Impact of Assessment Characteristics on Course Withdrawal: a Survival Analysis Approach

Juan Antonio Martínez-Carrascal [1] and Teresa Sancho-Vinuesa[1]

[1] *Universitat Oberta de Catalunya, Rambla Poblenou 156, Barcelona, 08018, Spain*

Abstract

Assessment characteristics are known to play a critical role in shaping student behavior and academic performance. However, and despite withdrawal is a key indicator of this performance, there is a lack of studies addressing the link between course-level withdrawal and course-specific assessment parameters.

This article explores the application of survival analysis as a statistical method to address this issue. Using an open dataset from a leading online university we analyze the impact of assessment characteristics on course dropout. Besides presenting the method, the results obtained show that reducing the time between assessments impacts on withdrawal decisions, leading to lower dropout rates. The effect of including computer-marked assessments is unclear according to the dataset used. The study also reveals that the effects may differ across different student groups and courses, encouraging the analysis on a per-course basis.

The study illustrates the potential of survival analysis as a tool for analyzing complex educational datasets and generating insights that can help evidence-based decision-making. The findings that emerge from its application have significant implications for designing effective assessment strategies that can help to reduce course dropout rates. By understanding the impact of assessment characteristics, educators can design assessments that promote engagement, motivation, and finally, academic success.

Keywords

Withdrawal, dropout risk, assessment characteristics, survival analysis

1. Introduction

Assessment activities are a nuclear part of the learning process, and their proper design contributes to improve that process [1]. The design of adequate course assessment models is an aspect where institutions invest efforts to increase in academic achievement. Assessment and academic achievement are closely linked as researched in a set of works [1]–[3]. Some assessment design parameters such as frequency of evaluations, level of difficulty of the exercises, and variety or types of assessment activities may impact academic performance.

At course level, the withdrawal ratio constitutes one of the most relevant indicators of academic performance. Withdrawal – understood as the decision to abandon an ongoing course – is one of the most challenging problems that the educational system faces at all levels. Its multifaceted structure [4] makes its reduction particularly challenging. It also results in a waste of resources such as time and money [5], which underscores the necessity to address this issue.

However, and to the best of our knowledge, no correlation has been yet quantified between withdrawal at course level and specific assessment characteristics linked to that course. The current literature points out a connection between assessment design and student engagement. However there is no research addressed to quantify the impact of assessment characteristics on course withdrawal as we do.

Learning Analytics Summer Institute Spain (LASI Spain) 2023, June 29–30, 2023, Madrid, Spain
EMAIL: jmartinezcarra@uoc.edu (A. 1); tsancho@uoc.edu (A. 2)
ORCID: 0000-0002-7696-6050 (A. 1); 0000-0002-0642-2912 (A. 2)

Our research focuses on establishing a methodological approach to determine the relevance of assessment parameters on the student's course withdrawal. Besides providing the method, we perform an application based on an open dataset provided by the Open University (OU) [6] which contains information from about 30,000 students enrolled in seven unique online higher education courses, summing up 22 course editions.

The dataset includes withdrawal information for those students who leave specific courses before reaching the end. Some of the courses implement different assessment models for each course edition, which allows to set up comparative analysis. In particular, the use of different time lengths between consecutive assessments and the inclusion of computer-marked assessments techniques (CMAs) - that automatically grade students without the need of instructor's intervention - are key components of the evaluation model. We analyze how tuning the assessment characteristics do have an impact upon the course withdrawal rates.

Our findings indicate the potential of the method. Based on the dataset used, the results indicate that longer periods without assessments generate a raise in course withdrawal rates. In addition, according to the same dataset, we cannot drive clear conclusions regarding the use of CMA and its impact on dropout. The survival analysis method efficiently evaluates the effects of specific assessment characteristics on withdrawal for the courses under observation, thereby contributing to the improvement of the student's retention by identifying those characteristics that perform the best in terms of engagement. It is important to notice that these results allow for additional analysis based on the reasons underlying the modifications to assessment strategies. However, this would require complete educational data, which is not included in the dataset. For this reason, we focus in the method itself, rather than delving deeper into specific findings, which is not the focus of this study.

The remainder of the paper is organized as follows. First, a theoretical framework introduces the analysis of the impact of assessment aspects on withdrawal, including a set of research questions. Afterwards we describe the study's methodological approach and the Open University dataset as the framework to conduct the analysis. Then, we present the results and discussion of the methodology including the main findings and their implications. We conclude the work with some remarks and future work.

2. Theoretical framework

2.1. Assessment characteristics and academic performance

Assessment can be defined as 'the systematic basis for making inferences about the learning and development of students'[7]. A proper design of an assessment model should contribute to improve students' learning and development. The role played by the assessment along the learning process has a relevant impact that justifies the need for more research on that direction. A compilation of assessment methods and modes is detailed in the work authored by Pereira et al. [8].

The effect of assessment in academic performance was already present in the early studies on the topic [2] and confirmed in recent studies, either considering their direct impact or more moderate influences [9], [10]. Even some authors consider assessment having a greater weight on the learning process than other regular teaching activities [11]. For example, [3] outlined the impact of assessment in the learning outcomes. The same authors, in later research [1] confirm their initial findings with a stronger statement 'there is more leverage to improve teaching through assessment than through anything else'. The authors also identify the best assessment strategies to effectively improve student's learning.

Other authors [12] point out that to optimize the potential of the assessment requires a thoughtful exercise. An assessment strategy may include the specification of a clear objective for each assessment item, the context in which the assessment takes place, and a detailed view of the tasks that comprise the assessment in order to build and then demonstrate the acquisition of learning. A suitable strategy does have a major impact on academic performance and its deployment can provide a set of teaching and learning indicators to follow up the effectiveness and overall quality [13].

Other authors like [14] evidence differences in patterns of assessment and mark distribution among different disciplines at course level. Those differences do not necessarily scale up to the degree level, where the discipline itself plays a greater influence. [15] focus on the specific role played by the assessment's characteristics that has a direct impact on the student's learning. [9] compile previous research and conclude that assessment's characteristics were combined to impact on student's grades. In that research, the analisys deals with the use of intermediate assessments, both mandatory and voluntary, the potential benefits of peer evaluation, and the effects of having rewards.

Richardson [16] analyzes the differences between continuous and final exam-based evaluation. Marks tend to be higher in environments when using a combination of coursework and examinations, rather that when using single examinations alone. Continuous evaluation also provides higher attainment [17]. As an additional difference, the type of assessment performed does not influence in continuous evaluation environments, while it does in scenarios where a single and punctual assessment is done [18].

The level of difficulty perceived by the student as a mirror of the assessment exigency is also found relevant [19]. Easy-to-pass assessments, sometimes based on predictable questions, derive in a lower dedication from the students which is also linked to less demanding learning outcomes [15]. From the student perspective, the same analysis indicates that learners welcome assessments built on their skill set. Assessments that demand more creativity were preferred instead. The perception of student workload is also a critical issue where the research shows that students prefer a more balanced effort along the whole course duration.

A more quantitative research has been carried on recently, such as [10] that reveals the impact of frequency, difficulty, and diversity of assessment on overall performance. While difficulty and diversity aspects have to do with the type of assessment, frequency is connected to time**.** By 'timeliness' the study refers to the importance of the timing in the publication of the assessments' guidance and the timing for the assessments' delivery and deadlines [15]. Most higher institutions implement evaluation models which include different assessments. A proper design of a continuous assessment strategy demands more dedication and effort from teachers and practitioners, however it provides benefits in terms of learning outcomes [20].

Gibbs et al. [1] show the negative effect between a poorly conceived assessment strategy in terms of assignments and its frequency. Two of the authors performed further research [21] stating that infrequent examinations tend to concentrate learning time and a more frequent assignment schedule is liked to more continuous study time and a better performance.

More recently, Vandist et al. [22] founded improvements in academic performance in environments with a higher volume of assessments. The authors indicate that this improvement is linked to the higher frequency of feedback the student receives. Day et al. [9] compile previous results and extend a similar finding, concluding that the proper number and frequency of assessments may depend on the course itself and also should take into account the whole duration of the course. However, the improvement in overall academic performance does not necessarily mean a reduction of dropout ratios [23] .

2.2. Connection between assessment and withdrawal

The impact in performance abovementioned is related to the influence on final course marks. However, the number of studies drops when considering the link between assessment and course withdrawal.

Course withdrawal is defined as ceasing studying a course without the intention to resume the study of that course later [24]. The concept is usually included within the broader term of dropout, which in higher education refers to 'situations where students leave the university study in which they have enrolled before getting a formal degree' [25].

Specific surveys analyze dropout in higher education [26] and specifically in online environments [27]. Those studies reflect the relevance of dropout both at degree and course level, and also the higher impact in online institutions. It is also noticeable that early works already indicated that course design and characteristics are present in 28% of withdrawal cases [28]. Assessment characteristics would be among those aspects, but are not explicitly analyzed.

The correlation between assessment aspects and dropout would be based on the mediating effect of assessment characteristics in student engagement reflected in [10], and the demonstrated effect of engagement in dropout [29]. However, the connection between engagement, in broader terms, and assessment aspects would still be unclear. Wang et al. [10] report a higher engagement when increasing frequency, difficulty, and diversity of assessments. However, Naude et al. [30] indicate that higher workload is not necessarily linked to higher engagement.

The current literature shows us that 'the evidence for what factors within assessments actually contribute to student engagement is not fully understood and more research is required' [15]. This is even more critical when focusing the impact of course dropout.

2.3. Research questions

Based on the theoretical framework described in the previous section, we propose an innovative method to determine how and to what extent specific assessment characteristics impact on course withdrawal decisions in higher education. To the best of our knowledge, no previous research has quantified the effect of a variation in specific assessment characteristics across various editions of the same course on its dropout rates.

The approach will be based on the use of survival analysis methodology, which is also used in other learning analytics scenarios [31]–[34]. Survival analysis has been tested to analyze withdrawal at course level in MOOC (Massive Open Online Course) environments [35], [36] or other specific programs [34]. More recent studies have explored the impact of demographics on withdrawal at course level through this technique [37]. However, none of those studies consider the specific impact of assessment characteristics and aspects.

In our analysis, we hypothesize that leaving the students for a long period of time within the course with no assessments forces disengagement, and this disengagement has an impact on course withdrawal. We formulate our hypothesis in a research question:

RQ1: How does the maximum period of time between assessments impact withdrawal decisions of students at a course level?

The approach will be translated into practice through an open dataset provided by The Open University [6]. In order to explore other aspects that might impact, the dataset also includes information regarding the use of computer marked assessments (CMA) through the course. Those CMA were automatically placed by the computer, without direct human intervention. This fact allows us to formulate a second research questions:

RQ2: How does the use of CMA, as part of the evaluation process, impact on the withdrawal decision?

Both the period of time between assessments and the use of CMA were explored using survival analysis, and our aim is to extend that method to any other assessment characteristics that may impact on withdrawal.

3. Materials and methods

The research methodology is based on the application of survival analysis tools to investigate the impact of specific assessment criteria on withdrawal to address both RQs. The second question demands a more practical approach where we use a specific dataset with a group of courses and their assessment features. We include both analysis in separated sections, and final section to merge both results by demonstrating how data is effectively modelled to fulfil the needs of that research and is prepared for further analysis.

3.1. Using survival analysis to study the impact of assessment differences on course dropout

Different research works [26], [35] use different techniques to identify students at risk of dropping out. Still a very limited number of works, encourages the use of longitudinal data techniques [31]. Other authors [38] indicate their ability to outperform traditional prediction algorithms using longitudinal data methods.

Among those techniques, survival analysis is particularly beneficial when addressing dropout. As indicated earlier, most studies at course level focus on MOOC scenarios [35], [36], [39], but quite recent research used survival analysis to identify at-risk populations and quantify the impact of differences among student's populations in regular higher education online courses [37]. Here, we extend this approach to analyze the impact based on differences in assessment characteristics among course editions.

The principles of survival analysis are outlined in [40]–[43]. The method is particularly suited for situations when the relevant component is the remaining time until a specified event occurs. The methodology comes from the research in medicine when examining the time to relapse or death as well as the significance and impact of various causes or treatments in the same period of time.

The described situations were directly transferable to learning analytics scenarios considering three key factors: a precise definition of the event of analysis (in this case, dropout), the observation period (which in this case will last from the time of enrolment to the end of the course), and survival as a whole. In the study of dropout, survival as a whole refers to the period of time during which the student remains engaged in academic activities without formally resigning. Once the mapping is complete, we can evaluate the effects of various assessment tactics.

Kaplan-Meier estimations will reflect the impact of this potential effect [40]. The estimations provide a graphical representation of the likelihood of surviving after a specific period, enabling comparison of the effects derived from a particular change. They also provide a clear view of whether life expectancy – time to dropout in our case – varies depending on the assessment method or implementation. In addition to the graphical representation, the analytical formulation of the curve can be used to ascertain the influence of a particular element on the survival probability or how tuning this factor may affect the total survival probability.

3.2. Dataset

The research conducted is based on an open dataset provided by the Open Univesity (OU) [6]. This dataset contains information from 22 editions, also called 'presentations' using OU jargon, of 7 distinct courses. Every course includes at least two editions. The total enrollment is of 32,593 students. Each course is referred as 'module' according to OU naming. The average length for each of these modules is around nine months. A remarkable aspect of the dataset is the inclusion of the dates of withdrawal from the course for each student.

All courses in the dataset are delivered through a virtual learning environment (VLE) and have more than 500 students enrolled. Courses have also been chosen based on the proportion of students that do not graduate the course, either by failing or dropping out. Table 1 provides a summary of academic performance in these courses in terms of pass, failure, and withdrawal.

Table 1.
Global view of academic achievement in the OU dataset

Concept	Enrolled	Fail	Pass	Withdraw
# Students	32.593	7,052	15,385	10,156
% of total	100%	21.64%	47.20%	31.16%

As Table 1 shows, 31,16% of the students enrolled withdraw before the end of the course which constitutes our target of analysis. All courses share a common framework for evaluation, which includes a set of tutor marked assessments (TMAs) and, optionally, some computer-marked assessments (CMA). Besides, there is commonly a final exam at the end of each course. Relevant variables related to the assessment characteristics are summarized in Table 2.

Table 2.
Meaning of the variables in the OU dataset linked to assessment characteristics

Variable	Meaning
assessment_type	TMA, CMA, or Final Exam.
date	information about the cut-off day of the assessment
weight	weight of the assessment as part of a summative assessment evaluation

In addition to the information in Tables 1 and 2 above, we have other course parameters such as presentation length, the enrolment date for each student and, as indicated before, the date when a student eventually withdraws. This data is particularly essential to map the problem in a survival analysis scenario.

3.3. Applying survival analysis to a specific learning dataset

Once the concepts of lifespan (time to dropout), event under consideration (dropout as such) and the period of observation (course duration) were clear, we focus on the impact on this event derived from specific assessment characteristics. Data pre-processing allows us to identify students that dropout before the course begins. Those students have been removed from our analysis because this decision is not based on assessment-related issues. As stated, our primary objective is to identify variations in the presentations that influence the decision to withdraw. Those early withdrawals may be associated to other factors that are out of the scope of our analysis.

The withdrawal decision is recorded in the *final_result* field in the student record under the tag of 'Withdrawn.' Regarding the observation period, we considered that the course starts in time equal to 0. Prior to that time, dates (especially enrolment dates) will be reported in days and deemed negative. In terms of lifespan, our analysis will be valid as long as the presentation is active. Students with life expectancies shorter than the course duration were more likely to dropout.

Despite our emphasis on two specific parameters – time and the potential inclusion of CMAs – we also searched for assessment-related parameters that could contribute to statistically significant variations in withdrawal rates. Taking into account the dataset, and specifically the data presented in Table 2, we consider the following parameters as part of our analysis:

- Maximum time between consecutive assessments
- Type of assessments –TMA or CMA- performed
- Number of assessments included
- Date of the first assessment
- Length of the course presentation

Making a parallelism with the medical research, here the idea is to segregate populations that have been treated with different methods – i.e. different assessment characteristics in our case – and benchmark the dropout expectancy. Due to the absence of A/B testing over the same presentation of a course in the provided dataset, we will search for courses with different assessment parameters in separate presentations. It may not be possible to establish the relevance of a given factor, as – for example – the number of tests conducted in all presentations of a given course could be the same for all courses, making it impossible to segregate populations based on this criterion. When feasible, we will establish distinct populations based on the analyzed courses and examine the evolution of withdrawal over time.

As it may be noticed, the survival approach could be done to analyze the impact of other parameters that are present in the database but not linked to assessments in the withdrawal decision. However, we focus on this subset to address our RQs.

4. Results

4.1. Population segregation based on differences in assessment characteristics

In a first set of results, we segregate populations that have been enrolled in presentations with a sundry of assessment parameters to evaluate the impact of the differences on dropout rates. This testing scenario clusters the assessment strategies followed by different groups. Table 3 summarizes the assessment parameters for the different course presentations in the dataset that may impact on dropout.

Table 3.

Presentation parameters linked to assessment characteristics for the different courses

Course	Presentation	Num. TMA	Num. CMA	Max Between TMA (days)	First TMA (date)	Duration
AAA	2013J	5	0	63	19.0	268
	2014J	5	0	63	19.0	269
BBB	2013B	6	5	42	19.0	240
	2013J	6	5	49	19.0	268
	2014B	6	5	42	12.0	234
	2014J	5	0	56	19.0	262
CCC	2014B	4	4	70	32.0	241
	2014J	4	4	77	32.0	269
DDD	2013B	6	7	49	25.0	240
	2013J	6	0	42	25.0	261
	2014B	6	0	42	25.0	241
	2014J	6	0	49	20.0	262
EEE	2013J	4	0	56	33.0	268
	2014B	4	0	49	33.0	241
	2014J	4	0	63	33.0	269
FFF	2013B	5	7	42	19.0	240
	2013J	5	7	49	19.0	268
	2014B	5	7	42	24.0	241
	2014J	5	7	63	24.0	269
GGG	2013J	3	6	63	61.0	261
	2014B	3	6	56	61.0	241
	2014J	3	6	63	61.0	269

Table 3 shows that some courses will not be suitable for the comparison of assessment strategies. For instance, no segregation can be performed for course AAA, as both presentations display the same values. However, the BBB course, allows us to segment the population due to the differences in CMA. Here, we focus on the behavior of populations that allows us to answer our RQs by considering the period of time between consecutive assessments and the potential use of CMA.

Table 4 displays the courses where the maximum time between consecutive assessments differs between presentations. For instance, presentation 2014B of course EEE has a maximum time without assessment of 49 days, while that time is of 63 days in 2014J. As minor differences may emerge, we include only courses with differences over two weeks.

Table 4.
Courses containing presentations with relevant differences in maximum time between assessments

Course	Presentations with a lower period of time between assessments (#days)	Presentations with a larger period of time between assessments (#days)
EEE	2014B (49)	2014J (63)
FFF	2013B,2014B (42)	2014J (63)

Regarding the course length, because all courses are annual (between 234 and 269 days) there were no differences to infer any additional result. The same applies to the time of the first test, as the differences are below a week.

Regarding the use of CMAs, courses BBB and DDD show the largest differences as stated in Table 5.

Table 5.
Courses containing presentations with differences in CMA

Course	Presentations with no CMA	Presentations including CMA (# CMA)
BBB	2014J	2013B, 2013J, 2014B (5)
DDD	2013J, 2014B, 2014J	2013B (7)

The information from Tables 4 and 5 allows us to segregate groups of students enrolled in the same course with different assessment characteristics. We will go deeper in the following sections to analyze its impact on dropout.

4.2. Maximum time between assessments: increasing time leads to higher dropout for some courses

Figure 1 illustrates the Kaplan-Meier curves indicating the survival function based on the differences in maximum time between assessments found in various presentations for courses EEE and FFF (as reflected in Table 4). The x-axis represents time, with t=0 signifying the course start. The y-axis indicates the percentage of students still enrolled in the course at each time point.

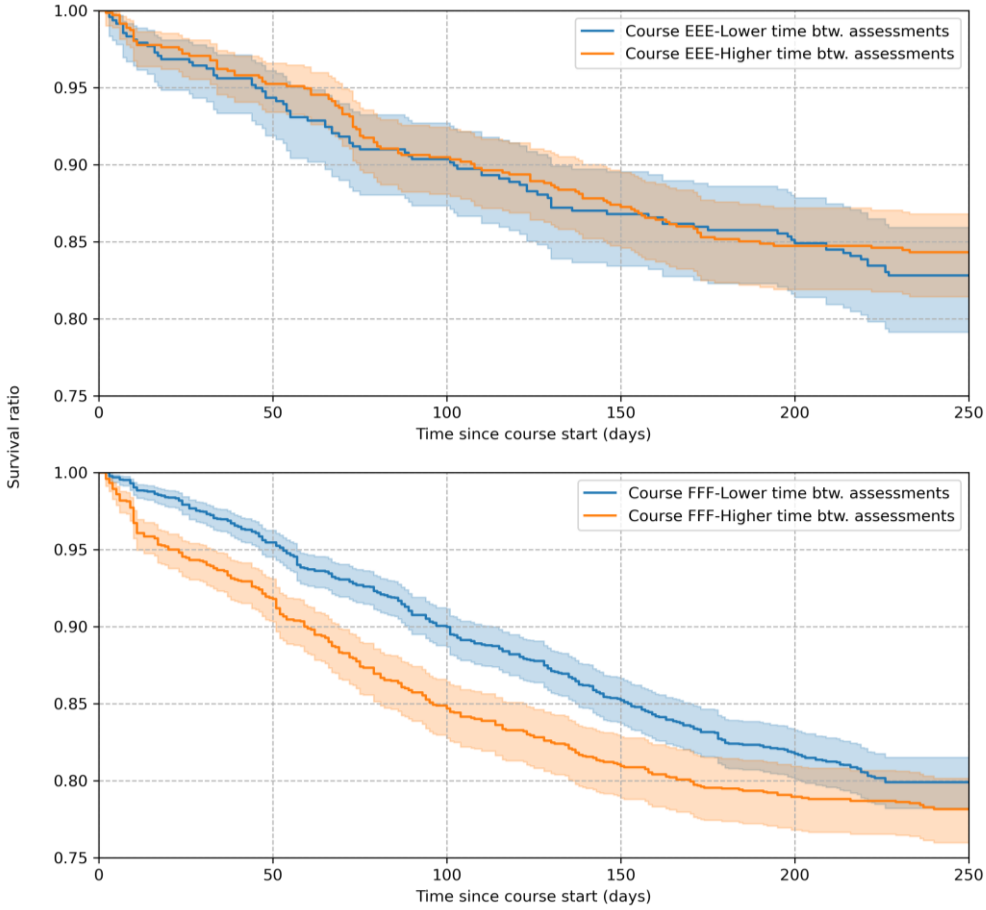

Figure 1. Survival evolution based on the largest time between assessments: courses EEE (above) and FFF (below)

Higher time intervals between assessments do not show statistically significant for the course EEE, while they do for course EEE ($p<0.005$). For this course, longer periods between assessments are associated to higher withdrawal rates. This finding will be examined in greater detail in the discussion section.

4.3. Impact of CMAs: no concluding results can be extracted

To compare the use of CMA in several courses, we made a population using the data from Table 5. In this survival analysis, shown in Figure 2, we included courses BBB and DDD as they correspond to different subjects with different global dropout characteristics.

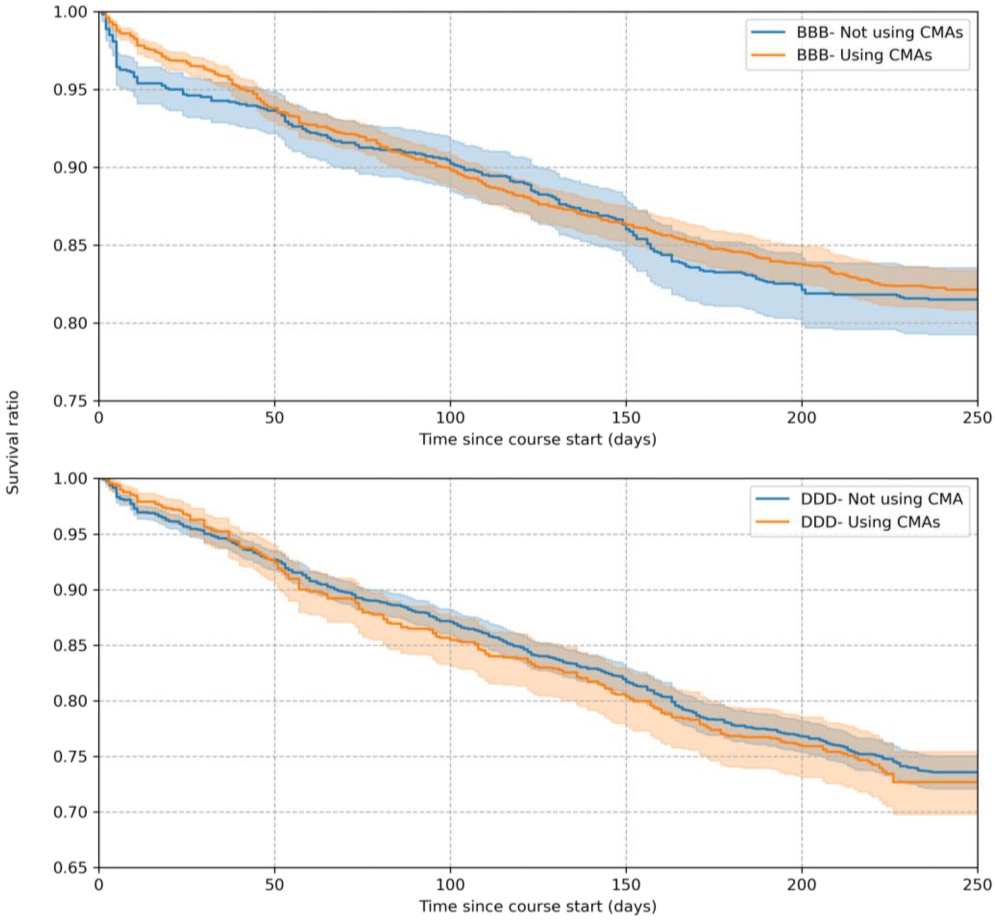

Figure 2. Survival evolution based on the use of CMAs: course BBB (above) and DDD (below)

Figure 2 depicts a clear difference in the behavior of the students for BBB course, however this difference does not show on DDD course. Statistical tests to support the null hypothesis (both groups having the same risk hazard curve) indicate $p<0.005$ for BBB course, making the difference statistically significant. As expected from the graph, differences were not significant for course DDD course with $p=0.95$

5. Discussion

In the results section, we have identified some assessment characteristics with a direct impact on the withdrawal decisions of students enrolled in higher education online courses. More precisely, an increase in the time between consecutive assessments enlarges the withdrawal probability on certain courses. The survival analysis does not reveal differences based on the inclusion of automatic assessment methods, such as CMAs.

Generally speaking, we reach two main conclusions from the research work carried on in this paper. First of all, some assessment strategies can provide better results than others when addressing dropout rates. Second, survival analysis provides an innovative method for a deeper analysis of dropout in general, and withdrawal at course level in particular.

While the main purpose of the study was to make a methodological contribution, we want to highlight the importance of having a thorough understanding of the pedagogical details underlying the courses under investigation. While utilizing an open dataset offers undeniable benefits, we lack information regarding the course's inner workings. Survival analysis indicates that raising the maximum time between assessments for course FFF has an effect on withdrawal that is absent for course EEE.

We believe that other course characteristics, such as difficulty, subject of study, etc., could explain this phenomenon.

Unfortunately, we lack the internal knowledge necessary to expand this analysis. This information would allow to analyze why this parameter impacts some courses but not others. It is important to note that we are transferring techniques. Survival analysis, when applied in medical science, can identify treatments that are effective only for specific populations. Similarly, when applied to learning analytics, it can reveal factors that are advantageous for some courses but irrelevant for others.

Our findings extend the impact on academic performance revealed in [10], showing that certain aspects explicitly impact on course dropout rates. In particular, frequency of assessments has shown an incidence in lower dropout rates. We cannot draw clear conclusions, based on the data used in our case study, regarding the potential impact of combining different types of assessment. While the use of CMAs has reduced dropout in one of the courses, it has not been so clear on another. This might suggest that the impact might depend on other variables related to the course.

Regarding the period of time between assessments, and considering in particular course FFF, larger times generate higher dropout probabilities. These findings are consistent with the results in [44] and also aligned with [45]. The first reference points out that successful students in online programs show higher time and study management capabilities. A reduction of the maximum time between tests forces students to keep more focused, adding additional control, and limiting potential disengagement due to longer periods without being assessed. This reduction would help students with lesser time management capabilities who may be adversely affected by long periods without assessment. The overall effect is a reduction in the number of withdrawal decisions. The second reference, with a different perspective, indicates that setting up constraints in the learning process – and so, reducing student freedom – would provide better academic results in online courses. As stated before, the use of CMAs do not raise any clear conclusion.

The impact of frequency of assessments is also aligned with the causes suggested in [17]. A higher frequency would increase attainment, which at the same time is related to dropout, so increasing attainment would also mean a reduction in dropout.

Further than the impact of specific factors, we would also like to focus on the methodology used. The method exposed been very effective to uncover the impact of factors related to dropout and opens an interesting approach to explore its connection with assessment parameters. Kaplan-Meier curves provide a graphical view of differences between groups of students and allows a much clear interpretation of the results. At the same time, tests can be performed to determine whether the results are statistically significant.

The survival analysis methodology could be used to extend some results such as those suggested in [9]. This study outlines the relevance of the appropriate number of assessments rather than just an increase in frequency. The lack of datasets to benchmark strategies based on differences in the frequency of assessments, the method provided helps to validate these particular results as well as other hypothesis linked to the evaluation of their impact. Specific datasets should be required for further results on that direction.

The results achieved need to be contextualized considering the theoretical framework for dropout. Internal factors are considered as one of the four components explaining dropout in the *composite persistence model* outlined in [46]. Pedagogical factors such as learning and teaching styles, are also included in this group of internal ones. In the particular case of online education, we infer that fitting course design into student's expectations increases academic integration and consequently reduces dropout. This was already reflected in [47] for traditional studies, while [26] indicates that part of the institution teaching quality is being able to make an impact on dropout.

To conclude, the statement from Gibbs et al. [1] "there is more leverage to improve teaching through assessment than through anything else" provides an observation about the power of assessment strategies when particularly addressing the reduction of dropout, and we advocate the analysis used with optimal assessment parameters to achieve reasonable levels of abandonment.

6. Final remarks and future lines of research

Our study has revealed the impact that the pace of assessment tasks may have on dropout at course level. Considering the importance of dropout reduction in higher education, and more precisely in online education, we encourage the use of the survival analysis methodology to extend this work to explore other aspects further than the ones covered in this research. To name a few, the overall number of assessment tests, their typology, level of difficulty, or the inclusion of gaming strategies may be part of the evaluation.

The use of an open dataset is clearly beneficial to perform an unbiased and reproducible study and allows to perform the methodological approach. However, the lack of deeper information on the dataset - including more details of the courses' pedagogy and the reasons behind the modification of assessment parameters - borders a deeper interpretation of results. With the information on the dataset, we have not been able to deepen into other parameters further than the impact of frequency or the use of CMA. We suggest recording potential changes in assessment policies that may lead to differences in course withdrawal, and even to perform A/B testing scenarios based on differences on assessment parameters.

Our study opens new research lines such as the execution of a comparative analysis between various types of learning disciplines and institutions. In addition, the investigation of specific applications oriented to design tailored assessment procedures, which could lead to personalized assessment plans addressed to lower dropout rates and ultimately boost academic achievement. The authors are open to collaborate in any of those research lines.

Acknowledgements

We would like to thank Professor Miquel Oliver from the Universitat Pompeu Fabra for his insightful feedback during the review process.

References

[1] G. Gibbs, C. Simpson, and R. Macdonald, “Improving student learning through changing assessment – a conceptual and practical framework,” in *EARLI Conference Padova*, 2003.

[2] L. R. B. Elton and D. M. Laurillard, “Trends in Research on Student Learning,” *Studies in Higher Education*, vol. 41, no. 1, pp. 87–102, 1979, doi: 10.1080/03075077912331377131.

[3] G. Gibbs, “Using assessment strategically to change the way students learn,” in *Assessment Matters in Higher Education*, Brown and Glasner, Ed., 1999, pp. 41–53.

[4] O. Aljohani, “A comprehensive review of the major studies and theoretical models of student retention in higher education,” *Higher Education Studies*, vol. 6, no. 2, pp. 1–18, 2016, doi: 10.5539/hes.v6n2p1.

[5] L. Aulck, N. Velagapudi, J. Blumenstock, and J. West, “Predicting Student Dropout in Higher Education,” in *ICML Workshop on #Data4Good: Machine Learning in Social Good Applications,* 2016. Accessed: Dec. 20, 2020. [Online]. Available: http://arxiv.org/abs/1606.06364

[6] J. Kuzilek, M. Hlosta, and Z. Zdrahal, “Data Descriptor: Open University Learning Analytics dataset,” *Sci Data*, vol. 4, no. 1, pp. 1–8, Nov. 2017, doi: 10.1038/sdata.2017.171.

[7] T. D. Erwin, *Assessing Student Learning and Development: A Guide to the Principles, Goals, and Methods of Determining College Outcomes*. Jossey-Bass Inc., Publishers, 1991.

[8] D. Pereira, M. A. Flores, and L. Niklasson, “Assessment revisited: a review of research in Assessment and Evaluation in Higher Education,” *Assess Eval High Educ*, vol. 41, no. 7, pp. 1008–1032, 2016, doi: 10.1080/02602938.2015.1055233.

[9] I. N. Z. Day, F. M. van Blankenstein, M. Westenberg, and W. Admiraal, “A review of the characteristics of intermediate assessment and their relationship with student grades,” *Assess Eval High Educ*, vol. 43, no. 6, pp. 908–929, Aug. 2018, doi: 10.1080/02602938.2017.1417974.

[10] S. Wang and D. Zhang, “Perceived teacher feedback and academic performance: the mediating effect of learning engagement and moderating effect of assessment characteristics,” *Assess Eval High Educ*, vol. 45, no. 7, pp. 973–987, Oct. 2020, doi: 10.1080/02602938.2020.1718599.

[11] C. Miller and M. Parlett, “Up to the Mark: A Study of the Examination Game.,” 1974.

[12] S. Bloxham and P. Boyd, *Developing Effective Assessment in Higher Education*, 1st ed. Open University Press, 2007.

[13] The European Association for Quality Assurance in Higher Education, “Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG),” 2015. Accessed: Sep. 01, 2021. [Online]. Available: https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf

[14] A. Simpson, “Assessment and its outcomes: the influence of disciplines and institutions,” *Assess Eval High Educ*, vol. 41, no. 6, pp. 917–937, Aug. 2016, doi: 10.1080/02602938.2015.1052369.

[15] S. Lynam and M. Cachia, “Students’ perceptions of the role of assessments at higher education,” *Assess Eval High Educ*, vol. 43, no. 2, pp. 223–234, Feb. 2018, doi: 10.1080/02602938.2017.1329928.

[16] J. T. E. Richardson, “Coursework versus examinations in end-of-module assessment: a literature review,” *Assess Eval High Educ*, vol. 40, no. 3, pp. 439–455, Apr. 2015, doi: 10.1080/02602938.2014.919628.

[17] L. Ellison and D. Jones, “First year law students: the impact of assessment type on attainment,” *Assess Eval High Educ*, vol. 44, no. 2, pp. 283–293, 2019, doi: 10.1080/02602938.2018.1496398.

[18] I. N. Z. Day, F. M. van Blankenstein, P. M. Westenberg, and W. F. Admiraal, “Explaining individual student success using continuous assessment types and student characteristics,” *Higher Education Research and Development*, 2018, doi: 10.1080/07294360.2018.1466868.

[19] R. Preston, M. Gratani, K. Owens, P. Roche, M. Zimanyi, and B. Malau-Aduli, “Exploring the Impact of Assessment on Medical Students’ Learning,” *Assess Eval High Educ*, vol. 45, no. 1, pp. 109–124, Jan. 2020, doi: 10.1080/02602938.2019.1614145.

[20] E. Trotter, “Student perceptions of continuous summative assessment,” *Assess Eval High Educ*, vol. 31, no. 5, pp. 505–521, 2006, doi: 10.1080/02602930600679506.

[21] G. Gibbs and C. Simpson, “Does your assessment support your students’ learning?,” *Learning and Teaching in Higher Education*, vol. 1, no. 1, pp. 1–30, 2004.

[22] K. Vandist, M. de Schryver, and Y. Rosseel, “Semisupervised category learning: The impact of feedback in learning the information-integration task,” *Atten Percept Psychophys*, vol. 71, pp. 328–341, 2009, doi: 10.3758/APP.71.2.328.

[23] C. Foster and P. Francis, “A systematic review on the deployment and effectiveness of data analytics in higher education to improve student outcomes,” *Assess Eval High Educ*, vol. 45, no. 6, pp. 822–841, 2020, doi: 10.1080/02602938.2019.1696945.

[24] T. Open University, “Changing Your Study Plans Policy 2022/23,” 2022. https://help.open.ac.uk/documents/policies/changing-your-study-plans/files/214/changing-your-study-plans-policy-2022-23.pdf (accessed Aug. 03, 2022).

[25] M. S. Larsen, K. P. Kornbeck, R. M. Kristensen, M. R. & Larsen, and H. B. Sommersel, *Dropout phenomena at universities: What is dropout? why does dropout occur? what can be done by the universities to prevent or reduce it? : A systematic review*. in Clearinghouse - forskningsserien. Copenhagen: Danish Clearinghouse for Educational Research, 2013.

[26] A. Behr, M. Giese, H. D. Teguim Kamdjou, and K. Theune, “Dropping out of university: A literature review,” *Review of Education*, vol. 8, no. 2, pp. 614–652, Jun. 2020, doi: 10.1002/rev3.3202.

[27] D. Orellana, N. Segovia, and B. R. Cánovas, “El abandono estudiantil en programas de educación superior virtual: revisión de literatura,” *Revista de la Educación Superior*, vol. 49, no. 194, pp. 45–62, Jun. 2016, doi: 10.36857/resu.2020.194.1124.

[28] S. R. Aragon and E. S. Johnson, “Factors influencing completion and noncompletion of community college online courses,” *Int J Phytoremediation*, vol. 22, no. 3, pp. 146–158, 2008, doi: 10.1080/08923640802239962.

[29] R. N. Landis and A. L. Reschly, "Reexamining gifted underachievement and dropout through the lens of student engagement," *Journal for the Education of the Gifted*, vol. 36, no. 2, 2013, doi: 10.1177/0162353213480864.
[30] L. Naude, L. Nel, R. van der Watt, and F. Tadi, "If it's going to be, it's up to me: first-year psychology students' experiences regarding academic success," *Teaching in Higher Education*, vol. 21, no. 1, pp. 37–48, Jan. 2016, doi: 10.1080/13562517.2015.1110788.
[31] S. Ameri, M. J. Fard, R. B. Chinnam, and C. K. Reddy, "Survival analysis based framework for early prediction of student dropouts," in *International Conference on Information and Knowledge Management, Proceedings*, 2016, pp. 903–912. doi: 10.1145/2983323.2983351.
[32] J. Cobre, F. A. C. Tortorelli, and S. C. de Oliveira, "Modelling two types of heterogeneity in the analysis of student success," *J Appl Stat*, vol. 46, no. 14, pp. 2527–2539, 2019, doi: 10.1080/02664763.2019.1601164.
[33] R. Villano, S. Harrison, G. Lynch, and G. Chen, "Linking early alert systems and student retention: A survival analysis approach," *High Educ (Dordr)*, vol. 76, no. 4, pp. 903–920, 2018, doi: 10.1007/s10734-018-0249-y.
[34] R. Spronken-Smith, C. Cameron, and R. Quigg, "Factors contributing to high PhD completion rates: a case study in a research-intensive university in New Zealand," *Assess Eval High Educ*, vol. 43, no. 1, pp. 94–109, 2018, doi: 10.1080/02602938.2017.1298717.
[35] W. Xing, H. Tang, and B. Pei, "Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs," *Internet and Higher Education*, vol. 43, p. 100690, 2019, doi: 10.1016/j.iheduc.2019.100690.
[36] S. Rizvi, B. Rienties, J. Rogaten, and R. F. Kizilcec, "Beyond one-size-fits-all in MOOCs: Variation in learning design and persistence of learners in different cultural and socioeconomic contexts," *Comput Human Behav*, vol. 126, no. C, p. 106973, 2022, doi: https://doi.org/10.1016/j.chb.2021.106973.
[37] J. A. Martínez-Carrascal, M. Hlosta, and T. Sancho-Vinuesa, "Using Survival Analysis to Identify Populations of Learners at Risk of Withdrawal: Conceptualization and Impact of Demographics," *The International Review of Research in Open and Distributed Learning*, vol. 24, no. 1, pp. 1–21, Feb. 2023, doi: 10.19173/irrodl.v24i1.6589.
[38] W. Xing, X. Chen, J. Stein, and M. Marcinkowski, "Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization," *Comput Human Behav*, vol. 58, pp. 119–129, 2016, doi: https://doi.org/10.1016/j.chb.2015.12.007.
[39] P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Munoz-Merino, and C. D. Kloos, "Prediction in MOOCs: A review and future research directions," *IEEE Transactions on Learning Technologies*, vol. 12, no. 3, pp. 384–401, 2019, doi: 10.1109/TLT.2018.2856808.
[40] T. G. Clark, M. J. Bradburn, S. B. Love, and D. G. Altman, "Survival analysis part I: Basic concepts and first analyses," *British Journal of Cancer*, vol. 89. pp. 232–238, 2003. doi: 10.1038/sj.bjc.6601118.
[41] T. G. Clark, M. J. Bradburn, S. B. Love, and D. G. Altman, "Survival analysis part IV: Further concepts and methods in survival analysis," *British Journal of Cancer*, vol. 89. pp. 781–786, 2003. doi: 10.1038/sj.bjc.6601117.
[42] M. J. Bradburn, T. G. Clark, S. B. Love, and D. G. Altman, "Survival analysis part II: Multivariate data analysis- An introduction to concepts and methods," *British Journal of Cancer*, vol. 89. pp. 431–436, 2003. doi: 10.1038/sj.bjc.6601119.
[43] M. J. Bradburn, T. G. Clark, S. B. Love, and D. G. Altman, "Survival analysis part III: Multivariate data analysis - Choosing a model and assessing its adequacy and fit," *British Journal of Cancer*, vol. 89. pp. 605–611, 2003. doi: 10.1038/sj.bjc.6601120.
[44] B. Holder, "An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs," *Internet and Higher Education*, 2007, doi: 10.1016/j.iheduc.2007.08.002.
[45] T. dong Kim, M. young Yang, J. Bae, B. a. Min, I. Lee, and J. Kim, "Escape from infinite freedom: Effects of constraining user freedom on the prevention of dropout in an online learning context," *Comput Human Behav*, vol. 66, pp. 217–231, 2017, doi: 10.1016/j.chb.2016.09.019.

[46] A. P. Rovai, “In search of higher persistence rates in distance education online programs,” *Internet and Higher Education*, vol. 6, no. 1, pp. 1–16, Jan. 2003, doi: 10.1016/S1096-7516(02)00158-6.
[47] V. Tinto, “Dropout from higher education: A theoretical synthesis of recent research,” *Rev Educ Res*, vol. 45, no. 1, pp. 89–125, 1975, doi: 10.3102/00346543045001089.