

# **Do early freshmen graduate earlier than late ones?**

## **Enrolment promptness as an indicator of academic success**

Marco Novarese<sup>a</sup>, Paolo Chirico<sup>a</sup>, Viviana Di Giovinazzo<sup>b</sup>

<sup>a</sup> DIGSPES, Alessandria, University of Eastern Piedmont, Italy

<sup>b</sup> Dept. of Sociology and Social Research, University of Milano-Bicocca, Italy.

### **1. Introduction**

Much of the existing literature on students' academic careers builds on studies concerning procrastination and enrolment patterns. Research on the effect of academic procrastination on student performance is abundant (Solomon and Rothblum 1984; Tice and Baumeister 1997; Ariely and Wertenbroch 2002; Schouwenburg, 2004; Howell et al. 2006; Burger et al. 2011; Bisin and Hyndman 2020), but findings are mixed (see, for example, Rotenstein et al. 2009). The literature on student enrolment patterns is also rich. This has been directed towards better understanding enrolment intensity (Stratton et al. 2008) and the relevance of gender and socio-economic factors (Brunello and Winter-Ebmer 2003; Bozick and DeLuca 2005; Rowan-Kenyon 2007). Much of the enrolment pattern literature relates to North American colleges and universities and is focused mainly on the transition between secondary studies and university.

In the Italian setting, little research has been done. Considering an earlier phase of the student's career, Azzolini et al. (2018) examine the effect of asset-building programmes for socio-economically deprived Italian high-school children on university enrolment and performance over the first academic year. In an exploratory analysis, Novarese and Di Giovinazzo (2013) show that date of enrolment can predict some important aspects of a student's career. Students who are quick to enrol have a lower probability of dropping out during their first academic year, and show both higher grades and a greater probability of graduating. On the contrary, those who enrol close to the deadline have a greater probability of withdrawing within the first year of studies.

De Paola and Scoppa (2015) consider the relations between university enrolment patterns and students' grades from the perspective of procrastination. Examining student enrolment patterns, they discovered that progress during the first two years of university career tends to be negatively correlated with delays in initial enrolment. They also found that late enrolment is positively correlated with the decision to drop out. However, in a related study, De Paola and Gioia (2017) find that students' impatience (determined using responses to a questionnaire concerning hypothetical intertemporal choices) is also correlated with the likelihood of their dropping out. The relationship between procrastination and impatience is explored in Reuben et al. (2015).

In this study we consider both promptness and procrastination in enrolment, with a view to understanding which factors determine these two opposed patterns of behaviour. Beyond that, the goal is to use enrolment data in order to develop a model/procedure that allows us to determine the effect of both promptness and procrastination on the risk of failure to graduate in freshmen. In Section 2 the sample under study is described, and a profile of early and late enrollers is made; in Section 3 a predictive model is tested to identify freshmen at risk; the last section concludes.

### **2. Early and late freshmen at the University of Eastern Piedmont**

The sample under study concerns 7,267 freshmen in the Bachelor degree programmes<sup>1</sup> of the

---

<sup>1</sup> More specifically, three-year Bachelor's degrees.

University of Eastern Piedmont (UPO) in the years 2001-2010. The freshmen meet the following requirements: (i) enrolled in open-access degrees; (ii) aged under 25 at the time of enrolment; (iii) not transferring from another degree. Freshmen in restricted-access degree programmes (typically in the health/nursing area) are excluded because these courses have short enrolment periods, which are generally respected. For similar reasons, freshmen transferring from another degree programme are excluded. The over-24s are excluded because they are working students whose career is significantly affected by the time they can devote to study. We focus on first-level (Bachelor) degrees, because enrolment in second-level degrees (Master's) is determined by the date on which the first-level degree is obtained.

In the decade in question, the period for regular enrolment at the University of Eastern Piedmont lasted 9-10 weeks. After this, there was an "overtime" in which enrolment was possible upon payment of an additional fee. Table 1 shows the distribution of enrolments in the regular period and in the overtime period. Since the length of the regular period was not the same every year, this length has been standardized to 1 for each year, so each time interval of width 0.1 corresponds to approximately one week.

**Table 1. Enrolment time of UPO's freshmen in the years 2001-2010**

time interval	frequency	% freq	cumulative
0.0 - 0.1	156	2.2%	2.2%
0.1 - 0.2	131	1.8%	4.0%
0.2 - 0.3	173	2.4%	6.3%
0.3 - 0.4	317	4.4%	10.7%
0.4 - 0.5	509	7.0%	17.7%
0.5 - 0.6	686	9.4%	27.1%
0.6 - 0.7	915	12.6%	39.7%
0.7 - 0.8	1121	15.4%	55.2%
0.8 - 0.9	1438	19.8%	74.9%
0.9 - 1.0	1520	20.9%	95.9%
$\geq 1.0$	301	4.1%	100.0%

Note: each interval of width 0.1 corresponds to approximately one week.

We can see that only a small percentage of freshmen enrol in the first week (2.2%) and around 10% in the first 4 weeks. Then the weekly enrolments increase until the peak is reached in the last week. There may be various reasons for such behaviour: (i) enrolment appears as a "heavy" choice, probably the first really important choice that these students have to face, a choice that frightens, and is perhaps not entirely a mature choice; (ii) the enrolment period generally begins in the middle of summer, when future freshmen are on vacation and very few are eager to enrol; (iii) enrolment involves an initial payment, which students may wish to postpone.

Based on enrolment time, each freshman is classified according to the following four categories: (i) "*early birds*" for enrolment in the first week of the regular period; (ii) "*ordinary*" for enrolment after the first week and before the last regular week; (iii) "*procrastinators*" for enrolment in the last week; (iv) "*latecomers*" if the student enrolls after the regular deadline but before the end of the "overtime" period. Table 2 reports some data on the careers of these categories. The rate of graduation among early birds is almost twice that of latecomers and about 50% higher than that of procrastinators. Moreover, the percentage of dropouts within the first year is significantly lower in early birds (16.4%) than in the other categories: it is almost a third of that of latecomers, who present a dramatic dropout rate: 48.5%! Why are early birds more successful than procrastinators and latecomers? Are they more motivated or do they have superior skills?

If we look at the high-school education of the freshmen (Table 3), we see that the early birds mainly come from classical or scientific Italian high schools, which are considered the most suitable for continuing studies at university; they have obtained a final high school grade of between 80 and 100 in 53.3% of the cases. On the other hand, procrastinators and latecomers have lower

percentages of classical or scientific high school students and higher percentages in the low-average grade classes. It is reasonable to deduce that: the share of "good" students is higher among early birds than among procrastinators and latecomers: good students generally decide in advance which degree to enrol in and do not like wasting time. The deduction is confirmed by the box-plots in Figure 1.

**Table 2. Enrolment time and graduation of UPO’s freshmen in the years 2001-2010.**

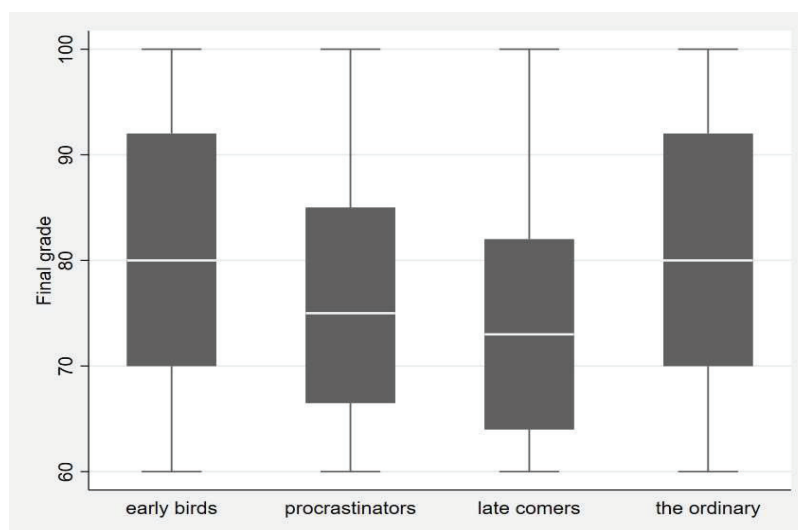
Carrier status	early birds	the ordinary	procrastinators	latecomers	All
not graduating	24.2%	35.0%	50.3%	60.5%	39.6%
- dropout1	16.4%	25.9%	40.1%	48.5%	25.9%
graduating	75.8%	65.0%	49.7%	39.5%	60.4%
- regular	60.6%	50.3%	35.2%	29.6%	45.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

Note: dropout1=dropout within the first course year”; regular=” graduate within the regular duration of the course”.

**Table 3. Enrolment time and high school of origin (with final grade).**

High School	early birds	the ordinary	procrastinators	latecomers	all
Classical-Scientific	43.0%	34.4%	32.1%	26.2%	33.7%
Technical	33.3%	38.2%	33.1%	36.5%	36.8%
Professional	5.5%	7.2%	10.3%	8.6%	8.0%
Other high school	18.2%	20.2%	24.4%	28.5%	21.5%
<i>Final h.s. grade:</i>					
60-69	20.6%	22.6%	31.9%	37.9%	25.5%
70-79	26.1%	25.5%	27.9%	26.2%	26.1%
80-89	24.2%	21.9%	19.7%	20.6%	21.4%
90-100	29.1%	30.1%	20.5%	15.3%	27.0%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

**Figure 1. Box-plot of the high-school final grade in the enrollment categories.**



However, the observed differences between freshman categories do not seem to fully explain why early birds, on average, enjoy a better student career. A logistic regression was performed to better understand the relationship between enrolment time and career success (Table 4).

With the exception of the final grade of the high-school diploma, the other explanatory variable

of the model is a dummy variable<sup>2</sup>. In particular, the variable *non-isee* detects who does not present the ISEE certificate at the enrolment, which allows the student to obtain a reduction in university fees. Some freshmen do not present the certificate because their financial state exceeds the limit for having the fee reduction, but others do not present it due to carelessness or misinformation. The variables *economics*, *law*, *sciences*, *humanities* indicate the fields of the degree programmes and are useful for controlling for the difficulty of the corresponding degrees relative to a benchmark degree in political science (which is excluded from the model in order to avoid the collinearity problem).

*Ceteris paribus* (high-school education, degree difficulty, etc.), early birds are more likely to graduate than “the normal” students: the odds ratio between the early birds and the ordinary is 1.65. This result reflects what the early birds have in comparison to “ordinary” students: stronger motivation and enthusiasm. Conversely, procrastinators and latecomers, as well as non-isee freshmen, are less likely to graduate because they generally have less motivation and conviction.

**Table 4.** Logistic regression results; dependent binary variable

	B	SE(b)	z	p	exp(b)
Constant	-1.781	0.247	-7.203	0.000	0.168
early birds	0.502	0.204	2.459	0.014	1.653
procrastinators	-0.545	0.063	-8.614	0.000	0.580
latecomers	-0.793	0.139	-5.715	0.000	0.452
non-isee	-1.367	0.083	-16.480	0.000	0.255
high school final grade	0.069	0.002	27.833	0.000	1.071
classical-scientific high school	1.158	0.063	18.347	0.000	3.183
field degree: - economics	-0.615	0.106	-5.811	0.000	0.540
- law	-1.052	0.131	-8.008	0.000	0.349
- sciences	-0.970	0.112	-8.690	0.000	0.379
- humanities	-0.629	0.114	-5.533	0.000	0.533
male gender	-0.120	0.059	-2.035	0.042	0.887

Note: Dependent variable is the career status (1=graduating; 0=not graduating).

### 3. Random forest classification for “failure” risk

The enrolment data allow us to have an idea of what the student's career will be like: we can use this data to identify students at “failure risk”, i.e. at a risk of not graduating, and monitor their career. For this goal there are several statistical tools known in the literature: *logit* and *probit* regression (Stratton et al., 2008), discriminant analysis and classification trees (Rai et al., 2014). In what follows, we briefly illustrate an application of a “machine learning” algorithm which in recent years has been the object of growing interest and widespread use: the random forest classification (RFC, see Breiman, 2001). An RFC algorithm classifies each new unit by taking the most frequently occurring classification provided by an ensemble of “reduced” classification trees (Breiman et al., 1984). Each reduced tree is generated using a random subset of units and explanatory variables. This procedure generally assures lower out-of-sample classification errors than the ordinary classification trees<sup>3</sup>.

The RFC algorithm was “trained” using a random sample of 5,100 freshmen extracted from the set of 7,267 freshmen under study<sup>4</sup>; The same variables were used as in the logistic regression described in the previous section. Then the algorithm was used to classify the 2,167 “out-of-sample” freshmen (Table 5).

The correct classification rate of student that fail to get the degree, i.e. “not graduating”, is

<sup>2</sup> These variables are equal to one when the student satisfies the condition of the name variable, zero otherwise.

<sup>3</sup> For more details, see Breiman (2001) and chapter 15 in Hastie et al. (2009).

<sup>4</sup> FRC was performed using the Stata module RFOREST (Schonlau and Zou, 2023) with 300 iterations, 3 explicative variables and unlimited deep.

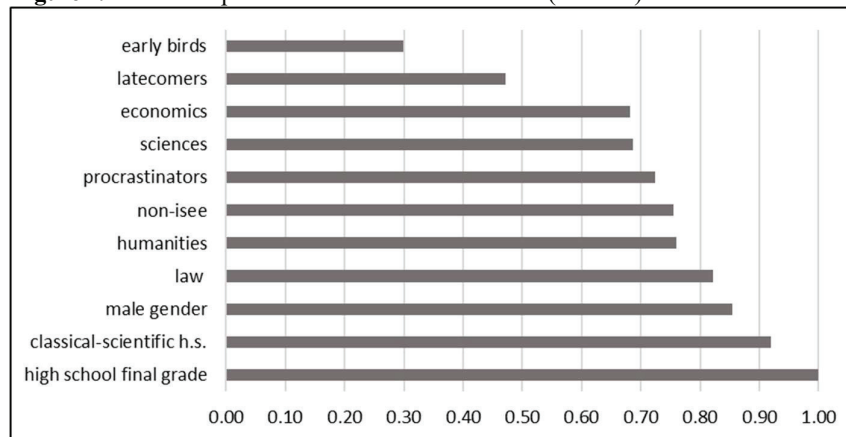
56.8%, which can be considered acceptable at the outset. The RFC classification procedure can be periodically repeated by adding the progressive data of the student's career, such as the number of exams and credits achieved after each exam session. This form of monitoring would certainly provide more and more accurate classifications of the cases at risk.

**Table 5. Prediction results of a “random forest” classification (OOB error = 31.6%).**

Career status Actual:	Predicted:		Total	% of correct predicted
	not graduating	graduating		
not graduating	497	378	875	56.8%
graduating	280	1,012	1,292	78.3%
Total	777	1,390	2,167	69.6%

Finally, Figure 2 reports the relative importance of the covariates (i.e. the importance of each covariate compared to the most important one) in classifying the freshmen. In this case, “early birds” is the least important covariate. This result does not mean that being early birds (i.e. being more motivated and enthusiastic) is not relevant for obtaining a degree, but that this variable is not particularly useful for classifying freshmen as graduates and non-graduates: it could not be because the early birds are only 2.2% of freshmen. Same explanation for "latecomers", who make up only 4.3% of freshmen.

**Figure 2. Relative importance of covariates in the RFC (max = 1).**



#### 4. Conclusion

Using data on student registration behaviour and their academic career, this paper showed that freshmen who register early also graduate early because they have better high-school education (i.e. attend a better high-school and achieve a higher final grade), motivation to study and enthusiasm.

No doubt, more research is required in order to strengthen our results, for example, by comparing our data with those on students registered at other universities. However, we believe this study is an important first step for devising aids to prevent dropouts and support students along their academic career.

**Acknowledgements** This work was supported by funding from the University of Eastern Piedmont.

#### References

Ariely, D., Wertenbroch, K. (2002). Procrastination, deadlines and performance: self-control by precommitment. *Psychological Science*, **13**, pp. 219–224.  
 Azzolini, D., Martini, A., Romano, B., Vergolini, L. (2018). Affording college with the help of asset

- building: first experimental impacts from Italy. *Economic Letters*, **169**: pp. 27–30.
- Bisin, A. and Hyndman, K. (2020). Present-bias, procrastination and deadlines in a field experiment. *Games and Economic Behavior*, **119**, pp. 339–357.
- Bozick, R., De Luca, S. (2005). Better late than never? Delayed enrollment in the high school to college transition. *Social Forces*, **84**(1), pp. 531–554.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth, Belmont, (CA).
- Breiman, L. (2001). Random forests. *Machine Learning*, **45**, pp. 5–32.
- Brunello, G., Winter-Ebmer, R. (2003). Why do students expect to stay longer in college? Evidence from Europe. *Economics Letters*, **80**(2), pp. 247–253.
- Burger, N., Charness, G., Lynham, J. (2011). Field and online experiments on self-control. *Journal of Economic Behavior & Organization*, **77**, pp. 393–404.
- De Paola, M., Scoppa, V. (2015). Procrastination, academic success and the effectiveness of a remedial program. *Journal of Economic Behavior & Organization*, **115**, pp. 217–236.
- De Paola, M., Gioia, F., (2017). Impatience and academic performance. Less effort and less ambitious goals. *Journal of Policy Modeling*, **39**, pp. 443–460.
- Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.
- Howell, A.J., Watson, D.C., Powell, R.A., Buro, K., (2006). Academic procrastination: The pattern and correlates of behavioural postponement. *Personality and Individual Differences*, **40**, pp. 1519–1530.
- Novarese, M., Di Giovinazzo, V. (2013). *Promptness and Academic Performance*. [https://mpra.ub.uni-muenchen.de/49746/1/MPPRA\\_paper\\_49746.pdf](https://mpra.ub.uni-muenchen.de/49746/1/MPPRA_paper_49746.pdf).
- Rai, S., Saini, P., Jain, A. K. (2014). Model for prediction of dropout student using ID3 decision tree algorithm. *International Journal of Advanced Research in Computer Science & Technology*, **2**(1), pp. 142–149.
- Reuben E., Sapienza, P., Zingales, L. (2015). Procrastination and impatience. *Journal of Behavioral and Experimental Economics*, **28**, pp. 63–76.
- Rotenstein, A., Davis, H.Z., Tatum, L. (2009). Early birds versus just-in-timers: the effect of procrastination on academic performance of accounting students. *Journal of Accounting Education*, **27**(4), pp. 223–232.
- Rowan-Kenyon, H.T. (2007). Predictors of delayed college enrollment and the impact of socioeconomic status. *The Journal of Higher Education*, **78**(2), pp. 188–214.
- Schonlau, M., Zou, R.Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, **20**(1), pp. 3–29.
- Schouwenburg, H.C. (2004). Trait procrastination in academic settings: An overview of students who engage in task delays, in *Counselling the Procrastinator in Academic Settings*, eds H.C. Schouwenburg, C. Lay, T. Pylchyl, and J. Ferrari, American Psychological Association, Washington (US), pp. 3–13.
- Solomon, L.J., Rothblum, E.D. (1984). Academic procrastination: frequency and cognitive-behavioral correlates. *Journal of Counseling Psychology*, **31**(4), pp. 503–509.
- Stratton, L.S., O’Toole, D.M., Wetzel, J.N. (2008). A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*, **27**(3), pp. 319–331.
- Tice, D. M., Baumeister, R. F., (1997). Longitudinal study of procrastination, performance, stress, and health: the costs and benefits of dawdling. *Psychological Science*, **8**: 454–458.