NIPE

Documentos de Trabalho

Working Paper Series

"Predicting the Performance of a First Year Graduate Student"

Luís F. Aguiar

NIPE WP 4 / 2002

NÚCLEO DE INVESTIGAÇÃO EM POLÍTICAS ECONÓMICAS

UNIVERSIDADE DO MINHO

Predicting the Performance of a First Year Graduate Student

Luís Francisco Aguiar

Universidade do Minho - NIPE

Abstract

In this paper, I analyse, statistically, if GRE scores are a good predictor of the performance of a first year Economics graduate student. I also test if the scholastic year of an Economics PhD program contributes to the harmonization of the skills of students with different backgrounds. To both questions the answer is positive. GRE scores are relevant and the first scholastic year is effective in bringing together students with different entering skills.

1. Introduction

Every graduate program in Economics seeks the best students it can find. When choosing the students the main objective of the university is to increase its own reputation. So the university tries to maximize is the probability that the accepted students will excel in and expand the research frontiers.

One of the main objectives of the first year of a graduate program is to harmonize students with very different backgrounds. If the program is well succeeded in this task then the pre-acceptance academic background should not be a very good predictor of the performance of the students by the end of the year. Thus, when predicting the students' performance in the end of the first year, their personal, non-academic, skills should be more important. We would expect GRE scores to be an important predictor of students' future performance, because they are supposed to measure those skills.

In this work, I will test if the GRE scores can be helpful predicting the success of a first year Economics graduate student, and also if the first year of courses does have the effect of harmonizing students with different backgrounds.

Generally, in the Economics PhD programs, by the end of the first year the students have to take two or three qualifier exams (Q exams) in order to pass for the second year. It is normal to have one Q exam of Microeconomics, another of Macroeconomics, and some universities also have a Q exam of Econometrics.

Since the Q exams are the ultimate goal of the first year student I will use them as a (crude) measure of the success of the student.

I was able to gather data for the Microeconomics Q exam of a top research university (consistently ranked in the top 20). To evaluate the evolution of the student along the year I will use the grades of the Microeconomics courses in the first and the second semester.

2. What do previous studies tell us?

There is no consensus in the literature about the effectiveness of the GRE sores. E.g. while Dawes (1971) argued that GRE scores were good predictors of the success a first year graduate student in psychology, Dunlap (1979) concluded that the GRE was a weak predictor of the performance of a graduate student in Social Work. The best predictors were faculty interviews and the undergraduate GPA.

In an analogous study for the undergraduate admissions, Crouse and Trusheim (1991) concluded that the intensive use of SAT scores in the undergraduate admissions process was spurious and there were no benefits for the school to use this criterion.

Finally two other studies are worth to be mentioned:

- Sternberg and Williams (1997) calculate the correlations between the GRE scores and several variables that measured the success of the students. All the correlation coefficients were significant.
- Krueger and Wu (2000) concluded that the GRE scores (specially the quantitative part) were useful predictors of future success in the job market of economics graduate students.

The main problem with the latter study is that if one of the admission criteria is the GRE then there will exist a sample bias, since highly ranked schools will be able to capture the students with higher GREs and reject the others. So when one observes that people with higher GRE scores perform better in the job market, one can argue that that happens because of the quality/ranking of the school.

Attiyeh and Attiyeh (1997) addressed the issue of whether GRE scores are a decisive determinant of graduate school admissions. The results showed that GRE scores (particularly the quantitative section) are highly significant for the probability of an applicant being accepted in the graduate program in Economics. These findings confirm the doubts rose in the previous paragraph about what really determines the quality of the job placement.

3. The Data

I was able to find data for all students who attended the first year of a Ph.D. program in economics from 1992 to 1996. Some observations were lost because some students dropped the program before taking the Q exams, while others were accepted without taking the GRE tests. It was also the case that some transferred students were waived of taking some classes, so, whenever that information is used those observations have to be dropped. It was also possible to gather some information about the students who

attended the program between 1997 and 1999. Unfortunately, for confidentiality reasons, some observations were censored.

The data available include country of origin, original interests (e.g. Macroeconomics, Microeconomics, Labour Economics, Development, etc.), if the student has any kind of post-undergraduate studies before studying at the PhD Program (e.g. MA, MSc, etc.), kind of financial support (internal fellowship, assistantship, external fellowship or self funded), general GRE test scores, grades of the first year courses (Microeconomics I & II), and, obviously, the grades in the Q exams of Microeconomics. In each model I estimate is this paper I have between 95 and 110 observations.

In table 1, we can observe the mean and the standard deviation of the GRE scores of the students in the first year of their Ph.D. program. It is clear that both the quantitative scores and the analytical scores have increased. But the low standard deviation of the quantitative scores indicates that this section of the GRE is extremely important in the decision of accepting students in this PhD program, confirming the results of Attiyeh and Attiyeh (1997). This fact creates a problem to my analysis because I may not have enough variation of this variable, and it will not be possible to predict the results in the Q exams of potential students with low scores in the quantitative part.

The Q exams grades are divided in six categories: fail, pass masters, pass minus, pass, pass plus, and distinction. Since from the point of view of a Ph.D. student pass minus is the minimum score to proceed to the next level of the program, I will make no distinction between fail and pass masters. So in my models the grades range, discretely, from 0 to 4, being 4 the highest grade.

In table 2, we can observe the Spearman rank correlation coefficient between each of the GRE sections and the results in the Microeconomics Q exam[1] (the values between parenthesis are the p-values). Doing this simple analysis I find similar results to the ones described in Sternberg and Williams (1997). The main difference is that they found simple correlation coefficients instead of a rank correlation coefficient. But this analysis is not very rigorous because it is taking into account the effect of only one variable each time, and so that variable will capture part of the information contained in other relevant missing variables

4. The Model

The characteristics of the students relevant to determine their skills to perform well in the Q exams, may be grouped in three categories:

i. Individual capabilities and research interests,
ii. Academic background, and
iii. Effort

We can interpret the fact that most universities require the GRE scores to evaluate their applicants as an attempt to measure the individual capabilities. The original interests of the student, revealed in the statement of purpose of the application give us some information about their research interests.

About the academic background I had not much information. One of the most relevant informations I have is if the student has any kind of a master degree or not. I will also include the nationality of the students, grouped in large groups: Europe, English native speakers, Latin America, China, India, Russia, Pakistan, other Asia, Africa, and Israel. This way, I hope to capture any effect that is specific to a particular educational system.

[1] The Rank Correlation test is a distribution free test that determines whether there is a monotonic relation between two variables.

In this PhD program all students are supposed to take a mathematical course. However a student can be exempted from taking this course it he/she passes in a math exam taken in the beginning of the first semester. I will consider passing in such an exam as a signal (proxy) of a good mathematical background of the student.

Finally, about the effort, I do not have a way to measure the individual level of effort, so I will have no way to evaluate the impact of this factor, but it seems natural to assume, *ceteribus paribus,* that the higher the level of effort (like the number of hours studied) the higher will be the students results[2].

I will assume that there is some an index I, measuring the skills of the students, for which:

$$I_i = Xb + u_i \qquad (1)$$

Where X includes GRE scores, variables measuring the academic background, dummy variables for the countries and dummy variables for the original research interests. The error term u_i, which will also capture the individual effort, is assumed to follow a normal distribution, and the estimated coefficients are scaled so that u_i is standard normal.

For example, when estimating the results in the Q exams we assume that:

$$\begin{aligned}
&grade = 0 \quad if \quad I_i \le m_1 \\
&grade = 1 \quad if \quad m_1 < I_i \le m_2 \\
&grade = 2 \quad if \quad m_2 < I_i \le m_3 \\
&grade = 3 \quad if \quad m_3 < I_i \le m_4 \\
&grade = 4 \quad if \quad m_4 \le I_i
\end{aligned}$$

So the probability of each grade is given by:

[2] Even if I had data on this variable I would expect it to be more or less constant across the students. According to informal conversations, all students seem to make a great effort when studying for the Q exams.

$$\Pr(grade = 0) = \Phi(m_1 - Xb)$$
$$\Pr(grade = 1) = \Phi(m_2 - Xb) - \Phi(m_1 - Xb)$$
$$\Pr(grade = 2) = \Phi(m_3 - Xb) - \Phi(m_2 - Xb)$$
$$\Pr(grade = 3) = \Phi(m_4 - Xb) - \Phi(m_3 - Xb)$$
$$\Pr(grade = 4) = 1 - \Phi(m_4 - Xb)$$

Where $\Phi(\ )$ is the standard normal distribution function.

5. The Results

5.1 Microeconomics Q

Table 3 shows the ordered probit estimates where the dependent variable were the results in the Microeconomics Q, using only data that is known by the admissions committee (although the admissions committee does not know if the student will waive the mathematical course, I use that variable as a proxy of the students mathematical background, which is an observable information for the comitee).

When estimating the Microeconomics Q results, the demographic variables turned out to be individually and jointly statistically insignificant, so they were excluded from the model[3]. As we can see in table 3, there are only three variables that are individually significant at a 5% significance level. Surprisingly it seems that students with interests in Field C[4] do not perform so well as the other students. Unexpectedly having a masters degree does not affect significantly the students' performance. For the GRE scores we can see that the only relevant part is the verbal part [5]. Another interesting result, not included in the table, is related to the kind of financial support of the student. If the three dummy variables are added to the regression of table 3, representing the financial status

[3] I used the Likelihood ratio to test jointly all the demographic variables. The p-value of the LR statistic was 0.794.
[4] At request of the Director of Graduate Studies I am not more explicit about the fields of interest.
[5] The LR joint test that the coefficients of the analytical and the quantitative part are both zero has a p-value of 0.44. So I could not reject the null of both variables being insignificant.

of the student (internal fellowship, Assistantship, and external funding), they turn out to be jointly insignificant[6].

In table 4 we can observe the marginal effects[7] computed for the variables that are not statistically insignificant. As we can see the probability that a student fails the Q exam increases drastically when that student's research interests are in the field C. It is also clear that students with a good mathematical background have much more changes of having an excellent grade in the Q exam. We can see that the effect of an increase in the GRE verbal scores, although statistically significant, has virtually no impact for small variations of the score. To have a relevant impact we would have to consider changes in the order of 100 points, so that the probability of failing would decrease about 5 per cent points.

5.2 Microeconomics course — first semester

The grades in the course of Microeconomics I range, in my sample, from 0 to 7 (from C to A+). I could treat this variable as a continuous variable and use the ordinary least squares method; but, to make the results easier to compare with those of the previous section, I will use the same technique as before (ordered probit estimation).

The results are shown in table 5[8]. Comparing these results with the results of the Q exams the two most important differences are that the research interests are no longer relevant, and that having a Masters became statistically significant. It is also interesting to note that the score in the analytical section of the GRE test is almost significant at 10% level. These results suggest that having a Masters degree is important in the first

[6] The LR test for redundant variables has a p-value of 0.44.

[7] $\frac{\partial \Pr(.)}{\partial x}$ stands for a discrete change of dummy variable from 0 to 1 for the variables "Field C" and "Waived Math", and for continuous change for the other variable .

[8] The demographic variables were insignificant (the p-value of the joint test is 0.58), and the variables representing the research interests were insignificant too (the p-value of the joint test is 0.67).

semester, but at the end of the first year, they become irrelevant. It seems that the first year program is doing a good job in harmonizing students from different backgrounds. I should also note that the same did not happen about the mathematical abilities. They are important for both the Microeconomics Q and the Microeconomics I course.

In table 6, we can see that when variables describing the financial support are included, the results do not change dramatically. The most relevant change is that GRE scores loose significance, but the joint test shows us that they are still significant at 10% level[9]. It is interesting to note that students that receive some kind of internal financial support (in the form of a fellowship or assistantship) have higher probabilities of having higher grades. Computing the marginal effects (not shown) one can see that the probability of having a grade of A-, A, or A+ increases, while the probability of having any other grade bellow A- decreases.

Another question that can be raised is if it is relevant the kind of funding that the university provides. Since having support in the form of an assistantship implies a 15-hour weekly work for the student, it would be expected students on fellowship to have better grades. Performing a Wald test, in which the null hypothesis is that the coefficients of both forms of funding are equal, I conclude that I cannot reject the null hypothesis[10]. So if there is any difference, that difference is not strong enough to be captured decisively by the standard tests.

5.3 Does the masters degree loose relevancy after the first semester?

As we can see in table 5 or 6 having a masters degree is an asset when students are in their first semester of study in a Ph.D. program. In table 4 we also saw that having a

[9] When testing if the verbal, quantitative and analytical scores are jointly significant he LR statistic has a p-value of 0.085.
[10] the p-value of the statistic is 0.664.

masters degree is statistically insignificant to the performance of students in the Q exams. This raises the idea that the first year of a Ph.D. program does a good job in harmonizing the skills of students with different backgrounds[11]. If that was the case then we would predict that having a masters degree is not as relevant in the grades of the second semester as it is in the first semester. In tables 7 and 8 I replicate the same regressions of tables 5 and 6, using as a dependent variable the grade of the second semester Microeconomics course.

We see that having a Masters is statistically non significant in both tables. The pattern is clear. Having a masters is highly significant in the first semester, in the second semester looses significance, and for the Q exam becomes completely irrelevant.

Interestingly having waived mathematics is not significant in table 8, while it is marginally significant (at 10% significance level) when the type of funding is not considered. A possible explanation for this difference is the fact that 72% of the students who waived math received a internal fellowship. So the weak relationship found in table 7 is lost when we introduce the financial variables.

5.4 Safety Clause

Although the results presented previously are statistically significant, I should note that there is a huge uncertainty when predicting the performance of an entering student. In figure 1, we can observe the estimated probabilities of the grades of an average student. By average student I mean an artificial student whose characteristics are precisely the average of the sample. As we can see, for any specific prediction that we make there is more than 50% chance that we are wrong.

[11] We did not observe the same harmonizing result with respect to the mathematical skills. Note that the fact that a student waived the math course is relevant in both situations.

6. Conclusions

Three main conclusions are achieved in this paper. First, the verbal part of the GRE scores is relevant in predicting the success of a first year student. It is useful both in the first semester, when there is the first cultural clash, and in the end of the year when the Q exams are taken. The most striking result is the fact that the quantitative section of the GRE is statistically insignificant. This may be explained by the importance of that score in the PhD application process. Since the University considers that score to be very important it will accept only students with high scores in the quantitative section of the GRE. This has two implications: there is not much variation in the scores of the students (see table 1), and students with low scores who are accepted have other strengths I could not observe (because I did not have access to the students application files). If this is the case it will be even harder to capture the effect of the GRE quantitative score.

The second conclusion is that the first year courses contribute to harmonize students with different backgrounds. We saw that students with a masters degree have a competitive advantage in the first semester that disappears in the second semester and in the Q exam. Interestingly the same effect is not obvious with the mathematical background. Students with a good mathematical background seem to perform better in the Q exams.

Third, there is considerable uncertainty in forecasting which students will be successful in the first year of studies.

References

Attiyeh, G. & Attiyeh, G. (1997), "Testing for Bias in Graduate School Admissions", *The Journal of Human Resources*, vol. 32, pp. 524–548.

Crouse, J., & Trusheim, D. (1991), "How Colleges Can Correctly Determine Selection Benefits from the SAT", *Harvard Educational Review*, vol. 61, pp 125–147.

Dawes, R. (1971), "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making", *American Psychologist*, vol. 26, pp 180–188.

Dunlap, W. R. (1979), "How Effective Are Graduate Social Work Admission Criteria?", *Journal of Education for Social Work*, vol.15, pp 96–102.

Krueger, A. & Wu, S. (2000), "Forecasting Job Placements of Economics Graduate Students", *Journal of Economic Education*, vol. 31.

Sternberg, R., & Williams, W., (1997), "Does the Graduate Record Examination Predict Meaningful Success in the Graduate Training of Psychologists?", *American Psychologist*, vol.52, pp. 630–641 .

	Years	Mean	St. Deviation
GRE Verbal	1992—1996	524.38	120.12
	1997—1999	530.87	127.36
GRE Quantitative	1992—1996	752.03	38.01
	1997—1999	767.33	27.10
GRE Analytical	1992—1996	636.09	119.05
	1997—1999	658.70	116.17

Table 1: Evolution of the GRE Scores

	Verbal	Quantitative	Analytical
Micro Q Exam	0.2646 (0.0057)	0.1959 (0.0422)	0.1783 (0.0648)

Table 2: Spearman Rank Correlation Coefficient

Dependent Variable: Grade of the Q exam of Microeconomics			
	Coefficient	z-Statistic	Prob.
Research interests			
FIELD A	-0.0882	-0.19	0.85
FIELD B	-0.3508	-0.88	0.38
FIELD C	-1.1105	-2.48	0.01
FIELD D	0.0395	0.10	0.92
FIELD E	-0.064	-0.16	0.87
FIELD F	-0.2349	-0.67	0.50
FIELD G	-0.0305	-0.08	0.94
FIELD H	0.1108	0.23	0.82
Academic background			
MASTERS	0.0872	0.36	0.72
WAIVED MATH	0.6012	2.14	0.03
GRE scores			
VERBAL	0.0022	2.11	0.04
QUANTITATIVE	0.0043	1.26	0.21
ANALALYTICAL	-0.00007	-0.07	0.94
Number of obs	108		
LR statistic (13 df)	29.57		
Probability (LR stat)	0.0054		
LR index (Pseudo-R2)	0.0873		

Table 3: Ordered Probit Estimates for the Micro Q exam

	$\frac{d\Pr(Fail)}{dx}$	$\frac{d\Pr(pass-)}{dx}$	$\frac{d\Pr(pass)}{dx}$	$\frac{d\Pr(pass+)}{dx}$	$\frac{d\Pr(distinction)}{dx}$
FIELD C	0.3433	0.0587	-0.1573	-0.1333	-0.1114
WAIVED MATH	-0.1098	-0.1187	0.0208	0.0810	0.1268
VERBAL	-0.0005	-0.0004	0.0002	0.0003	0.0004

Table 4: Marginal effects for the Micro Q exam

Dependent Variable: Grade of the Course of Microeconomics I			
	Coefficient	z-Statistic	Prob.
Academic Background			
MASTERS	0.5255	2.23	0.03
WAIVED MATH	1.1596	4.15	0.00
GRE scores			
VERBAL	0.0018	1.85	0.06
QUANTITATIVE	0.0029	0.91	0.36
ANALYTICAL	0.0016	1.56	0.12
Number of obs	105		
LR statistic (5 df)	37.89		
Probability (LR stat)	0.0000		
LR index (Pseudo-R2)	0.0994		

Table 5: Ordered Probit estimates for the Microeconomics I course

Dependent Variable: Grade of the Course of Microeconomics I			
	Coefficient	z-Statistic	Prob.
Academic Background			
MASTERS	0.5952	2.50	0.01
WAIVED MATH	1.0254	3.58	0.00
GRE scores			
VERBAL	0.0013	1.28	0.20
QUANTITATIVE	-0.0001	-0.03	0.97
ANALYTICAL	0.0016	1.52	0.13
Financial Variables			
FELLOWSHIP	0.8356	2.70	0.01
ASSISTANTSHIP	0.6893	1.69	0.09
EXTERNAL	0.2037	0.63	0.53
Number of obs	105		
LR statistic (8 df)	45.91		
Probability (LR stat)	0.0000		
LR index (Pseudo-R2)	0.1205		

Table 6: Ordered Probit estimates for the Microeconomics I course, including financial variables

Dependent Variable: Grade of the Course of Microeconomics II			
	Coefficient	z-Statistic	Prob.
Academic Background			
MASTERS	0.2893	1.24	0.22
WAIVED MATH	0.4514	1.67	0.10
GRE scores			
VERBAL	0.0019	1.91	0.06
QUANTITATIVE	0.0029	0.91	0.36
ANALYTICAL	0.0011	1.11	0.27
Number of obs	105		
LR statistic (5 df)	15.20		
Probability(LR stat)	0.0009		
LR index (Pseudo-R2)	0.0406		

Table 7: Ordered Probit estimates for the Microeconomics II course

Dependent Variable: Grade of the Course of Microeconomics I			
	Coefficient	z-Statistic	Prob.
Academic Background			
MASTERS	0.3560	1.51	0.13
WAIVED MATH	0.2604	0.94	0.35
GRE scores			
VERBAL	0.0009	0.83	0.41
QUANTITATIVE	-0.0006	-0.18	0.85
ANALYTICAL	0.0010	1.00	0.32
Financial Variables			
FELLOWSHIP	0.8606	2.82	0.01
ASSISTANTSHIP	0.6136	1.54	0.12
EXTERNAL	-0.1422	0.45	0.66
Number of obs	101		
LR statistic (8 df)	27.15		
Probability (LR stat)	0.0007		
LR index (Pseudo-R2)	0.0725		

Table 8: Ordered Probit estimates for the Microeconomics II course, including financial variables

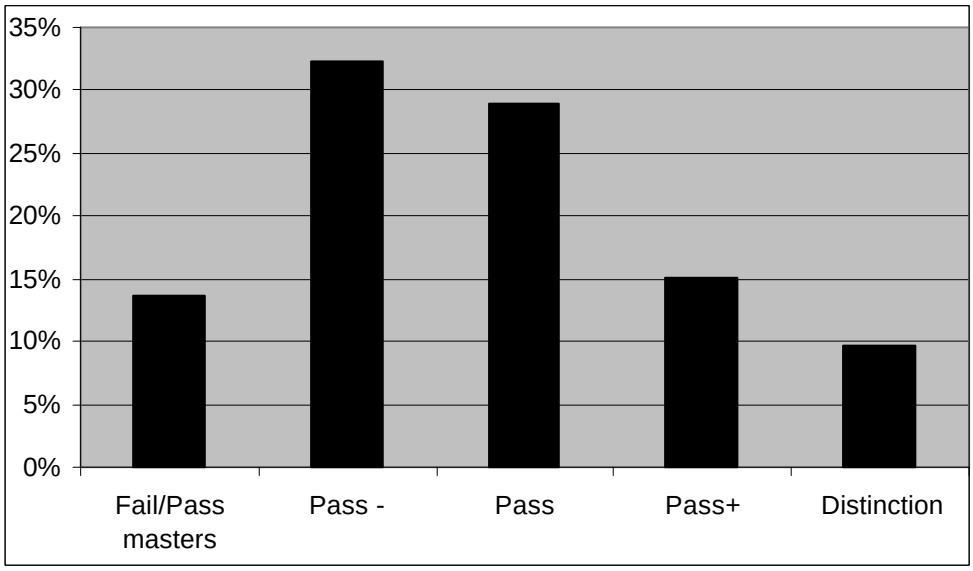

Figure 1: Estimates for the average student