UNIVERSITY OF CAMBRIDGE

A* at A Level as a Predictor of Tripos Performance: An initial analysis

Summary

This paper explores the interrelationship between achieved A* grades at A Level and University of Cambridge formal examination ('Tripos') performance at the end of students' first year of study in 2011. We believe that this is the first examination of the utility of A* as an indicator of potential for academic success at any university to have been made public.

The paper's key findings are as follows:

a) The more A* grades a student achieves at A Level, the likelier they are to do well in first-year examinations in Cambridge. This is true in Arts, Social Sciences and Sciences subjects;

b) Equally well qualified students (in terms of A* grades achieved) from state and independent schools and colleges are equally likely to prosper in the first year of Tripos, i.e. there is no 'sector gap';

c) In Sciences, students with achieved grades A*A*A* or better have significantly better prospects in first-year University exams than those with lower grades;

d) In Arts and Social Sciences the key divide is at grades A*A*A. Students whose record at entry is at or above this point are likely to outperform in Tripos students whose achieved grades are A*AA or lower at A Level.

Introduction

This paper tests two specific hypotheses about the relationship between A Level A* achievement and Tripos performance:

H_1 There is a positive relationship between number of A*s achieved at A Level and subsequent Tripos performance;

H_2 The relationship between A* count and Tripos performance differs with students' background in terms of educational sector, state against independent.

It is our expectation, based on previous analyses,[1] that H_2 will prove false – that is, that we will find educational sector to have no significant impact on the relationship between A Level A*s and Tripos performance.

To test these hypotheses, we will be using Analysis of Variance (ANOVA) to compare average Tripos performance by A Level achievement and by educational sector. This statistical method compares the difference in average scores between groups with the degree to which individual scores vary within each group, in order to determine whether the between-group differences are likely to have occurred by chance. If the result of the ANOVA is significant, it suggests that there is a genuine difference in average Tripos performance between some or all of the groups.

Where we find such a difference, we can use post-hoc tests to determine where (i.e. between which groups) the significant differences lie. Since the number of cases varies

[1] See, for example, G. Parks, 'Academic Performance of Undergraduate Students at Cambridge by School/College Background' (2011), and R. Partington *et al.*, 'Predictive Effectiveness of Metrics in Admission to the University of Cambridge' (2011), both at www.cam.ac.uk/admissions/undergraduate/research/.

substantially between groups, Hochberg's GT2 has been chosen as a robust test for these post-hoc comparisons.

Data used in the analysis

There are three variables under consideration in this analysis:

- *Educational sector*
  UCAS provides an indication of each applicant's 'apply centre'. For school-age applicants this is the school or college at which they are currently studying. We match the UCAS apply centre code to our schools database which conflates data from UCAS, DfE and other sources to determine whether the apply centre is a UK maintained or independent school or college, with all other apply centres designated 'Other and Overseas'. It is possible (although rare) for a school-age applicant to apply independently by not providing their school's apply centre; these applicants would be classified as 'Other and Overseas'.
- *A* Count at A Level*
  UCAS supplies information on applicants' A Level examination results through its Awarding Body Linkage; this data is sourced directly from A Level Awarding Bodies. This provides all A and AS Level grades taken in the summer sitting (2010 for the purposes of this paper), as well as any taken in the previous 18 months. A or AS Levels completed earlier than this would not be included.
- *Standardized Tripos percentage*
  Although overall percentage score is rarely used in isolation to determine a student's degree result, it provides a better measure than class marks for much statistical analysis, since it is continuous (rather than categorical) data and it is normally distributed for each subject. However, the means and standard deviations of these raw percentages vary substantially between subjects. To correct for this, we have standardized the percentages for each sitting of each Part of each Tripos (e.g. Natural Sciences Part IA Easter 2011). This provides standardized scores (z-scores) for each sitting with a mean of 0 and a standard deviation of 1.[2]

Because the A Level A* grade was only introduced in 2010, we are rather limited in the amount of data available to us. We can currently only consider 2010-cycle A Level students for whom we have first-year Tripos results from summer 2011. Four subjects are excluded from this analysis owing to a lack of first-year Tripos results: Anglo-Saxon, Norse & Celtic; English; History; and Modern & Medieval Languages. In addition, all students who did not sit A Levels, or who failed to achieve at least three As at A Level, have been excluded from the study.[3]

ANOVA for all subjects

The two-way independent ANOVA used for this study compares mean standardised Tripos percentage score between 15 groups, combining the three educational sector categories and the five A Level profile categories. It then tests three effects:

- Main effect of A Level profile;
- Main effect of educational sector;
- Interaction effect of A Level profile and educational sector.

[2] Scores have been standardized over all students sitting exams, not just A Level students, so the 'Total' means found in the final row of Table A1 in the appendix differ a little from the overall means of 0.
[3] A small number (N<20) of students are accepted by Cambridge each year without achieving three As at A Level. Since these acceptances usually involve consideration of special circumstances, it was not felt that this small group would provide an unbiased sample of A*A*B/A*AB students.

The two main effects measure whether there is variation in Tripos performance by each variable acting on its own, while the interaction effect measures whether there is additional variation when the two variables are considered together (i.e. whether the relationship between one variable and Tripos performance changes when we change the value of the other variable). The key results of the ANOVA are listed in Table 1 below, while group means are provided in Table A1 in the appendix.

Table 1: F-statistics for all subjects ANOVA[4]

	df	F	p	partial η^2
A Level profile	*4, 1837*	*12.16*	*<.0005*	*.026*
Educational sector	2, 1837	1.84	.159	.002
Interaction effect	8, 1837	1.79	.075	.008

Neither the main effect for educational sector, nor the interaction effect of A Level profile and educational sector, have reached the threshold for statistical significance ($p<.05$), and so we can reject H_2, since we have not observed a significant difference between groups either by educational sector alone or by a combination of sector and A Level profile.

The main effect of A Level profile has proven highly significant, however, suggesting that there is a genuine difference in A Level performance as A* count changes. To accept H_1, though, we must also demonstrate that this effect is positive and consistent – that is, that Tripos performance improves as A* count increases. For this detail, we must turn to the post-hoc analysis, detailed in Table 2.

Table 2: Hochberg's GT2 post hoc comparisons for A Level profile

Profiles compared		Mean Difference	Std. Err.	p
A*AA/AAA	*A*A*A*	*-0.247*	*0.067*	*.002*
	*A*A*A**	*-0.557*	*0.063*	*<.0005*
	*A*A*A*A**	*-0.939*	*0.064*	*<.0005*
	*A*A*A*A*A*+*	*-1.003*	*0.103*	*<.0005*
A*A*A	*A*A*A**	*-0.310*	*0.061*	*<.0005*
	*A*A*A*A**	*-0.692*	*0.062*	*<.0005*
	*A*A*A*A*A*+*	*-0.756*	*0.101*	*<.0005*
A*A*A*	*A*A*A*A**	*-0.381*	*0.058*	*<.0005*
	*A*A*A*A*A*+*	*-0.446*	*0.099*	*<.0005*
A*A*A*A*	A*A*A*A*A*+	-0.064	0.100	.999

Since the mean differences are all negative, it is clear that there is a positive relationship between A Level A* profile and mean Tripos performance: each group has a lower mean performance than all of the higher-ranked groups. There is only one case where this difference is not significant, which is the comparison between students with four A*s and those with five or more. We can accept H_1 as true, noting that average Tripos performance increases as number of A Level A*s increases, but with the proviso that there is no significant further improvement above four A*s.

ANOVA by subject group

In addition to identifying this overall trend, it may be instructive to briefly consider whether these patterns hold true when we separate the three subject groups – Arts, Science & Technology, and Social Sciences. The Arts group will, necessarily, be very incomplete,

[4] Statistically significant values are italicised here and throughout the paper.

since we are unable to include four subjects – Anglo-Saxon, Norse & Celtic; English; and History, which do not have first-year Tripos exams; and Modern & Medieval Languages, in which a single set of results are not provided. Together these comprise a large proportion of the Arts cohort.

For the subject-group analysis, we have also excluded students from the 'Other and Overseas' educational sector. As can be seen from the Table A1, there are very few students within this sector (73 total), leaving us with a small number problem when we further divide into separate subject groups, with the performance of only a handful of students being used to determine a group mean.

Table 3, below, summarises the ANOVA findings for each subject group. Details of mean values can be found in Table A1 in the appendix.

Table 3: F-statistics for subject group ANOVA

		df	F	p	partial η^2
Arts	*A Level profile*[5]	*3, 294*	*17.02*	*<.0005*	*.148*
	Educational sector	1, 294	1.10	.295	.004
	Interaction effect	3, 294	0.352	.788	.004
Science and Technology	*A Level profile*	*4, 1127*	*65.15*	*<.0005*	*.188*
	Educational sector	1, 1127	3.02	.082	.003
	Interaction effect	4, 1127	0.324	.862	.001
Social Sciences	*A Level profile*	*4, 330*	*9.99*	*<.0005*	*.108*
	Educational sector	1, 330	0.176	.675	.001
	Interaction effect	4, 330	1.08	.368	.013

Here we see the same pattern as for the all-subjects analysis. In each group, there is a significant main effect for A Level profile, but no significant main effect for educational sector nor a significant interaction effect. Exploring the post-hoc comparisons, however, indicates distinctions in the pattern of the A Level effect between subject groups.

For the Arts subject group, as can be seen from Table 4 below, we see the expected pattern of better Tripos performance being associated with a higher A* count, but there is no significant difference between students with two A*s and those with three A*s. Since there was only one student with five A*s (whose group has thus been merged with the four-A* group), we are unable to consider whether the distinction between four and five or more A*s is also insignificant.

Table 4: Hochberg's GT2 post hoc comparisons for Arts subjects

Profiles compared		Mean Difference	Std. Err.	p
A*AA/AAA	*A*A*A*	*-0.419*	*0.122*	*.004*
	*A*A*A**	*-0.678*	*0.141*	*<.0005*
	*A*A*A*A*+*	*-1.340*	*0.228*	*<.0005*
A*A*A	A*A*A*	-0.259	0.146	.379
	*A*A*A*A*+*	*-0.922*	*0.231*	*.001*
A*A*A*	A*A*A*A*+	-0.662	0.242	.038

For the Science & Technology subject group (Table 5), there is again a clear pattern of improvement in mean Tripos performance as A* count at A Level increases. This time, we

[5] There was only one Arts student with five or more A*s who took exams in Easter Term 2011. For the Arts group analysis, therefore, the four A* and five or more A* groups have been combined.

can see a repeat of the overall pattern for no significant difference between the four A* and five or more A* groups. However, in a break from the overall pattern, we can also see that there is not a significant difference at the bottom end of the scale, between the A*AA/AAA group and the A*A*A group.

Table 5: Hochberg's GT2 post hoc comparisons for Science and Technology subjects

Profiles compared		Mean Difference	Std. Err.	p
A*AA/AAA	A*A*A	-0.179	0.105	.599
	*A*A*A**	*-0.695*	*0.096*	*<.0005*
	*A*A*A*A**	*-1.177*	*0.094*	*<.0005*
	*A*A*A*A*A*+*	*-1.218*	*0.125*	*<.0005*
A*A*A	*A*A*A**	*-0.516*	*0.080*	*<.0005*
	*A*A*A*A**	*-0.998*	*0.078*	*<.0005*
	*A*A*A*A*A*+*	*-1.039*	*0.113*	*<.0005*
A*A*A*	*A*A*A*A**	*-0.482*	*0.066*	*<.0005*
	*A*A*A*A*A*+*	*-0.523*	*0.105*	*<.0005*
A*A*A*A*	A*A*A*A*A*+	-0.041	0.103	1.000

With Social Sciences (Table 6a), half of the post-hoc comparisons are not statistically significant, although the mean differences still show the trend we would expect. In particular, the five or more A* group has not proven significantly different from any group except the lowest (A*AA/AAA). Looking at Table A1 for the underlying data, it seems likely that this can be attributed to the small number of cases in the five or more A* group (N=8).

Table 6a: Hochberg's GT2 post hoc comparisons for Social Sciences subjects

Profiles compared		Mean Difference	Std. Err.	p
A*AA/AAA	*A*A*A*	*-0.486*	*0.131*	*.002*
	*A*A*A**	*-0.676*	*0.131*	*<.0005*
	*A*A*A*A**	*-0.953*	*0.156*	*<.0005*
	*A*A*A*A*A*+*	*-1.330*	*0.330*	*.001*
A*A*A	A*A*A*	-0.190	0.131	.800
	*A*A*A*A**	*-0.467*	*0.156*	*.030*
	A*A*A*A*A*+	-0.844	0.330	.105
A*A*A*	A*A*A*A*	-0.277	0.156	.548
	A*A*A*A*A*+	-0.654	0.330	.389
A*A*A*A*	A*A*A*A*A*+	-0.377	0.341	.955

Table 6b offers an alternative set of post-hoc comparisons where the five or more A* group has been combined with the four A* group, as in the Arts subjects analysis. Here we can see that there is a significant difference between students with only two A*s and those with four or more A*s, when taken together. However, both tables show that there is no significant difference between students with two A*s and students with three A*s, and also that there is no significant difference between students with three A*s and those with four or more.

Table 6b: Hochberg's GT2 post hoc comparisons for Social Sciences subjects (capped at 4+ A*s)

Profiles compared		Mean Difference	Std. Err.	p
A*AA/AAA	*A*A*A*	*-0.486*	*0.131*	*.001*
	*A*A*A**	*-0.676*	*0.131*	*<.0005*
	*A*A*A*A*+*	*-1.004*	*0.148*	*<.0005*
*A*A*A*	A*A*A*	-0.190	0.131	.620
	*A*A*A*A*+*	*-0.518*	*0.149*	*.003*
*A*A*A**	A*A*A*A*+	-0.328	0.149	.158

Conclusions

The ANOVA described in this paper has successfully demonstrated that:

a) there is statistically significant variation in Tripos performance dependent upon the number of A*s achieved at A Level;

b) Tripos performance does not vary by educational sector, either in isolation or as an interaction with number of A*s.

In addition, post-hoc comparisons have shown that there is a clear trend, across all subjects, for average Tripos performance to increase as number of A*s increases, although this increase ceases to be significant above four A*s.

Examination of subject groups has revealed the same basic trends for each of the three subject groups, but with fewer statistically significant differences between A Level performance groups. Specifically, for Science & Technology subjects there was no significant difference between students with A*AA/AAA and those with two A*s, while these groups were significantly different for other subjects. Both Arts and Social Sciences subjects showed a lack of significant difference between A*A*A and A*A*A*, with Social Sciences also showing no significant difference between A*A*A* and four or more A*s. From these findings, we might conclude that A*A*A* presents a useful lower threshold for higher-performing Science & Technology students, while A*A*A would be a more appropriate threshold in Arts and Social Sciences subjects.

When we look at all subjects together, A Level A* count, while statistically significant, does not seem to have a particularly large effect – the partial η^2 value of .026 suggests that number of A*s can only explain 2.6% of variance in Part 1/Part 1A Tripos results. It seems, however, that this low value is partly caused by the differences in effect between subject groups. The partial η^2 values for each subject group are substantially higher: .148 for Arts, .188 for Science and Technology, .108 for Social Sciences.

With only a single year's worth of results, and missing results for three large Arts subjects, it is too early to draw any firm conclusions about the relationship between A* count and Tripos performance. However, these preliminary findings certainly support our initial position – there is a positive relationship between A* count and Tripos performance, and this does not seem to vary significantly between educational sectors.

Dr Peter Chetwynd
Research Officer
December 2011

Appendix: Means table

Table A1: Mean standardized Tripos performance by educational sector and A Level profile

		All subjects			Arts			Science and Technology			Social Sciences		
A Levels	Educational Sector	Mean	St. dev.	N	Mean	St. dev.	N	Mean	St. dev.	N	Mean	St. dev.	N
A*AA/AAA	Independent	-0.568	0.931	108	-0.560	0.937	51	-0.855	0.808	30	-0.263	0.979	27
	Maintained	-0.527	1.002	224	-0.275	1.064	71	-0.736	0.957	85	-0.528	0.942	68
	Other and Overseas	-0.215	1.324	23									
	Total	-0.519	1.005	355	-0.395	1.018	122	-0.767	0.919	115	-0.453	0.955	95
A*A*A	Independent	-0.339	0.891	152	-0.095	0.922	54	-0.662	0.825	66	-0.083	0.770	32
	Maintained	-0.248	0.938	235	0.164	0.685	46	-0.550	0.916	128	0.095	0.927	61
	Other and Overseas	-0.065	0.889	16									
	Total	-0.272	0.918	403	0.024	0.828	100	-0.588	0.885	194	0.034	0.876	93
A*A*A*	Independent	-0.047	0.908	204	0.187	0.921	32	-0.162	0.867	124	0.094	0.972	48
	Maintained	0.080	0.906	291	0.386	0.791	30	-0.021	0.916	216	0.361	0.837	45
	Other and Overseas	0.325	0.689	18									
	Total	0.038	0.903	513	0.283	0.859	62	-0.072	0.900	340	0.223	0.914	93
A*A*A*A*[6]	Independent	0.442	0.907	205	1.032	0.724	9	0.403	0.923	175	0.518	0.768	21
	Maintained	0.435	0.848	260	0.859	0.451	9	0.415	0.857	222	0.488	0.861	30
	Other and Overseas	-0.189	1.179	14									
	Total	0.420	0.888	479	0.946	0.592	18	0.410	0.886	397	0.500	0.816	51
A*A*A*A*A*+	Independent	0.374	0.787	47				0.358	0.801	45	0.723	0.117	2
	Maintained	0.604	0.867	53				0.541	0.883	46	0.929	0.681	6
	Other and Overseas	-0.094	0.803	2									
	Total	0.484	0.834	102				0.451	0.844	91	0.877	0.585	8
All AAA+	*Independent*	-0.020	0.972	716	-0.126	0.997	146	-0.007	0.979	440	0.054	0.916	130
	Maintained	-0.006	0.990	1063	0.047	0.942	156	-0.029	1.00	697	0.030	0.986	210
	Other and Overseas	-0.041	1.058	73									
	Total	-0.013	0.985	1852	-0.037	0.971	302	-0.020	0.992	1137	0.039	0.959	340

[6] Figures for the Arts subject group include a single student who achieved more than four A*s.