Thematic Article

The interaction of meso and macro contexts with students' careers. Three applied analysis

Alessandra Decataldo[1], Andrea Amico[2], Giampiero D'Alessandro[3]

Abstract: This paper presents three applied analysis. Our principal interest is to understand how meso (university level) and macro (national level) contexts interact with micro phenomena such as students' careers. To achieve this aim, the paper shows a secondary analysis of longitudinal administrative data by means of Sequence Analysis and Event History Models. Furthermore, it shows the results of an additional analysis using a quasi-experimental research on longitudinal data.

The paper examines typical pathways of Italian students, by means of administrative information's on student enrollment from Sapienza University of Rome. The aim of this paper is triple: 1. describe students' careers in higher education by building a typology of pathways using Sequence Analysis. 2. Identify how socio-economic and macro level characteristics affect students' careers by multinomial logistic regression model where clusters are used as dependent variables and focus on the study of the "event" by using Event History analysis. 3. Evaluate the Italian Higher Education reform policies and their main outcomes throughout a quasi-experimental research.

Mainly, the outcomes show the importance of the warming-up period, the individual choice and the exogenous events after enrolment in determining the success/failure of each career. Our outcomes suggest that Italian universities should rethink the mechanisms available to manage failure and guide student choices.

Keywords: Administrative Databases; Longitudinal Analysis; Determinants of Students' Careers; Reform Evaluation

1. Introduction

University student attrition and retention represent two aspects of dropping out and persistence phenomena (Berger et al., 2012). These aspects are not complementary because one is more oriented to individual factors and the other on contextual ones.

In the course of the last twenty years, under pressure from Bologna process (http://ec.europa.eu/education/policy/higher-education/bologna-process_en) radical reforms have been introduced in Italy. Their aims are to control the phenomena of dropping out, perpetual students and low number of graduates. From AY 2001-02, the MD 509/1999 introduced the shift from a university education based on a single level degree to the three-cycle system (bachelor/master/doctorate).

We identified Sapienza University of Rome as a representative case because it is by far the largest university in Italy (http://statistica.miur.it/ustat/Statistiche/IU_home.asp). We constructed longitudinal datasets of information extracted from the administrative archives of this university.

---

[1] University of Milano Bicocca; alessandra.decataldo@unimib.it

[2] Censis; amico@censis.it

[3] Freelance researcher; giampiero.dalessandro@gmail.com

This paper shows the analysis of these longitudinal datasets in order to study how meso and macro level characteristics can affect students' careers. The longitudinal data structure gives us the possibility of an insight into the various events, such as re-orientation, stopping out, dropping out and graduation that might affect the duration of a student's career. The context changes could also be taken into account. To study the phenomena and to pay attention to the contextual variations occurred during the careers, we use the Sequence (SA) and Event History analysis (EH).

Furthermore, this paper presents the results of an additional analysis aiming at the evaluation of the MD 509/1999 and its main outcomes with the use of a quasi-experimental research on longitudinal data.

The paper consists of the following four main sections: the first one provides the research aims; the data and the methods, while the following section illustrates the analysis of the data. The paper ends with a discussion and conclusion section.

2. Research design and Methods

Our principal interest is to understand how meso (university level) and macro (national level) contexts interact with micro phenomena such as students' careers (Pike & Graunke, 2015). In order to achieve this aim, the paper shows a secondary analysis of longitudinal administrative data by means of SA and EH. These methods start from different assumptions. The first method rejects the concept of statistical causality that is the basis of the second one.

We use SA in order to 1. describe the phenomena of late graduation and late performance; 2. identify kinds of students who are dropping out from university; 3. evaluate other phenomena that could delay careers (such as mobility within and between faculties).

The hypothesis is that timing is the most relevant aspect, which determines the career outcome. For instance, a change of study course during the first part of a career could be more advantageous (it could be a useful career reorientation) than the one performed in the central part.

In SA we used only the information related to the students (micro level) and the university context (meso level), stressing out the relevance of temporal dynamics (Herzog, 2005). Part of our interest was also the inclusion of the external context (macro level) and its causal relation to career duration. For this purpose, we believed that EH could be useful for the integration of our analysis.

This EH model is focused on the context factors that could have impact on graduation or dropping out. This study supports the idea that some external conditions (e.g. low receptivity of the labour market and low rates of overall economic growth) could affect the decision of a student to use university as a "parking place" before attempting to enter the labour market (Häkkinen & Uusitalo, 2003).

The additional quasi-experimental research[4] (Shadish, Cook & Campbell, 2002) allows us to attribute mainly to the introduction of the MD 509/1999 (which we consider to be the experimental variable) the differences between the career outcomes of the post-2001 cohorts (the post-test population) and the ones of the pre-2001 cohorts (the pre-test population). In particular, the adopted research design represents a combination of the Time-Series and the Separate-Sample designs.

In the three analysis, for all the cohorts there are: 1. no students with a precedent university career; 2. no students with credits or exams performed at other universities or through a working activity; 3. only postgraduate students come from other universities (they enrolled for the first time in an MA at Sapienza).

Starting from Sapienza database, we created two longitudinal dataset (one for SA and EH, one for quasi-experimental research) that includes a data vector for each student, re-constructing his/her entire career, including the temporary stopping outs and mobility. These dataset are two "case by variables matrixes", in which every row corresponds to a student and every column to a variable about his/her university career at Sapienza University. The data referring to the socio-demographic and high school training variables are cross-sectional (registered at students' first enrolment),

---

[4] It is a quasi-experimental research because it does not use group randomization.

whereas those referring to university career and performance indicators are longitudinal (repeated for each semester of students' university careers).

For SA and EH, we monitored the cohorts of students until March 2014 (the end of AY 2012-13). Therefore, monitoring period for each cohort shortens every year.

With reference to SA, since the AY 2001-02 represents the turning point from the old to the BA/MA university system, the analysis focuses only on the students enrolled at Sapienza in this year (23,854). Each student is monitored up to 25 semesters. The used data-encoding format is the STS (State Sequence) (Ritschard et al., 2009).

With regard to EH, we extend this study to the students enrolled in BA courses from AY 2001-02 to AY 2010-11 (200,554). The career duration was determined from the first enrolment to the career conclusion: graduation, formal or informal dropping out[5]. The students that until March 2014 did not graduate or dropout were treated as right-censored (Box-Steffensmeier & Bradford, 2004).

For the additional analysis, information concerns the cohorts of the students enrolled (413,336) before (from AY 1991-92 to AY 2000-01) and after (from AY 2001-02 to AY 2007-08) the MD 509/1999. We monitored the pre-2001 cohorts until March 2008 and the post-2001 cohorts until March 2014. Therefore, the first cohort was observed for a period of sixteen years, the second one for fifteen years, etc. In order to guarantee a comparison between the pre-2001 and the post-2001 cohorts, we excluded students who have experienced the transition from a pre-reform course to a BA.

For SA the R package TraMineR (Gabadinho et al., 2011) was used both for the analysis and for the output graphs. We consider ten different administrative "states": 1.-4. four types of enrolment (due to the course length: 2, 3, 5 or 6-year course); 5. continuation (after the previous semester); 6. graduation; 7. formal dropping out; 8. informal dropping out; 9. change of faculty; 10. change of course (within the same faculty).

We used the "trate" method, a second generation of Optimal Matching measures, to measure the distance between careers (Studer & Ritschard, 2016). Therefore, a distance matrix between each pair of careers was calculated by OM algorithm (MacIndoe & Abbott, 2004; Studer & Ritschard, 2016). We subjected this matrix to two clustering algorithms (Wald and PAM methods, Studer 2013). The clusters are used as dependent variables in a multinomial logistic regression model (ML) for characterising them on the basis of students' information, the first year performance indicators and the university context variables.

Also for the EH analysis were used R packages: 1. for performing semi and fully parametric survival analysis *survival* (Therneau & Grambsch, 2000) and *eha* (Broström, 2012); 2) for elaborating the competitive cumulative hazard functions graph *cmprsk* (Gray, 2014).

The events that were analysed are graduation and dropping out. The probability that one of these events occurs in a certain moment of time is the core of this analysis (Box-Steffensmeier & Bradford, 2004). The pattern in the Schoenfeld residuals and the global significance (0.000) of the null hypothesis test show that the proportionality condition is not present in this case. Consequently, the log-logistic model is the best fully parametric model fitting with the data.

In the log-logistic parametric survival model, the administrative data of the careers (micro) were integrated with context indicators related to the meso and macro contextual time-varying covariates. This permits to consider the macro-economic and labour market contexts within which the students' careers take place. Both factors vary with the passage of time, but the implementation in the model as Time Varying Covariates (TVC) is different. In order to take account of the labour market circumstances, we assigned the relative[6] unemployment rate to each student. This factor was treated canonically as a TVC, by incorporating it into the model as three covariates, one per year for the first three years of enrolment. On the contrary, the macro-economic covariate (real GDP) was implemented into the model as a single variable, indicating the change in GDP occurred from the first enrolment year to graduation or dropping out.

---

[5] In Italy, there are two types of dropping out. The first one is formal and it implies that the student demands for an official study interruption. The second one is informal because the student interrupts the fee payment, but he/she does not forward any official communication to the university.

[6] According to sex, age group and area of residence.

Together, these two factors made it possible to control the effect of the general economic context over the length of the students' university careers. The period under analysis is perfect to control the impact of the economic crisis started in Italy in 2008.

Lastly, through the adoption of a quasi-experimental research design (Shadish, Cook & Campbell, 2002), we try to attribute to the transition from the old to the BA/MA university system any differences between the academic career outcomes of the post-2001 cohorts and those of the pre-2001 cohorts[7].

All the students are monitored for a period that is a year longer than the expected length of their study course (i.e. the BA students are monitored for 4 years), because in Italy the majority of graduates complete their studies later than expected.

3. Results

Figure 1 represents the index-plot of the 2001-02 cohort obtained by SA. The vast majority of the cohort enrolled in BA courses (92.8%); 5.1% in 5-year courses; 2.1% in 6-year courses. There is a preponderance of red colour, which indicates formal and informal dropping out. This outcome concerns particularly the BA students' careers. Although the majority of dropouts are registered in the transition between the first and second year of the study course, many students experience dropping out later (even after many years).

Through cluster analysis performed on the OM distance matrix six well-defined groups of students were identified. The groups differ in characteristics and dimensions (Figure 2). The first group consists of 12.3% of students. The majority (35.5%) is composed of students that are still enrolled at the end of the observation period; other part of the group are dropouts (informal, 32.3%, and formal 8.3%) or graduates after 10 years since the enrolment (22.2%).

The second (12%) and the fourth groups (20.9%) differ for the dropping out timing: the students belonging to the second group keep postponing the dropping out decision until the 9th semester, whereas for the fourth group the dropping out occurs in the first half of the observation period.

The second type of dropping out (the formal one) characterises the sixth group (15%), that drops out at the beginning of their career.

The other two groups (the third and the fifth) are composed of the students that end their career with graduation. Although the students of the third group (18.7%) graduate at the regular course length, the timing of graduation in the fifth group (21.1%) is slower.

These six groups are used as dependent variables in an ML (Table 1), that includes two types of independent variables: micro-level variables (social characteristics of students, high school information and first year performance indicators) and meso-level variables (university context information). The main result shows the importance of the warming-up period in determining the development of each career, as highlighted in other countries (Arias Ortiz & Dehon, 2013). As shown by Herzog (2005), the first year performance indicators are the best predictors of success. Furthermore, the university context information are important: in particular, students enrolled in a 5 or 6-year course stand more chances to graduate than the BA students do. Outcomes confirm that students coming from lower socioeconomic background are more vulnerable to dropout along the academic path (Arias Ortiz & Dehon, 2013).

---

[7] We analysed the factors that could jeopardize both the internal and the external validity of the quasi-experimental research.
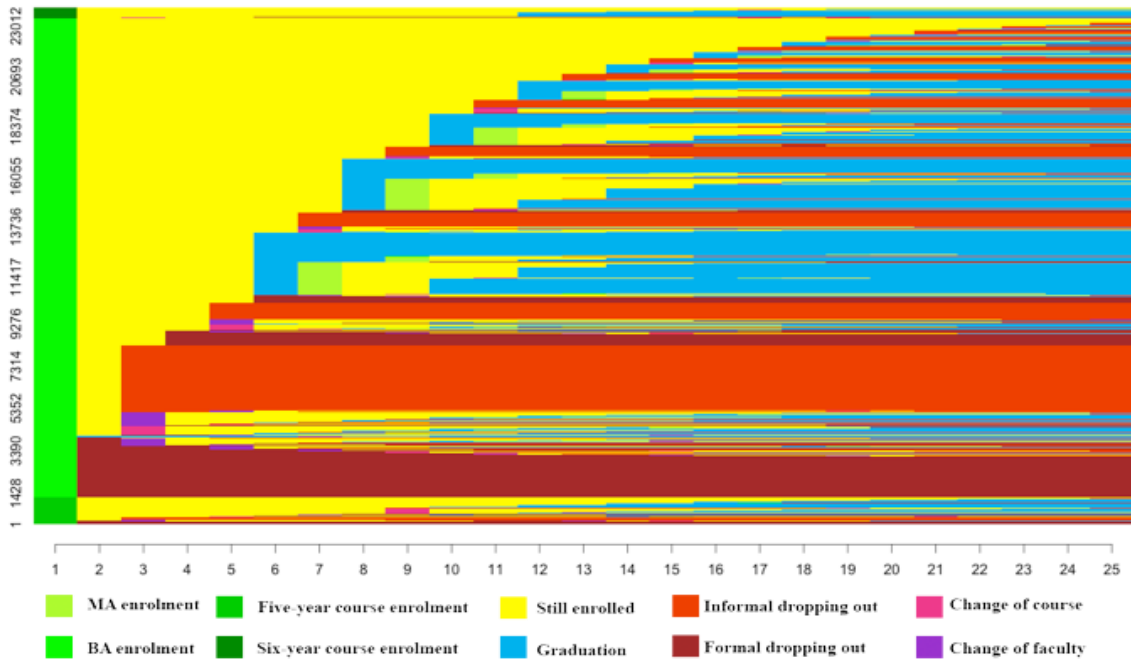
5

Central European Journal of Educational Research 1(1) 2019. 10 – 21.
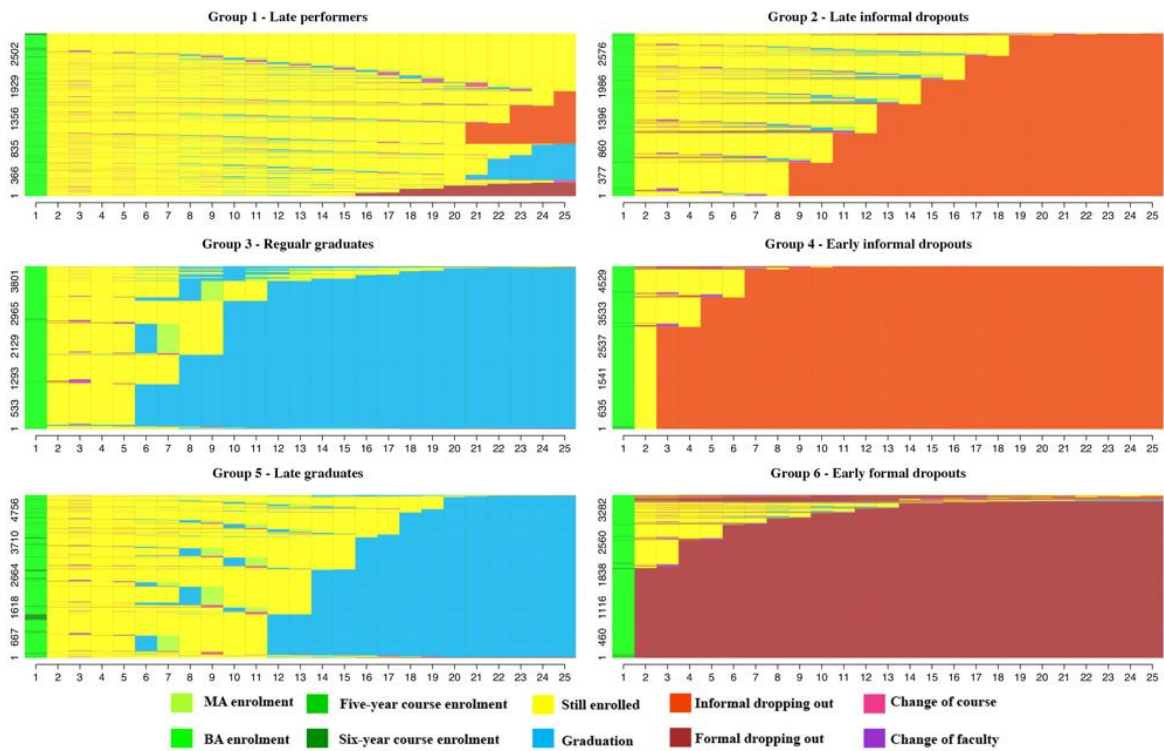
Figure 1. Index-plot of the 2001-02 cohort

**Figure 2**. Index-plot for the six groups

Table 1. Odds ratio of ML for 2001-02 cohort

Covariates		Early informal dropout OR	Late performer OR	Regular graduate OR	Late graduate OR	Early formal dropout OR	Late informal dropout OR
	Constant	0.673	0.048	0.014	0.051	0.497	0.077
Sex	Female*						
	Male	1.245		0.823			1.097^
Age at enrolment	Less than 19*						
	20 to 24	1.814	0.826		0.672	0.751	1.403
	Over 25	2.681	0.735	1.424	0.352	0.596	
High school type	Classic and Scientific*						
	Technical-Professional	1.499	0.845	1.224	0.773	0.909^	
	Other	1.500		1.357	0.732		
High school final mark	Low*						
	Medium	0.889		0.895^	1.182	1.183	0.870
	High	0.713	0.852	1.157	1.158	1.335	0.655
Residence	Local*						
	Regional	0.756		1.376			
	National	0.685				1.627	0.771
Family income	0€*						
	1-10,000€		1.305		0.864		1.195^
	10,001-20,000€		1.337			0.852^	
	20,001-30,000€		1.313				
	30,001-40,000€		1.189^				
	40,001€ or more	0.836^					
Length of study course	3 years*						
	5 years	0.758	1.643	0.160	2.620	0.711	
	6 years	0.295	1.854	0.013	5.081	0.571	0.306
Limited enrolment course	No*						
	Yes	0.788		1.239	0.836	1.255	
First year credits	0*						
	1-15 *Ects*	0.705^	3.184			0.570	2.588
	16-30 *Ects*	0.234	3.920	5.292	4.238	0.206	2.949
	31-45 *Ects*	0.067	2.426	15.837	6.204	0.083	1.769^
	46 *Ects* or more	0.025		36.346	4.836	0.028	
First year average mark	0*						
	Low	0.648			2.244^		
	Medium	0.582			2.447^		
	High	0.586					

* (reference); all values are significant at ≤0.05, except for the one marked with ^ (significant at ≤0.1)

Throughout SA we used the information related to micro and meso levels; EH could be useful to include the external context (macro level) and its causal relation to the duration of the careers (Johnson & Muse, 2012).

The cumulative hazard of the graduation and the dropping out events in Figure 3 shows the high dropout rates during the first year of the course. The "steps" in the dropout function are result of a high number of dropout students at the end of each AY. The dropout relative cumulative hazard is much higher in the first months, but it becomes slightly lower at around the 70th month. In a longer period, the probability that the students graduate is very similar to the probability of dropping out.
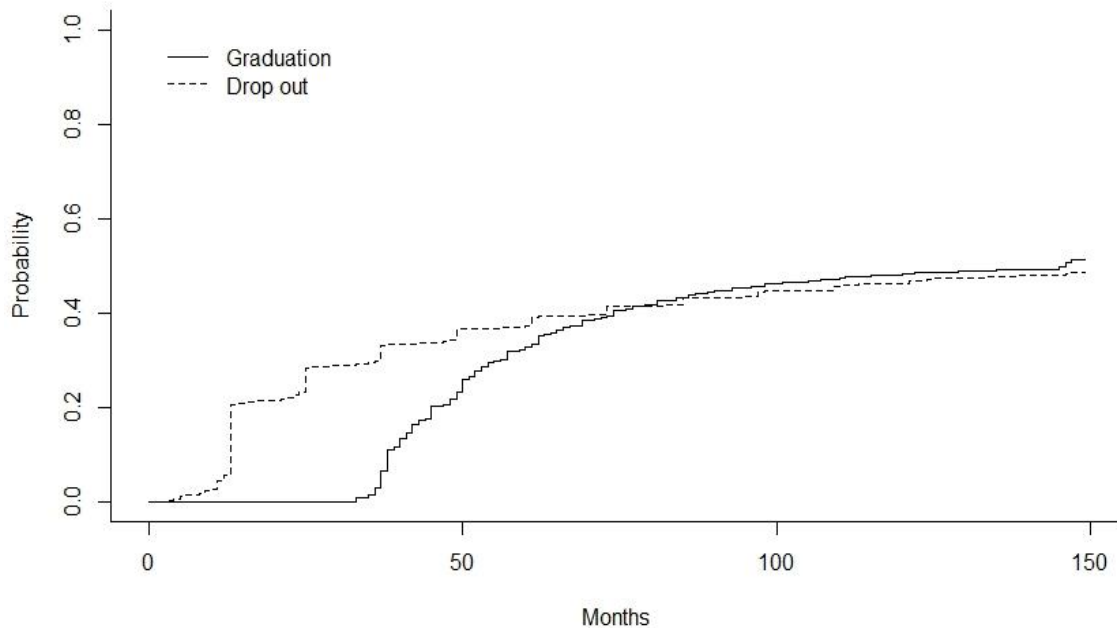
Figure 3. Competitive cumulative hazard functions

The Table 2 shows the final graduation and dropout log-logistic models that embed the macro TVC on the micro and meso covariates.

In the graduation model, the students with an age from 20 to 24 years old at their enrolment have a lower risk of reaching the event (graduation) in a short period (49% lower than the younger ones). Students with higher high school final mark and with scientific high school type have more probability to obtain a degree within a lower amount of time. Students with family residence outside Lazio (the Sapienza region) have 29% higher probability to have a short and successful university career: they live away from home and seem to have more motivation to conclude quickly their careers.

Low and middle family income affect duration of the careers in the same direction (by accelerating them) (such as in Arias Ortiz & Dehon, 2013; Pike & Graunke, 2015).

Meso (structural) covariates are mainly referred to the faculty of registration. Remarkable differences between learning areas are attributable to the enrolment selection of healthcare courses and to the different presence of the students with high school final marks. For almost all of these covariates in the dropout model, the coefficients have an opposite direction compared with the graduation one. It is coherent with the differential impact of these characteristics over the students' careers. For example, if an enrolled student who is over 25 years old has a lower probability to reach the graduation in a shorter period in comparison with the younger ones, on the contrary he/she has a higher probability of dropping out earlier.

A stationary[8] career is a faster and more successful one.

The variable type of course is used as a proxy to divide the cohorts in three periods. The students of the firsts 7 cohorts are enrolled in the "509" courses whereas in the last 3 cohorts there are students almost exclusively enrolled in "270" courses. In the first 7 cohorts 8.3% of the students choose to change their course. Students of the first cohorts has a higher risk of successful careers.

The contextual external covariates confirm that there is a lower risk of obtaining a degree in a shorter period with high unemployment rates during the first two years of the career and there is a higher risk when the GDP is rising. Consequently, if a career takes place during an economic crisis

---

[8] It is referred to the comparison between the fields of studies of the last registration and the one of the first enrolled year.

(negative variation of GDP) it is much slower. Even the dropout decision is postponed in a difficult macro-economic and labour market context.

A similar analysis conducted by Häkkinen and Uusitalo (2003) has shown that a policy of reducing financial aid has failed to speeding-up the university careers due to an increase in the unemployment rate. The problematic labour market situation caused longer students careers.

This model confirms our hypothesis that some external conditions (low receptivity of the labour market and lower overall economic growth rates) do affect students' decision to use the university as a "parking place" before attempting to approach the labour market.

Table 2. Parametric log-logistic models with TVC

Covariates		Graduation event model Exp(Coef)	Dropout event model Exp(Coef)
Sex	Female*	1	1
	Male	1.015^	1.200
Age at enrolment	Less than 19*	1	1
	20 to 24	0.703	1.546
	Over 25	0.907	2.225
School final mark		1.026	0.979
School type	Scientific*	1	1
	Classic	0.981	1.014^
	Technical	0.781	1.491
	Professional	0.795	1.795
	Others	0.860	1.332
Residence	Local*	1	1
	Regional	1.025	0.860
	National	1.239	0.963
Family income	Low*	1	1
	Medium	1.057	0.918
	High	0.861	0.678
Learning area	Scientific*	1	1
	Health care	2.433	0.485
	Humanities	1.148	1.044
	Social	0.975	1.232
Stationary	No*	1	1
	Yes	2.132	2.027
	509*	1	1
Type of course	270	0.783	0.889
	From 509 to 270	0.715	0.693
Unemployment rates 1 year		0.943	0.970
Unemployment rates 2 year		0.897	0.946
Unemployment rates 3 year		1.189	1.086
Var. GDP		1.159	1.113
Events		79,208	79,262
Total failure time		673,716	673,716
Max log-likelihood		-178,020	-223,465
LR statistic test		59,312.57	50,239.69
D.f.		22	22
Sig. total		0	0

* (reference); all values are significant at ≤0.05, except for the one marked with ^ (not significant)

In the additional analysis using a quasi-experimental research, if we consider the difference between the outcomes of the first post-2001 cohort and the last pre-2001 cohort (Figure 4), there can be noticed that Sapienza performance improves by increasing the number of graduates from 19.1% to 26%. The ability of attrition improves even more significantly with a drop of over 10 percentage points in the proportion of dropping out. However, already starting from the 1998-99 cohort the proportion of graduates had been increasing and that of dropping out decreasing.

By tracking the trend line with the best interpolation for the points that represent the percentage of graduates in the last 5 pre-2001 cohorts (the linear regression represents the best goodness of fit, r^2 0.91, Figure 5), we may notice – by projecting trends in the post-2001 cohorts – that Sapienza performance after the Reform is almost perfectly in line with the previous trend until at least the 2004-05 cohort. Later the MD 509/1999 seems to have lost its thrust because the proportion of graduates stops increasing and it settles at values close to 30%.
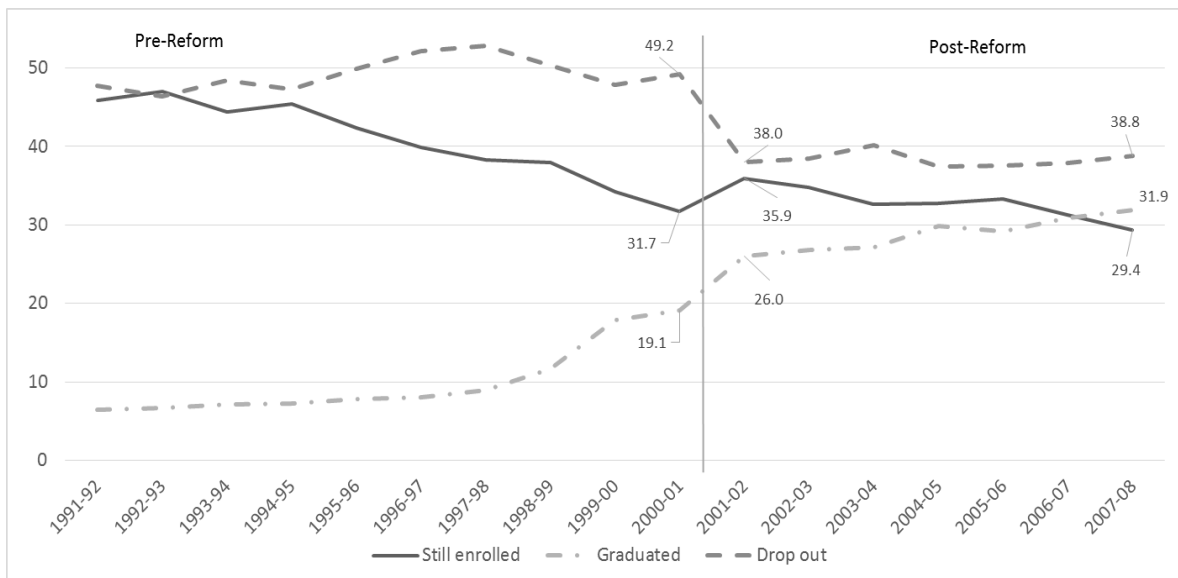
Graph 4. Overlook on students' administrative positions a year later than the length of their study course, %
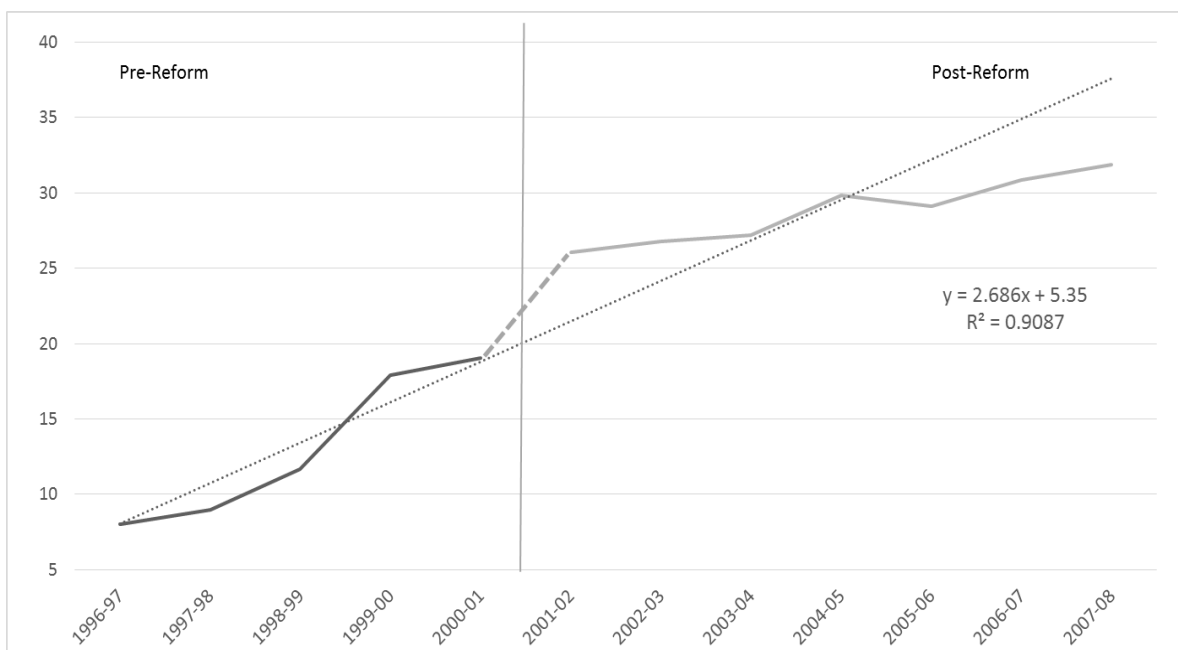
Figure 5. The trend line with the best goodness of fit with the pre-reform data

We implemented two models of multinomial logistic regression (Table 3), evaluating the probability of a persistent enrolled student to become a graduate rather than a perpetual student (still enrolled) or a dropout. Mainly, students with early enrolment compared to those with late enrolment has as a starting point a lower relative probability of being still enrolled respect to graduate in the pre-reform cohorts, and they reach a greater relative probability in the post-reform cohorts. These changes are signs that some of the Reform aims have been achieved.

Table 3. Multinomial logistic regression for the pre-2001 and the post-2001 cohorts

		Pre reform model	Post reform model			Pre reform model	Post reform model
	Sex				Sex		
	Male	0.994*	0.958		Male	1.144	1.253
	Female[b]				Female[b]		
	Age at enrolment				Age at enrolment		
	Early enrolled	0.654	1.063		Early enrolled	0.206	0.544
	Late enrolled[b].				Late enrolled.		
	Residence				Residence		
	Local	1.347	1.262		Local	1.221	1.471
	Regional	1.536	1.312		Regional	1.382	1.187
	National[b]				National[b]		
	High school type				High school type		
Still enrolled[a]	Classical and scientific	0.412	0.694	Dropout[a]	Classical and scientific	0.180	0.396
	Techincal-professional[b]				Techincal-professiona[b]		
	High school final mark				High school final mark		
	Low	4.071	3.037		Low	8.829	4.750
	Medium	2.024	1.941		Medium	2.904	2.292
	High[b]				High[b]		
	Learning area				Learning area		
	Scientific	1.585	1.031**		Scientific	1.644	0.743
	Healthcare	0.576	0.139		Healtcare	0.510	0.117
	Humanistic	1.687	0.782		Humanistic	1.640	0.793
	Social[b]				Social[b]		

Reference category: Graduated
*NOT significant
** Significant at .10 level

4. Discussion and conclusions

This paper shows the usefulness of a longitudinal dataset of students' careers to investigate how meso and macro level characteristics can affect students' careers.

The SA outcomes show the importance of the warming-up period in determining the success/failure of each career, in particular the first year performance indicators are the best predictors of success. Furthermore, the university context information, especially the length of study course, are important: students enrolled in a 5 or 6-year course stand more chances to graduate than the BA students.

Through the EH models, we show that later completion or early dropping out can be due both to the individual choice and to the exogenous events after enrolment. The changes occurred within the labour market and the macro-economic context seem to affect the duration of the careers, postponing the decision of dropping out and graduation.

The quasi-experimental design shows that the aims of the MD 509/1999 had already been reached for the last cohort of the pre-2001 and they had exhausted their potentiality at least for the 2004-05 cohort.

All these results can be immediately used to implement the improving policies. For example, once the warming up period has been identified as the most important meso level factor in a student's career, Sapienza (or another University) could implement a more effective tutoring activities during the first year of a course. Another example: in difficult macro-economic context the connection between students and labour market should be enhanced with curricular internship, even in the early stages of the careers.

Furthermore, our outcomes suggest that Italian universities should rethink the mechanisms available to manage failure and guide student choices.

Finally, our outcomes suggest the necessity that Italian main agencies of higher education evaluation use longitudinal rather than cross-sectional data in order to evaluate the reform achievements and to establish the Italian university standards.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arias Ortiz, E. & Dehon, C. (2013). Roads to Success in the Belgian French Community's Higher Education System: Predictors of Dropout and Degree Completion at the Université Libre de Bruxelles. *Research in Higher Education*, 54(6), 693-723.
2. Berger, J.B., Ramirez, B., & Lyon, S.C. (2012). Past to present: A historical look at retention. In A. Seidman (Ed.), *College student retention: Formula for student success*. Lanham, MD: Rowman & Littlefield.
3. Box-Steffensmeier, J.M., & Bradford, S.J. (2004). *Event History Modelling: A Guide for social Scientists*. Cambridge, UK: Cambridge University press.
4. Broström, G. (2012). *Event History Analysis with R*. Boca Raton, FL: CRC Press.
5. Gabadinho, A., Ritschard, G., Müller, N.S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37.
6. Gray, B. (2014), *Subdistribution Analysis of Competing Risks. Package 'cmprsk'*. Retrieved from: http://CRAN.R-project.org/package=cmprsk
7. Häkkinen, I., & Uusitalo, R. (2003). *The Effect of a Student Aid Reform on Graduation: A Duration Analysis*. Working Paper 2003:8. Department of Economics. Uppsala, Sweden: Uppsala University. Retrieved from: http://www.diva-portal.org/smash/get/diva2:129253/FULLTEXT01.pdf
8. Herzog, S. (2005). Measuring Determinants of Student Return VS. Dropout/Stopout VS. Transfer: A First-to-Second Year Analysis of New Freshmen. *Research in Higher Education*, 46(8), 883-928.
9. Johnson, I.Y. & Muse, W.B. (2012). Student Swirl at a Single Institution: The Role of Timing and Student Characteristics. *Research in Higher Education*, 53, 152-181.
10. MacIndoe, H., & Abbott, A. (2004). Sequence Analysis and Optimal Matching Techniques for Social Science Data. In M. Hardy, & A. Bryman (Eds.). *The Handbook of Data Analysis*. London–Thousand Oaks–New Delhi–Singapore: Sage.

11. Pike, G.R. & Graunke, S.S. (2015). Examining the Effects of Institutional and Cohort Characteristics on Retention Rates. *Research in Higher Education*, 56(2), 146-165.

12. Ritschard, G., Gabadinho, A., Studer, M., & Müller, N.S. (2009). Converting between various sequence representations. In Z.W. Ras, & A. Dardzinska (Eds.), *Advances in Data Management, Studies in Computational Intelligence* (155-175). Berlin-Heidelberg, Germany: Springer-Verlag.

13. Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin Company.

14. Studer, M. & Ritschard, G. (2016). What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures, *Journal of the Royal Statistical Society*, 179(2), 481-511.

15. Studer, M. (2013). WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social science. LIVES Working Papers, 24, 1-32.

16. Therneau T., & Grambsch, P. (2000). *Modeling Survival Data: Extending the Cox Model*. New York, NY: Springer-Verlag.