Predicting Academic Performance

Marcos Gallacher

Universidad del CEMA
(gmg@cema.edu.ar)

Resumen

El objetivo de este trabajo es analizar ventajas y desventajas asociadas a la utilización de "pruebas de admisión" como predictoras de performance en programas de estudio de grado. El trabajo analiza información relativa a performance de estudiantes de las carreras de economía y de dirección de empresas. Esta performance se vincula con resultados en examenes de admisión. El trabajo analiza aspectos adicionales relativos a performance de alumnos, en especial (i) ritmos diferenciales de progreso y (ii) diferencias en el grado en que los alumnos tienen "áreas de interés/habilidad" definidas. El trabajo concluye que los examenes de admisión constituyen una herramienta de valor aún cuando las predicciones que de ellos se derivan distan de ser perfectas.

Summary

This paper discussed advantages and disadvantages associated with the use of "admission tests" as predictors of performance in undergraduate studies programs. The paper analyzes performance of economics and business administration students. This performance is linked to admission tests results. The paper also analyzes aspects of performance related to (i) differential progress through time, and (ii) differences in the extent to which students have "areas of interest/ability". The paper concludes that admission tests are a usefull tool even when predictions derived from them are far from perfect.

Predicting Academic Performance

Marcos Gallacher

Universidad del CEMA
(gmg@cema.edu.ar)

I. Introduction

The prediction of individual performance attracts considerable interest. In the business world, decisions have to be made on who to hire and who to promote. For both of these decisions, information (of different kinds) is used to infer future performance. Track record in previous employment, academic training, interviews, references and in some cases personality and intelligence tests are some of the types of information that are used for deciding. Mistakes are made: in fact Lawrence Peter cogently argued that, sooner or later, everybody is promoted to his/her "level of incompetence" (Peter, 1979). More seriously, organizational theorists have pointed out some of the difficulties in predicting performance using (past) data: the highest-ranking employee may not necessarily be the employee with higher managerial potential (see Jensen page 213). Similarly, in the world of education decisions are required on admittance, financial aid, letters of recommendation and other matters. These also require prediction of the candidate´s future contributions.

Universities attempt to make "good" decisions regarding the potential of candidates for academic study. The quest for "rigor" points in the direction of only admitting those suficiently talented and motivated. Less desirable candidates - according to this viewpoint - are steered into non-professional jobs or into low-stress academic programs. Some universities may try to develop a reputation for rigor, as well as for "producing" graduates who promise future professional excellence. This may require some type of "screening" device at some point: typically at admission time, but also possibly during the program of study.

This paper analyzes the usefullness of admissions exams as a tool for predicting academic performance, and ultimately for deciding admission into academic programs. The "admission exam" considered here is administered by an Argentine university (Universidad del CEMA) as a requirement for entry into its academic programs. The paper discusses the value of these types of exams for decision-

making, and explores ways in which additional information might supplement test results.

II. Admission Tests and Academic Decisions

Admission to undergraduate education frequently requires the candidate to take some type of "admission test". These tests may be administered by the universities themselves, by a specialized test-taking organization (such as the ETS in the U.S), by a government body or by schools. In Argentina, the first option applies: universities requiring admission tests administer their own. This "institution specific" test is less usefull for the candidate, as it cannot be used to apply to a different institution. In some cases, test are used (succesivelly) for entry into more than one institution as a result of institutions waiving admission tests for students who have obtained admission to a "comparable" program elsewhere. Institution-specific tests are also less usefull for those administering them than tests such as the SAT (Scholastic Aptitude Test) that exists in the U.S. Indeed, the "SAT-type" tests allows prospective candidates to be benchmarked against a substantial data base instead of against a smaller sub-group of candidates of a particular institution. Decisions, then, are based on "how good" candidates are as compared to the whole population to which they belong.

Rosovsky (1990) describes admission decisions in U.S. universities. According to the author, "selectivity" varies widely between institutions: some 17-20 percent of applicants are accepted into Ivy League institutions such as Harvard, Princeton and Yale; in prestigious public universities such as Berkeley and Wisconsin the figure increases to 50 - 80 percent. However, some 95 percent of U.S. institutions of higher education show "comparative lack of selectivity" (Rosovsky, page 61-62). "Selectivity" in U.S. elite institutions, moreover, is not based exclusivelly in choosing students from a SAT ranking; this metric is "supplemented by subjective, qualitative, non-quantifiable and personal components" (Rosovsky page 63). In contrast, elite institutions in Japan and in France base their decisions on "objective" performance measures such as test results. According to Rosovsky the admission procedure in elite U.S. institutions is:

> ......an excercise in social engineering, involving high school grades, essays, interviews, recommendations from teachers, and above all a general vision concerning the composition of the ideal freshman class. That ideal is most easily defined as an optimum degree of diversity - hence my allusion to Noah´s Ark - within the framework of academic excellence, thereby maximizing the opportunity of the students to learn from each other.

In Argentina, the admission decision varies widely according to the type of university. Public universities - with the exception of some medical schools - are (as regards to entry) mostly non-selective. However, although no substantial barriers to entry exist, an important number of students have difficulties in completing the first year of study. Selectivity then is not a matter of entry, but of survival in coursework, particularly in the first and second year of study.

Argentine private institutions vary in their quality and rigor. As in the U.S., some institutions of higher learning offer little more than tertiary, vocational-type programs. Others, however, offer programs of high quality. Available data does not allow comparisons to be made on the basis of selectivity in admissions; however *entry-level* selectivity is probably low even in the high-quality programs. As in public institutions, students "are allowed in, but are not guaranteed of finally obtaining a diploma".

Admission policies face political and economic realities. Public institutions suffer congestion in many academic programs, but cannot implement "restrictive" admission policies because of opposition from student bodies (who participate in university governance). The term "congestion" as used above denotes a situation where addmission of an extra student causes a reduction in the quality of education of the rest of the students. Note that congestion may or may not be present in educational programs. Also note that if the students benefits not only from contact with professors, but with fellow students as well (see Rosovsky quote above) an extra student may result in an *increase* in the quality for the rest of the students.

Private institutions, on the other hand, may find that restricting admission limits possibilities for growth. Further, in these institutions congestion problems are normally not severe (the student supplies funds that can be used to expand buildings, buy computers and hire professors). Thus, a reasonable policy may be that of "liberal" admissions with increasing academic requirements over time. But: ¿Is this a "corrrect" policy ? ¿ Should "marginal" candidates be screened before entry ? Answer to the previous questions hinges on the extent to which an admission test can accuratelly predict future performance. If, for example, most students with "poor" admission grades will fail after the first semester of classes, or (particualry) if their presence will reduce the quality of educational services received by fellow classmates, the "correct" policy may be to deny entry (or insist that the candidate take remedial courses). If, on the other hand, admission tests are only weakly correlated with subsequent performance, either (i) the tests have to be revised, (ii) course grading criteria has to be studied or (iii) it has to be acknowledged that students are not "tied" to the good or bad

performance they had at test date; they can (perhaps with some effort) make up for deficiencies in high school education.

III. Empirical Analysis

Data set **G** is comprised of 90 graduates of 4-year undergraduate programs in economics and in business management. Data set **S**, on the other hand, is comprised of 91 students.[1] These are enrolled in the second, third and fourth year of studies. First-year students are not included because of lack of performance data. The empirical analysis is focused on answering the following questions. First, what is the predictive power of admission tests as relates to future academic performance. Second, do different dimensions of admission tests (quantitative, verbal) have differential value as predictors. And lastly, what alternatives - if any - can be devised to standard admission tests as relates to prediction of future performance.

III.1 Admission Tests and Academic Performance

Performance in coursework (as measured by graduating GPA) shows considerable variability. Moreover, the frequency distribution of GPA is quite different for students enrolled in the management as opposed to the economics program (Graph 1). In the former, the distribution is skewed to the right and shows large variance. Average GPA for these students is 5.9. In contrast, the distribution of GPA for economics students is left-skewed; moreover variability appears (at least visually inspecting the graph) lower than for the management program. For economics students average GPA is 7.7. This represents high academic achievement: a GPA of 8 entitles to a honors degree. The data contained in Graph 1 suggests that variation occurs *within* as well as *between* academic programs. Initial (pre-university) "human capital" may account for variations such as noted previously; however it is also possible that variations area caused not by "human capital" but by different learning objectives of students.

A superficial analysis of results in Graph 1 may sugges than many (30-40 percent) management-oriented students either lack the initial-level training to be able to perform at a high standard. However, it is also possible that - given their more applied orientation - they are not especially interested in achieving high grades in academic

---

[1] The total available records was 147 for data set G and 132 for data set S. Observations were discarded because of missing data on admission tests. Admission test data may be missing because: (1) admitance

study. Probably, a medium-effort completion of an academic program is deemed adequate for their future objectives.

Both of the above possibilities are explored in Graph 2. Distributions of "entry-level human capital" (as measured by the admission test) are shown. Candidates who subsequently completed an economics program appear to be "better" than those completing the management program (admission score 71 for economics, 64 for management). However differences in this metric between both groups of students appears smaller than differences in academic achievement as reported in Graph 1. The point to be made is the following: heterogeneity of motivation and objectives may be more important (in determining outcomes) than differences in simple measures of "human capital stock". Management students are not "worse trained in high school" than their economics counterparts, the difference seems to lie more in the area of motivation and personal objectives.

Linear regression model results (Annex 1) suggest that student performance (as measured by GPA) can be predicted using as independent variables the program that the student is enrolled in (economics or management) and admission test results in verbal and quantitative ability. Regression (I) uses data set **G** (graduates) and defines "performance" as graduating GPA. Program of study (Economics = 1) and mathematics test results are both significant (p 0 .10). Note, however, the low R-squared associated with this regressions: predictions can be made, but only with a large error. These results are similar to those found by the author in an analysis of academic performance in graduate programs (Gallacher, 1998).[2] Regression (II) uses both data sets **G** and **S**. Here the dependent variable is GPA for the first 8 courses taken (first year of studies). Regression (II) provides additional insights. In particular, test performance in verbal ability appears now a significant variable (p = .001). The lower R-squared in regression (II) as compared to (I) suggests that in a longer period of time fewer "suprises" are to be expected in performance.

Graph 3 shows additional information of the prediction error resulting from the use of the simple regression of Anex 1. Candidates classified as "somewhat weak" according to prediction (predicted GPA ranging from 6.0 to 6.5) subsequently at graduation showed GPA values ranging from "weak" (< 5.5) to "very good/excellent" (7.5 - 8.0). Further, 8 (out of a total of 19) students in this class achieved GPA values equal or higher than 7, a very satisfactory result. Similarly, "strong" candidates

---

test was waived (typically the case for transfer students) or (b) test data is not available for the student. Observations for students with courses transfered from other universities were also deleted.

[2] This research reports that admission exam test results is significant as an explanatory variable predicting overall GPA; however R-squared values of regressions (which included as additional variables age, sex, high school type and undergraduate education) were in the order of 0.25 - 0.50.

(predicted GPA´s in the 7.5 - 8.0 bracket) showed performance varying from "below average" (GPA < 6.5) to "excellent" (GPA > 8.5). Visual inspection of Graph 3 suggests the possibility that the simple regression model overpredicts performance when predicted GPA is "low" and undepredicts when predicted GPA is "high": i.e. a non-linear relationship between independent variables may more accuratelly reflect the underlying relationship. However, preliminary analysis regressing residuals against y-hats or against y-values did not yield significant results. This issue, however, can be explored in futher research.

Data points included in all graphs (and regressions) do not account for students abandoning their programs before completion. However, a non-negligible number of students abandon the program before finishing one year of studies, either voluntarily or because of not achieving minimum required GPA. The admission tests of these students show a GPA somewhat - but not markedly - lower than the class average. That is, it is not clear that they "fail" due to insufficient scholastic aptitude. Most probably, motivation plays an important part.

A relevant topic concerning performance relates to *improvement over time*. Take two students A and B. Student A shows a higher (graduating) overall GPA than B. However, B´s GPA in the last year of study is identical to A. That is, "B started badly but ´caught up´ with A". In a sense, B can be though of a compensating through time deficiencies in initial human capital. Alternativelly, B is a "late bloomer" that only later realises his interest in academic issues. This situation can also arise if A "fizzles out" while B maintains his steady plod forwards. Decreasing performance difference among students over time can also arise due to the fact that first and second years emphasize "hard" (mathematics and statistics) courses that are alien to more "practical-oriented" students.

Let GPA_1 denote performance in the first year of studies and (as before) GPA denote overall performance. The ratio GPA/GPA_1 (if > 1) is an estimate of "improvement" during an academic program. Graph 4 shows a negative relation between predicted performance at entry and the GPA/GPA_1 ratio. This negative relation is confirmed regressing the GPA/GPA_1 ratio on predicted GPA (Annex 2 Regression III). This suggests that "less atractive" candidates show a greater improvement in academic performance than those with better credentials (as measured by admission tests) at entry. Note that a greater rate of improvement through time may suggest - at least to an optimist - that "less desirable" candidates may yet attain performance levels comparable to those initially better endowed with academic talents. The turtle finally catches up with the hare.

"Academic progress" results should be interpreted with care due to the fact that "weaker" students may choose an academic program with (comparativelly) less stringent demands. If so, progress for these students might be overestimated, as "stronger" students who serve as a benchmark take courses characterized by a higher degree of difficulty. The bottom of the matter, however, is that comparisons of "progress" across programs can be difficult as the concept of "learning" may be interpreted differently for students being evaluated in "harder" as compared to "softer" subjects.

III.2 Abilities and Productive Specialization

The "Error Type I" and Error Type II" used in statistical theory are relevant for understanding the use of admission tests for purposes both of admissions as well as financial aid. In the context of this paper, the first type of error relates to accepting a student that should be rejected. In turn, rejecting a student that should be accepted results in Error Type II. [3] The "stricter" the admissions policy, the lower the probability of error Type I but the higher the probability of Error Type II.

Production of "substandard" graduates results in reputational loss for the institution. Future demand for educational services may be reduced. Reputational loss also affects graduates: their degree is "cheapened", with resulting decrease in income-earning opportunities. It is difficult, however, to come up with an adequate definition of "substandard". A graduate with modest academic performance may yet contribute in some area of business, art, public service or even academics. Witness, for example, Steve Jobs´ 2005 commencement speech in Stanford University: Jobs dropped out of an undergraduate program and only continued attending those classes (unofficially) in which he had personal interest. Among them caligraphy. But later on he founded Apple Computer.

A "substandard" graduate may be the one that finds it difficult or impossible to find the niche where his/her abilities might be productivelly used. If this is the case, the problem is one of matching abilities/preferences with productive specialization (niches) in order to bring about efficient exchange. The problem of "substandard" can then be approached by analyzing whether a graduate has found - during his academic program - some area of study where abilities may be used. Given two students A and B with identical overall academic performance (as measured, for example, by graduating GPA) the one with higher variability (across courses) in grades can be assumed more "narrowly focused". Arguably, in future exchange relationships this student can

specialize and thus achieve high personal productivity. In other words, students graduating with "low" average GPA have "better prospects" when the variability in their course grades is high than when the low GPA results from "substandard behavior" in all coursework. Thus, highest grades during coursework are an indicator of "potential area of interest/ability" of the student. In order to take into account overall performance of the student, the existence of this "potential area of interest" can be proxied by the ratio highest/lowest grades achieved by the student. For example, for two students having GPA of 6, student A whose best and worst performance are 8 and 4 appears to have some (potential) area interest/ability, while student B all of whose courses were graded 6 probably does not.

Graph 5 plots an index of "specialized interest" against predicted performance. The 80-percentile of grades attained by the graduate divided by the 20-percentile is taken as a proxy of "specialization" in ability and/or interest. The higher this index, the more focused is the student´s interest in some topic; presumably this helps in finding a professional specialization further along. As shown in Graph 5, the "initial attractiveness" of the graduate (as measured by predicted GPA) is inversely related to specialization in his interest/abilities. Regression results for this relationship are reported in Annex 2. This suggests that - as hypothesized previously - many "unattractive" candidates can well find when they eventually graduate an area of specialization that can productively make use of their talents. Milton Friedman (the reference is now lost to me) once remarked that young people engaged in frequent job rotation: "trial and error" in fact being the method used to finally discover where to settle down.

VI. Conclusions

Decisions regarding the design of undergraduate programs are not easy. Some observers feel that many universities maintain mediocre programs, offering little more than an "advanced high school" type of education. Higher-quality programs obviously require higher costs in terms of better professors, roomier classes, more powerfull computers and an up-to date library. But, they also require a higher input of student time, energy and in some cases individual talents or qualifications.

Admission tests are a tool for sending "signals" to candidates about the seriousness of the institution´s devotion to learning. The university that requires passing an admission test always has the option of turning down applicants. Results of this paper, however, caution on giving admission tests more weight than they deserve.

---

[3] Financial aid decisions can also be cast in the logic of these two types of errors.

The university always has the option of admitting a candidate, and later (if performance is substandard) forcing him to exit or take again all or part of the coursework.

A "liberal" admission policy may be reasonable if the costs of admitting a "wrong" candidate are low. Indeed, admittance allows a "wait and see" attitude. It allows decisions be made with better information about student potential (for example, performance in first year of coursework). This liberal policy, however, requires that those giving the "green light" relative to continuation of studies be free from pressures of students, university administrators and even parents. Note that a university that endows students with some decision-making rights in governance (such as public universities in Argentina) should be particularly carefull on deciding "who to let in". Paradoxically, not participating in governance has the advantage (for the student) of not forcing the university to erect higher barriers to entry.

Design of admission tests requires some thought on human performance and self-satisfaction in a complex world. Clearly, not all applicants to a business program are going to lead multinationals into the XXI´st century. Similarly, not all economics undergraduates will contribute to advance in knowledge of how markets work. It takes much more than hard work and access to education to achieve these ends. Richard Bolles (1978, page 129) presents a taxonomy of "work environments" where someone presumably can fit in. A wide variety of classes emerge. Among these "entreprising", "social", "investigative", "realistic" and "conventional". The point to be made is that human beings differ markedly in their interests and capabilities. University programs should - where possible - cater to these differences. This might require a different perspective of the concept of "rigor" and "excellence" in undergraduate education.

Probably the frase "scholastic aptitude tests" should be used instead of "admission tests". These tests are not perfect predictors of future performance. However, they are important pieces of information for the university making admission decisions, for the school from which the student graduated, as well as for the student. In summary, scholsatic aptitude tests used correctly can be a positive factor in higher education. If used incorrectly they can cause much frustration.

VII. References

Bolles, R.N.(1978), ***The Three Boxes of Life and How to Get Out of Them***. Ten Speed Press.

Jensen, M.(1998), ***Foundations of Organizational Strategy***. Harvard U.Press.

Gallacher, M.(1998), Predicción de performance academica: estudios de postgrado. Documento de Trabajo 129. Universidad del CEMA.
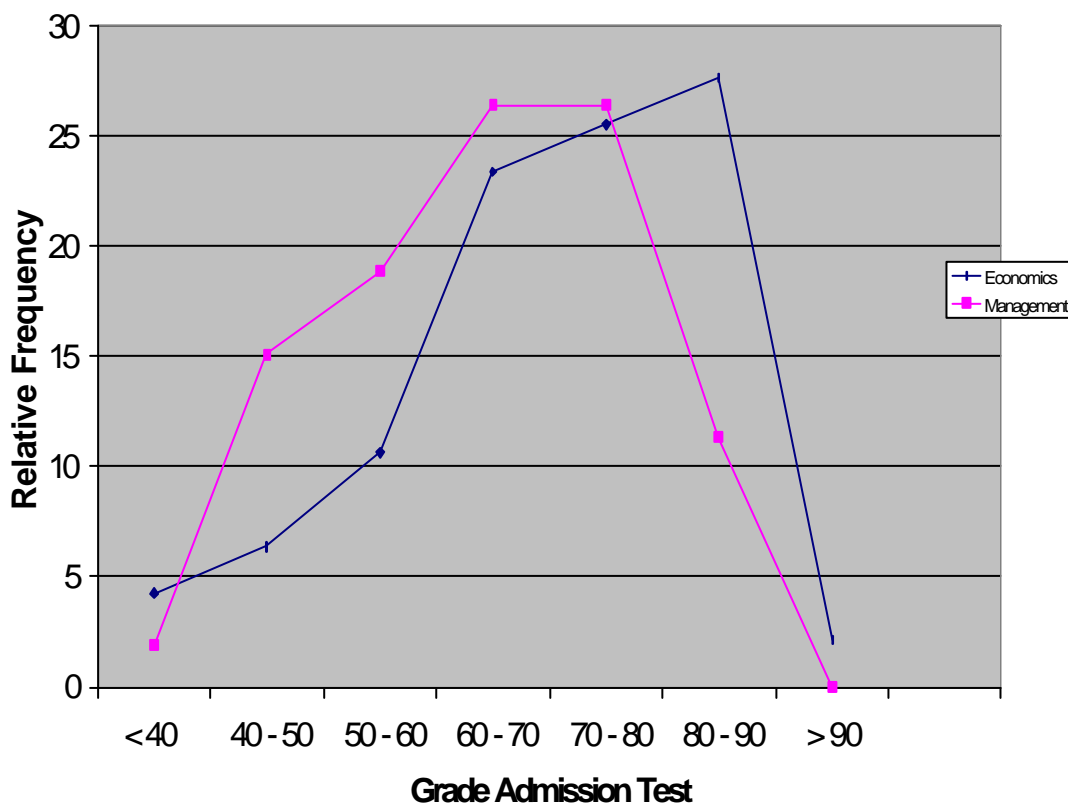
Peter, L.J.(1979), ***Peter´s People***. Tower Non-Fiction.

Rosovsky, H.(1990), ***The University: An Owners Manual***. W.W.Norton & Company.
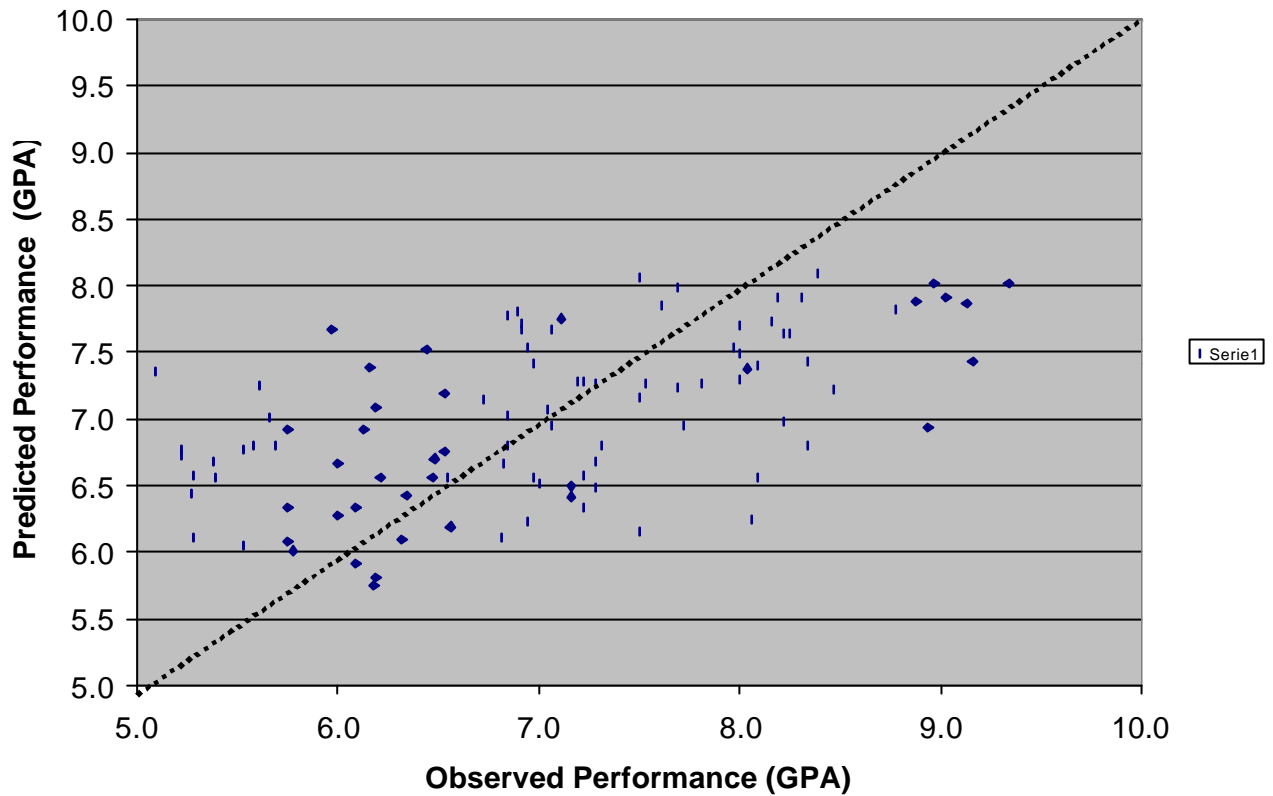
Graph 1: Graduates in Economics and in Management - GPA Distribution at Graduation (5 - 10)
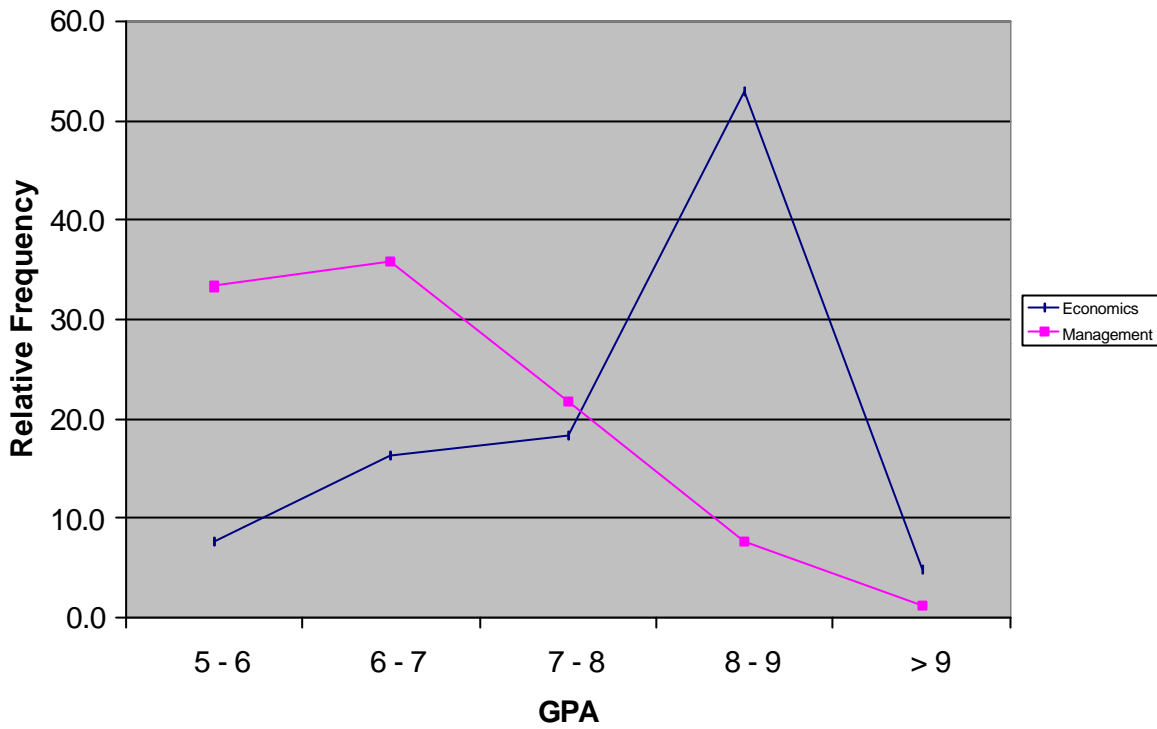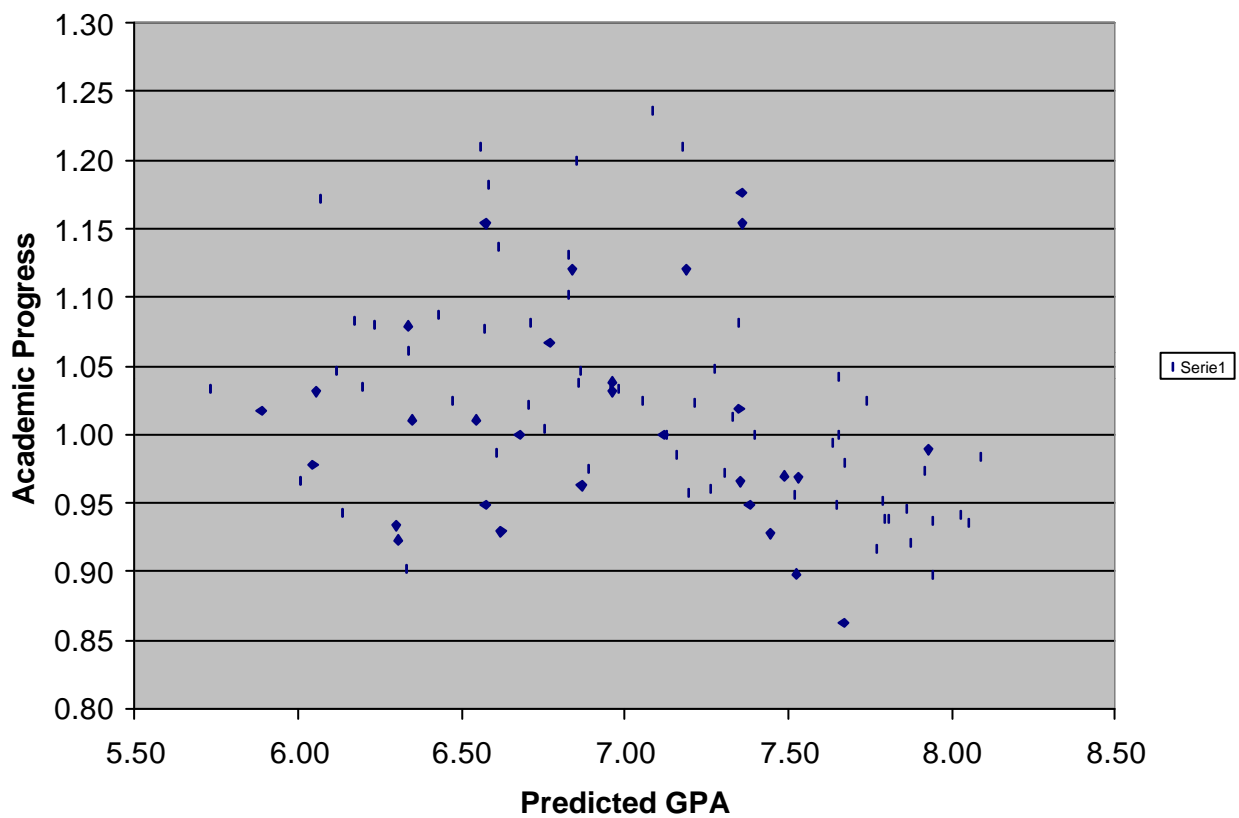

Graph 2: Candidates in Economic and in Management: Admission Test Results
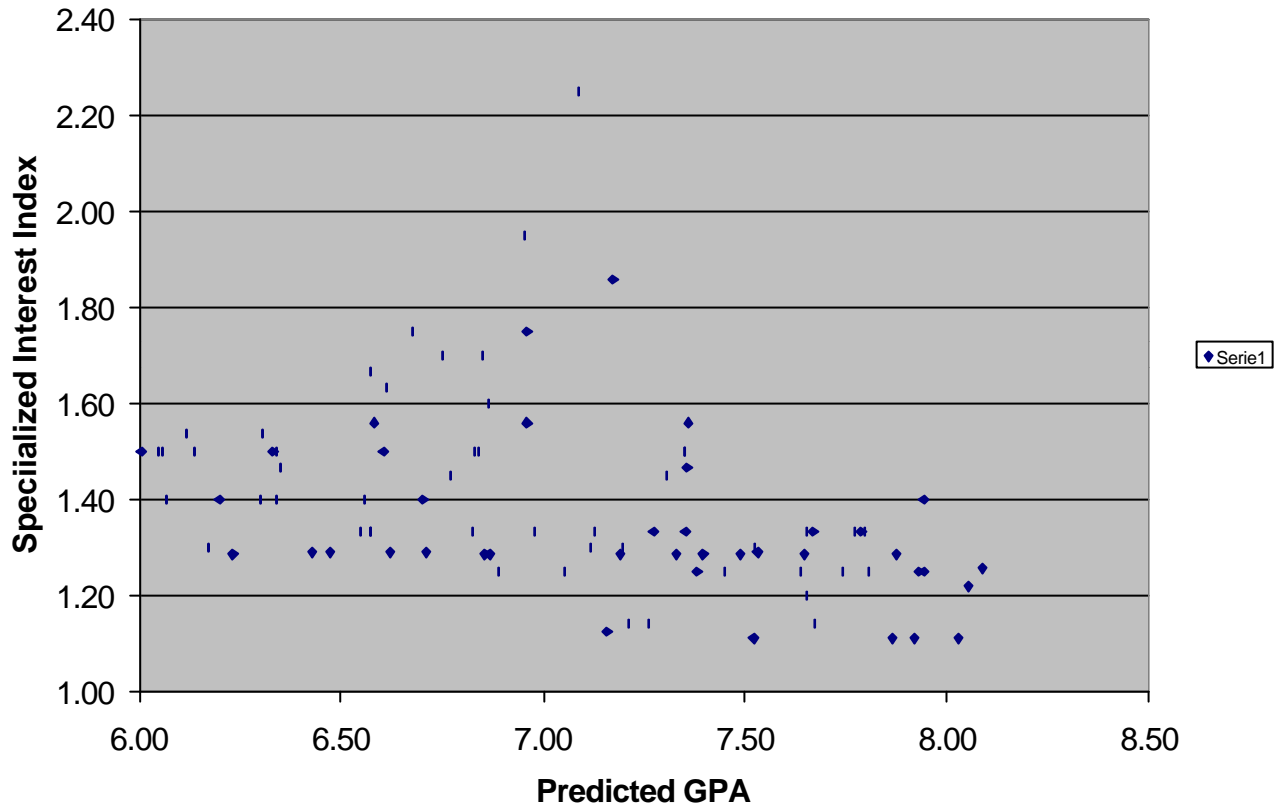
Graph 3: Performance - Observed vs Predicted

Graph 4: Academic Progress vs Predicted GPA

Graph 5: Specialized Interest Index vs Predicted GPA:

<u>Annex 1:</u> Regression Results (OLS)

	Regression (I)	Regression (II)
	Data Set G	Data Sets G + S
Dependent Variable	GPA	GPA_1
Constant	4.48 (78.47) *	4.20 (10.0) *
Program (a)	0.63 (3.15) *	0.90 (5.78) *
Verbal Ability	0.012 (1.39)	0.023 (3.92)
Mathematics	0.021 (3.08) *	0.012 (3.38)
R-sq adj	0.28	0.30
N	90	181

"*" = p 10 %

(a) Economics = 1, Management = 0

The variable to be predicted (GPA) is measured in UCEMA in the scale 5 - 10. A "5" correponds to a marginal pass, a 6 and 7 to "good", 8 to "very good" and 9 - 10 to outstanding.[4] Reported GPA statistics do not take into account "0" grades (failed courses) as for graduates these have always been replaced by the final (> 0) grade corresponding to succesfull completion of the course. Variable GPA_1 represents GPA after completing the first year of studies (8 courses).

---

[4] Courses graded "0" (fail) have to be retaken before graduation. Further, a student has to maintain a GPA of at least 5 in order to be permitted to continue his/her programa. Thus, while for an individual course grade may be 4, average GPA at graduation is at least 5.

Annex 2: Regression Results (OLS)

	Regression III	Regression IV
Dependent Variable	GPA/GPA$_1$	80Percentile/20 Percentile
Constant	1.34 (14.2) *	2.38
Predicted GPA	-0.465 (-3.47) *	-0.14 (-4.35)
R-sq adj	0.11	0.17
N	90	90