand scale). It is also true to say that the effect of strong A-level performance is rewarded differently at different universities.*

Discussion

This research set out to test the idea that the completion of advanced mathematics might predict the likelihood of attaining the highest degree outcomes in biology and chemistry in English universities, i.e. a first class degree. We used a recent high-quality dataset to investigate this issue for a national cohort of students. However, the NPD/HESA dataset cannot tell us everything about the students or the courses on which they were enrolled. These limitations are discussed below in an attempt to explain our findings, which though somewhat counterintuitive do have precedence as discussed in the introduction.

In summary, the analysis of biology and chemistry degree outcomes (for this cohort of 16-year olds in England's schools in 2005/6), in particular the probability of achieving a 'first class' award, suggests that:

- Having A-level Mathematics does not consistently predict degree outcomes in biology and chemistry. Only biology students with top grades in A-level Mathematics are more likely to achieve the better degree classifications than those without A-level Mathematics;
- Those with a low grades in A-level Mathematics (C-E), or in other subjects for that matter, achieve lower degree outcomes than those without A-level Mathematics. This is unsurprising and more to do with general ability than mathematical skills;

- High attainment in A-level Chemistry (i.e. grade A) is associated with higher chances of gaining a first class degree in biology and vice versa, though the evidence is less compelling for the effect of A-level Biology grade A on undergraduate chemistry;
- Ethnic minority students on biology and chemistry undergraduate programmes in England are substantially less likely to attain a first class degree than White groups in nearly all cases;
- In biology, females have a small increased chance (2.4%) of attaining a first class degree. There is not a clear gender difference in chemistry;
- Social class differences in attaining a first class degree are not particularly clear, though there is some evidence that for biology degrees those from lower socioeconomic status backgrounds do a little less well;
- General academic attainment, as evidenced in the average GCSE score, is a strong indicator of likely performance at degree level, but there is no evidence that GCSE Mathematics attainment can be distinguished from general attainment at that age as a predictor;
- The university attended has a much greater bearing upon a student's chances of attaining a first class degree than does their participation, or attainment, in A-level Mathematics. For our reference group this probability can range from around 7 to 51% in biology and 5-48% in chemistry degrees.

The key finding in relation to our research question is that there is apparently no positive association between A-level Mathematics completion and degree attainment in biology and chemistry (with one exception). These findings fit with some of the single-site studies discussed earlier. It is important to consider this more carefully in order to avoid jumping to the erroneous conclusion that mathematics skills are not important for undergraduate science. Based on our knowledge of higher education in England, we propose four possible explanations for this finding. These are not independent and further research is needed to understand whether any or all of these are important:

1. Mathematical aptitude seems to be more important than specific qualifications. Those with good general academic ability at GCSE (including mathematics) have sufficient intellectual resources to handle the mathematics that they encounter in their degree programmes (see Nicholas *et al.,* 2015);
2. A-level Mathematics skills and knowledge might be not relevant or not easily transferred to the undergraduate biology and chemistry context;
3. There are probably qualitative differences in undergraduate module pathways and course experiences, both within and between universities. Those who feel more mathematically confident chose, or are directed to, more mathematically demanding pathways and so the A-level Mathematics effect is neutralised;
4. Universities might well be better placed to teach discipline-specific mathematics and do it more effectively with better motivated students. Any initial differences in A-level Mathematics completion are thereby smoothed, but underlying mathematical competence and confidence remains important.

Taken together, these proposed explanations raise questions about the nature and usefulness of A-level Mathematics for the preparation of undergraduate chemistry and biology students. In order to assess their explanatory usefulness of these arguments, different kinds of longitudinal study would need to be undertaken. This problem is discussed by Donovan and Wheland (2009, 381) who, despite seeing a predictive connection between school mathematics and university chemistry outcomes suggest that “it is possible that the demonstrated connection is not based on actual mathematical knowledge, but on the premise that mastering mathematics develops the same higher-order cognitive skills as those needed in science”. They continue to say that such skills need not necessarily come from the mathematics classroom. Little is known about the specific cognitive skills that are developed through A-level Mathematics (for an example see Attridge & Inglis, 2013) but this matter is of growing interest in light of studies showing the economic returns to higher level mathematics qualifications (Authors, 2016).

A-level Mathematics is currently being reformed with the intention that it will incorporate greater amounts of problem solving and modelling. Similarly, A-level sciences in England are now required to incorporate more contextualised mathematics and statistics in synoptic assessments (20% in A-level Chemistry and 15% in A-level Biology). This *embedding* policy is probably a good move and reflects the argument 4 above that science students might learn mathematics more effectively as part of their disciplinary studies. Nevertheless, the relationship between mathematics education and science education remains unclear; the notoriously difficult question of knowledge transfer between disciplinary domains requires further research. When these qualification changes have embedded, a different relationship between school qualifications and university outcomes may emerge but it will be some years before the relevant data are available to assess this.

Our analysis, as with all research, is limited by the quality of the available data. In the case of England, high quality cohort administrative datasets do exist and this study represents one of the first non-governmental analyses of this linked NPD-HESA data. Our approach was to track a particular national cohort of aged-16 learners along their subsequent education pathways. Following the same approach over successive years would show whether the relationships between school qualifications and university outcomes are changing over time. But such a study would still not take account of course structures, student choices and attitudes, pedagogy, and so on. The use of a simple outcome variable – first class degree – is somewhat restrictive, but reflects the rather blunt categorisations presently used in England's universities and recognised by employers. A longitudinal study of such a cohort would allow for those important qualitative factors to be incorporated into the analyses and further improve understanding of the interface between schools and universities. All that can be said for now is that there is little compelling evidence in England that completion of advanced mathematics will impact on student's chances of attaining the best degree outcomes in biology or chemistry, albeit with the qualifications discussed above.

References

Armstrong, B., Fielding, M., Kirk, S., & Ramagge, J. (2014). Factors affecting success in CHEM101 at UOW. *Australian Mathematical Society Gazette,* 41, 91-98.

Attridge, N., & Inglis, M. (2013). Advanced mathematical study and the development of conditional reasoning skills. *PLoS ONE*, 8(7), e69399.

Authors. (2009). Research paper. Research in Mathematics Education.

Authors. (2016). Research paper. British Educational Research Journal.

Brogt, E., Sampson, K. A., Comer, K., Turnbull, M. H., & McIntosh, A. R. (2011). Using institutional research data on tertiary performance to inform departmental advice to secondary students. *Journal of Institutional Research,* 16*(2),* 26-41.

Brown, G. (2009). Review of Education in Mathematics, Data Science and Quantitative Disciplines: Report to the Group of Eight Universities. *Group of Eight (NJ1)*. Retrieved from https://go8.edu.au/sites/default/files/docs/go8mathsreview_0.pdf

Carpenter, J., Goldstein, H., & Kenward, M. (2011). REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types. *Journal of Statistical Software*, 45(5), 1-14.

CBI. (2015). *Inspiring growth: CBI/Pearson education and skills survey 2015*. London: Confederation of British Industry/Pearson.

Comer, K., Broght, E., & Sampson, K. (2011). Marked for Success: Secondary School Performance and University Achievement in Biology. *Journal of Institutional Research,* 16(2), 42-53.

Darlington, E., & Bowyer, J. (2016). How well does A-level Mathematics prepare students for the mathematical demands of chemistry degrees? *Chemistry Education Research and Practice*.

Donovan, W. J., & Wheland, E. R. (2009). Comparisons of success and retention in a general chemistry course before and after the adoption of a mathematics prerequisite. *School Science and Mathematics,* 109(7), 371-382.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian Data Analysis*, 3rd Edition. Boca Raton, FL: Chapman & Hall.

Gove, M. (2011). Michael Gove speaks to the Royal Society on maths and science. London:

Department for Education. Retrieved from http://www.education.gov.uk/ inthenews/speeches/a00191729/michael-gove-speaks-to-the-royalsociety-on-maths-and-science.

Groen, L., Coupland, M., Langtry, T., Memar, J., Moore, B., & Stanley, J. (2015). The mathematics problem and mastery learning for first-year, undergraduate STEM students. *International Journal of Learning, Teaching and Educational Research,* 11(1), 141-161.

Heck, A., & van Gastel, L. (2006). Mathematics on the threshold. *International Journal of Mathematical Education in Science and Technology,* 37(8), 925-945.

Hoban, R. A., Finlayson, O. E., & Nolan, B. C. (2013). Transfer in chemistry: a study of students' abilities in transferring mathematical knowledge to chemistry. *International Journal of Mathematical Education in Science and Technology,* 44(1), 14-35.

Hodgen, J., Pepper, D., Sturman, L., & Ruddock, G. (2010). Is the UK an outlier? An international comparison of upper secondary mathematics education. London: Nuffield Foundation.

Hulme, J., & Wilde, J. D. (2014). Tackling transition in STEM disciplines: Supporting the Science, Technology, Engineering and Mathematics (STEM) student journey into higher education in England and Wales. York: Higher Education Academy.

Little, R. (2006). Calibrated Bayes: A Bayes/frequentist roadmap. *The American Statistician*, 60(3), 213-223.

Nicholas, J., Poladian, L., Mack, J., & Wilson, R. (2015). Mathematics preparation for university: entry, pathways and impact on performance in first year science and mathematics subjects. *International Journal of Innovation in Science and Mathematics Education,* 23(1), 37-51.

Nicoll, G., & Francisco, J. S. (2001). An investigation of the factors influencing student performance in physical chemistry. *Journal of Chemical Education,* 78(1), 99-102.

Royal Society. (2011). Preparing for the transfer from school and college science and mathematics education to UK STEM higher education: A 'State of the Nation' report. London: Royal Society.

Rylands, L., & Coady, C. (2009). Performance of students with weak mathematics in first-year mathematics and science. *International Journal of Mathematical Education in Science and Technology,* 40(6), 741-753.

Sadler, P. M., & Tai, R. H. (2007). The two high-school pillars supporting college science. *SCIENCE,* 317, 457-458.

Shulruf, B., Hattie, J., & Tumen, S. (2008). The predictability of enrolment and first year university results from secondary school performance: the New Zealand National Certificate of Educational Achievement. *Studies in Higher Education,* 33(6), 685-698. doi:10.1080/03075070802457025

Spencer, H. E. (1996). Mathematical SAT test scores and college chemistry grades. *Journal of Chemical Education,* 73(12), 1150.

Stan Team. (2015). *Stan: A C++ library for probability and* sampling, Version 2.7.0. Retrieved from http://mc-stan.org/

Tai, R. H., Sadler, P. M., & Loehr, J. F. (2005). Factors influencing success in introductory college chemistry. *Journal of Research in Science Teaching,* 42(9), 987-1012.