Can Schooling and Socio-Economic Level Be a Millstone to a Student’s Academic Success?

Christopher Bruffaerts
SBS-EM, ECARES, Université Libre de Bruxelles

Catherine Dehon
SBS-EM, ECARES, Université Libre de Bruxelles

Bertrand Guisset
Université Libre de Bruxelles

ECARES working paper 2011-016
Can schooling and socio-economic level be a millstone to a student’s academic success?∗

Christopher Bruffaerts† Catherine Dehon‡ Bertrand Guisset§

Abstract

By using data collected through a survey containing newly enrolled student in Management Engineering at the University of Brussels, we show that even if all students do not come with the same chance of success at university, their working/studying behavior may lessen the burden of the past. Unlike the majority of the literature focusing on deterministic vision of success, we propose a more balanced view of the determining factors of academic success where success is explained both by what the student controls and what he does not control. We indeed take into account by means of a multivariate analysis the background of the student (personal characteristics, schooling and human capital of the family) as well as variables that are related to the study methods and habits of the student such as class attendance, the regularity of study and the study capacity during the exam period. Our results show that the work/studying pays off: the two most relevant factors explaining success are the work/study regularity as well as the number of hours the student studies/works during the exam period. In addition and in contrast with the common belief, both class attendance and guidance courses do not seem to be important to succeed but are the keys in successfully completing the year with a grade.

Keywords: Academic achievement, Management Engineering, multivariate models, socioeconomic factors, study methods

1 Introduction

The determining factors of academic success have always interested researchers. In 1960, Fishman and Pasanella [14] listed nearly 600 studies that were published after

∗Christopher Bruffaerts and Catherine Dehon gratefully acknowledges research support from FRFC (Fonds de Recherche Fondamentale Collective). Catherine Dehon gratefully acknowledges the ARC contract of the Communauté Française de Belgique.
†Université libre de Bruxelles, ECARES, Institut de Recherche en Statistique. E-mail: cbruffaerts@ulb.ac.be
‡Université libre de Bruxelles, ECARES, Institut de Recherche en Statistique. E-mail: cdehon@ulb.ac.be
§Université libre de Bruxelles. E-mail: Bertrand.Guisset@ulb.ac.be
1950 and which were all devoted to the prediction of success at both school and university. The considerable literature available nowadays shows the complexity and the scope of this topic. Nevertheless, most studies up to now have been based on univariate analyses, meaning that the impact of a single variable or a single family of variables on the success of students is studied. This approach therefore only considers a single facet of this complex reality.

In addition, the economic literature usually takes into account variables which the student no longer has control over such as his schooling, the human capital of his family, his personal characteristics, etc.. The literature thus favours a deterministic vision of success, which in fact neglects the student’s action regarding his own destiny. In this study, we propose a more balanced view of the determining factors of academic success where success is explained both by what the student controls and what he does not control. This will be done using a multivariate analysis that takes into account the interactions between different factors that can have an influence on academic success.

Specifically, we oppose the background of the student which cannot be changed to a set of variables concerning his life at university and under which he has control over. On the one hand, variables such as individual characteristics, his schooling and his socio-economic level are included. On the other hand, we consider variables related to the study habits of the student such as attendance, regularity of study and study capacity during the exam period. We enrich the studies developed in the literature by adding a dimension on which the student can act, namely his ability to work/study. As strange as it seems, the literature on the subject is not very prolific. The literature generally addresses the issue indirectly through the study of the ‘big five’, which are the five personality traits that impact a student’s success and on which researchers all agree (Trapmann et al. 2007 [29]). Among these, the conscientiousness of the student is associated with a tendency to self-discipline, respect for obligations and desire for accomplishment. Among the five personality traits, conscientiousness is the one that pushes the student to work. Various studies show that conscientiousness is positively correlated with academic success (Busato et al. 2000 [6]; Furnham, Chamorro-Premuzic and McDougall 2002 [16]; Farside and Woodfield 2003 [13]).

The framework of this study takes part in Belgium’s French Community (CFB) and has two features: first, free access to university which is based only on the certificate of higher education and secondly the low cost of entry (reduced tuition fees). Indeed, the CFB has opted for a university system which is mainly financed by public funds and therefore significantly reduces the share of private contribution (the second cheapest education costs in terms of tuition fees and study material costs among several developed countries, Usher and Cervenan 2005 [30]). This policy allows more than 60 % of students from general education schools to pursue university studies (Ministère de la Communauté Française de Belgique 2007 [23]). However, high failure and dropout rates are observed at the end of the first year, which considerably increase the cost of public funding for higher education which itself has a negative impact on students who fail.

The study of success factors for the first year at university is crucial in a system
that seems a priori very egalitarian. In this context the higher education system fulfills its mission of social mobility and is an instrument of redistribution. By giving free access to university, is it sufficient to ensure equality of opportunity, equality of results? Recent literature shows that the answer is rather negative. Indeed, teaching in the French Community of Belgium (CFB) was found by several studies (Dupriez and Vandenberghe 2004 [11]; Jacobs et al. 2009 [22]) to be highly unequal, and reproducing socio-economic inequalities in society instead of playing a role in social mobility. Unfortunately, this finding persists at higher education levels as shown in studies of De Meulemeester and Rochat (1995) [8] and Arias Ortiz and Dehon (2008) [1] where it is shown that socio-economic factors influence success in the first year at university thus confirming the analysis of the sociology of education on the theories of cultural capital.

This analysis is based on data, which was collected through a survey, from new students (who come directly from high school) in Management Engineering at the ‘Université libre de Bruxelles’ for the academic year 2008-2009. This study aims to complement the analysis of Arias-Ortiz and Dehon (2008) [1] by looking at the impact of the student’s background (family, educational background, gender, etc.) on his results at university and by adding an essential factor: the student’s work/study capacity. The focus of our study is on the one hand to show that even if all students do not come with the same chance of success at university, their working/studying behaviour may lessen the burden of the past. On the other hand, econometric models will be applied to find out whether the student’s characteristics (nationality, sex, field of study, etc.), his high school career, his socio-economic level and his work/study habits can have an impact on his success.

The main question we are addressing is the behaviour of the student if he wants to succeed? We take the two most prominent elements that were revealed by the econometric analysis. First, results show that the work/studying pays off: the two most relevant factors explaining success are the work/study regularity as well as the number of hours the student has studied/worked during the exam period. The message is clear: success and work/studying do not go well together and without considerable effort from the student, academic success is more than compromised. Second, results indicate that class attendance which is often considered essential does not seem to be that important in success at university. Provided that the student works regularly during the year and that he works when exams approach, class attendance is not essential. Nevertheless, the multinomial analysis concludes that class attendance can differentiate between success without a grade and success with a grade. Unfortunately, despite the fact that students’ working/studying habits have a considerable impact on their success at university, the impact of students’ schooling remains significant but has a lesser impact.

The article is structured as follows: in Section 2, we introduce the data, the variables selected and the methodology used. Section 3 presents empirical results of the analysis and we conclude in Section 4.
2 Data and methodology

2.1 Data

The data used in the analysis were collected in April 2010 via a survey with about forty questions on different families of variables considered (individual characteristics, socio-economic level, educational background, methodology and work/study habits, academic performance). The population of interest consists of 611 students enrolled in the first year in Management Engineering at the SBS-EM (Solvay Business School-Economics and Management) for the academic year 2008-2009. They are asked questions about their situation in 2008-2009 and the course of that year. Among these new students, we must distinguish three groups: first, those who have passed their first academic year in 2008-2009 and who enrolled in second year at the time of the survey ($N_1 = 178$), second, those who failed their first year in 2008-2009 but decided to re-enroll in the first year ($N_2 = 182$) and third, those who gave up either during their first year or after the failure of their first year ($N_3 = 251$). This last group of students who left the university is unfortunately difficult to contact. In addition, these students are usually students who have made the wrong choice of study which may explain their dropping out of university after one year. Because our goal is not to study the determining factors of a poor choice of study, we have decided to restrict ourselves to the first two groups of students. This choice could be problematic if students who have failed do not have the same profile depending on whether the student has decided to re-enroll or not. One could think that those students who decided to leave university after the first failure might have a lower socio-economic level than those who decided to re-enroll. Some students may not have a second chance at university because of the cost of the studies. If these students, who are worse off from a socio-economic point of view, are not well represented in the sample, the average socio-economic level of students who failed will be pulled up, thereby introducing a bias in the results. Nevertheless, the profile of new students in the whole sample is not statistically significantly different from the profile of newly arrived students of 2009-2010 in which the three groups are represented. This empirical analysis allows us to assume that if one reasonably considers the homogeneous nature of populations throughout the years, we can analyze only students who have made an informed choice and have therefore invested in their studies. Our sample thus consists of 201 students from the first two groups: 74 students passed their first year in Management Engineering and 127 failed the first year but were again registered BA1 management engineer for the 2009-2010 academic year.

2.2 The variables and some descriptive statistics

In this section we justify the choice of variables with the perspective of analyzing the determining factors of academic success. We will also investigate the relationship between the success of the first year at university and each variable separately. At this stage the
analysis is purely descriptive and univariate which means that caution must be taken in the interpretations which can afterwards be confirmed or not using multivariate models. In the first part, we will consider the success in first year through the use of a binary dependent variable which indicates whether the student passed or failed his first year in Management Engineering. In the second part, we also analyze the quality of success through an ordinal variable taking into account the failure of a student, success without grade or success with a grade in the first year at university.

**Individual characteristics**

The first group of variables summarises the individual characteristics of the student such as his gender (male or female), his native language (French or not) and his nationality (Belgian or not). The question of the impact of gender on success has been widely studied in the literature. To name a few recent studies, one can refer to De Meulemeester and Rochat (1995) [8], Droesbeke et al. (2001) [9], Bahar (2010) [2], Duckworth and Seligman (2006) [10], Arias-Ortiz and Dehon (2008) [1], Castagnetti and Rosti (2009) [7]. Studies that show an effect all go in the same direction, namely that girls perform significantly better than boys.

Regarding the second variable, intuition leads us to believe that mastering the teaching language facilitates the success of students because lessons are easier to follow, the study is less laborious, exam questions are clearer, etc. The literature confirms the importance of language proficiency (Odunze 1982 [26]; Ho and Spinks, 1985 [20]; Graham 1987 [19]). Specifically, Frischenschlager, Haidinger and Mitterauer (2005) [15] show that having the same mother tongue as the teaching language increases the chance of success. It is therefore expected that students whose mother tongue is French perform better in their first year in Management Engineering where the teaching language is French.

Finally, the third question that arises in relation to individual characteristics is whether foreigners and Belgians are equal with respect to academic success. A recent study (Arias-Ortiz and Dehon 2008 [1]) replied affirmatively, when one takes into account socio-economic factors, the impact of nationality disappears. Other studies conclude that nationality nevertheless plays a role regardless of the socio-economic variables taken into account (see in particular De Meulemeester and Rochat 1995 [8]). We must also consider that there is a natural link between nationality and mother tongue.

To assess the relationship between these potential determining factors and success, we compare the success rates associated with different modalities of the variables and we test through a test of proportion if differences are significant. A modality of reference is selected and other modalities are compared against it.

We note in Table 1 that the difference in success rates between girls and boys is 5.08% for girls, a result which confirms existing studies. However, the test does not confirm that this difference is significant (p-value = 0.470).

---

1. Both tests are performed: the asymptotic test and the exact Fisher test (based on the hypergeometric distribution).
2. However, one must remain critical: the test result depends on the sample size which is relatively small.
For the other two variables, the results are very marked. Students for which French is not their first language perform significantly (p-value = 0.024) less well than others. The difference in success rates is indeed very high (25.16 %), which is consistent with the fact that mastering the teaching language has a positive impact on success. Note also that the non-Belgian students are far behind the Belgian ones with a difference rate of around 17 % (p-value = 0.111). This result should obviously be interpreted with caution; it may mask other effects (language proficiency, socio-economic level, etc.).

**Socioeconomic background**

Central to this section is the abstract concept of transmission of human capital, which means that parents pass certain traits to their children by means of education a series of skills, knowledge and personality. The transmission of human capital has been subject to much research that suggests that the higher the human capital of parents, the greater that of their children. Studies followed to link human capital of parents to the academic performance of children (Gaston, Wolinsky and Bohleber 1976 [18], Murnane and Maynard 1981 [24]; Galagedera 1991 [17]; Saha 1992 [27]; De Meulemeester and Rochat 1995 [8]; Blau 1999 [5]; Ernisch and Francesconi 2001 [12], Behrman and Rosenzweig 2002 [3]; Black, Devereux and Salvanes 2005 [4]; Arias-Ortiz and Dehon 2008 [1]). The level of education, the type of job or the income are generally selected as proxy variables of human capital. Regarding the results in the literature, they are mixed: while the effect of parental status on access to university is clear, its impact on success is not yet that clear. Arias-Ortiz and Dehon (2008) [1], however, establish that a significant relationship exists in the case of students from the 'Université libre de Bruxelles'. In particular, they show that the mother’s education level and the father’s job influence the success of new students. The more prestigious the occupational status of the father, the more students will tend to succeed. Similarly, the higher the diploma of the mother, the easier it will be for students to succeed in their first year at university.

In our questionnaire, the 19 modalities that covered the different occupations of the parents were grouped into five classes: unemployed, low and middle class, upper class, teachers, liberals and self employed. Concerning the parents’ studies, we distinguished three levels: university level, higher education level and the high school or lower education level. As the income of the household in which the student lives is extremely difficult to measure accurately, the two income issues are indirect: how does the student pay for his studies? Is he getting a scholarship? As we know that in Belgium grants are generally awarded on the basis of income and not on the basis of secondary school marks, only students from modest backgrounds can benefit from a scholarship. On the basis of two questions, we created a variable called ”Financing” which contains the information about whether a student has a grant and whether the student has to work to

(n = 203). This is to say that a non-significant result does not mean that there is no effect of the variable on success but simply that the sample does not suggest a significant impact on this variable.
Table 1: Differential rate of success (DRS) in first year of university ($n = 201$).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Modality</th>
<th>Frequency</th>
<th>DRS</th>
<th>Proportion test</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>male (+)</td>
<td>126</td>
<td>5.08%</td>
<td>0.470</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRENCH SPEAKING</td>
<td>yes (+)</td>
<td>180</td>
<td>-25.16%</td>
<td>0.024**</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BELGIAN</td>
<td>yes (+)</td>
<td>178</td>
<td>-17.02%</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FINANCING</td>
<td>parents (+)</td>
<td>173</td>
<td>-34.48%</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>work/scholarship</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOTHER’S DIPLOMA</td>
<td>high school or lower</td>
<td>40</td>
<td>-23.74%</td>
<td>0.01**</td>
</tr>
<tr>
<td></td>
<td>higher education</td>
<td>68</td>
<td>-13.88%</td>
<td>0.076*</td>
</tr>
<tr>
<td></td>
<td>university (+)</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FATHER’S STATUS</td>
<td>unemployed</td>
<td>7</td>
<td>-17.41%</td>
<td>0.373</td>
</tr>
<tr>
<td></td>
<td>low/middle class</td>
<td>36</td>
<td>-26.53%</td>
<td>0.006***</td>
</tr>
<tr>
<td></td>
<td>upper class (+)</td>
<td>87</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>teacher</td>
<td>12</td>
<td>-4.31%</td>
<td>0.779</td>
</tr>
<tr>
<td></td>
<td>liberal/self employed</td>
<td>59</td>
<td>-12.08%</td>
<td>0.145</td>
</tr>
<tr>
<td>CULTURAL ACTIVITIES</td>
<td>never (+)</td>
<td>165</td>
<td>5.91%</td>
<td>0.505</td>
</tr>
<tr>
<td></td>
<td>once or more</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATHS</td>
<td>weak</td>
<td>42</td>
<td>-25.47%</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>strong (+)</td>
<td>159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LATIN OR GREEK</td>
<td>no (+)</td>
<td>136</td>
<td>25.17%</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td>yes</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAILURE HIGHSCHOOL</td>
<td>never (+)</td>
<td>182</td>
<td>-34.85%</td>
<td>0.003***</td>
</tr>
<tr>
<td></td>
<td>once or more</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATTENDANCE</td>
<td>low</td>
<td>56</td>
<td>-27.59%</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>43</td>
<td>-21.11%</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>high (+)</td>
<td>102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WORK HOURS EXAM</td>
<td>less than 4hrs</td>
<td>21</td>
<td>-23.30%</td>
<td>0.014**</td>
</tr>
<tr>
<td></td>
<td>4 to 8hrs (+)</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>more than 8hrs</td>
<td>90</td>
<td>35.56%</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>REGULARITY</td>
<td>low (+)</td>
<td>110</td>
<td>31.12%</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUIDANCE</td>
<td>low (+)</td>
<td>123</td>
<td>13.16%</td>
<td>0.059*</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
pay for his studies.

In addition, we also included in this category a variable on cultural and social activities of the student (scouting, music, theater, improvisation, etc..) which also measures part of the social capital of students and that was most likely transmitted by parents or by the family circle. The variable "Cultural and social activities" covers a wide voluntary field to allow the student to place extra-curricular activities which are neither athletic nor festive.

Let us consider the empirical results of our sample one variable at a time in the same order as the one given in Table 1. Students getting a grant, having reduced tuition fees or working to pay for their studies succeed on average significantly less well than other students, the difference in the success rate being 34.48 % (p-value = 0.001). This confirms the fact that students whose parents are unable or unwilling to bear the full cost of education for their children are disadvantaged in terms of their chances of succeeding. One can also suppose that those who work to pay for their studies have paradoxically less time to devote to studying, which would explain a lower success rate.

We turn now to the variables that are related to the highest degree obtained by each parent. We chose to analyze the degree of the mother in the sequel for several reasons. On the one hand, the literature shows that the effect of the mother’s degree is more significant compared to the father’s degree. On the other hand, one may wonder what the link is between the degree of both the mother and the father: do these two variables not give the same information? A multiple correspondence analysis (MCA) can provide an answer to this question. Without going into detail, the MCA is a generalization of the principal component analysis to a qualitative type of variables. By summarizing the information on a small number of dimensions, the MCA allows the study of graphic links that exist between the modalities of different qualitative variables. We therefore obtain a map of the variables by representing the modalities of the different variables in a plan. The proximity of two modalities results from a graphical proximity on the variables map. Figure 1 shows the map that summarizes the results of the MCA for the variables containing the degree of the father and the mother (captured inertia ratio: 74% by the two first axes). It is very clear that parents tend to form couples who are academically "homogeneous". From this analysis we conclude that the information provided by the two variables partly overlaps and in order to avoid problems of multicollinearity in the regression model, we decided to retain only the degree of the mother. 3 The descriptive analysis of our sample also shows that students whose mothers have higher education level (respectively, a high school or lower level) were significantly less likely to be successful on average in their first year at university, with a difference in success rate of 13.88 % (respectively 23.74 %). Obviously, there is a relationship between a student’s success and the human capital of his family background: the student whose mother has a university degree performs better than the student whose mother has a higher education

3Through a joint Wald test in the regression models, we find the same result as in the literature, that the mother has more influence than the father on the success of the student.
level. This last group of students themselves do better than those whose mother only has a high school or a lower level diploma.

The results concerning the professional status of the father are in line with the expectations: students whose fathers have highly qualified jobs have more chance of succeeding than other students. The difference with the children of teachers is not significant. In contrast, the difference between, on the one hand, the success rate of children whose father is from the upper classes and, on the other hand, the middle and lower classes is respectively 26.53 % (p-value 0.006) and 17.41 % (p-value 0.0373). The success rate concerning students whose fathers are unemployed is also significantly lower than the one of students whose father is from the upper class.

Finally, the results for the intensity of social and cultural activities of students show a difference in the success rate of 5.91 % (not significant with a p-value = 0.505).
**Prior schooling**

Two different aspects of the student’s schooling are examined: the choice of courses followed at secondary school as well as the number of times the student had to repeat a year since the beginning of his studies (Droesbeke et al. 2001 [9]; Superby, Vandamme and Meskens 2006 [28]; Arias-Ortiz and Dehon 2008 [1]). Can the choice of courses followed by the student during high school explain academic success? The literature has often emphasized that the study of ancient languages such as Latin and Greek prepares the student particularly well for higher education because it develops the spirit of analysis, synthesis and organization. The same positive effect was noted for the strong options in mathematics, and this should be especially true for Management Engineers for whom the first year programme is strongly focused on sciences. According to the literature, the impact of mathematics in our sample is positive, the average success rate of the option of additional maths is 25.47% (p-value = 0.002) which is higher than the success rate of students from the option of basic maths. The results for Latin and Greek go in the same direction. Students who chose Latin and/or Greek in high school see success rates increase on average by 25.17% (p-value < 0.001). However, we should keep in mind that these variables introduce a bias of selection which is due to the fact that the difficult options such as mathematics and classical languages were chosen by the brightest students.

We also examine the impact of past failures during primary and secondary studies on the success of students at university. Are the students who already failed a year at school and therefore go "later" to university penalized? Our empirical analysis confirms this hypothesis: the success rate of students with "academic delay" is on average lower by 34.85% compared to students who never failed a year before going to university, and the difference is significant (0.003).

**Study methods and habits**

University studies are demanding, their success thus requires work and effort. This section therefore focuses on how serious the student is about his efforts and his self-discipline. The Belgian university system offers an analytical framework for this interesting perspective. Firstly, the evaluation is concentrated on the periods of final examinations each semester, one in January and one in June, plus a catch-up session in August. Secondly, class attendance is not generally required. The student thus has considerable freedom of organization: some begin to work the first week of the semester while others discover the teacher on the day of the exam. This is why we think that the choice of the student’s organization is of particular importance.

Everyone has heard that class attendance is important and that it is better to work regularly. To what extent is this true? Is attendance as important as we say? Is regularity of work so crucial for success? To answer these questions, our survey asked students about the different facets of their work: did they attend lectures? To which extent do they work hard during periods of revision and for exams? Do they work regularly? Do they use the assistance offered through guidance?
The descriptive analysis of our variables goes in the same direction as in the literature: conscientious students are rewarded. Indeed, the student who works/studies regularly throughout the year has on average a higher success rate of 31.12% (p-value = 0.001) compared to other students. Similarly, students who attend courses regularly also have significantly higher success rates compared to those whose attendance is low (27.59%) and medium (21.11%). Finally, the year ends with the exam period for which the study time is very important. Indeed, a student who studies more than 8 hours per day during this period will on average have 35.56% more chance of succeeding than a student working between 4 and 8 hours per day. Finally, a student who works less than 4 hours per day will have a lower success rate of 23.30% compared to a student working between 4 and 8 hours per day.

Note also that the students who take advantage of guidance (learning support) during the year also succeed significantly better (13.16%). However, this univariate result cannot directly lead to the conclusion that this service is helping students to succeed. Indeed, it is very likely that we are dealing with a selection effect: the most enthusiastic students and those who work the hardest are also those using guidance. This is why the evaluation of such promotional services to success are extremely complicated to implement (Imbens and Wooldridge 2009 [21]). We are going to check using a multivariate model whether the most conscientious students are also good students regarding variables such as their work regularity, course attendance, the number of hours during the exam periods and attendance at guidance courses. To answer this question, we applied a MCA where we show a map of modalities that are associated to those variables (Figure 2.2). The result is quite clear: there are three groups of modalities for which the interpretation is relatively straightforward. The modalities that characterize the students that are not conscientious are on the right of the map: they are not taking advantage of guidelines, they are not regular in their work, they are not attending courses and do not work/study much during the exam periods. The modalities on the left of the graph characterize students who are very conscientious: they are very regular, attend courses, they are hardworking and take advantage of the guidance available to them. Finally, the intermediate modalities are found towards the centre of the first axis which can itself be seen as characterizing the intensity of the student’s work. This analysis confirms that there is a selection bias concerning the use of guidance: the characteristics of students who use it are very different from the characteristics of those that do not.

3 Empirical results

As already mentioned, the failure rate during the first year at university is extremely high. This first year is a crucial step in understanding the determining factors of success at university. The interest is obviously to identify the determining factors of success without neglecting the relationship between variables. To model the success, we created a dichotomous variable: 1 if the first year is successful, 0 if not. Because the dependent
Figure 2: Multiple Correspondence Analysis using variables concerning work/study habits
variable is binary, the logit model will allow us to simultaneously investigate the influence of various personal and socio-economic characteristics, schooling and work habits on the probability of success in the first year at university of a new student:

\[ P(Y_i = 1|X_i) = F(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}) \text{ for } i = 1, \ldots, n \]  

where \( Y_i \) is the dependent variable which is equal to 1 if the year was successful, \( n \) is the sample size, \( X_i = (1, x_{i1}, x_{i2}, \ldots, x_{ip}) \) is the matrix containing all the independent variables, \( \beta = (\beta_1, \ldots, \beta_p) \) is the vector of regression parameters and finally \( F(t) = \frac{\exp(t)}{1+\exp(t)} \) is the logistics distribution function used to obtain a probability between 0 and 1.

The intuition behind this model is to analyze the probability of succeeding at university given different variables. If the estimate of a regression parameter is positive, then the explanatory variable associated to this parameter will have a positive influence on the probability of succeeding. Similarly, if the estimate of a parameter associated to an explanatory variable is negative, then it will reduce the probability of succeeding.

The results of the logistic regressions are presented in Table 2. Four successive models are analyzed: model 1 contains only the personal characteristics and in model 2 the socio-economic variables are added. Model 3 includes the first two families of variables as well as the variables concerning the student’s schooling. Finally, we add the student’s work habits to the first three families of variables. Each model contains three columns: the first gives the coefficient estimate, the second provides the p-value resulting from the test of significance of the explanatory variable and the third contains the odds ratio. At the foot of the three columns is a measure of the quality of the model, namely the Nagelkerke \( R^2 \). When a categorical variable has more than two modalities, the first to appear in the table is the reference category, which is why no coefficients are shown for those modalities. The p-value that appears up to this reference modality is the result of the global test of Wald for the group of modalities that describe the variable; to reject the null hypothesis means that globally the variable has a significant impact on success.

The analysis of model 1 shows that gender does not play a role: although the literature suggests that girls have an advantage over boys at university, this is not the case for this sample. Moreover, we can see that some of the handicap of foreigners with respect to their chances of succeeding is due to their low proficiency in the language of instruction (French) which is a variable that is significant in this model in contrast to the variable “Belgian”.

13
Table 2: Estimation results for the logit models ($n = 201$).

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>p-val.</td>
<td>OR</td>
<td>Est.</td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>0.35</td>
<td>0.25</td>
<td>1.42</td>
<td>0.53</td>
</tr>
<tr>
<td>FRENCH-SPEAKING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>-1.27</td>
<td>0.05**</td>
<td>0.28</td>
<td>-1.03</td>
</tr>
<tr>
<td>BELGIAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>-0.64</td>
<td>0.24</td>
<td>0.52</td>
<td>-0.93</td>
</tr>
<tr>
<td>FINANCING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>work/scholarship</td>
<td>-2.21</td>
<td>0.004***</td>
<td>0.11</td>
<td>-2.07</td>
</tr>
<tr>
<td>CULTURAL ACTIVITIES</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>0.28</td>
<td>0.48</td>
<td>1.32</td>
<td>0.29</td>
</tr>
<tr>
<td>university</td>
<td>0.07*</td>
<td></td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>EDUCATION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>higher education</td>
<td>-0.42</td>
<td>0.23</td>
<td>0.65</td>
<td>-0.44</td>
</tr>
<tr>
<td>highschool or less</td>
<td>-1.01</td>
<td>0.026**</td>
<td>0.36</td>
<td>-0.89</td>
</tr>
<tr>
<td>MATHS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>basic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LATIN/GREEK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>0.85</td>
<td>0.014***</td>
<td>2.36</td>
<td>0.76</td>
</tr>
<tr>
<td>HIGHSCHOOL FAILURE(S)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>once or more</td>
<td>-1.71</td>
<td>0.11</td>
<td>0.18</td>
<td>-1.47</td>
</tr>
<tr>
<td>ATTENDANCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>-0.038</td>
<td>0.94</td>
<td>0.96</td>
<td>-0.25</td>
</tr>
<tr>
<td>average</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WORK HOURS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 to 4h</td>
<td>-19.57</td>
<td>0.99</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4 to 8h</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>more than 8h</td>
<td>0.005***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REGULARITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>0.69</td>
<td>0.09*</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td>GUIDANCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>0.13</td>
<td>0.75</td>
<td>1.13</td>
<td></td>
</tr>
</tbody>
</table>

Nagelkerke $R^2$ 0.055 0.189 0.293 0.476
We will now examine the impact of the socio-economic background by analyzing the second model. The type of financing plays a significant role: students funded through a grant or an additional job are at a disadvantage compared to students funded by their parents.

The impact of the level of the mother’s education is globally significant and very interesting: the student whose mother has at most a high school diploma is less likely to succeed than if he was the son or the daughter of a mother with a university diploma. This difference is not significant between students whose parents have a university diploma and students whose parents have a higher education diploma.

The above results are fairly stable when the student’s schooling is also considered (model 3). This model emphasizes the importance of math and classic languages at secondary school. A student with a strong mathematical background has a higher probability of succeeding than a student with a lower mathematical background. Similarly, a student who has studied Latin and/or Greek sees his chances of succeeding increase.

Finally, it is the analysis of the last model (model 4) when compared with previous models that proves to be the most interesting. One can notice for instance that by adding variables concerning work/study habits, the positive impact of math and latin decreases slightly both in terms of significance and impact. The effect "latin-math" is partly explained by how committed the student is. The weight of the student’s schooling decreases when we add work/study habits. Variables related to work/study habits do not however explain everything: the variables latin and maths remain significant in the complete model. It is not surprising to find out that a good mathematical background and good training in classic languages are assets in the first year at university although the eventual selection problem for these so-called “strong” orientations must be borne in mind.

The effect of daily working hours during revision and exam periods is clear especially for variables associated with work habits: the more a student works/studies, the more his chances of being successful. Those who work more than 10 hours a day increase their chance of success with respect to failure (ratio of success / failure) by almost 4 times (OR = 3.79) compared to those working between 4 and 8 hours a day. Similarly, students working between 0 and 4 hours will have much less chance of being successful with respect to those who work between 4 and 8 hours. The impact of the variable regularity is similar: the more regular students are in their work throughout the year, the less likely they will fail. More surprisingly perhaps, the variable attendance is not significant, in contrast to the regularity and the number of hours worked during revision and exam periods. Provided that the student is organized and is capable of working hard during and at exam time, class attendance is not required.

Concerning the assessment tools for the availability of guidance courses, one can say that using guidances or not does not influence success. This result confirms the hypothesis of selection bias: attending guidance courses certainly shows how committed the student is and which might reflect other characteristics but is not in itself a determining factor of success.
The analysis of the model’s quality in terms of predictions is done through the Nagelkerke $R^2$ coefficient which gives very positive information for students: variables that capture the commitment and the efforts of the student are the variables that best explain success. Indeed, the model containing only the individual characteristics of a student has a predictive value of 5.5%. When the socio-economic is added, this value reaches 18.9% and when the schooling is added to the model the $R^2$ increases to 29.3%. However, the most important jump in this progression of the ability to predict success is due to the introduction of variables that incorporate work/study methods (47.6%). Obviously the goal of this analysis is not to predict, a priori, whether a student will succeed or not but to highlight the determining factors of success. As can be observed, the ability of the student’s work is clearly an important determining factor in success. Nevertheless, it is obvious that other unmeasured variables will also impact on the success, such as motivation or the quality of work performed.

So far we have confined ourselves to a Manichean world: 0 or 1, failure or success. Yet the reality is more complex, we know in any case there are different ways to be successful. In this second section we decide to go a step further by distinguishing two levels of success: success without a grade and success with a grade. From a technical point of view, we use the multinomial logistic regression in which the principle is to compare each alternative to the dependent variable relative to a reference modality selected a priori. The dependent variable has three modalities (failure, success without a grade and success with a grade) and the reference category is success without a grade. The multinomial regression generates two sets of coefficients used to compare, on the one hand, success with a grade with success without a grade and on the other hand success without a grade with failure. As the results from this last comparison are similar to the results of the logistic model (see Appendix, Table 4), we will focus on the comparison of success with and without a grade (Table 3).

As we have seen previously, course attendance is not a key determining factor in success provided that the student works regularly and studies sufficiently during exam periods. In contrast, the multinomial model shows that attendance can differentiate between succeeding with or without a grade. Attending courses is the best way to discover the material that is not included in course material, or simply getting to know the professor, understand his reasoning, his manner of questioning, etc. It is not uncommon for professors to reward class attendance by asking a few specific questions in the exam which only diligent students can answer. We also note that the use of guidance courses has a significant effect on obtaining a degree; one can therefore ask oneself whether the guidance is targeting students who really need help, or if it is simply improving the knowledge of good students. This time the socio-economic factors that emerge from this comparison are emphasized through the participation in cultural activities and the status of the father. Students from the lower and medium classes have a significantly lower chance of getting a grade compared to students whose father is upper class. The variables associated with the method/work capacity differ more when considering failure and success than when considering success with or without a grade.
Table 3: Estimation results for the multinomial logistic regression model \((n = 201)\) comparing success with grade and success without grade (reference)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Modality</th>
<th>Est.</th>
<th>p-val.</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>female</td>
<td>-0.41</td>
<td>0.47</td>
<td>0.66</td>
</tr>
<tr>
<td>FRENCH-SPEAKING</td>
<td>no</td>
<td>-1.19</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>BELGIAN</td>
<td>no</td>
<td>0.18</td>
<td>0.86</td>
<td>1.19</td>
</tr>
<tr>
<td>FINANCING</td>
<td>work/scholarship</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CULTURAL ACTIVITIES</td>
<td>yes</td>
<td>-1.49</td>
<td>0.05**</td>
<td>0.22</td>
</tr>
<tr>
<td>FATHER STATUS</td>
<td>lower/middle class</td>
<td>-2.63</td>
<td>0.034**</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>teacher</td>
<td>-0.17</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>liberal/self employed</td>
<td>-0.16</td>
<td>0.8</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>unemployed</td>
<td>2.16</td>
<td>0.34</td>
<td>8.67</td>
</tr>
<tr>
<td>MOTHER EDUCATION</td>
<td>highschool or lower</td>
<td>-0.65</td>
<td>0.45</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>higher education</td>
<td>0.66</td>
<td>0.3</td>
<td>1.93</td>
</tr>
<tr>
<td>MATHS</td>
<td>basic</td>
<td>-0.36</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>LATIN/GREEK</td>
<td>yes</td>
<td>0.86</td>
<td>0.13</td>
<td>2.36</td>
</tr>
<tr>
<td>HIGHSCHOOL FAILURE(S)</td>
<td>once or more</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ATTENDANCE</td>
<td>low</td>
<td>-2.24</td>
<td>0.06*</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>-0.62</td>
<td>0.4</td>
<td>0.53</td>
</tr>
<tr>
<td>WORK HOURS</td>
<td>0 to 4h</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>more than 8h</td>
<td>0.55</td>
<td>0.38</td>
<td>1.73</td>
</tr>
<tr>
<td>REGULARITY</td>
<td>high</td>
<td>-1.01</td>
<td>0.11</td>
<td>0.36</td>
</tr>
<tr>
<td>GUIDANCE</td>
<td>high</td>
<td>1.3</td>
<td>0.02**</td>
<td>3.66</td>
</tr>
</tbody>
</table>

4 Conclusion

The results from the empirical analysis are in line with the literature and can also allow us to go one step further in explaining success at university. Schooling and socio-economic status of students are indeed determining factors of success but the variables related to work/study habits are even more important and in the same way they undermine the impact of other variables related to the past the student.

The two most striking elements revealed by the analysis with respect to the working/studying methods are firstly the fact that the regularity of work and the number of hours worked during the exam periods are significant factors of success in the first year at university, and secondly that both the attendance of courses and guidance courses are the keys in successfully completing the year with a grade.

While the student is an actor in his own success, the impact of his past should not be completely forgotten. The student’s past is indeed reduced when the academic behavior
is taken into account but nevertheless does not completely disappear. This means that students do not start university on an equal footing. Moreover, the student’s schooling also plays a role: students who studied languages such as Greek or Latin obtain better marks. The same effect can be observed for students with a strong mathematic background. The impact of the socio-economic background (the diplomas and the status of parents) remains, although the impact is smaller.

It would be interesting to repeat this analysis with several types of populations of university students in order to increase the sample size as well as increase heterogeneity among students’ characteristics. Despite having a small sample, several determining factors of success have been significantly emphasized.

We must also bear in mind that our data come from a questionnaire and there is no guarantee that respondents were perfectly honest. A social desirability bias (Social Desirability Bias) (Nederhof 1985 [25]) could be present in our sample and in particular concerning their work/study habits. It is indeed easier to say a posteriori that the failure of the year was due to a lack of work/study rather than a poor capability. Nevertheless, we believe that the bias remains small. Indeed, concerning the number of hours of work/study, rather than asking students to position themselves on a scale of 1 to 7 according to the effort they have made, the question is asked objectively in terms of hours spent studying. "Hard work" does not mean the same thing for all students, but having studied 1 hour is the same concept to everyone. In addition, the questionnaire is completely anonymous and in its introduction it is clearly explained that in exchange for anonymity it is expected that the student is as honest as possible.

The literature on the determining factors of success is strongly influenced by the study variables the student cannot control: his past, his parents, his background, his capabilities, etc. It encourages a deterministic view of success which can become very dangerous: can we imagine that a university refuses a student because its probability of success is a priori low? We are certainly not there yet but we nevertheless believe that the interest of a deterministic approach is limited—it certainly does not help the student in his course at university. We encourage research to focus on variables which the student can control. In particular, given the importance of work/study habits, our study examines only some of them and only considers the quantity of work/study by the student, not quality. Undoubtedly, there are different ways to study ten hours a day, just as there are different ways to attend classes.

5 Appendix

References

Table 4: Estimation results for the multinomial logistic regression model \((n = 201)\) comparing success without grade (reference) and failure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Modality</th>
<th>Est.</th>
<th>p-val.</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>female</td>
<td>-0.49</td>
<td>0.3</td>
<td>0.61</td>
</tr>
<tr>
<td>FRENCH-SPEAKING</td>
<td>no</td>
<td>1</td>
<td>0.27</td>
<td>2.71</td>
</tr>
<tr>
<td>BELGIAN</td>
<td>no</td>
<td>0.82</td>
<td>0.31</td>
<td>2.27</td>
</tr>
<tr>
<td>FINANCING</td>
<td>work/scholarship</td>
<td>1.13</td>
<td>0.19</td>
<td>3.09</td>
</tr>
<tr>
<td>CULTURAL ACTIVITIES</td>
<td>yes</td>
<td>-0.73</td>
<td>0.17</td>
<td>0.48</td>
</tr>
<tr>
<td>FATHER’S STATUS</td>
<td>lower/middle class teacher</td>
<td>-0.1</td>
<td>0.87</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>liberal/self employed</td>
<td>0.26</td>
<td>0.63</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>unemployed</td>
<td>0.22</td>
<td>0.94</td>
<td>1.24</td>
</tr>
<tr>
<td>MOTHER’S EDUCATION</td>
<td>highschool or lower</td>
<td>0.79</td>
<td>0.18</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>higher education</td>
<td>0.49</td>
<td>0.37</td>
<td>1.63</td>
</tr>
<tr>
<td>MATHS</td>
<td>basic</td>
<td>0.8</td>
<td>0.21</td>
<td>2.22</td>
</tr>
<tr>
<td>LATIN/GREEK</td>
<td>yes</td>
<td>-0.44</td>
<td>0.35</td>
<td>0.64</td>
</tr>
<tr>
<td>HIGHSCHOOL FAILURE(S)</td>
<td>once or more</td>
<td>0.99</td>
<td>0.39</td>
<td>2.69</td>
</tr>
<tr>
<td>ATTENDANCE</td>
<td>low</td>
<td>-0.59</td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>-0.6</td>
<td>0.91</td>
<td>0.54</td>
</tr>
<tr>
<td>WORK HOURS</td>
<td>0 to 4h</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>more than 8h</td>
<td>-1.14</td>
<td>0.02**</td>
<td>0.32</td>
</tr>
<tr>
<td>REGULARITY</td>
<td>high</td>
<td>-1.04</td>
<td>0.04**</td>
<td>0.35</td>
</tr>
<tr>
<td>GUIDANCE</td>
<td>high</td>
<td>0.47</td>
<td>0.35</td>
<td>1.6</td>
</tr>
</tbody>
</table>


