DISCUSSION PAPER SERIES

No. 6237

WHICH FACTORS DETERMINE ACADEMIC PERFORMANCE OF UNDERGRADUATE STUDENTS IN ECONOMICS? SOME SPANISH EVIDENCE

Juan José Dolado and Eduardo Morales

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP6237.asp**

ISSN 0265-8003

WHICH FACTORS DETERMINE ACADEMIC PERFORMANCE OF UNDERGRADUATE STUDENTS IN ECONOMICS? SOME SPANISH EVIDENCE

Juan José Dolado, Universidad Carlos III de Madrid and CEPR
Eduardo Morales, Harvard University

Discussion Paper No. 6237
April 2007

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6237

April 2007

ABSTRACT

Which Factors Determine Academic Performance of Undergraduate Students in Economics? Some Spanish Evidence*

This paper analyses the determinants of academic performance of first-year undergraduate students in Economics at Universidad Carlos III de Madrid, over the period 2001-2005. We focus on a few core subjects which differ in their degree of mathematical complexity. Type of school, specialization track at high school, and the grades obtained at the university entry-exam are among the key factors we examine. Our main finding is that those students who completed a technical track at high school tend to do much better in subjects involving mathematics than those who followed a social sciences track (tailor-made for future economics students) and that the latter do not perform significantly better than the former in subjects with less degree of formalism. Moreover, students from public schools are predominant in the lower and upper parts of the grade distribution while females tend to perform better than males.

JEL Classification: I21 and I29
Keywords: academic performance, multinomial logit, pre-university determinants and quantile regressions

Juan José Dolado
Departamento de Economia
Edificio Luis Vives
Universidad Carlos III
C/Madrid, 126
28903 Getafe, Madrid
SPAIN
Email: dolado@eco.uc3m.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=108304

Eduardo Morales
Department of Economics Littauer Center
Harvard University
1805 Cambridge St.
Cambridge, MA 0213
USA
Email: emorales@fas.harvard.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=166266

*We are grateful to Marisol Somolinos for help with the students´ administrative records at UC3M, and to Antonio Cabrales and Carmelo Nuñez for insightful comments on preliminary drafts of this paper. The second author is currently a graduate student in the Economics Ph. D. program at Harvard. Financial support from the European Commission (Dolado) under the project *The Economics of Education and Education Policy in Europe* (MRTN-CT-2003-50496) and Consejería de Educación de la Comunidad de Madrid (Morales) is gratefully acknowledged.

Submitted 12 March 2007

1. Introduction

It is often claimed that Economics is the discipline with the highest need for formalism in theory-building among the social sciences. Thus, modern undergraduate courses in Economics place a lot of emphasis on students acquiring a good mathematical background, in parallel with more specific training on economics-based topics. To check what kinde of prior qualifications leads to a better academic performance in this degree, this paper presents some evidence on the role played by a few pre-university determinants of the educational attainments of first-year undergraduates in Economics at a well-known Spanish university.

We are interested in examining whether the factors determining success in mathematics-based subjects differ from those which affect performance in other subjects with less mathematical content. To do so, we focus on the students´ performance in three subjects ordered in decreasing level of mathematical requirements: Mathematics, Introductory Economics and Economic History (henceforth, *Maths*, *Introecon* and *Econhist*, respectively).[1] In particular, we wish to analyze whether the social sciences track at high school, which is supposed to be the one tailor-made for students willing to do an Economics degree at university, enhances performance in all these subjects.

Our evidence relies upon individual-level data using sample of almost 400 first-year undergraduate students at Universidad Carlos III de Madrid (UC3M) who took exams in these subjects from 2002 to 2005. [2] The students in our sample were enrolled in the above-mentioned subjects during the first semester of the first year in their four-year BA degree (*Licenciatura*) in Economics or Business Administration (LADE).[3] Information is available on the type of high school (public, charter[4] and private) they attended during their upper-secondary education (two years of *Bachillerato*), the training received during this period (i.e., the *bachillerato* specialization track) and the grade they obtained in

[1] The subjects taught in *Maths* are limits, differentiation and integrals. *Introecon* is a basic Microeconomics course (consumer theory and theory of the firm) and *Econhist* deals with the long-run development process in Western Europe.

[2] The choice of first- year undergraduates is dictated by data availability. However, the fact that the withdrawal rate is rather high (about 35%) during the completion of a degree in the Spanish university system implies that the estimates obtained for this restricted group of students is less likely to be affected by selection biases (due to omission of those who withdraw) than those obtained from students in later courses.

[3] Students in Business Administration (LADE) have the same subjects as those in Economics in their first course. Thus, we will refer in the sequel to all of them as undergraduates in Economics.

[4] A charter school (*concertado*) is a school subsidized by the public sector, typically run by religious orders.

the national entry- exam to the university (*Selectividad* exam) which is bound to control somewhat for unobserved skills.[5] These variables, plus some additional individual characteristics, are the controls we use to explain outcomes (marks awarded in the final examination of the three subjects) using an achievement production-function approach. It is worth noticing that, although we lack parental background, the entry-exam grade and type of school are bound to be highly correlated with this missing piece of information (see Calero, 2006).

Marks in the Spanish education system are numerical, ranging from 0 to 10. A grade below 5 implies a *Suspenso* (Fail in the anglosaxon system), between 5 and 7 is an *Aprobado* (Third), between 7 and 9 is a *Notable* (Lower Second), between 9 and 10 is a *Sobresaliente* (Upper second) and 10 (or very close to that grade) is a *Matrícula de Honor* (First or Distinction). However, the only marks appearing in the official records delivered by the university to the students at the end of each course are the categorical ones discussed above. To simplify notation in the sequel, we will label these categories with the acronyms SUS, AP, NOT, SOB and MH, respectively.

Given these characteristics of the grade scale in Spain, we rely empirically on three different econometric approaches. First, we use the continuous support of the dependent variable (marks) to run OLS regressions. Secondly, we measure the impact of the determinants on the dependent variable at different points in its conditional distribution, by means of quantile regressions (QR). In this fashion, we will be able to provide a sense of how the impact of the explanatory variables may differ throughout the marks distribution. For example, one may find that attending a certain type private school or having completed a particular type of *Bachillerato*, while seemingly important at the mean as a determinant of the outcome, may in fact have different impacts across students with high or low marks. Finally, we use the categorical nature of the marks to estimate multinomial logit models distinguishing among three marks categories which broadly signal pre-labour market human capital attainments.

Our paper falls into a large literature that examines the determinants of university students ´academic performance (see, e.g., Dearden et al., 1998, and Smith and Naylor, 2001, for the UK, and Eide and Showalter, 1998, for the US). However, to our knowledge, none of these studies have analyzed academic performance distinguishing by specific subjects within a given degree, as we do here. In this sense, our contribution could be useful to shed some light on the ongoing debate within the higher education sector on whether the currently available system of high-school tracks is adequate to complete some degrees, like e.g., Economics.

---

[5] The overall grade is a weighted average of the school grade (with a weight of 60%) and the centralized exam grade (40%). Because school grades can be inflated, we only use the exam grade in this paper.

The rest of the paper is structured as follows. Section 2 discusses the data used for the estimation. In Section 3, we present the econometric results. Section 4 contains a discussion about how representative is the sample at hand and checks for some potential selection biases. Finally, Section 5 concludes.

2. Data and Descriptive Statistics

The data we have is made up of a survey among the four cohorts of students enrolled in the bilingual group of the BA degree (*Licenciatura*) degree in Economics during the academic courses from 2002/03 to 2005/06. [6] All these students were taught *Maths* by the same instructor (one of us) in classrooms with a maximum enrollment rate size of 100 students per group. We solicited information from them about the type of school (public, charter and private) they attended during high school (two years of *Bachillerato*), the kind of training they received during this period (there are four types of *Bachilleratos*: technical, natural sciences & health, social sciences and humanities which are chosen by high-school students in the year when they become 16, i.e., in *Primero de Bachillerato*) and the mark they obtained in the entry exam to the university (*Selectividad*) at 18. This last information was cross-checked with the administrative records of UC3M in order to avoid measurement errors. The response rate to the survey was 96.5%, yielding a sample of 386 individuals.

A brief description of the relevant variables is provided in the sequel. Tables 1a and 1b below present the conditional distribution of student and school characteristics given marks and the converse conditional distribution, respectively. For expository purposes, we have grouped these marks into three broad categories: S (SUS; category 0), AN (AP+NOT; category 1) and SM (SOB+MH; category 2). Thus, the frequencies in the relevant rows (columns) of Table 1a (Table 1b) add up to 100. Overall, 49.2% of the students are male whilst 89.3% are Spanish. By type of school, 42% of the sample comes from a public school, 21% from charter schools and 37% from private schools. By type of *bachillerato* track, 67% have done social sciences, 26% the technical one and the remaining 7% did natural sciences & health (3%) and humanities (4%). It should be noticed that the high school training in mathematics is more intense in the technical and natural sciences & and health *bachilleratos* than in social sciences. These are the three tracks where students take two compulsory yearly courses in mathematics, whereas it is only an optional subject for those enrolled

[6] Being a student in the "bilingual" group means that, except for a few subjects (e.g. those related to Law), all teaching takes place in English. Admission to this group is conditional on passing an English exam. Courses are organized on a semester basis and there are ten subjects in their first year (five each semester), with exams taking place in February and June.

in humanities. [7] Likewise, education quality (student/teacher ratios, computer facilities, foreign languages) in non-public schools is considered to be higher that in most public schools, in exchange for annual tuition fees of about € 6,000 in private schools and around € 2,000 in charter schools.

This descriptive information shows that male students are less successful in passing the subjects than female students (except in *Introecon*). Likewise, those coming from public schools with a social sciences or humanities *bachilleratos* tend to do worse. Interestingly, however, students from public schools (mostly with a technical track) do very well in achieving the highest category (SM=2) in all subjects. Thus, students from public schools seem to have a U-shaped distribution across grades. The lower tail contains those students who did social sciences whilst those in the upper tail did a more scientific-oriented *bachillerato*. Given that a majority of high-school students (66%) are enrolled into the public education system, the latter effect could be explained by higher competition among the best students in public schools, particularly in the technical track. Thus, comparing the best students in this track (equivalent in all observable characteristics but with different school backgrounds), the ones coming public schools are likely to be drawn from a higher point in the underlying ability distribution. It is also worth noticing that foreign students exhibit a much higher variability in marks than natives. Finally, the last row in Table 1a presents the correlations between the numerical marks in each of the subjects and the marks in the *Selectividad* exam. These correlations range between 0.50 and 0.67, being largest in the case of *Maths*.

Figure 1 depicts the (kernel) densities of the (numerical) marks in the three subjects. As can be observed, the distributions in *Econhist* and *Introecon* are unimodal with the former being the one more shifted to the right (i.e., higher probability of a pass grade). By constrast, the density of *Maths* is bimodal (not too different from a uniform distribution) and is the one more shifted to the left (i.e,. lower probability of a pass grade).[8] To achieve comparability across subjects in the estimation of the effects of the different pre-college determinants, we use the standarized marks in the empirical analysis Hence, the estimated effects are measured in terms of the corrresponding standard deviations (henceforth, s.d.´s). To convert them into numerical grades, one should multiply these estimates by the grades. Figure 2, in turn, displays the density of the

[7] All the students in our sample coming from the humanities confirmed to us that they had taken in high school at least one of these optional subjects in mathematics.

[8] The moments (mean and s.d.) of the three distributions are as follows: *Maths* (5.17, 2.50), *Introecon* (5.70, 1.57) and *Econhist* (6.48, 1.44)

university entry-exam grades which is similar to a conventional skills distribution.[9]

3. Econometric approaches

Our modeling point of departure is based on an extensive literature analyzing schooling outcomes in developing and developed countries using a production function approach; cf. Hanushek (1995), Case and Deaton (1999), Bjorklund et al. (2003), and Glewwe and Kremer (2005). According to this approach, grades are explained as a function of several inputs in the following manner:

$$A = f(x) \tag{1}$$

where A represents some metric of academic performance (continuous or categorical marks) and x is a vector containing the determinants described above. In order to estimate (1), we adopt three different econometric approaches discussed below.

3.1 Least-Squares

The dependent variable is the numerical grade of individual i in a pooled regression model where the controls are the variables alluded to before. Specifically, the model estimated by OLS has the following generic form:

$$A_{is} = a + \beta' x_{is} + u_{is} \tag{2}$$

where A_{is} is the grade of student I in subject s and x_{is} is a vector of controls formed by a G_i dummy (female=1), N_i is a nationality dummy (foreigner=1), UG_i (numerical), ST_i (dummies for charter and private schools), BT_i (dummies for natural sciences & health, technical and humanities) and T_i are three dummies for cohorts. Thus, the reference group corresponds to male students from public schools with a high school track in social sciences who took the (February) exams in the 2002/03 course. The constant and the (*i.i.d.*) disturbance terms are captured by a and u_{is}, respectively.

In the first columns of each subject in Table 2, the OLS estimates of the coefficients in (2) are reported together with their robust standard errors. The largest effects are found for the entry-exam grade and the technical degree covariates. An extra point in the entry-exam gives rise to about 0.60 extra s.d.´s (1.50, 0.85 and 0.85 points, respectively) relative to the reference group in each subject, with slightly larger coefficients in both *Maths* and *Introecon*. Likewise, having completed the technical track in *bachillerato* leads to 0.68 extra s.d.´s (1.7

[9] Notice that some of the students have a mark below 5 (the pass grade) because, as mentioned in footnote 5, the centralized exam grade only accounts for 40% of the overall mark.

points) in Maths and 0.45 s.d.´s (0.7 points) in *Introecon*, without any significant gain in *Econhist*, whereas 0.5 extra s.d.´s (0.75 points) in *Maths* are achieved by those who followed the natural sciences & health track. By contrast, the humanities track has a penalty of 0.6 s.d´s (1.5 points) in that subject. Regarding type of school, having attended a non-public school seems to ensure a better grade on average. For example, coming from a private school leads to about 0.25 extra (0.7 points) in *Maths* and *Introecon*, relative to coming from a public school. As regards gender, female students get 0.15 s.d.´s (0.38 points) more than their male classmates in *Maths*, without significant differences in the remaining subjects. Finally, with the exception of the 2004/05 course, the cohort dummies are significantly negative. Despite the short sample period, this gives some support to the extended opinion among several education pundits that training in high schools has been deteriorating over time although this effect might be contaminated by the presence of different instructors in two of the subjects.

Next, in order to analyze the impact on grades of having followed a given type of *bachillerato* at a given school, the second columns in Table 2 present the estimated coefficients of a similar regression augmented with interaction terms between school type and high-school track. The results are similar to those discussed above with the only statistically significant coefficients being the ones on the interactions between the natural sciences & health track and both private and charter schools. For example, the students who come from private schools and followed this track get 0.33 s.d.´s (=1.09-0.76; equivalent to 0.82 points) more than those who did that track in a public school.

3.2 Quantile Regressions

The fact discussed earlier that we may not have well-behaved distributions in the outcome and in some of the other variables, implies that least-squares coefficients may yield partial information. Accordingly, in line with a growing literature on the application of this technique to achievement production functions, we use quantile regressions (QR).[10] Following the well-known methodology first proposed in Koenker and Bassett (1978), the model of QR in the setup of the achievement production function described in (2) can be described as follows. Let (A_i, z_i) be a random sample, where z_i =(1, x_i) and $Q_\theta(A_i | z_i)$ is the conditional θ^{th} quantile of the distribution of A_i given z_i. Then, under the assumption of a linear specification, the model can be defined as

$$A_i = z_i'\beta_\theta + u_{\theta i}\,, \qquad Q_\theta(A_i | x_i) = z_i'\beta_\theta \tag{3}$$

[10] For examples of the use of QRs in the literature on schooling outcomes, see, e.g., Eide and Showalter (1998), Levin (2001) and Bhorat and Oosthuizen (2006)

where the distribution of the error term $u_{\theta i}$, $F_{u\theta}(\cdot)$, is left unspecified, just assuming that $u_{\theta i}$ satisfies $Q_\theta(u_{\theta i}|z_i) = 0$. The estimated vector of QR coefficients, $\hat{\beta}_\theta$, is interpreted as the marginal change in the conditional quantile θ due to a marginal change in the corresponding element of the vector of coefficients on z, and can be obtained using the optimization techniques described in Koenker and Bassett (1982).

In order to facilitate the comparison of results across subjects, we choose different quantiles for each subject so that the percentiles become similar in terms of both numerical and categorical grades. These happened to be: θ=0.25 (grade: 2.8, SUS), 0.75 (7.0 NOT) and 0.95 (9.5, SOB) for *Maths*; θ= 0.10 (3.8, SUS), 0.80 (7, NOT) and 0.98 (10, MH) for *Introecon*; and, finally, θ= 0.10 (4.5, SUS), 0.70 (7.2, NOT) and 0.98 (9.3, SOB) for *Econhist*. Tables 3a, b and c report the estimated coefficients at the relevant quantiles (together with the regression at the median, i.e., at θ= 0.50) using the specification without interaction terms. For convenience, we reproduce the OLS estimates in the first column (*average*) in order to compare the coefficients at the mean as opposed to the coefficients at the chosen quantiles of the conditional distribution of (numerical) grades. [11]

The QR offer valuable additional information to the one discussed above. [12] The key result in *Maths* is that the impact of private and charter schools (in the range of 0.2 to 0.4 extra s.d.´s or 0.5 to 1 points relative to public schools) is much weaker at top quantiles, in line with the prevalence of students coming from public schools at the higher part of the grade distribution. A similar effect is observed for the entry-exam grades (the most significant variable, together with the technical track), whose effect decreases throughout the distribution. The opposite effect takes place for the humanities track. As regards the other subjects, the results are similar, with the only exception that having completed the more science-based tracks does not seem to help.

3.3 Multinomial Logits

Finally, we use the categorical grades to estimate two multinomial logit models determining the odds-ratios of obtaining a grade in a certain category relative to a comparison one. Notice that, despite the fact that our dependent variable has a natural ordering (from higher to lower grade category), the "proportional odds" assumption underlying the estimation of ordered logit models was rejected in our dataset. [13] We distinguish among the three grade categories

[11] For the sake of brevity, we do not report the estimated coefficients on the cohort dummies. However, the pattern of negative coefficients for the 2003/04 and 2005/06 cohorts remains across quantiles.

[12] An F test on the joint equality of all the coefficients across the chosen quantiles yields a p-values close to zero, therefore rejecting the null.

[13] We used Wald and LR chi-squared tests to check the "parallel regression" or "proportional odds" assumption under which an ordered logit is a more appropriate model (see, e.g., Winkelmann and Boes,

defined in the Introduction, denoted as Cat_*i* (3) with *i*=1 (AN), 2 (SM). Category 0 (S) is the base category with the following dummies omitted: male, Spanish, public school, social sciences, and 2002/03.

The columns in Table 4 report the corresponding odd-ratios, that is, the odds of obtaining a grade in a given category with respect to the base one.[14]. In line with the evidence above, it is found that that the highest (and more significant) odds ratios are those pertaining to a technical *bachillerato*, females and private and charter schools whereas the humanities track have the lowest odd-ratios. In particular, the technical track and the entry-exam grades present increasing odds ratios in all subjects as we move up from Cat_1 to Cat_2, while the opposite result holds for the natural sciences & health track. Having attended a charter/private school seems to help a lot in obtaining AN (Cat_1) in the three subjects as reflected by odds-ratios larger than unity. By contrast, this type of schools have odd-ratios smaller than unity in Cat-2 of *Maths* and *Introecon*, pointing out again that, among those get the best grades in these subjects, a large fraction come from public schools. Finally, it is worth noticing that some of the estimated odd-ratios are very large. This is the case, for instance, of the coefficients on the natural sciences & health track dummy in Cat_1(3) in two of the subjects (albeit with large s.d.´s) and on coefficient on the technical track in Cat_2(3) in all subjects. The explanation for these anomalous estimates is that a very large fraction of the students in these tracks obtained grades in the above-mentioned categories, blowing up therefore corresponding odds-ratios.

4. Selection bias

Our sample of students has two characteristics which could lead to (favourable) sample selection biases. The first one is that UC3M is considered to be one of the Spanish universities with the highest reputation in Economics. [15] That, in principle, could lead to attracting better students than other universities with a lower ranking in this field. Unfortunately, we do not have any control group in order to test for selection. However, there is ample evidence that the

2006). These tests have the null of identical coefficients in the different categories of the multinomial logit. If the assumption is rejected, the estimates of the latter model are consistent while those of an ordered logit are inconsistent. In both multinomial logits the assumption was strongly rejected with p-values of both tests being very close to zero. We also estimated a generalized ordered logit model, which does not impose the previous restriction. Although not reported, the results were similar to the ones displayed in Table 4 yet with lower precision of the estimates in category 2, due to the smaller number of observations.

[14] Thus, for example, the coefficient on Female in the first column indicates that the odds of a female getting a grade in category 1 compared to category 0 is about 2 times (1.985) the odds of male students doing the same.

[15] According to the rankings published in the newspaper EL MUNDO (CAMPUS magazine) since 2003, UC3M is one of the two best universities in Spain to complete *licenciaturas* in Economics or LADE, together with UPF.

mobility of students across regions is very low and the entry-exam grade requested by UC3M to get admission in the Economics is a low pass (5.0), despite being somewhat larger in LADE (6.0). These acceptance grades are similar to those requested by most universities. Thus, we conjecture that biases are bound to be minor in this respect.

The second potential bias stems from the fact that students in our sample belong to a group is taught in English. Given that Spain is one of the European countries with the lower share of the population speaking foreign languages (44%), it could be the case that the students enrolling in this group are a (favourably) selected group relative to the population of students taking lectures in Spanish, which is an ample majority. An indication that this bias could be present is that the proportion of students coming from public schools in the bilingual group (42%) is significantly lower than the corresponding share in the total population of students completing higher-secondary education (66%).

In order to check for this bias, we have used another dataset regarding two groups (taught in Spanish) of first-year undergraduate students in Economics at UC3M during 2002-206. The aggregate sample size for these control groups is 572 students. Information on these students was obtained from the university archives and relates to gender, nationality, grades at the *Selectividad* exam and on whether students completed high school in the region of Madrid (CM) or in other Spanish regions. Unfortunately, we lack the remaining individual information that was used before in analyzing the determinants of outcomes for the bilingual group.

To control for sample selection biases we estimate a participation equation in the bilingual group as the first step in the conventional two-stage Heckman approach for selection correction. We use the pooled sample of all students (both from the Spanish and bilingual groups) which includes 958 students (=572+386). Given the scarce information available, we use the residence in CM (which is also available for the students in the bilingual group, but has not been used as a covariate in the previous sections) as the identifying variable. The insight for this choice is as follows: if the bilingual group is a (favourably) selected group from the population of students enrolled in Economics degrees at UC3M, then it is likely that a larger share of students from other Spanish regions will enrol in this group, given that there are very few universities in Spain offering bilingual courses.[16]

[16] The fractions of students living outside the region of Madrid are 18.3% and 12.2% in the bilingual and Spanish groups, respectively. The means of the entry-exam grades are 6.8 and 6.2 respectively, though a test for equal means does not reject the null with a p-value of 0.13.

The first column in Table 5 presents the results from a first-stage probit model where the dependent variable equals 1 if a student belongs to the bilingual group and 0 otherwise. The covariates are gender, nationality, a dummy variable on residence (CM=1), (numerical) grade at the *Selectividad* exam and the cohort dummies (not reported). Results point out that being a foreigner and living outside CM increase the probability of belonging to the bilingual group whilst the other covariates do not have significant effects. Thus, our identifying strategy seems to work well. The next three columns in Table 5 report the results the OLS estimation of the linear model in Table 2, this time augmented with the inverse Mills ratio (*lambda*) from the participation equation. This last term turns out to be always insignificant and, despite some minor quantitative changes in the estimated coefficients, none of the qualitative results stressed above change with the selection correction. Hence, although we cannot completely discard selection biases with respect to the overall population of Spanish first-year undergraduate students in Economics, our results seem to be valid inference the context of UC3M undergraduates and, possibly, in relation to the overall population of similar students completing an Economics degree in other universities in Madrid.

5. Conclusions

Our main finding in this paper is that, conditional on our proxy for skills, the most important determinant of academic success in the subject of *Maths* during the first course of an Economics degree is to have followed a technical track in the last two years of *bachillerato*. Interestingly, we also find that having completed a social science track instead of a technical track does not improve performance in other two subjects with less (*Introecon*) o very little (*Econhist)* mathematical content but which require more prior economic training. Given that the social sciences track was designed by the Spanish education authorities to provide the appropriate training for those students willing to become economists, it is fairly striking that a school background in, e.g., mathematics, physics or chemistry leads to a better academic performance. One possible policy implication of the above results is that high- school students who intend to do an Economics degree should take the *maths* courses of the technical track rather than the ones taught in the social sciences specialization.

Another interesting finding is that, among the best students, there is a majority of those coming from public schools. One possible explanation for this finding is that public schools seem to exert higher competition among the best students than non-public schools which helps them adapt better to the university environment. Finally, on average, females tend to do better than males.

References

1. Bhorat, H. and M. Oosthuizen (2006), "Determinants of Grade 12 Pass Rates in the Post-Apartheid South African Schooling System ", Cahiers du SISESA, WP.2006/6.
2. Bjorklund, A. Edin, P-A., Fredriksson, P. and Kruger, A. (2003), "Education, Equality and Efficiency: An Evaluation of Swedish School Reforms" SNS, Stockholm, Sweden.
3. Calero, J. (2006) "Desigualdades tras la Educación Obligatoria: Nuevas Evidencias", DT 23/2006. Fundación Alternativas, Madrid.
4. Case, A. and A. Deaton (1999), "School Inputs and Educational Outcomes in South Africa ", *Quarterly Journal of Economics,* 114, 1047-1084.
5. Eide, E. and M.H. Showalter (1998), "The Effect of School Quality on Student Performance ", *Economics Letters,* 58, 345-350.
6. Dearden, L., Ferri, J., and C. Meghir (1998), " The Effect of School Quality on Educational Attainment and Wages" Institute of Fiscal Studies W.P. 98/3.
7. Glewee, P. and M. Kremer (2005), "Schools, Teachers and Education Outcomes in Developing Countries", forthcoming in *Handbook on the Economics of Education*. Elsevier.
8. Hanushek, E. (1995), "Interpreting Recent Research on Schooling in Developing Countries", *World Bank Research Observer,* 10, 227-246.
9. Koenker, R. and G. Bassett (1978), "Regression Quantiles", *Econometrica*, 46, 33-50.
10. Koenker, R. and G. Bassett (1982), "Robust Tests for Heteroskedasticity on Regression Quantiles", *Econometrica*, 50, 43-61.
11. Levin, J. (2001), " For Whom the Reductions Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement", *Empirical Economics,* 26, 221-246.
12. Smith, J. and R. Naylor (2001), "Determinants of Degree Performance in UK Universities: A Statistical Analysis of the 1993 Student Cohort", *Oxford Bulleting of Economics & Statistics*, 63, 29-60.
13. Winkelmann, R. and S. Boes (2006), *Analysis of Microdata*, Spinger-Verlag.

Table 1a: Distributions of Students and School Characteristics by Grade

	Maths			***Introecon***			***Econhist***		
Grades	*S=0*	*AN=1*	*SM=2*	*S=0*	*AN=1*	*SM=2*	*S=0*	*AN=1*	*SM=2*
Frequency	37.05	53.63	9.32	26.62	68.92	4.66	10.36	85.24	4.40
Male	42.63	49.47	7.89	24.74	71.05	4.21	11.58	85.79	2.63
Female	31.63	57.65	10.71	28.06	66.84	5.10	9.18	84.69	6.12
Public	55.21	35.58	9.21	41.10	52.76	6.14	13.50	85.23	4.40
Charter	25.93	65.43	8.64	16.05	81.48	2.47	7.41	90.12	2.47
Private	22.54	67.61	9.85	15.49	80.28	4.23	8.45	86.62	4.93
Social Sc.	45.53	50.97	3.50	31.52	66.54	1.95	12.06	84.82	3.12
Tech.	9.09	63.64	27.27	9.10	77.78	13.13	5.05	85.86	9.09
NSc &Health.	7.69	92.31	0.00	23.08	76.92	0.00	0.00	100.0	0.00
Hum.	94.12	5.88	0.00	52.94	47.06	0.00	23.53	76.47	0.00
Spanish	36.23	54.78	8.99	26.38	69.57	4.06	9.86	86.38	3.77
Foreigner	43.90	43.90	12.20	26.83	63.41	9.76	14.63	75.61	9.76
Entry Ex.*		0.668			0.610			0.498	

Note: (*) The figures in the last row correspond to the correlations between the (numerical) grades in each subject and the university entry-exam grades

Table 1b: Distributions of Grades by Students and School Characteristics

	Maths			***Introecon***			***Econhist***		
Grades	*S=0*	*AN=1*	*SM=2*	*S=0*	*AN=1*	*SM=2*	*S=0*	*AN=1*	*SM=2*
Male	56.64	45.41	41.67	46.08	50.75	44.44	55.00	49.54	29.41
Female	43.56	54.59	58.33	53.20	49.25	55.66	45.00	50.46	70.59
Public	62.94	28.02	41.68	65.69	32.33	55.56	55.0	40.43	47.07
Charter	14.70	25.60	19.44	12.75	24.81	11.11	15.00	22.19	11.76
Private	22.36	46.38	38.88	21.56	42.86	33.33	30.00	37.38	41.17
Social Sc.	81.81	63.28	25.00	79.41	64.28	27.77	77.50	66.26	47.05
Tech.	6.29	30.43	75.00	8.82	28.95	72.22	12.50	25.83	52.94
NSc&Health	0.70	5.81	0.00	0.30	3.75	0.00	0.00	3.95	0.00
Hum.	11.19	0.48	0.00	8.82	3.00	0.00	10.00	3.95	0.00
Spanish	87.41	91.30	86.11	89.22	90.22	77.77	85.00	90.57	76.47
Foreigner	12.59	8.70	13.89	10.78	9.78	22.23	15.00	9.43	23.53

Table 2: Grades Production Function Estimates

Dependent variable: Grades (numerical; tipified)

Variable	***Maths* Linear**	***Maths* Int. Terms**	***Intrecon* Linear**	***Intrecon* Int. Terms**	***Ecohist* Linear**	***Ecohist* Int. Terms**
Female	0.137** (0.066)	0.123* (0.067)	0.079** (0.039)	0.085** (0.043)	0.102 (0.089)	0.103 (0.090)
Foreigner	-0.068 (0.068)	-0.069 (0.111)	0.204 (0.130)	0.193 (0.132)	0.071 (0.147)	0.062 (0.149)
Charter	0.219*** (0.089)	0.322*** (0.113)	0.104 (0.107)	0.240* (0.136)	0.183 (0.121)	0.342** (0.154)
Private	0.290*** (0.077)	0.338*** (0.091)	0.240*** (0.091)	0.312*** (0.110)	0.118 (0.103)	0.259** (0.124)
Nat. Sc & Health	0.523*** (0.184)	1.088*** (0.326)	-0.044 (0.220)	0.082 (0..390)	0.031 (0.249)	0.460 (0.443)
Technical	0.676*** (0.081)	0.759*** (0.132)	0.449*** (0.097)	0.632*** (0.158)	-0.050 (0.110)	0.232 (0.180)
Humanities	-0.569*** (0.163)	-0.513*** (0.220)	-0.262 (0.194)	-0.113 (0.267)	-0.121 (0.220)	0.056 (0.301)
Entry grade	0.616*** (0.040)	0.615*** (0.042)	0.605*** (0.048)	0.590*** (0.050)	0.576*** (0.055)	0.553*** (0.056)
Course_0304	-0.469*** (0.099)	-0.490*** (0.102)	-0.234** (0.118)	-0.230* (0.121)	-0.251* (0.134)	-0.266* (0.137)
Course_0405	0.005 (0.095)	0.00 (0.096)	-0.155* (0.113)	-0.157 (0.115)	-0.226* (0.128)	-0.231* (0.130)
Course_0506	-0.375*** (0.100)	-0.390*** (0.102)	-0.302*** (0.120)	-0.302*** (0.122)	-0.384*** (0.136)	-0.747 (0.568)
Chart* N.Sc. & Hth..		-0.762* (0.432)		-0.288 (0.518)		-0.747 (0.587)
Chart*Tech.		-0.241 (0.202)		-0.412* (0.241)		-0.429 (0.273)
Chart*Hum.		0.012 (0.441)		0.048 (0.530)		-0.188 (0.600)
Priv.*N.Sc. & Hth.		-0.989** (0.449)		-0.102 (0.600)		-0.482 (0.6678)
Priv*Tech.		-0.069 (0.181)		-0.195 (0.121)		-0.447* (0.242)
Priv*Hum.		-0.196 (0..371)		-0.549 (0.445)		-0.494 (0.504)
Nº Obs.	386	386	386	386	386	386
R^2	0.607	0.614	0.435	0.443	0.275	0.287

Note: ***, **, * represent significance at 99, 95 and 90% respectively.

A constant term is included. Omitted group: males, Spanish, public school, social sciences, cohort 2002/03.

Table 3a. QR (and OLS). Maths

Dependent variable: Grades (numerical; standarized)

Covariates	Average	θ=25	θ=50	θ=75	θ=95
Female	0.137***	0.176*	0.166	0.079	0.066
	(0.066)	(0.100)	(0.163)	(0.071)	(0.133)
Foreigner	-0.068	-0.305**	-0.042	0.007	0.187
	(0.068)	(0.145)	(0.163)	(0.175)	(0.204)
Charter	0.219***	0.170	0.286***	0.320***	-0.234***
	(0.089)	(0.106)	(0.116)	(0.128)	(0.084)
Private	0.290***	0.260***	0.342***	0.337***	-0.061*
	(0.077)	(0.082)	(0.106)	(0.116)	(0.036)
NSc. &Health	0.523***	0.597***	0.534	0.480	0.457*
	(0.184)	(0.129)	(0.217)	(0.345)	(0.273)
Technical	0.676***	0.680***	0.614***	0.520***	0.638***
	(0.081)	(0.111)	(0.1233)	(0.097)	(0.148)
Humanities	-0.569***	-0.366***	-0.634***	-0.746***	-0.796***
	(0.163)	(0.080)	(0.182)	(0.194)	(0.228)
Entry grade	0.616***	0.704***	0..637***	0.593***	0.444***
	(0.040)	(0.056)	(0.065)	(0.051)	(0.060)
Nº Obs.	386	386	386	386	386
Pseudo-R^2	0.607†	0.393	0.425	0.428	0.416

Note: As in Table 2. Cohort dummies have also been included.

Table 3b. QR (and OLS). Introecon

Dependent variable: Grades (numerical; standarized))

Covariates	Average	θ=10	θ=50	θ=80	θ=98
Female	0.079***	0.027	0.166**	0.020	-0.040
	(0.029)	(0.105)	(0.083)	(0.081)	(0.111)
Foreigner	0.204	0.039	0.175	0.188	1.187**
	(0.130)	(0.125)	(0.163)	(0.213)	(0.521)
Charter	0.104	0.178	0.213***	0.176*	-0.183**
	(0.107)	(0.206)	(0.287)	(0.103)	(0.092)
Private	0.240***	0.357***	0.387***	0.181*	-0.152**
	(0.091)	(0.143)	(0.156)	(0.106)	(0.074))
N Sc. &Health	-0.044	0.277	0.534	-0.268	-0.757*
	(0.220)	(0.229)	(0.678)	(0.245)	(0.143)
Technical	0.449***	0.231	1.413***	0.483***	0.538*
	(0.224)	(0.211)	(0.243)	(0.158)	(0.283)
Humanities	-1.412***	-0.536**	-1.684***	-0.17	-0.096
	(0.097)	(0.278)	(0.282)	(0.164)	(0.148)
Entry grade	0.605***	0.578***	1.637***	0.718***	0.753***
	(0.048)	(0.099)	(0.108)	(0.081)	(0.096)
Nº Obs.	386	386	386	386	386
Pseudo-R^2	0.435†	0.264	0.425	0.375	0.506

Note: As in Table 2. Cohort dummies have also been included.

Table 3c. QR (and OLS). Econhist

Dependent variable: Grades (numerical; standarized)

Covariates	Average	θ=10	θ=50	θ=70	θ=98
Female	0.102	0.089	0.126	0.061	0.084
	(0.089)	(0.165)	(0.163)	(0.0841	(0.211)
Foreigner	0.071	-0.035	0.102	0.217*	0.018
	(0.147)	(0.225)	(0.163)	(0.125)	(0. 201)
Charter	0.183	0.211	0.213	0.186	-0.169**
	(0.121)	(0.306)	(0.187)	(0.133)	(0.074)
Private	0.118	0.363*	0.387***	0.111	-0.052*
	(0.103)	(0.193)	(0.126)	(0.085)	(0.030)
NSc. &Health	0.031	0.257	0.193	0.097	-0.785***
	(0.249)	(0.442)	(0.378)	(0.214)	(0.243)
Technical	-0.050	0.133	0.014	0.093*	0.048*
	(0.110)	(0.232)	(0.163)	(0.056)	(0.253)
Humanities	-0.121	-0.324**	-0.164*	-0.090	0.365
	(0.220)	(0.314)	(0.282)	(0.154)	(0.286)
Entry grade	0.576***	0.575***	0.637***	0.608***	0.581***
	(0.055)	(0.133)	(0.088)	(0.051)	(0.109)
Nº Obs.	386	386	386	386	386
Pseudo-R^2	0.275†	0.204	0.425	0.241	0.367

Note: As in Table 2. Cohort dummies have also been included.

Table 4: Multinomial Logit (Odds ratios)

Dependent variable: Grade category

Variable	Cat_1(3) Maths	Cat_2(3) Maths	Cat_1(3) Introeco	Cat_2(3) Introeco	Cat_1(3) Econhist	Cat_2(3) Econhist
Female	1.985***	1.881*	0.610***	0.564	1.033***	1.914*
	(0.582)	(1.143)	(0.169)	(0.383)	(0.373)	(1.109)
Foreigner	0.638**	0.820	1.427	2.268	0.522*	1.845
	(0.286)	(0.795)	(0.632)	(2.238)	(0.287)	(1.711)
Charter	2.757***	0.617*	3.123***	0.763	1.328*	0.706*
	(1.062)	(0.355)	(1.202)	(1.058)	(0.689)	(0.403)
Private	3.493***	0.731	3.849***	0.962	1.290***	0.965
	(1.187)	(4.119)	(1.436)	(1.623)	(0.542)	(0.713)
NSc. & Health	27.37	0.000	1.242*	0.000	13.90	0.006
	(33.82)	(0.001)	(0.984)	(0.015)	(21.34)	(0.012)
Technical	7.519***	73.67**	1.931**	13.94**	2.376*	3.173*
	(3.487)	(37.25)	(0.828)	(12.10)	(1.290)	(2.683)
Humanities	0.068	0.000	0.720	0.000	0.379	0.000
	(0.073)	(0.001)	(0.395)	(0.001)	(0.248)	(0.001)
Entry grade	4.550***	53.57***	5.429***	31.71**	2.316**	12.67**
	(1.258)	(26.41)	(1.630)	(15.51)	(0.673)	(5.636)
Nº Obs.	386		386		386	
Pseudo-R^2	0.430		0.293		0.219	
Log-lik	-203.24		-205.23		-153.38	

Note: As in Table 2. Cohort dummies have also been included.

Table 5: Probit and Grades Production Function Estimates (with selection correction)

Dependent variable: Grades (numerical; standarized)

Variable	**Participation Probit (Bil=1)**	**Variable**	***Maths* Linear**	***Intrecon* Linear**	***Ecohist* Linear**
Female	0.024 (0.033)	Female	0.126** (0.062)	0.081** (0.041)	0.098 (0.092)
Foreigner	0.094** (0.045)	Foreigner	-0.072 (0.0600)	0.195 (0.146)	0.076 (0.156)
Entry grade	0.125 (0.247)	Charter	0.189*** (0.089)	0.112 (0.107)	0.212 (0.145)
Residence (CM)	-0.168** (0.083)	Private	0.312*** (0.084)	0.274*** (0.102)	0.126 (0.112)
		N Sc & Health	0.496*** (0.195)	-0.041 (0.220)	0.027 (0.261)
		Technical	0.694*** (0.086)	0.473*** (0.106)	0.005 (0.131)
		Humanities	-0.554*** (0.181)	-0.293 (0.212)	-0.093 (0.242)
		Entry grade	0.592*** (0.046)	0.572*** (0.052)	0.556*** (0.059)
		Lambda	0.051 (0.047)	0.023 (0.062)	0.026 (0.075)
Nº Obs.	958	Nº Obs.	386	386	386
Pseudo- R^2	0.178	R^2	0.607	0.435	0.275

Note: As in Table 2. Cohort dummies have also been included.

Figure 1: Distributions of Grades

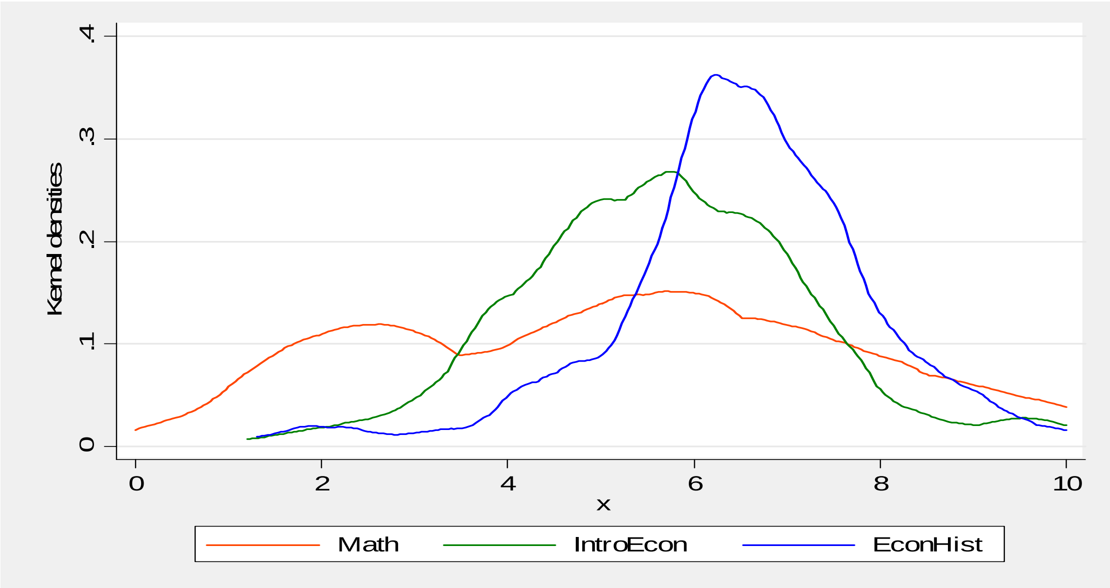

Figure 2: Distributions of Grades in the *Selectividad* Exam