Assessing the effect of Massive Open Online Courses as remedial courses in higher education

Tommaso Agasisti, Giovanni Azzone and Mara Soncin*
School of Management, Politecnico di Milano, Milan, Italy

Postal address
School of Management, Politecnico di Milano, Via Lambruschini, 4, 20156, Milan (Italy).

This is a post-print of a paper published on Innovation in Education and Teaching International. Suggested citation: Agasisti, T., Azzone, G., & Soncin, M. (2022). Assessing the effect of Massive Open Online Courses as remedial courses in higher education. *Innovations in Education and Teaching International*, 59(4), 462-471.

Abstract

The current study assesses the effect of using Massive Open Online Courses (MOOCs) with the specific goal of providing remedial education. The data refer to an Italian flagship university, Politecnico di Milano, where a MOOC platform was launched following the strategy "MOOCs to bridge the gaps". Hence, the study aims at assessing the effect of students completing a MOOC taken as a foundation course in physics on their ability to pass the subsequent on-campus exam in physics (N=2,830). The research used Propensity Score Matching (PSM), basing the propensity scores on personal and academic information about the students. The results show that completers are 7 to 16 percentage points more likely to pass the related exam than the other students enrolled in the same MOOCs. These findings support the idea that using MOOCs for remedial purposes is effective in terms of student achievement within a formal education context.

Keywords: Higher education; Online learning; MOOC; Remedial education; Propensity score matching

Introduction

The trend for using Massive Open Online Courses (MOOCs) in higher education is firmly-established and gaining ground worldwide. Since the 'year of the MOOC', as The New York Times defined the year 2012 (Pappano, 2012), academics and practitioners have been debating on how MOOCs are able to disrupt higher education as we know it (Al-Imarah & Shields, 2019). MOOCs followed a 'fast cycle of hype and disappointment' soon after their introduction, mainly because of the low number of people actually completing courses among the high number of those who had initially registered to a class (Banerjee & Duflo, 2014, p.

* Email: mara.soncin@polimi.it

514), however we are now approaching a ‘second wave of MOOC hype’, as MOOC-based degrees are considered (Shah, 2018).

A relevant point to be considered is what place MOOCs hold alongside traditional campus-based education. Research into online education showed that online learning has either a non-significant or a negative effect compared to face-to-face education (Figlio et al., 2013; Bowen et al., 2014). Nevertheless, as long as blended education is considered, mixing online and traditional education has positive effects on student outcomes in terms of their achievement and engagement (Bernard et al., 2014). MOOCs have been blended with traditional education into various formats with mixed results (Hoxby, 2014). The present research contributes to this topic by analysing MOOCs used as foundation courses for prospective engineering students, with the specific aim of providing remedial education. The study analyses the case of the MOOC platform developed by Politecnico di Milano (PoliMi), the first Italian university to develop its own portal. The specific rationale underpinning the MOOCs provided on this platform is encapsulated in the statement or soundbite ‘MOOCs to bridge the gaps’. In the context of the MOOCs offered by PoliMi through its POK (PoliMi Open Knowledge) platform, this study wants to explore the following research question:

What effect has taking a MOOC foundation course on a student’s subsequent academic performance, after taking account of their personal characteristics and academic abilities?

The paper is organised as follows. Section 2 sets out a review of the related literature, and Section 3 introduces the data used. Section 4 presents the methodology, Section 5 shows the results, and the discussion and conclusions are given in Section 6.

Related literature

Using MOOCs as an integral part of on-campus classes is one of their most promising developments (Billington & Fronmueller, 2013) and can be included into the broader literature concerning the effect of online education on student achievement. The current study aims to explore two closely interlinked areas of research: the effect of online learning on student performance in HE and the combination of MOOCs and traditional education.

With reference to the effect played by online learning on student performance, mixed results emerged from the literature. Adopting a randomised experimental approach, Figlio et al. (2013) compared two courses, one online and one face-to-face, at a research university in the USA, demonstrating that the online learning has a negative effect on performance and that this is particularly damaging for low ability students. Bowen et al. (2014) applied a randomised experiment across six university campuses in the USA to compare blended education with face-to-face education, finding that no statistically significant difference could be detected in the students' results. Alpert et al. (2016) compared learning outcomes for traditional, blended and purely online formats in the USA, finding that online courses compared negatively with traditional education, most especially for disadvantaged students. In a quasi-experimental setting, Krieg and Henson (2016) assessed the effect of students taking a required course online or face-to-face on the corresponding follow-up course at a university over a ten-year time span and found that taking the class online had a negative effect compared to the traditional method. Overall, the literature points at demonstrating the negative effects of online learning compared to traditional face-to-face models. The current research contributes to this latter group of studies by applying a quasi-experimental design to assess the effect of online education.

Researchers investigated also the integration between MOOCs and traditional education. Online education was integrated into traditional type teaching either by replacing these traditional courses or by complementing the other courses. The first stream of applications has been covered in several studies investigating a format that can be traced back to the evaluation of blended courses in higher education (Bruff et al., 2013; Griffiths et al., 2014). Israel (2015) summarised several cases where MOOCs were blended with face-to-face classes. On reviewing five integration models, she found a small positive effect on the students' outcomes independently of student demographics, but lower student satisfaction compared to traditional classes. When MOOCs are blended with a traditional class in a course, Bralić and Divjak (2018) observed that students especially liked being able to study at their own pace and to test themselves regularly to check on what they had learnt. As Bruff et al. (2013) observed, 'instead of flipping one's course by producing online lecture videos or leveraging textbooks, instructors can wrap their courses around existing MOOCs'. In this respect, students raised some concerns about the increased workload they faced when MOOCs were brought into the design of a course (Bruff et al., 2013, p. 189). There is, however, evidence showing that MOOCs and online courses can complement traditional education, acting as a form of remedial education to ensure that students gain prerequisite or post-requisite skills. This leads to the area of main interest for the present study and the one to which we are specifically contributing. Most of literature has focused on assessing the effect of different forms of learning methods. As Wisneski et al. (2017) reported, the issue is connected to how and if knowledge can be transferred from an online environment to a traditional class. In their analysis of six universities in the USA, the authors assessed the effect of taking prerequisite and post-requisite courses online versus face-to-face. They found no difference in impact between the two in terms of the students' outcomes. Overall, little evidence has emerged from the literature analysing the specific application of MOOCs for remedial purposes.

Data

The data analysed in the study came from two sources. Information on the POK platform was provided by the University teaching innovation centre. Data on the students' characteristics were provided by the student services office, with particular reference to the first-year students who matriculated in the academic years 2014-15 and 2015-16.

Among the nineteen MOOCs offered on POK, the two MOOCs examined in our analysis – FIS101 and FIS102 – are in physics. We choose to focus on these MOOCs in part because physics is one of the subjects where students struggle the most, so supplementary material can be especially beneficial, in part because the syllabus of the MOOCs is aligned with that of the on campus course. Within this study, the two MOOCs in physics will be treated as one for two reasons. The first is that they are closely interconnected in terms of content, as FIS101 covers experimental physics relating to mechanics and thermodynamics, and FIS102 covers electromagnetism and optics, and both these subject areas are evaluated in the university exam in experimental physics. Second, the two MOOCs have similar features, with a common design and matching learner characteristics. In detail, FIS101 and FIS102 are divided into four or five weeks of different basic physics topics. Thus, five weeks is the shortest period of time to conclude the MOOCs, which are made of 649 and 450 minutes of videos and 18 and 22 tests, respectively. The courses are paced (i.e. material is initially released on a weekly basis), but once the content is available, students can decide to study only a selection of modules and not finish the rest of the course. All the materials remain available for a few months before the course stops and a new wave starts.

The MOOCs show a considerably high percentage of university students enrolled (70 to 75%, against an overall average of 54%), in line with the hypothesis that these MOOCs are mainly taken by (prospective) students willing to enrol or already enrolled at the university.

There is no requirement to take these courses in any programme (and no credits are awarded for completing a MOOC), although at on-campus pre-course sessions (free sessions offered to all prospective students on the basic topics), students are strongly encouraged to watch the relevant MOOCs in order to keep up with their lectures. Thus, the communication that students receive is both related to pre-courses, within which the MOOCs are embedded, and to the MOOC platform per se. The two MOOCs are specifically designed for remedial purposes and (together with mathematics) are part of the courses designed to fill possible gaps of knowledge between the high school and the Bachelor of Science (as stated on the platform). In this respect, we did not select students for the purpose of our study, but we investigated the learning profile of students who decided to enrol in the MOOCs beforehand.

We considered as our reference population only POK learners who were also university students and thus our results refer to the comparison between students enrolled in POK who did and did not complete the MOOCs. This choice partially addressed the issue of self-selection of students to the platform, as their personal and academic variables are used to match students at a subsequent stage of analysis. The aim was to compare learners who are as similar as possible in terms of observable personal characteristics, relative ability and motivation to enrol, but who differ in whether or not they completed the MOOC. In fact, by comparing the performance of students who completed the MOOCs with the entire student population, we would be likely to incur in a selection bias, as enrolment in MOOCs is not mandatory for any students and thus learners would differ both in our metric of interest (i.e. course completion) and in their motivation to enrol, making it impossible to disentangle the two effects, while we are solely interested in the first one.

Descriptive statistics comparing the PoliMi students enrolled in selected MOOCs to all other first-year students are given in Table 1, where students enrolled in the physics MOOC (N=2,830) are compared to the remaining population of first-year students (N=9,338). PoliMi students enrolled in one of the editions of FIS101/102 were more likely to be female (30% of students) than the remaining population (20%), they are slightly more disadvantaged in terms of their socio-economic status (6.59 to 10, while the average for the remaining students is 6.84) and they scored slightly less in the admissions test (2.46 to 5, against an average 2.53 for the other students).

[Table 1 around here]

With reference to data from the POK platform, we were not given information about the learners' clickstream, but we have data about the MOOC marks they scored by correctly completing the quizzes. Students enrolled in FIS101/102 show a peculiar distribution in terms of their marks, as the average MOOC completion rate (i.e. the number of students with a score of at least 60% and awarded the certificate over the number of students initially enrolled) is the lowest of all the MOOCs offered on POK at around 5%. Students completing the MOOC are labelled from here on as passer students. The completion rate is in line with the literature on this topic (Perna et al., 2014) and raises a point about the distribution of the marks of all the remaining students who did not complete the MOOC. Hence, alongside the definition of passer students, we added other labels to describe the learners' profiles. Active users are students whose final score was at least 10% and make up 8% of students in the MOOCs of interest, while quasi-active users are those with a final score of less than 10% but more than zero and make up 21% of the students. These students gave a correct answer to no more than 5 or 6 questions in the quizzes (equivalent at most to the test for having completed one week's worth of the course), but they were still somehow active in the course. The last group are the inactive students who enrolled but never completed a test, and these are around 40-60% of all students enrolled in the MOOCs of interest.

Empirical strategy

Defining the treatment and control group

Starting from the original population of students enrolled in one of the editions of FIS101/102, 'treatment' is defined as the student being in the position to obtain the certificate stating that he/she has completed the MOOC. We then considered three different control groups. The first is that of the inactive students, those who enrolled but did not take any tests. The second group consists of the quasi-active students, those who started at least one of the intermediate or final tests, and who engaged with the platform over and above merely registering to the MOOC. The third group is a combination of the two (inactive and quasi-active students). Active students were not selected into a control group in order to avoid comparing learners achieving scores just above and just below the threshold (that is, passing 60% of the quizzes). In our data, treatment is expressed through a binary variable (D_i) equal to 1 when the student *i* is treated, and 0 otherwise. The underlying assumption is that completing the MOOC would have a positive/negative effect on student's capacity to pass the relative university exam, which is our outcome variable.

Methodology: Propensity score matching

To investigate our research question, we applied a Propensity Score Matching (PSM) approach (Rosenbaum & Rubin, 1983), where the treatment effect T for a student i can be expressed as

$$T_i = Y_i(1) - Y_i(0) \qquad (1)$$

where $Y_i(D_i)$ is the potential outcome, given the treatment D_i. The final parameter of interest, defined as the 'average treatment effect on the treated' (ATT) is then defined as

$$T_{ATT} = E[Y(1)|D = 1] - E[Y(0)|D = 1] \qquad (2)$$

As outcome Y_i we consider a dummy variable equal to 1 when the student passed the physics exam and 0 otherwise. The effect of the treatment can, therefore, be read as the effect on the probability that the student will pass the exam. Propensity scores are defined through both a logit or a probit regression, where the dependent variable is whether the student gained the MOOC certificate (dummy = 1) or not (dummy = 0). To estimate this parameter, we included a vector for the student's personal characteristics (X_1) and information about his/her educational career (X_2), in order to predict the conditional probability of a student receiving the treatment on the basis of their pre-treatment characteristics (3). By computing the propensity scores through this methodology, our aim was to include as many observable individual characteristics as possible, so that the matching procedure only allowed comparisons between students who differ in terms of their exposure to treatment.

$$\text{p (X)} = Pr\ (D = 1|X_{1,2}) \qquad (3)$$

There are two key assumptions underlying the use of matching methods. The first is the conditional independence assumption, according to which, after controlling for a set X of observable characteristics, the outcome of interest is independent of treatment status, as expressed in (4).

$$[Y(1), Y(0)] \perp D|X_{1,2} \qquad (4)$$

The conditional independence assumption is also known as selection on observables, as it requires all the variables affecting the treatment probability to be included as covariates. As such, it cannot be directly tested on data and as such represents a possible threat to validity. Nevertheless, as proposed by Becker and Ichino (2002), we balanced the pre-treatment variables, given the propensity scores, in order to ensure that the correct matching procedure was run. If the balancing hypothesis is satisfied, observations with a given propensity score are assumed to be comparable in terms of observable and unobservable characteristics, independently of the treatment.

The second is the common support assumption, according to which there is a positive probability of being both treated and untreated for each value of X, as given in (5).

$$0 < Pr\left(D = 1 \middle| X_{1,2}\right) < 1 \tag{5}$$

If the assumptions are satisfied, then treatment assignment is ‘strongly ignorable’ (Rosenbaum & Rubin, 1983, p. 43). Hence, all the algorithms are restricted to the common support region to ensure the validity of results.

To ensure the consistency of results across models, different specifications of the PSM were applied (Caliendo & Kopeinig, 2008). In particular, the following matching methods were used:

- A one-to-one matching with replacement (one nearest neighbour, Model_1, or three nearest neighbours, Model_2), where each individual in the treatment group is compared to the ‘most similar’ individual(s) in the control group in terms of propensity score, and the same observation in the control group is allowed to be the best match for more than one unit in the treatment group.

- A one-to-one matching without replacement (Model_3), which is the most restrictive, and allows each unit in the treatment group to be matched to only one observation in the control group.

- A Radius matching method (Model_4), which matches control units where the propensity scores fall within a radius r of the treated observations; a default value of 0.1 is used as the radius.

- A Kernel matching method (Model_5), where a Gaussian is selected as kernel function, with a default bandwidth of 0.06.

- A stratification method (Model_6), which stratifies propensity scores into blocks, matching observations between treatment and control groups within each block.

The different set of results were used to compare the outputs of the different methodological approaches in order to avoid the findings being affected by the empirical specification.

Results: The effect of taking a MOOC

The main results of the study are presented in Table 2, which shows the treatment effect of taking a MOOC on the student's performance in university exams. Additionally, our results comply with the balancing condition, as available from the corresponding author on request. As previously described, a number of matching procedures were carried out for each specification of the control group, in order to ensure the validity of the results. The variable identifying the treatment effect is equivalent to the difference, between the treated and the control group, in the probability of a student passing the university exam in physics. As a result, students successfully completing the MOOC in physics are more likely to pass the exam by 7 to 16% than the other students enrolled in the same MOOC. The effect is significant regardless of the model specified and across control groups. Limited exceptions are the probit model specification for control group 2 (quasi-active students) and the Nearest Neighbour (NN) specifications for control group 3 (inactive and quasi-active students).

Hence, the treatment is significant when comparing completers (treated) and quasi-active or inactive students. The results are robust compared to other matching procedures as available from the corresponding author on request. In general terms, the results show that, with personal and academic characteristics being the same and self-selection in an online learning environment remaining constant, students who actively engage in the MOOC have higher probabilities of passing the university exam. These findings have a number of practical implications that are discussed in detail below.

[Table 2 around here]

Discussion and conclusions

In this study, we found that taking a foundation MOOC had a positive effect on key competencies in physics. However, for the treatment to be effective, students should persist in completing the course, getting through (at least) 60%. This is a relevant point in the MOOC-related debate, as one of the emerging directions of development is ‘bite-sized’ learning (de Freitas et al., 2015), involving small, self-contained information nuggets within online courses. The contribution of this study is that, when combined with formal education, this learning approach is not successful, and a ‘see to completion’ approach is in fact more effective than a ‘grab-and-go’ learning style (which is how we can interpret the approach taken by the quasi-active students).

Hence, using these tools in a formal education context should be combined with actions that encourage students to persist and complete the course. It can be the case that prospective university students have not yet learnt how to approach learning flexibly, while the method can work in graduate education or life-long learning. It can also be the case that differences can depend on the subject or type of education, and a shorter exposure to treatment can be effective for less technical material or in informal learning contexts. This kind of heterogeneity suggests a first direction for future extensions to this study.

Thinking of university managers and policy-makers, our findings suggest that there is room for MOOCs alongside face-to-face classes, specifically as foundation, pre-requisite courses. However, student persistence throughout the course should be encouraged, and this is a point for open discussion in the MOOC debate where completion rates are critically low (Perna et al., 2014; Gregori et al, 2018).

Findings support increasing evidence about the positive effects of online learning when it is combined with traditional education (Bernard et al., 2014), particularly to cover pre-requisite topics (Krieg & Henson, 2016; Wisneski et al., 2017). What this study adds is an insight into how to carry out this combined approach. Students completing at least 60% of a course have a higher probability of passing the relative university exam in the same subject by 7-16%, especially in comparison with students who completed only a small part of the MOOC (i.e. less than 10%, as in the case of the quasi-active students). Hence, our findings suggest that there is space for implementing online learning as a form of remedial education or foundation course, in order to provide an equal opportunity of success to all students starting tertiary education.

As a limitation of the study, we are aware of a possible bias caused by students self-selecting to the platform and by the consequent effect that unobservable variables (like motivational aspects) can have. However, we put in place several strategies to limit these issues, as we control for the relative achievement of students and we compare only students registered on the platform.

Acknowledgments. Authors would like to thank Susanna Sancassani, Federica Brambilla, Luigi Bissolotti and all the staff from PoliMi who helped and supported us in data collection.

Disclosure statement

No potential conflict of interest was reported by the authors

Notes on contributors

Tommaso Agasisti is Professor at Politecnico di Milano, School of Management and Associate Dean for Internationalization, Quality and Services at MIP Politecnico di Milano Graduate School of Business. His studies are in the field of Public Economics and Finance, Public Management and Policy, with particular reference to the educational sector.

Giovanni Azzone is Professor of Business Economics and from 2010 to 2016 he was Rector at Politecnico di Milano. His research activities are mainly in the field of organizational analysis and accounting and management control in private and public organisations with more than 65 publications on international journals and books.

Mara Soncin is Junior Assistant Professor at Politecnico di Milano, School of Management. Her main field of research concerns the evaluation of the performances of digital learning initiatives in Higher Education. Her research interests also involve educational management and the use of econometric models for the evaluation of educational policies.

References.

Al-Imarah, A. A., & Shields, R. (2019). MOOCs, disruptive innovation and the future of higher education: A conceptual analysis. *Innovations in Education and Teaching International*, *56*(3), 258-269.

Alpert, W. T., Couch, K. A., & Harmon, O. R. (2016). A randomized assessment of online learning. *American Economic Review*, *106*(5), 378-82.

Banerjee, A. V., & Duflo, E. (2014). (Dis) organization and Success in an Economics MOOC. *American Economic Review*, *104*(5), 514-18.

Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal, 2*(4), 358-377.

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87-122.

Billington, P. J., & Fronmueller, M. P. (2013). MOOCs and the future of higher education. *Journal of Higher Education Theory and Practice*, *13*(3/4), 36.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six‐campus randomized trial. *Journal of Policy Analysis and Management*, *33*(1), 94-111.

Bralić, A., & Divjak, B. (2018). Use of moocs in traditional classroom: blended learning approach. *European Journal of Open, Distance and E-learning*, *21*(1), 1-9.

Bruff, D. O., Fisher, D. H., McEwen, K. E., & Smith, B. E. (2013). Wrapping a MOOC: Student perceptions of an experiment in blended learning. *Journal of Online Learning and Teaching*, *9*(2), 187.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31-72.

De Freitas, S. I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology*, *46*(3), 455-471.

Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, *31*(4), 763-784.

Gregori, E. B., Zhang, J., Galván-Fernández, C., & de Asís Fernández-Navarro, F. (2018). Learner support in MOOCs: Identifying variables linked to completion. *Computers & Education*, *122*, 153-168.

Griffiths, R., Chingos, M., Mulhern, C., & Spies, R. (2014). *Interactive online learning on campus: Testing MOOCs and other platforms in hybrid formats in the University System of Maryland*. Ithaka S+ R.

Hoxby, C. M. (2014). The economics of online postsecondary education: MOOCs, nonselective education, and highly selective education. *American Economic Review*, *104*(5), 528-33.

Israel, M. J. (2015). Effectiveness of integrating MOOCs in traditional classrooms for undergraduate students. *The International Review of Research in Open and Distributed Learning*, *16*(5), 102-118.

Krieg, J. M., & Henson, S. E. (2016). The educational impact of online learning: How do university students perform in subsequent courses?. *Education Finance and Policy*, *11*(4), 426-448.

Pappano, L. (2012, November 2). *The year of the MOOC*. The New York Times, ED26. http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in massive open online courses. *Educational Researcher, 43*(9), 421-432.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Shah, D. (2018, June 3) *The second wave of MOOC hype is here and it's online degrees.* https://www.class-central.com/report/second-wave-of-mooc-hype/

Wisneski, J. E., Ozogul, G., & Bichelmeyer, B. A. (2017). Investigating the impact of learning environments on undergraduate students' academic performance in a prerequisite and post-requisite course sequence. *The Internet and Higher Education, 32*, 1-10.

Table 1. Descriptive comparison between first-year PoliMi students enrolled in the MOOCs of physics and the other first-year PoliMi students.

Variable	Polimi first-year students (enrolled in FIS101or102)		Other Polimi first-year students	
	Mean	*N*	Mean	*N*
Female student (dummy=1)	0.30	*2,830*	0.20	*9,338*
Italian citizen (dummy=1)	0.95	*2,830*	0.94	*9,338*
Region (North=1; Center=2; South=3)	1.34	*2,778*	1.25	*9,008*
Socio-economic status (ordinal variable from 1 to 10)	6.59	*2,829*	6.84	*9,101*
Scholarship granted (dummy=1)	0.08	*2,830*	0.05	*9,338*
Score in the admission test (up to 5)	2.46	*2,820*	2.53	*8,713*
Scientific high school diploma (dummy=1)	0.73	*2,830*	0.75	*9,338*
High school grade (up to 100)	83.17	*2,820*	82.43	*9,063*

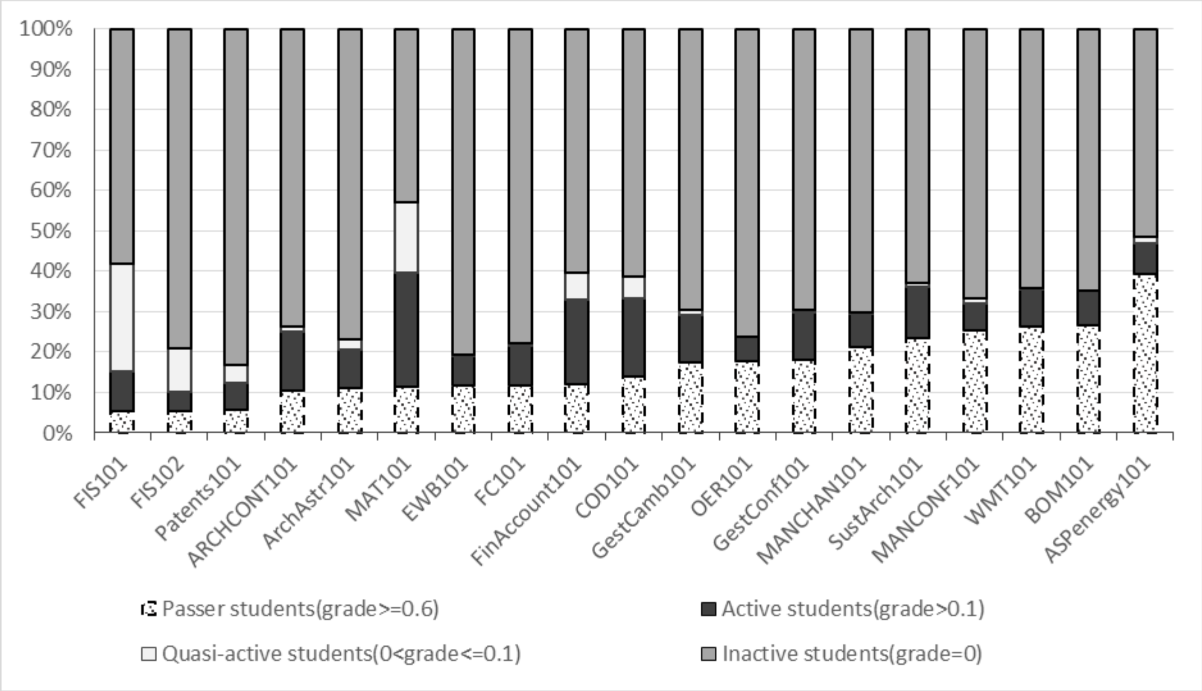
Note: Mean and number of observations (N) given for each individual-level variable. FIS101&102 indicates the MOOCs of physics.

Table 2. Propensity score matching: results of the treatment effect.

Variable	Logit estimation					Probit estimation				
	Treated (T)	**Controls (C)**	**Difference (T-C)**	**S.E.**	**T-stat**	**Treated (T)**	**Controls (C)**	**Difference (T-C)**	**S.E.**	**T-stat**
Sample: Passer (treated) vs. Inactive (control 1, grade=0)										
NN(1) with replacement	N=164	N=1,590	**0.091**	0.049	1.87	N=164	N=1,590	**0.122**	0.049	2.49
NN(3) with replacement	N=164	N=1,590	**0.073***	0.036	2.02	N=164	N=1,590	**0.085***	0.037	2.31
NN(1) without replacement	N=164	N=164	**0.110***	0.044	2.52	N=164	N=164	**0.122***	0.044	2.78
Radius matching method	N=164	N=1,582	**0.159***	0.030	5.35	N=164	N=1,582	**0.154***	0.030	5.18
Kernel matching method	N=164	N=1,582	**0.116***	0.028	4.07	N=164	N=1,582	**0.116***	0.026	4.39
Stratification method	N=164	N=2,585	**0.079***	0.028	2.84	N=164	N=2,585	**0.076***	0.03	2.51
Sample: Passer (treated) vs. Quasi-active (control 2, 0<grade<=0.1)										
NN(1) with replacement	N=164	N=742	**0.110***	0.049	2.09	N=164	N=742	**0.037**	0.049	0.74
NN(3) with replacement	N=164	N=742	**0.091***	0.043	2.26	N=164	N=742	**0.055**	0.040	1.36
NN(1) without replacement	N=164	N=742	**0.103***	0.043	2.40	N=164	N=742	**0.043**	0.041	1.05
Radius matching method	N=162	N=740	**0.135***	0.033	4.12	N=162	N=740	**0.130***	0.033	3.99
Kernel matching method	N=164	N=740	**0.092***	0.031	3.01	N=164	N=740	**0.093***	0.030	3.13
Stratification method	N=162	N=2,590	**0.070***	0.029	2.37	N=162	N=2,590	**0.071***	0.030	2.41
Sample: Passer (treated) vs. Quasi-active & Inactive (control 3, 0<=grade<=0.1)										
NN(1) with replacement	N=164	N=2,332	**0.067**	0.045	1.50	N=164	N=2,332	**0.073**	0.045	1.61
NN(3) with replacement	N=164	N=2,332	**0.065**	0.034	1.87	N=164	N=2,332	**0.049**	0.034	1.42
NN(1) without replacement	N=164	N=164	**0.073**	0.042	1.74	N=164	N=164	**0.079**	0.042	1.87
Radius matching method	N=164	N=2,332	**0.162***	0.029	5.61	N=164	N=2,332	**0.160***	0.029	5.54
Kernel matching method	N=164	N=2,332	**0.135***	0.027	5.09	N=164	N=2,332	**0.136***	0.026	5.24
Stratification method	N=164	N=2,558	**0.082***	0.027	2.98	N=164	N=2,558	**0.081***	0.027	2.94

Note: NN stands for Nearest Neighbour. *Passer* students: students getting a grade of at least 60% in the MOOC. *Inactive* students: students getting a grade of 0 in the MOOC. *Quasi-active* students: students getting a grade lower than 10% but different from 0. The column "Difference" refers to the main effect. *T-stat <>2.

Figure 1. Descriptive statistics of passer, active, quasi-active and inactive students for all the MOOCs offered on the POK platform, ordered by percentage of passer students.



Note: *Passer students*: students getting a grade of at least 60% in the MOOC. *Quasi-active students*: students getting a grade lower than 10% but different from 0. *Active students*: students getting a grade lower than 60% but higher or equal to 10%. *Inactive students*: students getting a grade of 0 in the MOOC. All POK users included in the analysis (N=60,608).