UNIVERSITY OF CAMBRIDGE

Predictive Effectiveness of Metrics in Admission to the University of Cambridge

This report summarises key findings of a recent study into indicators at undergraduate admission for entry to the University of Cambridge. The study, which involved both linear regression and multiple regression analysis, was completed by the Admissions & Data Services team at Cambridge Admissions Office (CAO), reporting to the Admissions Research Working Party (ARWP).[1]

Method

The primary relationship examined was that between metrics at admission and Cambridge University examination ('Tripos') percentages – in Part I exams in the years 2006-9.[2] A linear regression analysis produced correlation coefficients, on the Pearson -1 to +1 scale, between a range of indicators at point of admission and subsequent performance in Tripos.

The indicators examined were:

AS unit scores (UMS)[3]
GCSE results
STEP Mathematics[4]
BMAT (Bio-Medical Admissions Test)[4]
TSA (Thinking Skills Assessment)[4]

Other factors examined by the study included gender, school background, and the Cambridge Colleges' long-established use of an agreed weighting system by which English school performance data had been used to 'adjust' applicants' GCSE performance. The opportunity was taken to examine whether, in seeking to correlate AS performance with Tripos, a student's three 'best' AS subjects (in terms of UMS) provided better or worse correlations than their three 'most relevant' subjects.[5] The study also examined whether three AS subjects correlated better with Tripos than did four, and whether performance in individual, 'key' AS subjects (e.g. in Mathematics among Economics students) was especially pertinent.[6]

Finally, the study sought to identify, by Cambridge subject, the best 'model' in which various metrics at admission combined to predict Tripos performance. The idea here was that a combination of, say,

---

[1] Work was completed by Dwayne Carroll and Peter Chetwynd. Statistical advice was provided by William Peterson of the Faculty of Economics and John Bell of Cambridge Assessment. The study's findings were independently checked by statistician Eurof Walters.

[2] Cambridge undergraduate degrees are divided into two 'Parts': I and II. Part I, which is ordinarily sat at the end of the first and/or second years of study, is often subdivided into Parts IA (first year) and IB (second year). There are no 'finals' in Cambridge in the conventional sense, as different degree Parts are equally weighted.

[3] By 'AS' what is meant is all AS or A2 UMS achieved at point of application. Cambridge applicants provide these via a questionnaire. In the vast majority of cases such scores derived largely from AS units. Whether we included or excluded any A2 scores made virtually no difference to the study's findings. This was helpful, as excluding A2 scores is difficult in practice because of the nature of Maths A Level, in which many AS and A2 units are effectively interchangeable. No Critical Thinking or General Studies marks were included as these subjects do not form part of our offers.

[4] Information about STEP, BMAT and TSA is here: www.admissionstests.cambridgeassessment.org.uk/adt/

[5] Mathematics was treated as 'one' subject, even where both Mathematics and Further Mathematics were being taken. 'Best' subjects means the subjects in which the student achieved their highest overall marks. 'Most relevant' subjects means the subjects that are identified in the University undergraduate prospectus as essential, highly desirable or desirable to the proposed course of study. Where a student did not complete three 'relevant' subjects, 'relevant' subjects were included first and then 'best' subjects completed the tally of three.

[6] An additional study into the relationship between scores achieved in the International Baccalaureate (IB) and Tripos has shown a positive relationship between the two.

AS and GCSE performance might predict Tripos more effectively than a consideration of each element independently of the other. This was the multiple regression component of the study.

The study looked initially at all of Cambridge's larger Tripos subjects. The subjects examined (and the number of students in each study) were as follows:

Economics – Part I	(420 students)	Law – Part IA	(515)
Engineering – Part IA	(718)	Mathematics – Part IA	(565)
English – Part I	(472)	Medicine and Vet. Medicine – Parts IA & IB	(988 & 893)
Geography – Part IA	(304)	Modern & Medieval Languages (MML) – Part IA	(502)
History – Part I	(492)	Natural Sciences (NST) – Parts IA & IB	(1655 & 1457)

In a second phase of analysis, five smaller Cambridge subjects were also studied: Archaeology and Anthropology; Asian and Middle Eastern Studies (AMES); Computer Science; Music; and Politics, Psychology and Sociology (PPS). The size of the dataset in each case meant that the results obtained were, at best, at the margins of statistical significance. Nonetheless, it was striking that, with the exception of AMES, they reflected the pattern of results in the larger subjects.

In judging the correlation coefficients that the study identified, the yardstick that Cambridge Assessment generally uses has been applied. Judgments depend upon circumstances, but, broadly speaking, in terms of predicting future success, exams achieving correlations above 0.35 are deemed good, those above 0.4 very good, and those above 0.5 excellent. Such correlations are by no means easy to establish and sustain in any educational context.

Correlation coefficients and optimal models from the linear and multiple regression analyses

The summary table (Table 1) on the next page shows the study's key findings, which derive from both the linear regression analysis and the multiple regression analysis. Overall key conclusions are listed below, but first the last two columns in the table should be explained. These are provided essentially for information and by way of context. The factors at admission listed in the penultimate column are those that, together, constitute the best predictive model for Tripos performance in the subject in question. However, it should be strongly stressed that, for all subjects other than Mathematics, AS UMS is overwhelmingly the best indicator in these models. Where other factors are listed, they contribute to producing the highest predictive validity, but only at the margins. In other words, were AS the only factor included, the model would predict almost as effectively; only marginal improvement is brought by the inclusion of additional factors. For Mathematics, the key indicator is STEP III but STEP II and GCSE also contribute to the best model. The final column, adjusted R^2, indicates (where available) the confidence that one may have in the effectiveness of the various statistical models. Broadly speaking, a score of ~ 0.4 or above indicates that a model is one in which high confidence may be placed, and any score above ~ 0.2 is significant. Scores below ~ 0.15 are not.

It is also worth commenting at this point on questions of school background and gender, as these factors appear in many of the optimal predictive models. Although they do appear, they are not statistically significant, making (as has been said) only a marginal difference, and the pattern – particularly in relation to school background – is in any case inconsistent. Using school and college A Level performance data to 'weight' exam results at point of application did little to improve correlations with Tripos performance. The study also found that Cambridge's long-established weighting system for GCSE had now become ineffective in improving correlations between GCSE and Tripos performance. In general, GCSEs were best assessed simply by counting the number of A*s, with no cap on the total. However, it was further discovered that, among state-sector students from lower performing schools (those whose school GCSE capped score was below 40), GCSEs correlated with Tripos more strongly than was generally the case, the strength of the correlation mounting as school performance declined. This suggests that the Cambridge admissions process should pay particular attention to students with very good GCSE results from lower performing schools. They are likely to have the potential to do well. (It does not suggest that allowance should routinely be made for weaker GCSE results.)

One other issue is probably best commented on here: the TSA correlations. Although the correlations between the 'problem solving' component of the TSA and performance in Engineering and Natural Sciences are, on the face of it, encouraging, the overall coefficient depends upon four years of data, and this masks a significant difficulty with using the TSA in admissions. This is that, on a year-by-year basis, its performance has actually been somewhat inconsistent. There are years in which it has correlated well and others in which the correlations have essentially disappeared.[7] It is not known why this inconsistency occurs.

Table 1: correlation coefficients and optimal models[8]

Subject	Part	AS UMS	GCSE	STEP II	STEP III	BMAT Section 1	BMAT Section 2	TSA critical thinking	TSA problem solving	Components in best multiple regression model	Best model adjusted R^2
Economics	I	0.36	0.31					0.19	0.1	AS, TSA, school, gender	0.29
Engineering	IA	0.42	0.13					0.18	0.25	AS, TSA prob, gender	0.44
English	I	0.40	0.29							AS, gender, school	0.17
Geography	IA	0.36	0.28							AS, gender, school	0.2
History	I	0.39	0.24							AS, GCSE, gender, school	0.37
Law	IA	0.32	0.32							AS, gender, school	0.16
Mathematics	IA	0.22	0.31	0.47	0.54					STEP III, STEP II, GCSE	0.4
MedVet	IA	0.38	0.27			0.19	0.26			AS, BMAT 1, GCSE, school	0.29
MedVet	IB	0.37	0.25			0.14	0.21			AS, GCSE, BMAT 1, gender, school	NA
MML	IA	0.47	0.19							AS, gender, GCSE, school	NA
NST	IA	0.48	0.26					0.14	0.23	AS, GCSE, TSA, gender, school	0.38
NST	IB	0.40	0.29					0.08	0.11	AS, gender, GCSE, TSA, school	NA

Key conclusions of the study

1) AS UMS have provided a sound to verging on excellent (mean = 0.38) indicator of Tripos potential in every major subject Cambridge offers, with the exception of Mathematics. In Mathematics AS has been a much less effective predictor than STEP.

2) GCSEs have mostly correlated reasonably with Tripos (mean = 0.26) but have largely been a less effective predictor than AS UMS and have generally added little to the predictive validity of a model once the best indicator has been included. In short, AS has trumped GCSE for every subject other than Mathematics. For Mathematics, GCSEs have correlated better with Tripos than have AS scores – though they have remained a less effective indicator than both STEP III and STEP II.

3) STEP III and STEP II, in that order, have predicted the Mathematical Tripos very well (mean = 0.5).

[7] The same conclusion has been reached by a study completed by Cambridge Assessment, who provide the TSA.

[8] In Table 1 the main body of the table displays correlation coefficients between, on the one hand, Tripos results in the subjects appearing as row headings and, on the other, performance variables at time of admission as indicated by the column headings. The final column gives the square of the multiple correlation coefficient. This represents the proportion of the variability in the Tripos results accounted for by the best model as described in the penultimate column.

4) Aptitude tests have been a less effective predictor of Tripos performance overall (mean = 0.18) but the BMAT has had a positive utility, especially in the absence of AS and/or GCSE, and the TSA would arguably have been useful in a similar way had it been more consistent.

5) In general, for Arts subjects the correlations for AS have worked best if one has taken into account the 'best' (in terms of UMS achieved) subjects, whereas for Sciences the highest correlations have been achieved by counting the 'most relevant' subjects first, and the 'best' ones thereafter. The 'best' three AS subjects have produced strong correlations with Tripos on the Arts side even in 'technical' subjects, such as Economics and MML – though in these two cases the difference between 'best' and 'most relevant' three has been marginal.

6) Three AS subjects (Mathematics being counted as one subject, even where both Maths and Further Maths have been sat) have generally provided better correlations with Tripos than have four AS subjects, though in most cases the difference has been small.

7) The effectiveness with which achieved AS/A2 UMS have predicted Tripos has not varied according to gender, or to school or college background. Given the same examination results at admission, students from different schools and colleges, and from the state and independent sectors, have been equally likely to perform well in Cambridge.

8) In general, GCSEs have best been assessed simply by counting the number of A*s, with no cap on the total. However, very good performance at GCSE among students from schools whose GCSE capped score is below 40 has indicated potential in Tripos.

Some caveats and future direction

The major limitation of this study is that it has only been able to look at students who have made it through the Cambridge admissions process. They are, by definition, a select group, the data range of whose qualifications at point of entry is relatively narrow. Nonetheless, the consistency of this study's findings is striking.

An additional limitation to the study is the range of subjects and year-groups examined. The dataset enabled the study to look at Part IB as well as Part IA in two subjects: Medicine and Natural Sciences. The results of this are arguably encouraging in that the correlations for IA held up well for IB. Any study like this will show that correlations with qualifications at point of entry decline across time – for obvious reasons. It is intended that the existing dataset will continue to be used and will be added to as new students matriculate in the University and duly sit Tripos.

As we move ahead, plans are in place to run a similar analysis in the summer of 2011 of A2 results achieved by the end of Year 13. In this, our new First Year's school/college exam results will be set against their performance in Part IA of Tripos. The question that we will be seeking answers to in particular is whether achievement of A*s at A Level predicts Tripos success. A secondary question that should be addressed is how far performance at AS predicts performance at A2.

Richard Partington
Senior Tutor, Churchill College
Chair, Admissions Research Working Party
February 2011