How many attempts until success in some core 1st. year disciplines?[1]

Graça Leão Fernandes
ISEG-Technical University of Lisbon

João Andrade e Silva
ISEG-Technical University of Lisbon

Margarida Chagas Lopes
ISEG-Technical University of Lisbon

ABSTRACT

Due to a general development in education brought about by democracy, Portugal has witnessed tremendous development in Higher Education (HE) since the beginning of the 1980s. Nevertheless, the percentage of graduates among the Portuguese population still ranks far below most European countries. This is why academic performance in HE 1st cycle matters so much and warrants a careful analysis of the main determinants for failure.

Life cycle theories are taken as our main theoretical framework.

This paper worked with Individual semi-longitudinal data on ISEG students retrieved from the School Pedagogic Observatory Database. The number of attempts until the successful completion of two 1st year disciplines (Mathematics I and Economics I) was the chosen *proxy* for failure.

A discrete-time survival model was used. To deal with the unobserved heterogeneity a random variable was introduced - frailty - in the linear part of the model. A normal distribution was assumed in the estimation of a logit panel data model with random effects.

As expected, students took longer to achieve success in Mathematics 1 than in Economics 1. Gender, degree course, the track followed in upper-secondary, and the application mark to enter HE are among the significant determinants of success in Mathematics 1 and Economics 1.

Keywords: Higher education, Academic performance, Life cycle theories, Survival models, Determinants of success

---

[1] This paper has been developed in the framework of ISEG Pedagogic Observatory Studies and utilizes its data base.

INTRODUCTION AND PURPOSE

Due to a general development in education brought about by democracy, Portugal has witnessed a huge development in Higher Education (HE) since the beginning of the 1980s. In the decade 1980 to 1990 the Portuguese HE enrolment rate reached the corresponding value for Greece (1996), surpassed that of Belgium and grew until 2003. The trend then began to reverse due to demographic evolution (OECD 2006). A very noticeable feature in Portuguese HE is the high feminisation rate: 59.3% in 2009. This is higher than the corresponding rate for most European Union (59.2%) central, southern and Anglo-Saxon countries (OECD 2009). The increase in the overall demand for HE during the recent decades has indeed largely depended upon the enrolment of Portuguese women.

Nevertheless, despite this very positive evolution, the percentage of graduates among the Portuguese population is still far below most European countries, and the gap tends to widen when we consider the oldest age groups, as we can see in figure 1(Appendix 3). This is the reason why academic performance in HE 1st cycle matters so much and warrants a careful analysis of the main determinants of failure.

Failure in academic performance in the HE 1st cycle has become more important now that the Bologna Chart has been implemented for tertiary education. In fact institutional arrangements have become stricter under Bologna; namely there has been a reduction in the duration of 1st cycle degrees (from 4 to 3 years in most Portuguese social science programmes). This reduction in time was accompanied by a cut in the number of curricular units per programme, but it hardly affected the core curricular units, whose syllabuses remained practically unchanged in terms of extension and complexity.

Research recently carried out on four Portuguese higher education institutions' MSc. and PhD programmes (Alves, M. G, Cabrito, B., Lopes, M.C., Martins, A. M. & Pires, A.L. 2008) revealed that there are still great differences among institutions in the average time required to complete identical degrees. This outcome shows that under the three-year programme introduced by the Bologna Chart the student success rate in post-graduate courses varies widely among HE institutions (Chagas Lopes 2007). These results suggest that the main factors affecting students' performance at the beginning of higher education programmes are important.

Given the fact that syllabuses in the second and third years of the degrees are interdependent, we seek to avoid the effects of this interdependency on student performance by analysing two 1st year core course units and investigating the main features behind the failure rates of these. These units provide the basic main qualifications upon which further knowledge will settle during the course of the undergraduate degree.

In this paper an individual semi-longitudinal data on ISEG students retrieved from the School Pedagogic Observatory was used as the database. The length of time needed to successfully complete two 1st year units (Mathematics I and Economics I), which are common to the different programmes, was used as a *proxy* for failure. The main hypotheses are:

- success in completing core 1st year graduation subjects, measured by the length of time required by each individual to complete a subject, will be negatively affected by lower Social Economic Status (SES) of the family of origin, poor performance during previous schooling, present family demands, and possible income shortages;
- male and female outcomes will most probably differ either relative to success rates and to time patterns induced by the above determinants.

A discrete-time survival model was used. To deal with the unobserved heterogeneity a random variable was introduced - frailty - in the linear part of the model. A normal distribution was assumed in the estimation of a logit panel data model with random effects was .

After the Introduction and the General Purpose (Section 1), the guidelines for the Theoretical Background are presented in Section 2. The data and work sample are explained in Section 3. Section 4 explains the methodology. Some conclusions and policy implications are emphasized in Section 5.

GENERAL FRAMEWORK

In this paper life cycle theories are the main theoretical framework since dynamic transitional processes instead of single turning points are the main concern here, and to assess the interaction between schooling, labour market status and own family the main purpose of this paper. Educational success and failure are the outcome of dynamic and complex interacting features that spread quite diversely along individuals' trajectories and whose effects show up in the length of time taken to conclude a degree. Applying life cycle theories to education appears to be warranted whenever research concerns the effects on learning and schooling of factors which affect the amount of time needed to complete given course units. As previously mentioned, failure in tertiary education is measured by the length of time required to complete two core subjects.

Most research carried out on higher education success and failure rates still relies upon cross section methodologies supported by synchronic data. But learning is by itself a rather complex multidimensional and time dependent process, especially when it coincides with transitions to adult life (Bidart & Lavenu, 2005, Cunha, F., Heckman, J.J. & Lochner, L. 2006). Likewise, analyses on school success and failure which do not take into account dynamics risk neglecting a great deal of the corresponding major determinants, namely most of those which characterize transition to adulthood for women and men.

Individual longitudinal trajectories have long deserved increased attention from labour economics research[2]. This research has been developed within the framework critiques of human capital theories, and it takes a broader modern approach in which the role played by life cycle theories attracts greater concern. Its main purposes encompass the identification of the major interactions that take place between education/training and work/earnings (and family, sometimes) trajectories along individual life cycles.

Quite diverse impending restrictions can be at stake by the time one attends higher education: self-motivation and resilience, programme scheduling and general accessibility, and even employment and income restrictions, sometimes combined with family responsibilities among many others.

The OECD Examiners' Report on higher education in Portugal stresses that "(…) price is a major determinant of student choice (…)" (OECD 2006: 28), an expected outcome given the current average level of tuition fees and the increasing constraints on public social policy. Most Portuguese students studying for their first degree have indeed to depend on a meagre scholarship or a place in the labour market. So, income restrictions and the need to cope with them by means of a paid part time or even full time job should be taken in account when researching time allocation by Portuguese HE students.

Other factors often mentioned by literature and research are the parents' school level, own previous schooling patterns and the role played by education institutions successively

---

[2] See, among other, Ben-Porath 1967; Heckman & Macurdy 1980; Albrecht, J., Per-Anders, E., Sundström, M., Vroman, 1991.

attended. Education sociology gives these factors special attention when trying to approach multiple interaction effects of the interplay between individual and structural factors along life cycle trajectories. Dynamic analyses have been enlightening the meaningful role usually played by previous school trajectory upon future studying and later employment and career opportunities (Hanushek & Welch 2006).

Both education sociology and economics of education research have been shedding light on the influence of origin (father's and/or mother's) and present families' social and educational background (SES) upon school and employment success. For upper secondary education, previous research using semi-longitudinal data as well as official reports based on synchronic data confirm that SES actually exerts a meaningful impact on Portuguese students' opportunities (Chagas Lopes, M. & Medeiros, J. 2004; ME-GEPE 2007). But does it affect tertiary education students? The literature provides empirical evidence that it does although the pattern of influence changes with students' age, gender and school path (Vandenberghe 2007; Hassink & Hanna 2007). We try to answer the above question for the Portuguese HE.

Among the determinants, fathers' and mothers' education level is one of the most important ones; OECD (2007) reports that Portuguese HE students have one of the highest social immobility rates as the share of those with tertiary education whose father's education level is tertiary too is extremely high. Nevertheless, father's and mother's school level may well influence their children success opportunities differently, depending upon children's academic path, gender and other features (Pronzato 2008).

Tracing the main gender differences is another objective of this research. The parents' situation in terms of activity, employment and occupation are also factors to be addressed. Most 1st year students are still living with their parents and are dependent upon their family of origin's income. Income failure or budget constraints in the family of origin may affect children's studying opportunities, namely by forcing them to search for a paid job in the labour market, therefore affecting the average length of time needed to successfully complete core subjects. The pertinence of studying these determinants is increased now due to the economic crisis. In fact the effects of the family of origin's socioeconomic status (SES) in an economic crisis context have been well established; for youngsters in their late teens the influence of SES tends to increase with a bad economic situation (Belley & Lochner 2007).

All these determinants interplay to foster not only educational access and success (or failure) but also background values, beliefs and motivation, which shape life cycle trajectories (Plug 2002; Black, Devereaux & Salvanes 2004). Background values, beliefs and motivation are not dealt with here although there is the awareness of their impact on individual school trajectories.

School trajectories and success prior to transition to higher education have been extensively investigated by the reference literature. For Portuguese upper secondary students ME-GEPE shows that girls exhibit lower age deviations relative to the expected age and higher scores in prior trajectories than boys. The report stresses that girls' socialization is more prone to school values than boys', which is associated with the higher school expectations generally developed by girls when compared to boys. Another line of argumentation emphasises the fact that girls invest more in school to overcome discrimination in other fields like the labour market (ME-GEPE 2007, *op. cit.*).

More recently, on the eve of Bologna agreement, research concerning higher education has been developed in some EU countries (Noyes 2003; Ammermüller 2005). No study of the kind has yet been developed for Portugal. Therefore retention episodes and their frequencies either during basic or in secondary education have to be investigated in the framework of research on success rates in the Portuguese HE 1st cycle. Mobility between school establishments in school cycles prior to tertiary education must be addressed as well.

Research on the Portuguese upper secondary and tertiary patterns has been providing evidence on the influence upon school success of some main individual characteristics, such as age and gender (Chagas Lopes, M. & Medeiros, J. 2004, *op cit*; Chagas Lopes, M., Medeiros, J. & Pinto, A. 2005; Amâncio (2005) and Perista, H. & Silva, A. (2004)). As previously mentioned, the feminisation rate among most Portuguese HE programmes is consistently increasing, despite an overall enrolment downturn trend over the last years. Will the time restrictions behind failure equally impart upon women's and men's trajectories?

There is the aim to assess the joint effect of the above mentioned determinants on the amount of time required by each individual to complete a set of two core 1st year course units, as well as, to systematically investigate gender patterns associated with the time spells. Finally, time required to successfully complete each subject was set as a *proxy* for the failure rate. To derive the main determinants of women's and men's success rates is an expected result of the research here.

DATA AND WORK SAMPLE

Semi-longitudinal data on ISEG students retrieved from the Pedagogic Observatory was taken as database. This database has observations on 1758 students over the three years and the four graduation programmes (Economics, Management, Finance and Mathematics Applied to Economics and Management-MAEG).

The above mentioned database has information about:

- each individual's characteristics (gender, place of birth and nationality);
- Social Economic Status (SES) of the family of origin (fathers' and mothers' education level and their occupation and employment status;
- each individual's previous school trajectory (upper secondary track, HE application mark);
- each individual's employment status, civil status, occupational situation on entry to HE.

There is also information about each student's enrolment and success/failure in all 7 semesters from the first semester of 2006 till the first semester of 2009 for Mathematics 1 and in all 4 semesters (1st, 3rd, 5ª and 7º) for Economics 1 since this subject is only taught in 1st semester of 1st year. The database excludes students following the MAEG major for Mathematics 1[3] but not for Economics 1. The missing values high frequency is a problem in some of the variables considered for analysis, in particular, Civil Status, Employment status, Mother's and Father's Education Level [Appendix 1].

The main characteristics of the sample are the following:

- The feminization rate is 45.6%, 97% are Portuguese, 99.6% are single, and 91.2% perform no regular paid occupation.
- With regard to the field of study, 38.1% are enrolled in Economics, 47.1% in Management, around 7.5% in Finance, and 7.3% in MAEG.
- Concerning the SES of the family of origin, fathers' and mothers' school level, figure 2(Appendix 3) shows that most parents have at least a secondary certificate.

The majority of parents were employed in the beginning of the 2006/2007 academic year; more mothers than fathers suffered from unemployment and were inactive or retired

---

[3] These students curricula have Analysis 1 instead of Mathematics 1.

from the labour market as showed in Figure 3(Appendix 3). The above outcomes replicate quite closely the average Portuguese employment situation in the corresponding period of time, except for fathers' unemployment share, which appears to be lower than the average.

- Self-employment is higher among fathers (30.1%) than among mothers (15.2%). Most fathers and mothers are employees (64% and 67.6%, respectively). We must notice the meaningful share of family non paid workers among mothers (16.3%).
- With regard to students' previous school trajectory 90.9% - e.g. Portuguese and foreign students – completed upper secondary in Portugal.

METHODOLOGY

To answer questions like "how many attempts are necessary to complete Economics I successfully? And Mathematics 1?" a discrete-time survival models is used. This is a natural choice for two main reasons. First, it is clear that there is a survival problem since each student enters the experiment when he/she enrolls in the course for the first time and leaves (a) when he/she passes, (b) if the observation period ends before passing, or (c) if he/she drops out of ISEG. Here the success event is, obviously, to pass the course unit. Second, the use of discrete-time models is mandatory since accreditation with a passing grade can only be obtained twice a year, at the end of the fall semester or at the end of the spring semester. As the number of observed failures is limited and relatively small due to the period of observation (in our analysis this period is less than or equal to 7 semesters), it makes no sense to approximate the number of attempts by means of a continuous variable.

Since the pioneering approach of Cox (1972), which was mostly devoted to continuous models but sowed the seeds for discrete-time models, some authors have a developed discrete-time approach (Allison (1984), Singer and Willet (1993), Box-Steffensmeier and Jones (2004), among others). Special mention should be made of the paper by Singer and Willet (1993), which gives a very simple presentation of such models and uses them to model educational data. Note, however, that the problem discussed in Singer and Willet (1993) is quite different from our situation since they are dealing with the duration of the career of special educator.

Let us first consider a homogeneous population to define the key concepts. Let the random variable T represent the number of attempts until success whose support is the set of positive integers. To characterize this random variable the probability mass function can be used, $f(t) = \Pr(T = t)$, the survival function, i.e. the probability of not experiencing success until attempt t, $S(t) = \Pr(T > t)$ or, as is most commonly the case, the hazard function, $h(t) = \Pr(T = t \mid T \geq t)$, which represents the probability of success at attempt t given that success has not been observed yet. Using the fact that $S(0) = 1$, it is straightforward to see that $f(t) = h(t) \prod_{j=1}^{t-1}(1 - h(j))$ and that $S(t) = \prod_{j=1}^{t}(1 - h(j))$. Although these functions can be estimated using the well-known Kaplan-Meier estimators, the same results can be obtained using an approach based on the logistic regression, which will be more convenient to deal with the heterogeneity problem in our data.

As the hazard rate is a probability, one solution is to model the logit transform of the hazard rate by means of a time dependent coefficient, i.e. $\mathrm{logit}\left(h(t)\right) = \ln\left(\dfrac{h(t)}{1 - h(t)}\right) = \alpha_t$, which is to say that $h(t) = \dfrac{\exp(\alpha_t)}{1 + \exp(\alpha_t)} = \left(1 + \exp(-\alpha_t)\right)^{-1}$.

It is well-known (see Singer and Willett (1993) for instance) that maximum likelihood estimates of α_t can be obtained by means of a logistic regression using a person-period data set, i.e. where each person generates as many records as attempts to get approval. Each record refers to the order of the attempt and if this attempt results in a success or a failure. The later variable (value 1 for a success and 0 otherwise) will be the dependent variable, and as many dummy variables as there are attempts in the sample (let k be that number) will be introduced. For instance, if a maximum of 7 attempts is observed in the sample and if a given individual is accredited with a passing grade at his third attempt, it will generate 3 records, the first one presenting value 0 for the dependent variable, value 1 for the dummy D_1 representing the first attempt and value 0 for variables D_2 to D_7, record 2 will also have value 0 for the dependent variable, value 1 for D_2 and value 0 for D_1 and for D_3 to D_7 and, finally, in record 3 we will get value 1 for the dependent variable, value 0 for D_1 and D_2, value 1 for D_3 and again value 0 for D_4 to D_7. If we define Y as being the dependent variable, the logistic regression will be given by

$$\Pr(Y_i = 1 \mid \mathbf{D}_i) = \frac{\exp\left(\sum_{t=1}^{k} \beta_t D_{it}\right)}{1 + \exp\left(\sum_{t=1}^{k} \beta_t D_{it}\right)} = \left(1 + \exp\left(\sum_{t=1}^{k} -\beta_t D_{it}\right)\right)^{-1}$$

Note that, having introduced as many dummy variables as possible attempts, the intercept in the linear combination has to be dropped to avoid perfect multicollinearity. After estimating the model, $\hat{\alpha}_t = \hat{\beta}_t$ and $\hat{h}(t) = \left(1 + \exp\left(-\hat{\beta}_t\right)\right)^{-1}$ the Kaplan-Meier estimates, will be obtained.

If the same data set is used but a different formulation for time dependencies is choosen, different results will be found. In many situations there can be a pattern for time dependency that allows to get a more parsimonious estimation of the model.

Now observed heterogeneity in the population can be introduced by means of a set of explanatory variables, $\mathbf{x}_i$, for individual i. The generalization is easy, and, for instance, one gets $h_i(t) = \Pr(T = t \mid T \geq t, \mathbf{x} = \mathbf{x}_i)$ instead of $h(t)$. Note that for simplicity it is assumed that the explanatory variables are "time" independent, but it can also be considered that some or all of them are time dependent (in this case we must consider $\mathbf{x}_{it}$ instead of $\mathbf{x}_i$). Following Singer and Willett (1993), the previous formulation will be generalized, using the same data set enlarged to include explanatory variables. If all explanatory variables are time independent, the different records generated by the same person will differ on 2 points only: the value of the dummy variable, D, which represents whether a student attempts to conclude a course unit or not; and for the last recorded attempt, its outcome, which is represented by a dummy variable, Y (value 1 for success, or 0 for failure). For all records but the last one the value of the dependent variable is 0.

The logit model using the person-period data set is then defined by

$$\Pr(Y_r = 1 \mid \mathbf{x}_r, \mathbf{D}_r) = \frac{\exp\left(\sum_{t=1}^{k} \beta_t D_{rt} + \sum_{j=1}^{r} \beta_{k+j} x_{rj}\right)}{1 + \exp\left(\sum_{t=1}^{k} \beta_t D_{rt} + \sum_{j=1}^{r} \beta_{k+j} x_{rj}\right)}$$

Standard statistical procedures for the logistic regression can be used to analyze the results and to get the "best" possible model. To comment on the parameters estimates it is sometimes easier to think in terms of the logit function,

$$\ln\left(\frac{\Pr(Y_r = 1 \mid \mathbf{x}_r, \mathbf{D}_r)}{1 - \Pr(Y_r = 1 \mid \mathbf{x}_r, \mathbf{D}_r)}\right) = \sum_{j=1}^{k} \beta_j D_{rj} + \sum_{s=1}^{r} \beta_{k+s} x_{rs} .$$

Broadly speaking, each parameter β can be interpreted as a proportional change in the odds ratio per unit of change of the corresponding variable. As in the previous case, time dependency can follow a given pattern and the model should be adapted to incorporate such a pattern. Thinking in terms of the hazard and assuming that the characteristics are time independent, one gets

$$\text{logit}\left(h_i(t \mid \mathbf{x}_i)\right) = \ln\left(\frac{\Pr(Y_i = 1 \mid \mathbf{x}_i, D_t = 1, D_u = 0, u \neq t)}{1 - \Pr(Y_i = 1 \mid \mathbf{x}_i,, D_t = 1, D_u = 0, u \neq t)}\right) = \beta_t + \sum_{s=1}^{r} \beta_{k+s}\, x_{is} \ ,$$

that is

$$h_i(t \mid \mathbf{x}_i) = \left(1 + \exp\left(\beta_t + \sum_{s=1}^{r} \beta_{k+s}\, x_{is}\right)\right)^{-1} .$$

Until now it have been assumed that all the heterogeneity can be captured by the observed explanatory variables, but in empirical work this is rarely the case. Most of the time relevant variables cannot be observed (or sometimes the variables are observed with error), and, consequently, there will be unobserved heterogeneity in the model.

Ignoring unobserved heterogeneity will bias downward hazard probabilities and underestimate time-independent covariates effects. The rationale behind this conclusion is that the observations with a higher hazard rate for the same values of the independent variables will obtain a passing grade (experience the event) faster, so the survivors at any given time are increasingly composed of observations with a lower hazard rate. If there are two groups of students where the first group experiences a higher proneness to obtain a passing grade, then the remaining individuals tend to form a more or less selected group with lower proneness. An estimate of the individual hazard rate, without taking into account this situation would therefore underestimate the true hazard function, and the bias would increase as times progresses.

To deal with the unobserved heterogeneity a random variable, called the frailty, will be introduced in the linear part of the model.

$$\ln\left(\frac{\Pr(Y_i = 1 \mid \mathbf{x}_i, \mathbf{D}_i)}{1 - \Pr(Y_i = 1 \mid \mathbf{x}_i, \mathbf{D}_i)}\right) = \sum_{t=1}^{k} \beta_t\, D_{it} + \sum_{j=1}^{r} \beta_{k+j}\, x_{ij} + e_i$$

where e is an unobservable random variable with mean 0 (to avoid an identification problem) and variance σ^2. The problem of choosing a specific distribution for the random variable remains. The literature about the implications of an erroneous choice of this distribution is not completely conclusive, but it seems that a flexible formulation for the time dependency helps to obtain more robust estimates. Once the distribution is chosen, the random variable has to be "integrated out" in order to obtain the marginal distribution of the dependent variable.

In the following models a normal distribution is assumed and the models will be estimated using a logit panel data with random effects, taking advantage of the Stata software. Each individual in the sample will correspond to an observation of the random variable.

CONCLUSION AND POLICY IMPLICATIONS

The analysis carried in this paper allows to identify the main determinants behind the number of attempts until success for two core first cycle disciplines, Mathematics 1 and Economics 1 (Appendix 2).

It should be borne in mind that the databases which supported the study of these two course units are not exactly the same. For Mathematics 1 there is information for all majors but MAEG for 7 semesters, since most students enrolled in each semester until they successfully completed the subject. So, even attempts match even semesters. For Economics

1 there is information for all four majors but only for 4 semesters because Economics 1 is taught only in odd semesters along the period that was at stake. Therefore, it was not correct to compare them on the basis of the number of attempts. In any case, the factors responsible for the number of repetitions are not the same for Mathematics 1 and for Economics 1.

The probability of success, within a shorter number of attempts in Mathematics 1 is positively affected by the HE application mark. This result is not surprising since entry requirements directly depend on the scores obtained during upper secondary combined with final examinations scores in Mathematics among other subjects. Moreover, it seems that the HE access mark also seems to encompass, at least partially, the effects from the interaction between ability and socio-economic background. As a matter of fact, when the HE application mark was inserted in the model, mothers' education level lost significance although the pattern evidenced by mother's education level was not linear. Although the literature states a positive interaction between parents' and especially mothers' educational capital and students' performance, there is some evidence that this is not so clear for HE students[4].

The impact of HE application mark on academic success reinforces the need for a more adequate articulation of Mathematics syllabuses between upper secondary courses and degree programmes.

One interesting result of the present study was that students whose application took place in the years 2007 and 2008 faced higher probabilities of failure, i.e. they needed, *ceteris paribus*, a higher number of attempts to complete Mathematics 1. As there has been a widespread conviction (followed by experts' and public discussion) that in the above two years selectivity in upper secondary final examinations was abnormally low, and that consequently examination scores were artificially inflated, the outcomes relative to those two years become fully justified.

For Mathematics 1 the attempts coefficient is increasingly negative with the number of attempts, which may be seen as a sign that students not only do not learn from past failure experiences but also that there is increasing lack of motivation as times goes by. In the 4th attempt the coefficient is slightly less negative probably due to the fact that this coincides with the end of the 2nd academic year, and so students make an effort to succeed.

For the above curricular unit an even semester effect was observed. When the even semester covariate was introduced, the variables attempts lose statistical significance. This effect can be explained by the fact that "second semester" is the one associated with year transition, which may be an incentive for students to intensify their study effort in order to complete an educational year. But it may also be that teachers and faculty will tend to relax selectivity criteria in second semesters, since they are aware that in even semesters they are dealing with a pool of students that are comparatively worse than the 1st semester one. This trend might be accompanied by less demanding pedagogical methodologies. These hypotheses would need to be confirmed by the analysis of unit scores in further research studies.

Gender and course were also shown to exert a positive influence on the probability of obtaining a passing grade more quickly. Women display better performances than men, an outcome which is in line with evidence in most studies on gender effect on academic performance. The Finance major also appears to be associated with a smaller number of attempts to complete Mathematics 1. Nevertheless, none of the other degree courses display a meaningful influence.

---

[4] Hassink,W. & Kiiver, H., 2007; Pronzato, C..

Considering now the results on Economics 1, a much lower number of attempts was observed compared to Mathematics 1, which is not surprising because it is a subject with a lesser degree of formalization. Moreover, since Economics 1 is not offered in both semesters but only in the fall one, a semester effect is obviously not found.

For Economics 1, neither sex nor degree programme, except for Finance, exert a significant effect. Now, this degree programme appears to be associated with a larger number of attempts. This result may be associated with the dissimilarity between Economics 1 and the other disciplines which make up the curriculum for Finance, but only the analysis of the subjects' scores would allow us to draw a robust conclusion on this feature.

As to the variables representing students' socio-economic background, for students with the same HE application mark, only mother's school level displayed a significant yet negative effect, which means that students whose mothers completed higher education levels make more attempts to complete Economics 1. This astounding outcome needs to be approached very carefully. One possible explanation lies in the fact that some students coming from wealthier families extend youth time at the expense of their families and neglect studying chores and responsibilities. This effect was found to be statistically more significant in males than in females in a study based on the ISEG database (Chagas Lopes & Fernandes 2010). In fact, when broken down by gender, mother's school level effect has a significant negative effect for boys but remained insignificant for girls. Male children from more educated mothers take longer to successfully complete the subject, which is in line with the explanation given above.To overcome this and other kinds of problems associated with students' immaturity and irresponsibility, ISEG is preparing the implementation of a system of prescriptions according to which a maximum number of attempts to conclude a discipline will be fixed.

For Economics 1, the track followed during upper secondary education seems to influence the number of trials to complete a discipline. Specifically, the model revealed that the area of Arts negatively affects success. The area of Science and Technology, on the contrary, has a positive influence although it is not statistically significant.

The latter outcomes emphasize once again the importance of a good articulation between skills and learning acquisition in upper secondary and higher education. For students in this sample, any one of the four tracks offered in upper secondary could lead to the four degree majors in ISEG albeit by means of different combinations of HE access disciplines. Naturally, students whose upper secondary trajectories did not comprise Mathematics and/or introduction to Economics, which is the case for those who followed Arts, would have more difficulties in successfully completing ISEG disciplines. A better conciliation between learning trajectories in upper secondary and specific higher education courses would be advisable too.

No "degree programme effect" on Mathematics 1 was found for either Economics or Management programmes, probably because the results are so bad that the discriminatory power of this variable is lost.

The "degree programme effect" associated with the Finance programme that was found on Mathematics 1 as well as on Economics 1 needs further development. As the institutional HE access requirements are the same for the four degree programmes at ISEG, we must investigate further students' characteristics and programme organization in order to shed light on this influence.

Given the burden of the high number of attempts to successfully complete some core subjects on their 1st year of HE, it seems advisable to implement supplementary classes and a mentoring system especially addressed to 1st year students and to the learning of those disciplines. Mathematics 1 is surely one of them.

REFERENCES

Albrecht, J., Per-Anders, E., Sundström, M., Vroman, (1991), "Career interruptions and subsequent earnings: a reexamination using Swedish data", *Journal of Human Resources* 34, 249-311;

Allison, P.D. (1984) Event history analysis. Regression for longitudinal event data – Quantitative applications in the social sciences, nº 46, Sage, USA;

Alves, M. G, Cabrito, B., Lopes, M.C., Martins, A. M. e Pires, A.L. (2008), *Universidade e Formação ao Longo da Vida*, Oeiras: Celta Editora;

Amâncio, L. (2005), "Reflections on science as a gendered endeavour: changes and continuities", *Social Science Information*, vol. 44, nº1, 65-83;

Ammermüller, A. (2005), "Educational Opportunities and the Role of Institutions", Research Memoranda 004, Maastricht, ROA;

Belfield, C. R. (2000), *Economic Principles for Education: Theory and Evidence,* Cheltenham, UK and Northampton, MA, USA, Edward Elgar Publishing, Inc.;

Belley, Ph. & Lochner, L. (2007), "The changing role of family income and ability in determining Educational achievement", *Journal of Human Capital* **2**(1), 1-31;

Ben-Porath, Y. (1967), "The production of human capital and the life cycle of earnings", *Journal of Political Economy* 75; 353-367;

Bidart, C. & Lavenu, D. (2005), "Evolutions of personal networks and life events", LEST, Université Aix-en-Provence, SN Working Papers, 27;

Black, S., Devereaux, P. & Salvanes, K. (2004), "Fast Times at Ridgemont High? The Effect of Compulsory Schooling Laws on Teenage Births", IZA DP, nº 1416 ftp://repec.iza.org/RePEc/Discussionpaper, (Ac. September 2006);

Box-Steffensmeier, J.M. and Jones, B.S. (2004) *Event History Modeling – A guide for social scientists*, Cambridge University Press, UK.

Chagas Lopes, M. & Medeiros, J., (2004), "School failure and intergenerational "human capital" transmission in Portugal", Paper presented at the European Conference on Educational Research, University of Crete, 22-25 September (http://brs.leeds.ac.uk/cgi-bin/brs_engine);

Chagas Lopes, M., Medeiros, J. & Pinto, A. (2005), "Does school improve equity? Some key findings from Portuguese data", Paper presented at the European Conference on Educational Research, University College Dublin, 7-10 September (http://brs.leeds.ac.uk/cgi-bin/brs_engine);

Chagas Lopes, M. (2007), "Time to complete a post-graduation: some evidence on "school effect" upon ISCED 7 trajectories", Paper presented at the European Conference on Educational Research, University of Ghent, 19-21 September (http://brs.leeds.ac.uk/cgi-bin/brs_engine);

Chagas Lopes, M. & Leão Fernandes, G. (2010), Success/Failure in Higher Education: How Long does it Take to Complete some Core 1st. Year Disciplines? *Journal of Research in Educational Sciences*, vol 0(1), pp. 32-50;

Cox, D.R. (1972) Regression models and life tables, *Journal of the Royal Statistical Society*, B, 34, 187-220.

Cunha, F. , Heckman, J.J. & Lochner, L. (2006), Interpreting the Evidence on Life Cycle Skill Formation, in Hanushek, E &Welch, F. (2006), *Handbook of Economics of Education*, Vols I, North Holland.

EUROSTAT (2009), *Key Data on Education in Europe 2009*, EURYDICE.

Hanushek, E & Welch, F. (2006), *Handbook of Economics of Education*, Vols I, North Holland

Hassink, W. & Kiiver, H. (2007), Age-dependent Effects of Socio-economic Background on Educational Attainment – Evidence from Germany, Tjalling Koopmans Research Institute, Discussion Papers nº 07-26;

Heckman, J. & Macurdy, T. (1980), "A life cycle model of female labour supply", *The Review of Economic Studies*, vol 47(1), 47-74;

Jenkins , S. P. (2005) Survival analysis, Umpublished Manuscript, Institute for Social and Economic Research, University of Essex, Uk. Available at http://www.iser.essex.ac.uk/teaching/degree/ stephenj/ec968/pdfs/ec968lnotesv6.pdf[5]

Lawless, J. (1982), *Statistical Models and Methods for Lifetime Data*, New York, Wiley & Sons;

ME-GEPE (Ministry of Education – GEPE), 2007, "Estudantes à Entrada do Nível Secundário de Ensino - Resultados Globais do Questionário Piloto" « http://www.gepe.min-edu.pt/np4/136.html» (Acc. 13th. July 2007);

Noyes, A. (2003), "School transfer and social relocation", *International Studies in Sociology of Education*, XIII (3);

OECD (2006), *Thematic review of tertiary education – country background report: Portugal*, http://www.oecd.org/dataoecd/23/1/37745972.pdf (Ac. 26th. June 2008);

OECD (2007), *Education at a glance 2007*, http://www.oecd.org/dataoecd/17/15/39245059.xls (Ac. 1st. July 2008);

Perista, H. & Silva, A., (2004), *Science Careers in Portugal*, MOBISC-EURODOC, Bruxelas, Comissão Europeia;

Plug, E. (2002), *How do Parents Raise the Educational Attainment of Future Generations?* Bonn, IZA DP nº 652, http://www.iza.org/index_html (Ac. July 2004);

Pronzato, C. (2008), "Why educated mothers don't make educated children ?", ISER Working Papers 11/2008, http://d.repec.org/n?u=RePEc:ese:iserwp:2008-11&r=edu (Ac. 20th June 2008);

Singer, J.D. and Willett, J.B. (1993) It's about time: Using discrete-time survival analysis to study duration and the timing of events, Journal of educational Statistics, 18, 2, 155-195.

Vandenberghe, V. (2007), "Family Income and Tertiary Education Attendance across the EU: An empirical assessment using sibling data", London, CASE-LSE W/P, nº123

Appendix 1

Variables	Valid	Missing	Missings (%)
Sex	1758	0	0
Civil Status	1645	113	6,87
Nationality	1647	111	6,74
Situation towards employment	1604	154	9,60
Application mark	1757	1	0,06
Mother education level	1591	167	10,50
Father education level	1585	173	10,91
Upper secondary track	1755	3	0,17

[5]

Appendix 2

Attempt1	Dummy variable assuming value 1 when it is student's first attempt and 0 otherwise. Similar definitions for attempt2, attempt3 and attempt4.
even_sem	Dummy variable assuming value 1 when the attempt corresponds to an even number
female	Dummy variable assuming value 1 for a girl
HE_appl_mark	HE application mark on a scale from 95 to 200 (for practical purpose students get an HE_appl_mark greater than 95)
mark07	When the enrolment year at ISEG is 2007, mark07 is equal to HE_appl_mark and 0 otherwise. Similar definition for mark08.
finance	Dummy variable assuming value 1 when the major is Finance and 0 otherwise
arts_track	Dummy variable assuming value 1 if student followed arts track at the high school level and 0 otherwise.
moth_EL_boys	Dummy variable assuming value 1 if a girl's mother has completed higher education level and 0 otherwise.

```
Estimated model for Mathematics 1
Random-effects logistic regression          Number of obs    =    3138
Group variable (i): np                      Number of groups =    1628

Random effects u_i ~ Gaussian               Obs per group: min =       1
                                                           avg =     1.9
                                                           max =       7

                                            Wald chi2(6)     =  192.99
Log likelihood = -1934.6795                 Prob > chi2      =  0.0000

------------------------------------------------------------------------------
        ap |    Coef.   Std. Err.     z    P>|z|    [95% Conf. Interval]
-----------+------------------------------------------------------------------
  Even_sem |  .7952345  .0995099    7.99   0.000    .6001986   .9902703
    female |  .3840315  .0852215    4.51   0.000    .2170003   .5510626
HE_appl_mark|  .0491858  .0041966   11.72   0.000    .0409606    .057411
   finance |  .4782928  .1596527    3.00   0.003    .1653793   .7912063
    mark07 | -.0022992  .0007106   -3.24   0.001   -.0036919  -.0009065
    mark08 | -.0037014   .000812   -4.56   0.000    -.005293  -.0021098
     _cons | -7.819057  .5939871  -13.16   0.000    -8.98325  -6.654864
-----------+------------------------------------------------------------------
  /lnsig2u | -1.510763  .5875419              -2.662324   -.359202
-----------+------------------------------------------------------------------
   sigma_u |  .4698314  .1380228               .2641701   .8356036
       rho |  .0628784  .0346207               .0207717   .1750791
------------------------------------------------------------------------------
      Likelihood-ratio test of rho=0: chibar2(01) =    3.39 Prob >= chibar2 = 0.033
```

Estimated model for Economics 1

```
Random-effects logistic regression        Number of obs    =    1844
Group variable (i): np              Number of groups   =    1588

Random effects u_i ~ Gaussian           Obs per group: min =      1
                                             avg =     1.2
                                             max =      4

                         Wald chi2(8)    =   417.16
Log likelihood  = -908.28757          Prob > chi2    =   0.0000

-----------------------------------------------------------------------------
      ap |   Coef.  Std. Err.    z   P>|z|   [95% Conf. Interval]
-------------+---------------------------------------------------------------
   attempt1 | -1.347347  .6947099   -1.94  0.052   -2.708953   .0142595
   attempt2 | -1.855672  .6920967   -2.68  0.007   -3.212157  -.4991876
   attempt3 | -2.747681  .7541021   -3.64  0.000   -4.225694  -1.269669
   attempt4 | -2.333712  .903665    -2.58  0.010   -4.104862  -.5625607
HE_appl_mark |  .0211396  .0047389    4.46  0.000    .0118515   .0304277
   finance | -.837328   .189942    -4.41  0.000   -1.209607  -.4650485
 arts_track | -.5786378  .2621857   -2.21  0.027   -1.092512  -.0647633
moth_EL_boys | -.4369171  .1441883   -3.03  0.002   -.7195209  -.1543133
-------------+---------------------------------------------------------------
   /lnsig2u | -1.280215  .5224467            -2.304192  -.2562387
-------------+---------------------------------------------------------------
   sigma_u |  .5272356  .1377263             .3159738   .8797484
      rho |  .0779118  .0375334             .0294537   .1904504
-----------------------------------------------------------------------------
Likelihood-ratio test of rho=0: chibar2(01) =     0.90 Prob >= chibar2 = 0.171
```

- Without frailty

```
Logistic regression                  Number of obs   =     1844
                         LR chi2(8)    =      .
Log likelihood = -908.73758           Prob > chi2   =      .

-----------------------------------------------------------------------------
      ap |   Coef.  Std. Err.    z   P>|z|   [95% Conf. Interval]
-------------+---------------------------------------------------------------
   attempt1 | -1.266179  .6522051   -1.94  0.052   -2.544477   .0121196
   attempt2 | -1.943211  .6488603   -2.99  0.003   -3.214954  -.6714685
   attempt3 | -2.931899  .7047398   -4.16  0.000   -4.313164  -1.550635
   attempt4 | -2.568054  .8423718   -3.05  0.002   -4.219073  -.9170359
HE_appl_mark |  .0199309  .0044234    4.51  0.000    .0112611   .0286007
   finance | -.7682583  .1740029   -4.42  0.000   -1.109298  -.427219
 arts_track | -.5465725  .2455562   -2.23  0.026   -1.027854  -.0652911
moth_EL_boys | -.4133283  .1354542   -3.05  0.002   -.6788136  -.1478429
-----------------------------------------------------------------------------
```
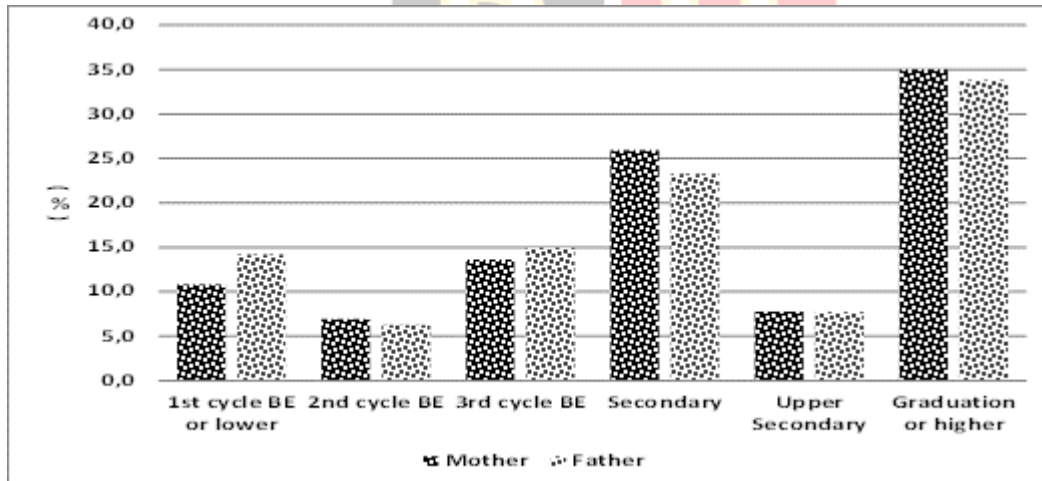
Appendix 3 – Figures

Figure 1
Percentage of population with tertiary qualifications (ISCED 5 and 6) in the population aged 30-64 by age group, 2007
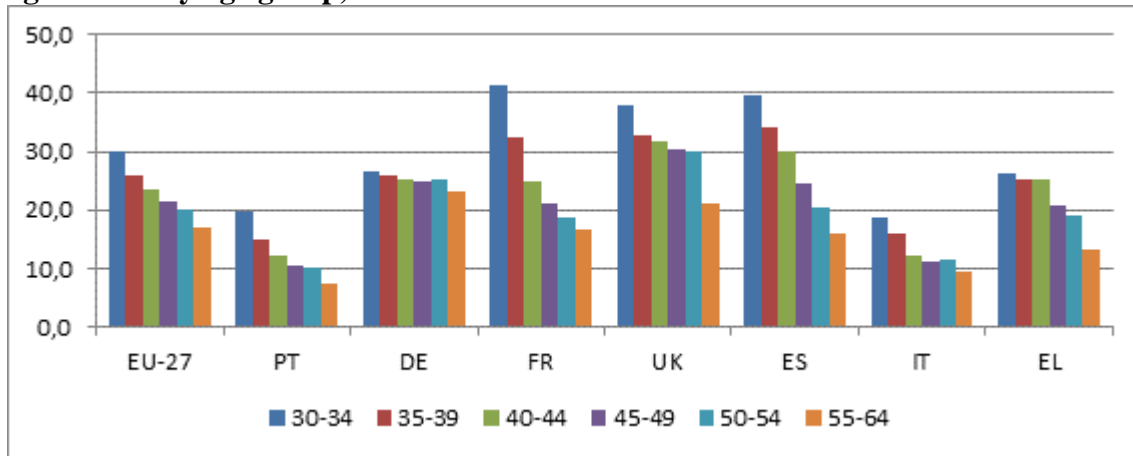
Source: EUROSTAT (2009).

Figure 2
Parent's school level

Figure 3
Parent's employment status