Why are some online courses more open than others?

Jacqmin, Julien

10 November 2018

Why are some online courses more open than others?

Julien Jacqmin[*]

November 9, 2018

Abstract

This paper studies the determinants of the intellectual property license of the Massive Open Online Courses (MOOCs) provided on a French platform. While they can opt for a traditional intellectual property right license, the vast majority chooses among creative commons licenses. Considering the ordering in terms of degree of openness of these licenses, we study the course-, teacher- and institution-related factors driving this decision. We observe that the field of study of the course and the characteristics of the institution providing the course play a role. We explain this result by the teachers' variable awareness and preferences with regard to open practices. We also find support for the idea that those willing to raise revenues by selling goods complementary to the MOOC (e.g. a textbook, certificate of completion or bundle of courses) are less likely to choose an open license.

Keywords: Massive Open Online Courses, intellectual property rights license, creative commons, open educational resources .

JEL codes: O34, I29, L17

1 Introduction

Massive Open Online Courses (MOOCs) first attracted media attention when two Stanford professors decided to put online their 'Introduction to artificial intelligence course'. In the wake of its huge success with more than 160,000 students enrolled, several online platforms have sought to spur similar initiatives, mostly from higher education institutions. To date, Coursera and edX are the most popular MOOC platforms, hosting respectively 33 and 14 million students, enrollment figures which are rarely heard in the higher education context.

Various efforts to take advantage of information technologies to provide online courses predate these initiatives. Launched in 2001, OpenCourseWare published via internet and for free the video captions of courses taught in various universities, including MIT. As of the mid-2000s, a first initiative also called MOOCs originated from the open educational resource movement. These MOOCs are now referred to as cMOOOcs in opposition to the more recently developed xMOOCS available on online platforms like Coursera or edX (Daniel [2012]). All these initiatives have in common that they use internet to provide cost- and selection-free education to students. Apart from their relative confidentiality in terms of student figures, a key distinctive feature between these two previous attempts to provide online courses and the more recent xMOOCs is that the content of the courses provided is protected by open-source licenses. These licenses, such as creative commons licenses, concern the rights to share, transform and redistribute creative works. In other words, they provide an explicit legal framework to construct innovations built up on the shoulders of previous works. The objective of this paper is to understand why online courses tend to differ in terms of openness, as defined by their copyright license.

In order to study this question, we analyze data from FUN, a French MOOC platform jointly initiated by the French government and French universities. Since its creation in 2013, the platform has attracted 3 million students enrolled in more than 400 courses, mostly provided by French higher education institutions. One key feature of the platform is that it is organized in a decentralized way, leaving a lot of freedom to the teachers authoring courses. For example, it is up to the pedagogical team behind the course to determine how to license its content. Hence, some courses are protected by intellectual property rights, as is the rule with other, mostly U.S.-based, xMOOC platforms. Other courses, however, are also protected by an 'open' license, and more precisely

by one of the creative commons licenses. We distinguish three categories of factors that can explain this heterogeneity: course-specific, professor-related and institutional factors

We find that science and health MOOCS tend to have a more open license. This is true also of courses originating from foreign institutions. Top-ranked higher education institutions, though, prefer to choose a relatively 'closed' license. We also observe that courses trying to raise money by selling a product complementary to the MOOC, e.g. a certificate of completion or textbook, tend to restrict the redistribution of the course content to non commercial purposes, which is made possible by creative commons licenses. We argue that these results can be explained by (1) the heterogeneity between academic disciplines of the peer recognition concerning open practices, (2) the greater 'taste for openness' of teachers from foreign institutions, and (3) the higher foregone potential earnings of choosing an open license when a textbook or a certificate of completion is sold to the students enrolling freely in the MOOC.

After introducing the concept of open educational resources in Section 2, we explain what open source licenses are, especially by discussing the literature that has looked at why innovators might choose a more or less open licensing scheme. Section 4 presents our data and methodology. The results of our empirical analysis are presented in Section 5. Section 6 concludes.

2   Open educational resources

According to the OECD (Hylen et al. [2012], P. 19), "Open Educational Resources (OER) are digital learning resources offered online (although sometimes in print) freely and openly to teachers, educators, students, and independent learners in order to be used, shared, combined, adapted, and expanded in teaching, learning and research. They include learning content, software tools to develop, use and distribute, and implementation resources such as open licenses". Initiatives made possible thanks to the emergence of information and communication technologies such as the MIT OpenCourseWare or the Khan Academy are all examples of such resources. It is the central element promoted by the open education movement, which can be seen as the equivalent to the education context of the, better known, open science movement.

OER have been established as a key priority in the "Renewed EU agenda for higher education" of the European Commission (European Commission

[2017]). Recently, the U.S. Department of Education has introduced an open licensing rule for educational resources produced by the winners of some of its competitive federal grants (Lieberman [2018]). From a normative perspective, two key benefits of OER make them desirable. First, their openness reduces the costs of education for its main users, i.e. students, by making them costless. Second, by choosing open licenses, they facilitate the improvement and personalization of the learning material by other teachers.

The commitment to non excludability created by the decision to use an open license makes it difficult, for its initiators, to appropriate direct monetary returns from OER as it is costless for its users. This issue is conceptually similar to the difficulty of privately providing public goods, especially as internet also reduces the extent of rivalry. Hence, higher education institutions as well as educators might refrain from investing in OER. These impediments have been studied in the education policy literature. Based on a survey of higher education representatives worldwide, Murphy [2013] argues that inhibitors are mainly at the institutional level and come from the lack of knowledge and committed resources. Analyzing the interviews of 30 higher education educators, Kaatrakoski et al. [2017] rather argue that the lack of recognition from investing in OER projects explains the limited uptake by teachers. Both these individual and institutional explanations are also acknowledged by Castana Munoz et al. [2016], who surveyed close to 200 administrators from various European higher education institutions.

This paper also tries to understand the factors accounting for the heterogeneous behavior of professors in open educational practices. We focus more specifically on the Massive Open Online Courses provided on a platform called FUN, which stands for France Université Numérique.[1] All the courses provided are free of charge for students and there is no requirement needed in order to participate. However, as opposed to the MOOCs provided on platforms like edX or Coursera, courses differ with respect to the license protecting what is explicitly allowed with the reuse, remix and redistribution of the pedagogical materials in other contexts. Apart from this different focus, our methodological approach also differs from the one used in the education policy literature. First, we rely on revealed preference data rather than stated preference data, as we look at the chosen license posted on the course webpage available on the FUN platform. Second, we use a multivariate regression approach rather than a bivariate analysis. Thanks to this approach, the

---

[1]For a further discussion of this institutional context, see Jacqmin [2018]

presence of omitted variable bias is reduced.

3   Open source licensing

The copyright legal framework grants exclusive rights to the creator of an original work concerning its usage by others. Acts of reproduction, adaptation and the distribution thereof can also be specified, but only on a case by case basis, negotiated by the two parties concerned. The goal of open source licenses is to enable innovators to explicitly specify these rights and duties for all users (Vanbrabant [2016]). Hence, they provide a legal framework to facilitate innovations built upon others by decreasing transactions costs between the parties concerned. Creative commons licenses are the most famous example of such public copyright licenses. Initiated in the early 2000s by Lawrence Lessig, they have been widely used by Youtube, Wikipedia and numerous open-source software.

Creative commons licenses differ based on the combinations of four distinct conditions enabling the innovator to manage the circulation of their innovation. Attribution (BY) requires to credit the author of the work. No derivative (ND) gives permission to copy and distribute copies of the work in its original format but not to create derivative works and remixes based on it. Share-alike (SA) is a clause requiring copies and derivative works to circulate under a license similar to that of the original work. Non-commercial (NC) gives permission to copy, display and redistribute the work and its derivatives but only for non-commercial purposes. In the context of online courses, videos of the professors discussing the topic covered tend to be the backbone of the course. BY requires, when reusing this content for another purpose, to stipulate the paternity of the video. ND means that it can be re-broadcast in another context but only in the same format. For example, the video cannot be translated or adapted by introducing interactive content, such as by integrating a quiz, in the video. When SA applies, it is possible to rework the video to adapt it to the context of the new usage, only if it is redistributed under the same condition as the original work. With NC, this can only be done for non commercial purposes. Hence, it is not possible to take the video as an input of an online learning program for which students would have to pay.

By mixing together (some of these) four conditions, we end up with seven different creative commons licenses.[2] As pictured in Figure 1 (Creative Com-

---

[2]Note that some of the conditions cannot be mixed together. For example, it is not possible
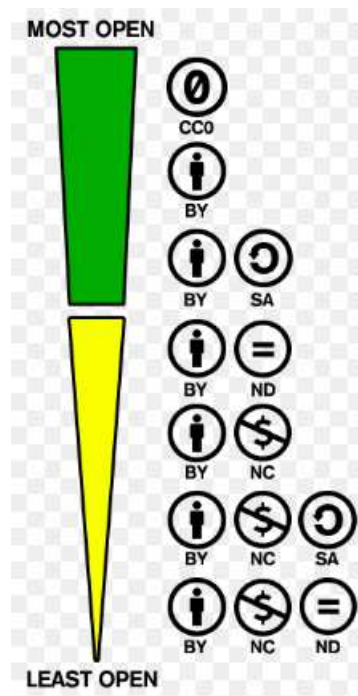
Figure 1: Ordinal ranking of Creative Commons licenses from least to most open (Creative Commons [2018])

mons [2018]), which comes from the official website of the creative commons licenses, these licenses can be ranked from the least to the most open ones. The license BY NC ND is the least open creative commons license, as it does not authorize the re-adaptation of the original work and the reuse for commercial purposes. This license is ranked right above "all rights reserved" in term of openness which only protects the work following the sole traditional intellectual property system and does not provide further guidelines about its reuse. The most open license CC0 sets the work in the public domain. In between, the other possible combinations of BY, SA, ND and NC, can be ranked depending on the ability to share the work, remix it and use it for commercial purposes.

The objective of this paper is to explain why we observe heterogeneous licensing practices among the Massive Open Online Courses provided on an online platform. While all these courses are freely accessible to students, they variously condition their reuse and distribution. Hence, compared to most of the literature on open source issues, we do not focus on the incentives to contribute/produce to the project but on the specific licensing decision that takes place once the project is completed (Lerner and Tirole [2002]).

---

to choose the condition ND and SA together

Economic theory can help us shed light on the licensing decision of MOOC teachers. Balancing the costs and benefits of the licenses available, they will choose the degree of openness that provides them the largest net benefit. The main cost incurred when choosing a more open license relates to the foregone potential earnings from selling close complements to the course (Lerner and Tirole [2005]). In our context, we can think of a certificate of completion, a textbook, a tutorial or a bundle of complementary online courses (Belleflamme and Jacqmin [2016]). This way the free online course is a strategy chosen within a freemium business model where the initial goal is to attract a large user base that is then tempted to pay for a different, but related, good or service. In the case where a relatively open license is chosen, especially if it does not include the conditions NC and ND, these revenues are more difficult to appropriate. Another innovator could freely enter the market by using the course content, adapt it and commercialize it. For example, as the courses are aired according to a specific timing on the MOOC platform, e.g. usually a timespan of a few months, another person could put up a website to air the course under a more flexible timing at almost no production cost and charge access to it. Hence, under this freemium business model, the course providers face potential additional competitive pressures when a more open license is chosen.

Teachers might also differ with respect to the non-monetary benefits they derive directly or indirectly from the chosen license. A pre-requirement lies in the heterogeneous awareness of the potential benefits that can be obtained from choosing an open license. The concept of openness in research practices is heterogeneously used and known across academic disciplines and institutions (Eger et al. [2015] and Zhu [2017]). This is likely correlated with their awareness of the opportunities created by open educational practices.

Teachers can first derive immediate intrinsic benefits from choosing an open license. They can have stronger preferences towards openness or some form of altruistic motives that impact their decision. In the context of research, as discussed in Roach and Sauermann [2010], the "taste for science", which includes the "taste for openness" is likely to differ between people working in the higher education system and in other sectors, due to the self-selection of workers with this characteristic into academia.

The benefits derived can also be indirect as the license choice can act as a signaling device. This signal can be valued both by his employer or peers. As theoretically discussed by Holmstrom [1999], the licensing choice can be influenced by these dynamic form of incentives. However, it is unclear whether

career concern motives do play a role as we have no example in mind of employers valuing the license chosen for a MOOC for future job prospects. However, social motives in the form of peer recognition are more likely to play a role. Hence, the collegiate reputation-based reward system present in research activities discussed a.o. in Dasgupta and David [1994]) can potentially explain why, despite the absence of direct private appropriability, some teachers choose very open creative commons licenses to protect their course content.

This paper is closely related to the literature focusing on the empirical determinants of licensing choices in academia, which uses mainly faculty-level data. While many papers have focused on explaining the patenting decision of faculty members (see for example Thursby et al. [2009]), recent works have looked at why the degree of openness of the inputs and outputs of academic research tends to be heterogeneous. These works aim at understanding the movement towards more open science. Analyzing a survey of, respectively, German and UK academics, Eger et al. [2015] and Zhu [2017] look at the explanations behind an open access publishing decision. They find that aspects such as the discipline or the personal characteristics of the faculty members are relevant. Andreoli-Versbach and Mueller-Langer [2014] study the characteristics of a selected sample of empirical economists to see why some voluntarily share their data while others do not. They observe that personal characteristics, such as age, academic tenure and attitudes towards the principles of open science as well as the mandatory data-disclosure policies set by journals are relevant. However, in contrast to this literature, a key feature of this work is that it studies an intellectual property rights decision related to a teaching rather than a research activity.

One peculiarity of our work is that we look at the degree of openness of the creative commons licenses chosen by the teachers of the courses provided on a MOOC platform. This decision is made after producing the course content, when organizing the host webpage of the course on the platform.[3] To our knowledge, Bazen et al. [2015] is the only other paper that has investigated the rich information available about the various dimensions of openness that lie behind creative commons licenses. For this purpose, they use data from Jamendo, a free online music platform, where artists can choose the licensing protection of their music using these licenses. After analyzing a survey of some of the artists active on the platform, one of their observations is that

---

[3]Note that there is no default option provided and that the platform refers to the website of the creative commons licenses for further information, remaining neutral with respect to the licenses available.

younger artists, not deriving income from their music, tend to choose more open licensing rights.

4 Data and Methodology

This research draws on web scrapped data of all the 410 courses provided on the FUN platform, up to September 2018. This information is completed by additional data collected from various internet sources. Hence, one key aspect of this approach is that revealed information is used, while most works on open practices use survey data. Descriptive statistics are shown in Table 1.

Table 1: Descriptive statistics (N=410)

	Mean	Standard deviation	Min	Max
Dependent variables:				
CC license (rank)	1.371	1.053	0	6
ND	0.565	0.496	0	1
SA	0.268	0.444	0	1
NC	0.846	0.361	0	1
Course-related independent variables:				
Business/law	0.246	0.43	0	1
Science	0.373	0.484	0	1
Health	0.115	0.319	0	1
Humanities/social sciences	0.266	0.442	0	1
Paid certificate	0.049	0.216	0	1
Book on sale	0.093	0.29	0	1
External financing	0.371	0.484	0	1
Grading	0.224	0.418	0	1
Teacher-related independent variables:				
# of teachers	4.485	3.937	1	24
H-index	8.185	12.543	0	82
Male	0.693	0.462	0	1
PastMOOCs	0.122	0.328	0	1
Lecturer	0.341	0.475	0	1
Assistant professor	0.215	0.411	0	1
Professor	0.443	0.497	0	1
Age	13.463	12.167	0	44
Webpage	0.307	0.462	0	1
Institution-related independent variables:				
Foreign institution	0.051	0.221	0	1
Grande école	0.336	0.473	0	1
University	0.48	0.5	0	1
Ranking (log)	2.034	2.869	0	6.633

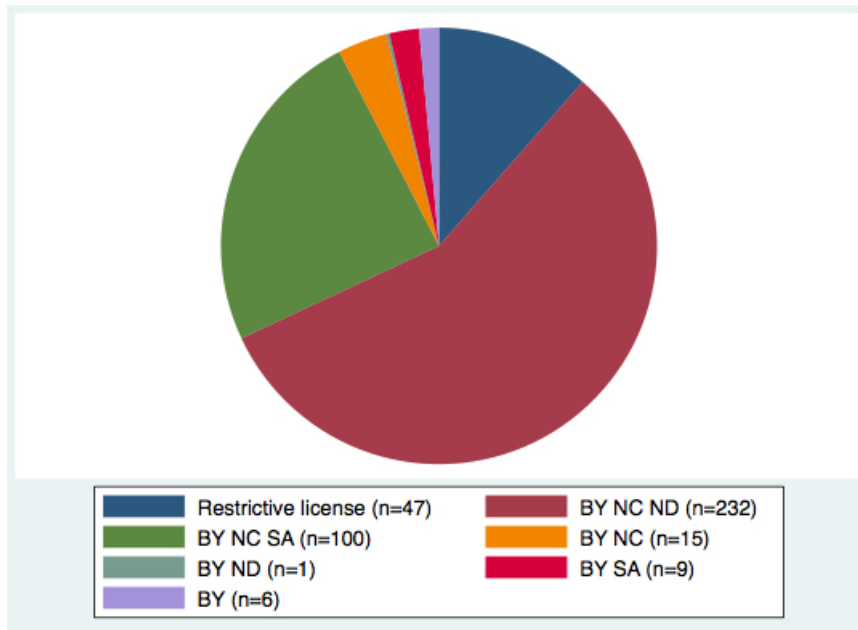Our main dependent variable *CC license* is the degree of openness of the

Figure 2: Course content license choice on the FUN MOOC platform

license chosen for the course content by its teacher. It is an integer between 0 and 6, which mirrors the ordinal ranking of the different licenses from the least to the most open, pictured in Figure 1. The most restrictive license "all rights reserved" receives the score 0. From 1 to 6, we have the different categories of the creative commons licenses, as no course has chosen the most open license CC0. As shown in Figure 2, about 10% of the courses have not chosen a creative commons license. A majority of the courses have set the most restrictive creative commons license by choosing CC BY NC ND. The remaining third of the courses have chosen more open creative commons licenses. We further analyze separately the determinants each of the three conditions: *ND*, *SA* and *NC*.

Our independent variables are grouped into 3 separate categories: Course-, teacher- and institution-related factors.[4] Note that a yearly time trend is also considered in each of our regressions. The first course-related factors concern the discipline in which the course is taught. There are 4 categories (*business/law*, *science*, *health* and *humanities/social sciences*, the last one being our reference category) which are each represented by a dummy variable. *Paid certificate* and *book on sale* are dummies whether or not complements to the

---

[4]Various other variables have been considered e.g. the number of classes taught, the estimated effort required for students to complete the course, the language of instruction (not all courses are taught in French), or the presence of prerequisites. Neither tend to play a role and, their inclusion does not affect the quality of our results.

course are advertised on the course webpage. Two of them are considered: a certificate of completion of the course and a textbook related to the course being taught. This is the case for, respectively, 5% and 10% of the courses. *External financing* is a dummy equal to one if it is mentioned that the initiators of the course have received funding from external sources, such as regional governments or private enterprises. *Grading* is a dummy equal to one if the course grading system is based on more than multiple choices. Using essays and problem sets (graded automatically or by fellow students) or valuing the participation to forum discussions can be seen as a proxy of the teachers' pedagogical investment.

As for the teacher-related variable, we first control for the number of teachers of the course with *# of teachers*. For the other teacher-related factors, we only take into account the main teacher of the course, identified as the person in charge of it. If not explicitly assigned, the one cited on top of the list is taken and, if this list is in alphabetical order, we take the first appearing in the trailer of the MOOC. *H-index* is the h-index of the main teacher as found on Scopus. We include a dummy for *male*. *PastMOOCs* is a dummy equal to one if the teacher has already taught a MOOC on the FUN platform. We also control for the academic rank. Three categories of academic positions are considered, each represented by a dummy: *lecturer* (for non-tenured teachers), *associate professor* (for tenured teachers with the lowest tenured rank position, also known as maître de conférence) and *full professor* (for teachers with a senior/full professor position, our reference group). As information about the age of the main teacher is not always available, we use, as a proxy, the *age* of the first publication available on Scopus. A dummy whether or not the teacher has their own *webpage* is also included. This last variable can be seen as a proxy of the teacher's attitude towards the openness of other academic information related to their research and other teaching activities.

Finally, we also consider institutional factors.[5] First, as 5% of the MOOCs are taught by teachers affiliated to foreign institutions, we consider the dummy *foreign institution*. We also take into account a dummy *grande école* to control whether the course is provided by a French public higher education institution allowed to select students at the entrance. If not by such an institution,

---

[5]Online courses are provided by 90 different institutions. One issue relates to whether or not these institutions have provided licensing guidelines to their teachers. Although we do not have representative information about this, our data suggests that this is the case only in a minority of places. From the institutions that have put up more than three MOOCs, only 20% of them have courses with an invariable license choice. Hence, this variation among providers is a sign that binding guidelines are far from being the norm in our setting.

we control whether it is provided by an *university*. Finally, we control for the Shanghai *ranking* of the institution providing the course. We take the 2017 ranking and, to facilitate the interpretation of the sign of the parameter, we subtract the rank of the institution from 800, which corresponds to the last rank. We also take the log of ranking.

Due to the ordinal nature of the dependent variable *CC license*, we estimate ordered probit regressions. Note that an ordered probit model only takes the rank into consideration while the value assigned to the dependent variable is meaningless. As some teachers teach several classes, we cluster standard errors at the professor level. We also analyze the three creative commons conditions separately, using a probit model. We focus on the sign of the coefficient rather than on the size, as their interpretation is far from trivial with this non-linear approach.

5 Empirical analysis

The main results are presented in Table 2, where each group of control variables are introduced step-by-step. Focusing on regression (3), we first observe that the discipline in which the course is taught plays a significant role. *Science* and *health* MOOCs tend to pick a license that is more open than those taught in *business/law* and *humanities/social sciences*.[6] This result is reminiscent of those of Migheli and Ramello [2013], Eger et al. [2015] and Zhu [2017] in the research context. Science and health scholars also tend to be more proactive with respect to open access publishing practices, e.g. by publishing in open access journals or by using self-archives/repositories. A first theoretical explanation behind this result is linked to the idea that, teachers of these disciplines are more aware of and tend to prefer these open licenses for their MOOC. A second plausible explanation, bound to the first, relates to the greater peer recognition received from choosing a more open creative commons licenses in these disciplines, which suggests that open licenses there tend to be the social norm.

---

[6]Specifying ten disciplines instead of four provides similar results, with MOOCs offered in the humanities, law, business and social sciences being less open and those offered in applied science, computer science and health being more so.

Table 2: Creative commons license of the MOOC content

Ordered probit model	(1)	(2)	(3)
Dependent variable	CC license	CC license	CC license
Business/law	-0.232	-0.263	-0.223
	(0.17)	(0.178)	(0.188)
Science	0.306*	0.373**	0.393**
	(0.166)	(0.179)	(0.187)
Health	0.256	0.594**	0.679**
	(0.239)	(0.271)	(0.267)
Book on sale	0.046	0.103	0.16
	(0.188)	(0.204)	(0.21)
Paid certificate	0.033	0.036	-0.098
	(0.145)	(0.157)	(0.17)
External financing	0.095	0.074	0.078
	(0.136)	(0.142)	(0.141)
Grading	0.331**	0.338**	0.336**
	(0.158)	(0.158)	(0.163)
# of teachers		0.005	0.012
		(0.019)	(0.019)
H-index		-0.01	-0.008
		(0.008)	(0.008)
Male		0.191	0.229
		(0.141)	(0.143)
PastMOOCs		0.279	0.303*
		(0.187)	(0.184)
Age		-0.002	-0.003
		(0.009)	(0.009)
Lecturer		0.229	0.235
		(0.178)	(0.183)
Assistant Professor		0.274*	0.289*
		(0.158)	(0.159)
Webpage		0.067	0.119
		(0.162)	(0.163)
Foreign Institution			0.632**
			(0.26)
Grande école			0.17
			(0.175)
University			0.181
			(0.196)
Ranking (log)			-0.057**
			(0.026)
Yearly trend	0.021	0	-0.002
	(0.041)	(0.044)	(0.045)
Cut1	-0.993***	-0.79***	-0.63*
	(0.195)	(0.289)	(0.332)
Cut2	0.735***	0.977***	1.156***
	(0.183)	(0.282)	(0.332)
Cut3	1.74***	2.009***	2.219***
	(0.182)	(0.279)	(0.333)
Cut4	2.084***	2.365***	2.589***
	(0.209)	(0.287)	(0.34)
Cut5	2.115***	2.398***	2.622***
	(0.21)	(0.285)	(0.339)
Cut6	2.526***	2.836***	3.073***
	(0.254)	(0.306)	(0.338)
N	410	410	410
Pseudo R^2	0.025	0.043	0.055

Statistical significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The variable *grading* also plays a positive and significant role in explaining the degree of openness of the license chosen. Using more complex mechanisms to evaluate the participation of students, especially using peer-graded assessments, is in line with what the first waves of MOOCs did. These so-called cMOOCs, which emerged from the early open educational resource movement fostered a student-centered pedagogical approach and the use of open licenses, such as creative commons licenses. Hence, we can expect that a teacher using a more elaborate grading system than just multiple choice questions to assess students has a greater awareness/preference for more open creative commons licenses.

Foreign institutions also tend to choose more open licenses. One explanation for this result lies in the fact that French institutions choose to provide MOOCs on the FUN platform by default as it is an initiative of the French government. However, foreign institutions who prefer more openness are likely to choose FUN as a partnering platform, compared to other international platforms which put less emphasis on openness. For example, edX, Coursera, Udacity, Iversity and Futurelearn do not allow for a creative commons license for the courses offered on their platform, as content is by default protected by an "all rights reserved" license.

A final significant variable is *ranking (log)*. We have that better ranked higher education institutions tend to choose significantly more closed licenses. The most likely explanation for this result finds its roots in the fact that top and mostly anglophone higher education institutions, recycle some of their (former) MOOCs by bundling them together to provide a fully online program. A much-talked about example is the Master in computer science program offered by the Georgia Institute of Technology in partnership with the MOOC platform Udacity. This program was offered both on campus and online, for a reduced fee in the latter format (Goodman et al. [2019]). Micro Masters, also known as nanodegress, bundle together a handful of courses on a specific topic and are another recent initiative from platforms like Coursera or edX. Hence, top-ranked institutions prefer to choose a closed license to limit the foregone potential earnings from selling a related complement to the course, as otherwise other institutions could rebroadcast their pedagogical content via another channel for a limited production cost.

Note that two other variables have a mildly significant impact. The fact that teachers with more experience of MOOCs are more likely to value openness can explain why *pastMOOCs* has a positive coefficient. However, as we are not aware of cases where the choice of an open license is valued for future

job prospects, it is unclear why the dummy *assistant professor* is significant, at the 10% threshold, while those of *lecturer* and *age* are not. All the other variables do not seem to significantly impact the license choice.

As a further robustness check, we next look at the determinants of each creative commons condition separately: *ND*, *SA* and *NC*. These results are provided in Table 3. As they cannot be chosen together among the existing creative commons licenses, we analyze side-by-side regression (4) and (5), where *ND* and *SA* are (respectively) used as dependent variables. As ND, contrary to SA, does not promote the creation of content built on the one initiated by others, one can question whether choosing the ND condition can be considered an "open" practice (Mishra [2017]). Note as well that, with the licenses BY and BY NC, it is possible to choose neither of these two conditions. The main result is that institutional factors do play a significant role. Teachers at foreign institutions are more likely to choose the condition Share-Alike as opposed to No Derivative. As for our main results, this preference can be explained by the self-selection into the FUN platform of foreign institutions which prefer relatively more openness. In addition, we observe that teachers at public higher education institutions, universities or "grande écoles", also tend to opt for Share-Alike. Together, more than 80% of the MOOCs are provided by higher education institutions. The other are developed by research centers, the French public administration, government bodies or a few private schools. Here also, this result can be explained by the self-selection of teachers with a "taste for openness" into academic institutions, as also observed by Roach and Sauermann [2010] in the research context.

Finally, assistant professors are more likely to choose *SA*, even though we have not heard of examples where choosing this condition was seen as a valuable signal for promotion. We also have that MOOCs in the field of *business/law* are less likely to choose *SA* and that *lecturers* are also less likely to choose *ND*. However, these last two results are only significant at the 10% threshold.

Table 3: Determinants of each creative commons condition

Probit model	(4)	(5)	(6)
Dependent variable	ND	SA	NC
Business/law	0.202	-0.403*	-0.044
	(0.206)	(0.224)	(0.226)
Science	0.209	-0.147	0.52*
	(0.21)	(0.211)	(0.27)
Health	0.042	-0.271	0.393
	(0.288)	(0.296)	(0.339)
Book on sale	0.066	0.124	0.61**
	(0.264)	(0.266)	(0.308)
Paid certificate	0.512	-0.007	
	(0.358)	(0.345)	
External financing	0.085	-0.174	-0.039
	(0.156)	(0.17)	(0.198)
Grading	-0.27	0.277	-0.125
	(0.165)	(0.17)	(0.197)
# of teachers	-0.016	0.028	-0.012
	(0.021)	(0.021)	(0.031)
H-index	-0.004	0	-0.008
	(0.009)	(0.009)	(0.01)
Male	-0.046	-0.068	-0.142
	(0.162)	(0.171)	(0.209)
PastMOOCs	-0.06	0.009	-0.226
	(0.19)	(0.206)	(0.233)
Age	-0.013	0.011	-0.017
	(0.01)	(0.01)	(0.011)
Lecturer	-0.38*	0.34	-0.271
	(0.221)	(0.245)	(0.245)
Assistant Professor	-0.262	0.476**	0.041
	(0.211)	(0.222)	(0.251)
Webpage	-0.168	0	-0.192
	(0.176)	(0.181)	(0.217)
Foreign Institution	-0.585*	0.706**	0.21
	(0.327)	(0.325)	(0.376)
Grande école	-0.943***	0.857***	-0.118
	(0.243)	(0.258)	(0.32)
University	-1.215***	0.748***	-0.826***
	(0.253)	(0.278)	(0.317)
Ranking (log)	0.038	0.007	0.046
	(0.029)	(0.031)	(0.035)
Yearly trend	0.035	-0.052	0.082
	(0.053)	(0.055)	(0.06)
Constant	1.38***	-1.545***	1.635***
	(0.409)	(0.455)	(0.469)
N	410	410	410
Pseudo R^2	0.099	0.0941	0.096

Statistical significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In regression (6), we look at what explains the choice of not allowing the commercialization of the original content or its derivative. First, it is important to note that *paid certificate* has been dropped because it perfectly predicts the choice of *NC*. The twenty MOOCs offering to buy a certificate of completion have all chosen the condition Non Commercial. In addition, as *book on sale* has a positive coefficient, significantly different from zero, teachers selling a textbook related to the content of the MOOC are more likely to choose *NC*. These two results can be explained by the cost arising from choosing *NC* in the form of foregone potential earnings. Otherwise, anyone could take (part of) the MOOC, especially if allowed to remix and sell it as an online course (at a more flexible timing for instance) or in a textbook format for little additional effort. This would facilitate the emergence of competitors in the segment where the MOOC provider aims to raise money by taking advantage of the presence of a large number of students enrolling free of charge in the online course on the FUN platform. In addition, we have that *university* has a negative and significant coefficient, meaning that university scholars providing MOOCs are quite favorable to others using the content they create for commercial purposes. Finally, *science* MOOCs are more likely to choose *NC* but this is only marginally significant.

6 Conclusion

In this paper, we try to understand why the courses provided on a MOOC platform tend to differ with respect to their authors' choice of intellectual property licenses. We find that course-, teacher- and institution-related aspects tend to play a role. As we observe that science and health MOOCs distinguish themselves by having more open licenses, we believe that peer recognition does play a role in the licensing choice. As these disciplines are more used to open access practices in their research activities, awareness of the benefits of open licenses is another potential explanation. As foreign institutions do opt in for FUN as a host platform, they are far more likely to value its greater care for openness. This is why these institutions' teaching staff tend to choose more open creative commons licenses. As is the case for other highly-ranked higher education institutions active on other platforms, top institutions active on FUN are also likely to provide MOOCs as a means to finance other activities of theirs (like research or educational programs taught in front of a class). This is the idea of the freemium business model where these free

courses attract a large, partially captive, user base, which is offered to buy some complementary good, e.g. a textbook, a certificate of completion or a bundle of complementary courses. As top institutions have a higher outreach, the sale of such goods also makes for higher potential revenues. Choosing a less open creative commons license bares potential competitors from the reuse of their content. Hence, by creating this barrier of entry, they can protect their monopoly power on these complementary market segments.

Even though this analysis is positive in nature, our results can provide empirically-sound policy advices to increase the adoption of open practices by the mean of online courses. They concern MOOC platforms, MOOC providers as well as governmental authorities. A first set of policies aims to improve awareness of openness. Due to its intermediary role, MOOC platforms are the best suited to be proactive in informing them about open educational resources and the licenses available to support the production of open practices. For example, awareness-raising activities can be included in the training programs of future MOOC teachers. MOOC providers, and especially those in the higher education world, also have a role to play. Their interventions can target their teachers by better informing them ex-ante but also by promoting, ex-post, open educational practices, e.g. in the context of internal promotion decisions. Higher education institutions can train students at using open educational practices. However, standardized citation practices should first emerge for this purpose. First and foremost, initiatives from governmental authorities are crucial. Some forms of funding, like those supporting the development of pedagogical materials, can be provided with strings attached, relative to the degree of openness of the practices financed. Finally, as we have seen that some institutions use MOOCs as an indirect means of financing other activities by selling complementary goods, adapting the way higher education is financed by public sources is key in encouraging these practices.

We must mention one key shortcoming of our approach; this concern relates to our sample. While we consider all courses provided on the FUN platform, it is important to keep in mind that MOOCs, and online distance education at large, are still a rather recent phenomenon. Our sample of teachers is made of early-adopters of these new methods of teaching; for putting their course online in an easily accessible way. These pioneers already adopt very open practices as compared with the average teacher. Hence, further extrapolations must be made with caution.

References

ANDREOLI-VERSBACH, P. AND F. MUELLER-LANGER, "Open access to data: An ideal professed but not practised," *Research Policy* 43 (2014), 1621–1633.

BAZEN, S., L. BOUVARD AND J. ZIMMERMAN, "Musicians and the Creative Commons: A survey of artists on Jamendo," *Information Economics and Policy* 32 (2015), 65–76.

BELLEFLAMME, P. AND J. JACQMIN, "An economic appraisal of MOOC platforms: Business models and impacts on higher education," *CESifo Economic Studies* (2016), 148–169.

CASTANA MUNOZ, J., Y. PUNIE, A. INAMORATO DOS SANTOS, M. MITIC AND R. MORAIS, "How are higher education institutions dealing with Openness," Technical Report, JRC science for policy report, 2016.

CREATIVE COMMONS, "www.creativecommons.org, last checked: 22 October 2018," (2018).

DANIEL, J., "Making sense of MOOCs: Musings in a maze of myth, paradox and possibility," *Journal of interactive Media in education* 3 (2012).

DASGUPTA, P. AND P. DAVID, "Toward a new economics of science," *Research Policy* 23 (1994), 487–521.

EGER, T., M. SCHEUFEN AND MEIERRIEKS, "The determinants of open access publishing: survey evidence from Germany," *European Journal of law and economics* 39 (2015), 475–503.

EUROPEAN COMMISSION, "on a renewed EU agenda for higher education," Communication from the Commission to the European Parliamant, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2017) 247 final, European Commission, 2017.

GOODMAN, J., J. MELKERS AND A. PALLAIS, "Can online delivery increase access to education?," *Journal of Labor Economics* 37 (2019).

HOLMSTROM, B., "Managerial incentive problems: A dynamic perspective," *Review of Economic Studies* 66 (1999), 169–182.

HYLEN, J., D. VAN DAMME, F. MULDER AND S. D ANTONI, "Open educational resources: Analysis of responses to the OECD country questionnaire," *OECD Publishing* (2012).

JACQMIN, J., "Providing MOOCs: A fun way to enroll students?," *Information Economics and Policy* (2018).

KAATRAKOSKI, H., A. LITTLEJOHN AND N. HOOD, "Learning challenges in higher education: An analysis of contradictions within Open Educational Practice," *Higher education* 74 (2017), 599–615.

LERNER, J. AND J. TIROLE, "Some simple economics of open source," *Journal of industrial economics* 50 (2002), p. 197–234.

———, "The Scope of Open Source Licensing," *Journal of Law, Economics, & Organization* 21 (2005), 20–56.

LIEBERMAN, M., *Feds Come Around to OER – Slowly* (Inside Higher Ed, 2018).

MIGHELI, M. AND G. B. RAMELLO, "Open access, social norms and publication choice," *European Journal of law and economics* 35 (2013), 149–167.

MISHRA, S., "Open educational resources: removing barriers from within," *Distance education* 38 (2017), 369–380.

MURPHY, A., "Open educational practices in higher education: Institutional adoption and challenges," *Distance education* 34 (2013), 201–217.

ROACH, M. AND H. SAUERMANN, "A taste for science? PhD scientists academic orientation and self-selection into research careers in industry," *Research Policy* 39 (2010), 422–434.

THURSBY, J., A. FULLER AND M. THURSBY, "US faculty patenting: Inside and outside the university," *Research Policy* 38 (2009), 14–25.

VANBRABANT, B., *La propriété intellectuelle - Nature juridique*, volume 1 (Larcier, 2016).

ZHU, Y., "Who support open access publishing? Gender, discipline, seniority and other factors associated with academics' OA practice," *Scientometrics* 111 (2017), 557–579.