Christoph Meinel, Thomas Staubitz, Stefanie Schweiger, Christian Friedl, Janine Kiers, Martin Ebner, Anja Lorenz, George Ubachs, Catherine Mongenet, José A. Ruipérez-Valiente, Manoel Cortes Mendez (Eds.)

EMOOCs 2021

2021 – xii, 295 p.
ISBN 978-3-86956-512-5
DOI https://doi.org/10.25932/publishup-51030

Suggested citation:
Julien Jacqmin: What Drives Enrollment in Massive Open Online Courses?, In: Christoph Meinel, Thomas Staubitz, Stefanie Schweiger, Christian Friedl, Janine Kiers, Martin Ebner, Anja Lorenz, George Ubachs, Catherine Mongenet, José A. Ruipérez-Valiente, Manoel Cortes Mendez (Eds.): EMOOCs 2021, Potsdam, Universitätsverlag Potsdam, 2021, S. 1–16.
DOI https://doi.org/10.25932/publishup-51689

What Drives Enrollment in Massive Open Online Courses?

Evidences from a French MOOC Platform

Julien Jacqmin

NEOMA Business School
1 Rue du Maréchal Juin, 76130 Mont Saint-Aignan, France

The goal of this paper is to study the demand factors driving enrollment in massive open online courses. Using course level data from a French MOOC platform, we study the course, teacher and institution related characteristics that influence the enrollment decision of students, in a setting where enrollment is open to all students without administrative barriers. Coverage from social and traditional media done around the course is a key driver. In addition, the language of instruction and the (estimated) amount of work needed to complete the course also have a significant impact. The data also suggests that the presence of same-side externalities is limited. Finally, preferences of national and of international students tend to differ on several dimensions.

1 Introduction

Massive open online courses platforms provide to higher education institutions a place where they can interact with students by providing them online courses. To date, various platforms have enrolled several million students, numbers which are rarely heard in the higher education context. They offer an unusual setting to understand the behavior of students with respect to various aspects of their learning decisions in an online context.

Economically-speaking, one key peculiarity of MOOCs is that they can be defined as global public goods [12]. Thanks to online technology that limits congestion, the "consumption" of the course by one student does not negatively impact the "consumption" possibility of other students. Courses are designed such that feedbacks to students throughout their learning experience take place via automatically- or peer-graded assessments or via interactive *Q&A* video sessions [17]. Hence, MOOCs have the property of being non-rival in consumption. They are also non-excludable for two reasons. First, digitization eliminates the issue of capacity constraint prevalent in traditional higher education programs due to the limited physical size of classrooms. As a consequence, enrolling an additional student online has a marginal cost equal to zero. Second, MOOC platforms have committed to remain open, as attested by the second "O" of the MOOC acronym.[1] Courses can be accessed free of charge and students may not be excluded from them following a selection procedure.

Hence, it is not just because they take place online and that they are free that set MOOCs apart from traditional programs. Openness is also a defining feature, as translated in the student allocation mechanism to courses. Traditional programs are relying on two-sided allocation technologies. Students do not only have to apply to enroll a program but they also have to be accepted by the school side. This last stage takes place in various forms, sometimes in a decentralized way or via a procedure common to various programs (see for example [1]). To be accepted, students might be required to pass some (centralized) exam or to fulfill some more or less specific criteria. Instead, MOOCs use a one-sided allocation mechanism, where students can unilaterally decide to enroll or not. In addition, educational programs are unbundled and the student's enrollment decision takes place on a course per course base.

The goal of this paper is to study three issues. First, the determinants of enrollment in courses organized online and open to anyone for free and without administrative requirements are analyzed. Thank to this analysis, we are able to know more about why some of these global public goods attract more beneficiaries than others. Second, we look at whether the preferences of French and international students differ, as in this setting they are not discriminated by administrative procedures or by the distance that separates them from the institution providing the course. Finally, the timing of the launch of the first run of the course and its impact on enrollment is analyzed. For this purpose, we analyze, by the mean of a multivariate regression approach, course level enrollment data from FUN (for France Université Numérique), a French MOOC platform, using data covering all the MOOCs launched during its first four years of existence.

First of all, we find that (social and traditional) media coverage of the MOOC is a key driver of enrollment for both national and international students. This result suggests the importance of well-designed media campaigns organized by the institution providing the MOOC in order to inform potential students as a way to increase enrollment. Second, course related factors tend to have an impact on enrollment but factors related to the teacher of the MOOC or to his/her host institution have a more limited influence. Among others, the instructional language,

[1] Over the recent years, some platforms like Edx and Coursera have limited the access to some of the content of their courses (quizzes, access to forum, certificate of completion, etc.). However, there still remains some content open to all users, without paying.

the possibility to pay to obtain a certificate of completion and the expected amount of work required to finish the course all have a positive and significant impact on enrollment. French students tend to differ from international ones on several dimensions. For example, they prefer courses taught in French, while international students rather enroll courses taught by scholars from foreign institutions. Finally, this analysis does not confirm the existence of (positive or negative) same-side externalities, as enrollment is not impacted by the launch of other courses at the same time.

After a presentation of the literature in section 2, the methodology used is presented in section 3. The results of the empirical analysis are presented in section 4. section 5 concludes.

2 Literature Review

Distance education has recently made a comeback under the spotlight of the economics literature, thanks mostly to the emergence of MOOCs (see [4, 6], for reviews of the phenomenon). Two questions have been central in the literature so far. The first concerns whether online educational programs provide a good alternative to their brick-and-mortar version in terms of student learning outcomes [3]. The second analyzes the reasons why higher education institutions decide to innovate by organizing online programs [5].

There is a large literature in the field of economics of education studying the factors driving enrollment decision in non-compulsory tertiary education which is more closely related to the research question of this paper. A first strand of the literature has analyzed the demand for higher education programs by looking at the role of fees and of financial aid (see for example [11]). Others as reviewed in Ehrenberg [7] have rather looked at the role of proximity between students and the higher education institutions using gravity models. How students are allocated to higher education institutions and questions related to the impact of the application procedure on enrollment is a question that has also recently attracted a lot of attention. One key feature of this setting is that neither the tuition expenditure, nor the transport cost nor the selection procedure plays a role in determining the students' demand as courses are free, take place online and are open to all after the few clicks required to log in.

Hence, this paper is closely linked to the literature analyzing the demand for online educational programs. For example, Ortagus and Yang [15] study the impact of decreasing state funding on the enrollment of students in online programs provided by U.S. public universities. In another work, Goodman, Melkers, and Pallais [8] study whether providing in parallel an online and an in-class version of a program can boost the enrollment in the latter. These two works have the peculiarity that, despite taking place online, students need to be selected to attend the program and to pay a tuition fee. Closer to us, Tong and Li [19] evaluate the factors impacting the demand for Massive Open Online Courses using cross-sectional data of OECD countries and China. Due to the absence of aggregate enrollment data at the country level, they construct indices using various sources from the internet to proxy the evolving demand for MOOCs. They observe that the quality of the broadband connection and the level of unemployment are key drivers of this demand.

One key distinguishing feature of this approach is that we are looking at the demand for online courses, using course level data rather than usual program or institution level data, as courses are offered in an unbundled format.[2] Hence, we can look at other additional factors related to the course and his/her teacher. To our knowledge, there are few papers with access to this kind of proprietary data.[3] Hansen and Reich [9] and Ruiperez-Valiente, Halawa, Slama, and Reich [18] are notable exceptions. Using data from 68 MOOCs launched by Harvard and MIT, Hansen and Reich [9] study the background characteristics of enrollees and their influences on student dropout. They find that student's socioeconomic status play a central role in explaining gaps in educational outcomes. Ruiperez-Valiente, Halawa, Slama, and Reich [18] analyze and compare data from Edx, a U.S. MOOC platform jointly launched by Harvard and MIT, and Edraak, a MOOC platform targeting students located in Arab countries, to study whether student's preferences differ from one platform to the other. The authors observe that younger and less educated learners tend to enroll the Edraak platform. These papers have in common that they focus on correlations linked with the enrollment of students. Using a multiple regression analysis, we aim at obtaining more robust measures of controlled correlations by taking into consideration the presence of various omitted variables, i.e. variables both related to our dependent and independent variables, that would lead to otherwise biased estimates. In addition, we do not focus on a sample of courses provided by some institutions on a MOOC platform but on all the courses provided on the platform during its first years of existence.

[2] Remark that enrollment must not be confused with a learning outcome, it is a required step to pursue a training on a MOOC platform. Hence, it can be seen as a key signal of the student's willingness to learn, even if many of them are only there to sample a part of the course.

[3] Using course level data from brick-and-mortar courses, Budish and Cantillon study the course allocation mechanism in place in Harvard where a draft mechanism allocates students to seats available due to the limited capacities of classes. This analysis departs from this approach as in this setting there is no limitation and no administrative barrier to the number of students enrolling the MOOC. Hence, there is no need to develop a market design for this allocation problem.

3 Methodology

3.1 Context

This empirical analysis is based on proprietary data from FUN, a French MOOC platform. This platform was launched jointly by a consortium of French universities and the French government. It is built upon the Open Edx software. As of 2020, it has enrolled in total more than 6 million students in its various courses. The data is organized at the course-level. Courses aired first online from January 2014, the launch of the platform, to November 2017 are included in this sample. Only data about the first run of the course taught is considered. In total, the final sample includes 284 different courses. This information is completed by data gathered from each of the course webpages available on the FUN platform and various additional sources such as Scopus, the Shanghai University's Academic Ranking of World Universities (better known as the Shanghai ranking) and the Europresse database. Descriptive statistics are shown in Table 1.

3.2 Measures

The main dependent variable is total enrollment. It is the number of students who have subscribed to the first organized run of the course. As pictured on Figure 1, we observe that the distribution is skewed to the right. In addition to be also able to interpret the coefficients, this variable is log transformed. Next, data about the origin of the students is used, i.e. whether they are French or international students. This information is collected via FUN, when subscribing to the platform. Unfortunately, only a minority of students mention this information. National and international enrollments are extrapolated assuming that they form the same share of the student course attendance as the one that has filled in this question of the survey.

Independent variables are grouped into four separate categories: Course, teacher, institution related and other factors. The topic of the course is considered in this analysis by the mean of four binary variables: *Mgmt and law*, *humanities*, *STEM* and *health*, the last being the reference category. They are included as the popularity of the course is likely to be related to the topic being presented. If instead the course categories are refined in 10 categories, the results remain the same. *Sequel* is the number of times the course was taught, whether with similar or additional content but with the same course title. This variable is similar to the one used in the economic literature on movies [13]. A higher value can be interpreted as a measure available *ex-post* of the success of the MOOC, as if the teaching team

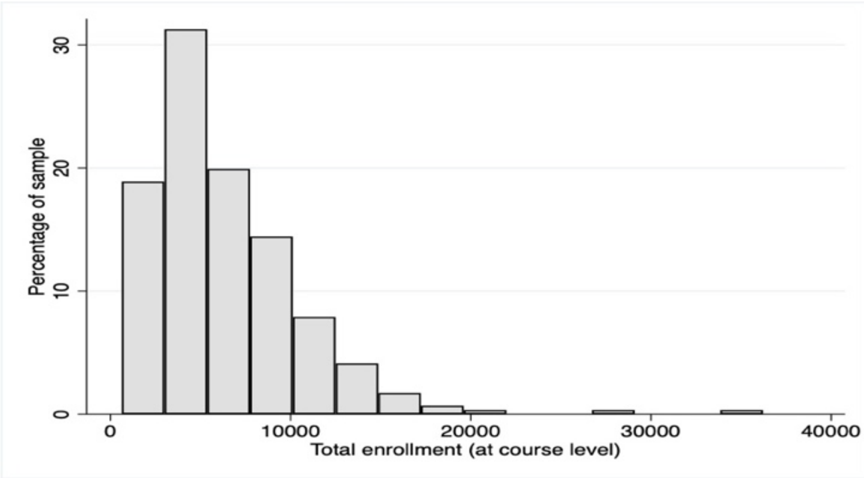

Figure 1: Distribution of the total enrollment of students MOOCs

had a bad MOOC experience they would not organize additional sequels.[4] *Work* is the estimated amount of hours needed to fulfill the course requirement, as advertised on the course webpage. This variable can be seen as a proxy of the depth of the course. *Prerequisite* is a binary variable whether an informally required prerequisite is stated on the course webpage, even if there is whatsoever no control of their fulfillments at the enrollment stage. As they can put off some students from enrolling, it is likely to drive down enrollment. The presence of a *paid certificate*, represented via a binary variable in this analysis, can attract some students, as they might think that some employers value this information signal on their CV. Only 5% of the courses provide this possibility. *Grading* is a binary equal to one if the course grading system is based on more than multiple choice questions. Using problem sets or essays graded by fellow students or valuing the participation to forum discussions can be seen as a proxy of the teachers' pedagogical investment, which can be valued by students at the time of their enrollment decision. *French* is a binary whether the course is taught in French. As more than three fourth of the students are French and 94% of the courses are taught in French, it can be expected to have a positive coefficient.

[4] Remark that this variable can potentially be interpreted as a bad control as the decision to put up a sequel can depend on the success of the initial course, as measured by enrollment figures. However, not including this variable in this model does not impact the quality of these results.

Tenured teacher is a first teacher-related factor included in the model.[5] It is a binary variable equal to one if the main teacher of the course is a full or an associate professor from a higher education institution. *MOOC experience* is a binary equal to one if the main teacher of the MOOC has had beforehand an experience teaching a MOOC. It is the case for one fourth of the courses provided on FUN. *MOOC experience* can be seen as a proxy of the experience of the teacher with this format of learning practice. As obtained via Scopus, *H-index* is a measure of the reputation within the scientific community of the teacher of the course. This variable can be seen as a way to test whether "superstar" researchers are more able to cater students to enroll their courses. The Hirsch index is defined as the number of publications x that have ever been cited at least x times.

Four institution-related independent variables are also included. *University* and *Grande Ecole* are two binary variables, the latter being French public higher education institutions traditionally allowed to select students based on their abilities at the entrance of their traditional programs. Note that the reference category is an heterogeneous group of private higher institutions, not-for-profit organizations, government and administrative bodies. As 4% of the courses are provided by institutions based outside of France, we control for this with a binary called *foreign institution*. Finally, we control for the international prestige of the institution using its ranking in the Shanghai University's Academic Ranking of World Universities. *Ranking* is the ranking of the year 2018 and, to facilitate the interpretation of the sign of the parameter, we subtract the rank of the institution from 1000, which corresponds to the last ranked institution. We also take the log of ranking. This variable will allow us to test whether MOOCs hosted by prestigious institutions tend to enroll more students.

In addition, we control for two variables related with the communication campaigns made around the course: *Twitter* and *Media coverage*. The first is a binary variable whether there is a Twitter profile associated with the course in order to communicate about it. The second variable measures the number of times the course was mentioned on the Europresse database, a database covering most French and international newspapers and magazines.[6] As it is highly skewed, we take the logarithm of the number of times the course was mentioned, plus one unit due to the presence of zeros. We also include a monthly time trend in order to take into consideration the increasing affiliation of students across the first

[5]For these factors, we take into account the main teacher of the course, identified as the person in charge of it. If not explicitly assigned on the course webpage of the FUN platform, the one cited on the top of the list is taken into consideration and, if this list is in alphabetical order, we take the first appearing in the trailer of the MOOC.

[6]If we instead use appearances in Google News, we end up with a variable that is highly correlated with the one we use. We also obtain very similar results using this source to proxy the media exposure of the MOOC.

47 months of the existence of the MOOC platform. Finally, we study the impact of the timing of the course launch on enrollment using two approaches. First, we consider monthly binary variables, using the month of November as a reference category as it is the one with the highest number of course launched. Then we control for the number of courses launched on the platform with a three-month window around the course considered. This approach will help characterize the same-side/within group externality on the content-provider side of the platform. If this externality is negative, it means that course compete one with another while if it is positive, launching a course at the same time as other courses is a good thing for your own enrollment. We define in three different ways the number of courses: the total number of MOOCs, the number of MOOCs on the same theme and the number of MOOCs taught in the same discipline. FUN does not play an editorial role in selecting the courses aired on its platform and it takes at best 6 to 12 months to repurpose a course to the MOOC format [10]). Hence, the decision to launch a course is as good as exogenous from the starting time of other courses. Summary statistics are presented in Table 1.

Table 1: Summary statistics

N = 284	Mean	Standard Deviation	Min	Max
Dependent variables:				
Total enrollment	6 397.10	4 205.76	629.00	36 217
National enrollment	4 446.21	3 194.71	361.15	27 592
International enrollment	1 950.89	1 485.69	127.18	9 607
Course related independent variables:				
Humanities	0.24	0.43	0	1
Mgmt and law	0.24	0.43	0	1
STEM	0.42	0.49	0	1
Health	0.10	0.30	0	1
Sequel	2.41	1.46	1	11
Work	17.75	9.20	1.5	59.5
Prerequesite	0.52	0.50	0	1
Paid certificate	0.05	0.22	0	1
Grading	0.24	0.43	0	1
French	0.94	0.24	0	1
Teacher related independent variables:				
Tenured teacher	0.70	0.46	0	1
MOOC experience	0.25	0.44	0	1
H-index	9.10	13.45	0	82

Continued on next page…

Summary statistics (cont.)

N = 284	Mean	Standard Deviation	Min	Max
Institution related independent variables:				
University	0.48	0.50	0	1
Grande Ecole	0.31	0.46	0	1
Foreign Institution	0.04	0.20	0	1
Ranking (log)	2.32	3.09	0	6.87
Twitter	0.39	0.49	0	1
Media coverage (log)	0.63	0.80	0	3.58
Monthly time-trend	27.17	13.24	1	47
# of MOOCs (total)	22.47	7.74	1	36
# of MOOCs (theme)	3.23	1.74	1	8
# of MOOCs (discipline)	7.25	4.17	1	19

3.3 Proposed Approach

For this multiple regression analysis, Ordinary Least Square (OLS) estimators are computed on this sample of 284 courses using the software STATA. Due to the skewness of the enrollment data and the presence of outliers, we apply a log transformation of this dependent variables. Multicollinearity leads to unstable and unreliable estimates. However, the computation of the Variance Inflation Factor (VIF) values show no reason to suspect it is an issue in our analysis. We present heteroskedasticity-robust standard errors. Overall, these coefficients should be analyzed with caution, as controlled correlations rather than causal implications.

4 Results

In order to analyze these results, we proceed in two steps. First, we will analyze the drivers of total enrollment, as shown in regression (1) in Table 2. Second, we will look at whether the factors affecting the enrollment of French students differ from the ones of international students, as respectively shown in regression (2) and (3). Finally, in regression (4) to (9) shown in Table 3, we study the issue of the timing of the launch of the course.

Table 2: Results (1): The drivers of total, national and international enrollment

Dependent variable:	(1)	(2)	(3)
Enrollment	Total	National	International
Mgmt and law	−0.035 (0.1)	−0.027 (0.111)	−0.104 (0.113)
STEM	0.034 (0.088)	−0.016 (0.095)	0.076 (0.096)
Health	−0.156 (0.136)	−0.161 (0.142)	−0.111 (0.144)
Sequel	−0.016 (0.025)	−0.013 (0.027)	−0.01 (0.029)
Work	0.012*** (0.004)	0.012*** (0.004)	0.011** (0.006)
Prerequesite	−0.034 (0.07)	−0.104 (0.075)	0.094 (0.081)
Paid certificate	0.334** (0.15)	0.327** (0.164)	0.366** (0.155)
Grading	−0.044 (0.083)	−0.087 (0.091)	0.023 (0.091)
French	0.304* (0.161)	0.59*** (0.187)	−0.114 (0.167)
Tenured teacher	0.096 (0.083)	0.093 (0.088)	0.14 (0.093)
MOOC experience	0.025 (0.079)	0.005 (0.086)	0.055 (0.086)
H-index	−0.01*** (0.003)	−0.01*** (0.003)	−0.011*** (0.003)
University	−0.098 (0.113)	−0.158 (0.119)	0.009 (0.13)
Grande Ecole	−0.04 (0.105)	−0.117 (0.111)	0.092 (0.127)
Foreign Institution	0.254 (0.156)	0.322** (0.161)	0.11 (0.167)
Ranking	0.013 (0.013)	0.026* (0.013)	−0.007 (0.014)

Continued on next page. . .

Results (1): The drivers of total, national and international enrollment (cont.)

Twitter	0.177***	0.174**	0.227***
	(0.068)	(0.074)	(0.077)
Media coverage	0.235***	0.281***	0.171***
	(0.041)	(0.045)	(0.047)
Monthly time trend	−0.001	−0.001	0
	(0.003)	(0.003)	(0.003)
Constant	7.966***	7.342***	6.984***
	(0.241)	(0.273)	(0.254)
N	284	284	284
Adj. R^2	0.272	0.326	0.167

Statistical significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Heteroskedasticity robust standard-errors in brackets

Concerning course-related characteristics, we first observe that the theme of the course does not have an influence on students' enrollment. Courses requiring more *work* hours from students in order to fulfill the course requirements tend to attract more students. The likely reason is that these courses provide more content and are better organized compared to others due to their length. Having the option to obtain a *paid certificate* is valued by students, as they attract on average 33% more students. If this creates few additional burdens to the MOOC provider, this option can be a fruitful way to attract more students and create revenue sources. Even if the coefficient is only marginally significant, teaching a course in *French*, instead of another language, helps increase enrollment as well, as they attract 30% more students, all else equal. Hence, the platform predominantly cater students willing to follow courses in French.

Only one teacher-related factor has a significant influence on student enrollment: *H-index*. Surprisingly, it has a negative sign. While the distribution of enrollment pictured in Figure 1 highlights the presence of superstar courses, as already mentioned by Acemoglu, Laibson, and List, this result emphasizes the fact that the reward in enrollment numbers of online education is not captured by the same persons as the ones in the research context.

Institutional factors have no significant impact on enrollment. However, both *Twitter* and *media coverage* have a positive and significant influence on enrollment. These results point towards the importance of putting up a media campaign in parallel to setting up of the course as a way to alleviate potential informational asymmetries on the students' side.

Based on regression (2) and (3), it will now be possible to answer the second research question: Do national and international students differ in their enrollment decision? While the number of *work* hours of the course, *paid certificate*, *H-index* and the two communication factors influence significantly the two student populations, two factors tend to highlight differences in the preferences of national and international students. First, having the course taught in *French* drives more national students towards subscribing but this is not true for international students. Hence, international students are not attracted by the fact that the platform offers French courses. Second, we observe that *foreign* institutions attract on average more international students. One potential explanation is that courses cater students from the local and national environment of the institution sponsoring the course as they are likely to be better informed about their existence.

Interestingly as well, the model is better able to explain the decision of national than of international students, with respectively a R^2 of 0.33 and 0.17. Hence, other factors not accounted for are likely key drivers of international student enrollment. This is another sign that the two student crowds tend to differ in their preferences to enroll in a specific MOOC. There are two possible explanations behind these observations: (1) national and international students have different profiles and (2) there are different motives behind their enrollment decision. Unfortunately, more detailed data would be needed to make further claims on this topic, more precisely, data about the students' socioeconomic background and about their course completion.

One final point relates to the timing of the course launch, both concerning the other MOOCs being provided at the same time on the platform and the month when the course starts. The fact that the presence of other successful courses at the same time on the platform could impact enrollment relates to the existence of same-side externalities, also known as direct or within-group network effects. They can be negative, meaning that competition from additional courses organized at the same time drives down enrollment. Negative externalities are more likely in the presence of congestion, for example when users face a limited amount of time and have to give up an option when choosing another one. Compared with traditional courses, positive externalities are more likely online as the "consumption" of the course is more flexible, as it can be done on demand at the preferred time rather than at a predefined time slot [16]. Knowledge about these externalities is not only important for the course providers for their release decision, even if it is unlikely that they can anticipate the courses launched by other higher education institutions. It is also key for the MOOC platform and the MOOC market, as negative externalities contribute to facilitate the coexistence of several platforms and positive externalities create concentration forces leading to a limited number of platforms in the market. It is also more likely in this setting as we only look at the enrollment decision, which takes a very limited amount of time, compared with course completion for example. From regression (4) to (6) of Table 3, we consider three definitions of the number of courses: the total, the amount taught within the same theme and within the same discipline. # *of MOOCs* is only significant when considering the total amount of courses provided and it is positive. Based on this result, it is complex to precise further the shape of this externality, except by saying that they are less likely to be negative.

Table 3: Results (2): The drivers of total enrollment and the timing of the course launch

Dependent variable	(4)	(5)	(6)	(7)	(8)	(9)
Total enrollment				–0.005		
# of MOOCs	0.009**			(0.006)		
(total)	(0.004)				0.007	
# of MOOCs		0.023			(0.022)	
(theme)		(0.023)				–0.012
# of MOOCs			0.006			(0.012)
(discipline)			(0.01)			
Constant	7.852***	7.943***	7.946***	8.494***	9.321***	8.435***
	(0.249)	(0.249)	(0.251)	(0.296)	(0.26)	(0.249)
Monthly time trend	–0.004	–0.001	–0.001			
	(0.003)	(0.003)	(0.003)			
Month fixed effects	NO	NO	NO	YES	YES	YES
All indep. Variables	YES	YES	YES	YES	YES	YES
N	284	284	284	284	284	284
Adj R^2	0.279	0.275	0.272	0.429	0.427	0.43

Statistical significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Heteroskedasticity robust standard-errors in brackets

So far, we have considered a monthly time trend. In regression (7) to (9), we instead consider monthly binary variables, for the month when the course started. We observe that courses launched in April tend to attract significantly more students, and those in June and September significantly less students. In Figure 2, where the month fixed effects are pictured when national and international students are considered separately, we see that national and international students have similar preferences except in April, June and December.

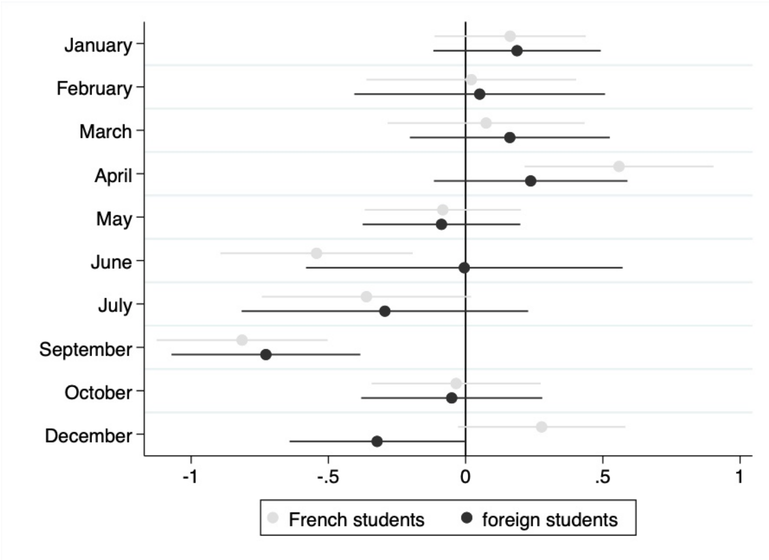

Figure 2: Month fixed effects for model (7) for national/international students

5 Conclusion

Taking advantage of the open access to students of online courses, without tuition fee nor a selection procedure to enroll, this paper analyzes what can explain the differences in enrollment numbers of online courses provided on a French MOOC platform. The results reveal that course-level factors are key, such as the amount of study time needed to fulfill the coursework, the instructional language and the possibility to pay to obtain a certificate of completion. The data analysis shows that communicating about the course on social media and in the traditional press helps decrease informational asymmetries and further on improves enrollment. We also observe several differences in the preferences of national and international students concerning the language of instruction and the starting time of the course. Finally, there is little evidence of the presence of same-side externalities between the courses launched on the FUN platform.

These results suggest two implications related to the scale-up of MOOC platforms and the timing of courses. First, if MOOC platforms want to attract more students, we have shown that one key factor is to be pro-active in the media. How precisely these media campaigns should be organized remains an open question. For example, recent data from the U.S.-based platform Coursera suggests the central role of marketing expenses, as 37% of their total revenues are concerning marketing and sales expenses [14]. However, it is unclear whether this approach should be pursued by the FUN platform as well to be able to scale-up further. Second, from our results concerning the timing of courses, we have that it is more important to consider when to launch a course than to think about whether other, potentially competing, online courses will be aired first at the same time.

Despite the relatively large sample of courses considered, any extrapolation should be done cautiously. The methodology used is suited to understand the key factors behind the course demand but is not adequate to make predictions. In addition, coefficients have to be interpreted as controlled correlations rather than causal impact. Finally, it is important to have in mind that enrollment is not by definition a learning outcome. However, it is still a key measure of the intention to learn, as this step is necessary to further pursue an online learning experience on the MOOC platform.

6 Acknowledgement

The author gratefully acknowledges Marco Gazel for his helpful comments and FUN for providing him access to the data.

References

[1] A. Abdulkadiroğlu and T. Sönmez. *School choice: A mechanism design approach. American economic*. 2003.

[2] D. Acemoglu, D. Laibson, and J. A. List. "Equalizing superstars: The Internet and the democratization of education". In: *American Economic Review Papers and Proceedings* 104.5 (2014), pages 523–27.

[3] W. T. Alpert, K. A. Couch, and O. R. Harmon. "A randomized assessment of online learning". In: *American Economic Review* 106.5 (2016), pages 378–82.

[4] P. Belleflamme and J. Jacqmin. "An economic appraisal of MOOC platforms: business models and impacts on higher education". In: *CESifo Economic Studies* 62.1 (2016), pages 148–169.

[5] Q. Chen, Z. D. Ozdemir, and Y. Liu. "The diffusion of internet- based distance education technology among US associate colleges". In: *Economics of Innovation and New Technology* 18.6 (2009), pages 587–605.

[6] P. Diver and I. Martinez. "MOOCs as a massive research laboratory: Opportunities and challenges". In: *Distance Education* 36.1 (2015), pages 5–25.

[7] R. G. Ehrenberg. "Econometric studies of higher education". In: *Journal of econometrics* 121 (2004), pages 1–2, 19–37.

[8] J. Goodman, J. Melkers, and A. Pallais. "Can online delivery increase access to education?" In: *Journal of Labor Economics* 37.1 (2019), pages 1–34.

[9] J. D. Hansen and J. Reich. "Democratizing education? Examining access and usage patterns in massive open online courses". In: *Science* 350.6265 (2015), pages 1245–1248.

[10] F. M. Hollands and D. Tirthali. *MOOCs: Expectations and Reality, Technical Report*. Teachers College, Columbia University, 2014.

[11] T. J. Kane. "Public intervention in post-secondary education". In: *Handbook of the Economics of Education* 2 (2006), pages 1369–1401.

[12] S. Marginson. "Higher education and public good". In: *Higher education quarterly* 65.4 (2011), pages 411–433.

[13] J. McKenzie. "The economics of movies: A literature survey". In: *Journal of Economic Surveys* 26.1 (2012), pages 42–70.

[14] L. McKenzie. "Coursera IPO filling reveals a company successfully monetizing MOOCs". In: *Inside Higher Education* (2021).

[15] J. C. Ortagus and L. Yang. "An examination of the influence of decreases in state appropriations on online enrollment at public universities". In: *Research in Higher Education* 59.7 (2018), pages 847–865.

[16] M. Peitz and M. Reisinger. "The economics of internet media". In: *Handbook of media economics*. Volume 1. North-Holland, 2015, pages 445–530.

[17] J. Reich. *Failure to Disrupt: Why technology alone can't transform education*. Boston, USA: Harvard University Press, 2020.

[18] J. A. Ruiperez-Valiente, S. Halawa, R. Slama, and J. Reich. "Using multi-platform learning analytics to compare regional and global MOOC learning in the Arab world". In: *Computers and Education* 146 (2020), page 103776.

[19] T. Tong and H. Li. "Demand for MOOC: an application of Big Data". In: *China Economic Review* 51 (2018), pages 194–207.