The Journal of Higher Education

ISSN: 0022-1546 (Print) 1538-4640 (Online) Journal homepage: http://www.tandfonline.com/loi/uhej20

Persistence Patterns in Massive Open Online Courses (MOOCs)

Brent J. Evans, Rachel B. Baker & Thomas S. Dee

To cite this article: Brent J. Evans, Rachel B. Baker & Thomas S. Dee (2016) Persistence Patterns in Massive Open Online Courses (MOOCs), The Journal of Higher Education, 87:2, 206-242, DOI: 10.1080/00221546.2016.11777400

To link to this article: http://dx.doi.org/10.1080/00221546.2016.11777400

Published online: 31 Oct 2016.

Brent J. Evans
Rachel B. Baker
Thomas S. Dee

Persistence Patterns in Massive Open Online Courses (MOOCs)

Using a unique dataset of 44 Massive Open Online Courses (MOOCs), this article examines critical patterns of enrollment, engagement, persistence, and completion among students in online higher education. By leveraging fixed-effects specifications based on over 2.1 million student observations across more than 2,900 lectures, we analyzed engagement, persistence, and completion rates at the student, lecture, and course levels. We found compelling and consistent temporal patterns: across all courses, participation declines rapidly in the first week but subsequently flattens out in later weeks of the course. However, this decay is not entirely uniform. We also found that several student and lecture-specific traits were associated with student persistence and engagement. For example, the sequencing of a lecture within a batch of released videos as well as its title wording were related to student watching. We also saw consistent patterns in how student characteristics are associated with persistence and completion. Students were more likely to complete the course if they completed a pre-course survey or followed a quantitative track (as opposed to qualitative or auditing track) when available. These findings suggest potential course design changes that are likely to increase engagement, persistence, and completion in this important, new educational setting.

Keywords: student persistence, online higher education, MOOC

Brent J. Evans (corresponding author) is an Assistant Professor of Public Policy and Higher Education at Peabody College, Vanderbilt University; b.evans@vanderbilt.edu. Rachel B. Baker is an Assistant Professor of Educational Policy at University of California, Irvine's School of Education; rachelbb@uci.edu. Thomas S. Dee is a Professor of Education at the Stanford Graduate School of Education; tdee@stanford.edu.

The Journal of Higher Education, Vol. 87, No. 2 (March/April 2016)

MOOCs have become a critical topic of conversation and debate in major media outlets, in state and federal policy communities, and in departmental and faculty senate meetings. This rapid rise to prominence is due to the exceptional number of students enrolling and the elite institutions involved in their growth. Many observers in the popular media believe MOOCs have the potential to revolutionize higher education (Friedman, 2012; Webley, 2012) and some even believe we will rapidly approach the time when a professor is relegated to a "glorified teaching assistant" (Open letter from San Jose State Department of Philosophy, 2013, p. 2). Other journalists have discussed the tradeoffs between MOOCs' potential to provide free and easy access to higher learning to a wider audience and concerns about the possible unintended consequences of this new endeavor (Kim, 2012). Many perspectives exist, although few are grounded in data. This article presents novel evidence on the patterns of student engagement and persistence by examining data from the more than 2 million students who registered in a large and diverse array of 44 Coursera MOOCs.

The growing popularity of MOOCs is evident in the millions (e.g. over thirteen million on Coursera in spring 2015) of students across the globe who have registered for the courses, in the growing number of courses offered (e.g. over 1000 on Coursera and over 575 on edX as of summer 2015), and in their breadth of subject areas. MOOCs are distinct from most other forms of online higher education in that they are free, simultaneously reach tens of thousands of students, and have support from top tier institutions which grants them an air of legitimacy that online courses have never previously achieved. Both Coursera and Udacity are the brainchildren of Stanford University faculty, and edX began as a partnership between Harvard University and MIT. More than one hundred universities across the world now collaborate to offer courses on these platforms.

As MOOCs have risen in prominence, scale, and scope, there has also been limited, but widely publicized, descriptive evidence that surprisingly large numbers of registrants fail to finish these courses. Describing patterns of student behavior in MOOCs requires a different vocabulary and framework than examining student behavior in traditional classes. We defined three main terms as follows: engagement refers to any instance when a student interacts with the course (in this article, downloading or watching any course lecture); persistence is prolonged engagement—watching a number of lecture videos over a number of weeks; and completion is defined as engagement with the course until the end of the course—watching lecture videos through the last week or earning a certificate.

A more systematic and large-scale examination of student engagement and persistence in MOOCS is important for three reasons. First, much of the prior literature relied on student survey data with extremely low response rates (approximately 5%) while we relied on the complete universe of students. Second, there may be simple course and lecture design features that lead to increased student engagement, persistence, and completion (thus leading to greater learning). There have been attempts to improve design features through student surveys, but again, response rates are low (Johanes & Lagerstrom, 2014). If we better understand how students respond to these design features, MOOC instructors and platforms can implement them at low cost with positive effects on student learning. Third, many optimists believe that MOOCs can provide a successful pathway to a college degree by offering traditional college credit through MOOC platforms. As news outlets report, several such efforts are underway: Colorado State University—Global Campus offers transfer credit for a Udacity computer science course (Mangan, 2012), the American Council on Education supports several Coursera MOOCs for college credit (Lederman, 2013), San Jose State University partnered with Udacity to offer introductory and developmental math classes in MOOC format for credit (Kolowich, 2013), and Udacity has partnered with AT&T and Georgia Tech to offer an entire master's degree program in computer science for only $6,600 (Chafkin, 2013; Lewin, 2013). If the trend to expand MOOC credit continues, examining engagement, persistence, and completion in this modality is imperative.

The article's main research question is: What factors at the course, lecture, and student levels best predict in-course engagement, persistence, and completion? We answered this question by employing several econometric specifications to analyze an exceptionally large dataset with over 2.1 million student level observations across more than 2,900 lectures in 44 courses. By employing fixed-effects techniques on panel data, we controlled for many unobserved differences across courses and time, and we identified significant effects of course and lecture features on student engagement and persistence. Throughout these analyses, we examined multiple definitions of course persistence and success, accounting for the fact that students have different end goals (e.g. learning a particular topic, watching all course videos, or earning the certificate of completion offered in most Coursera courses).

We found patterns of engagement, persistence, and completion that fall into five broad areas. First, several course features are predictive of patterns of student engagement and persistence. For example, subsequent offerings of a course have lower rates of completion than the original offering, and courses with prerequisites have lower rates of certification among students with demonstrated engagement and persistence. Second, temporal patterns are very strong and nearly universal. Across all courses, participation falls throughout the course in a manner similar to exponential decay. Third, lecture-level design matters nonetheless. Specific words in lecture titles are significantly associated with levels of student engagement, and students watch the first lecture released in a particular batch more than any other lecture in that batch regardless of its length. Fourth, early, significant engagement is the strongest predictor of completion. For example, students who completed a pre-course survey were roughly three times more likely to earn a certificate than students who did not in one STEM MOOC. Finally, students who are motivated to enroll in a MOOC by the course's connection to a prestigious university are more likely to persist.

Our findings lead us to suggest several design features that course designers and instructors can put to immediate use to improve engagement and persistence. In contrast to most of the extant academic literature on MOOCs, which is focused primarily on learning analytics and describing MOOC users' demographics, our study addresses broader educational and policy issues surrounding the ability of MOOCs to engage students and provide a viable pathway to credit and degree attainment. As far as we are aware, this article is the first to predict engagement and persistence using course, lecture, and student characteristics across one of the largest MOOC data sets in the literature. These descriptive and correlational findings can provide an important foundation for future research in MOOCs, and many instructors and education researchers can use these patterns of behavior and persistence predictors as guidelines for designing interventions and improving the curriculum to increase course engagement, persistence, and completion.

Persistence Theory

Much as DeBoer, Ho, Stump, and Breslow (2014) argue that many traditional conceptualizations of variables within education must be rethought when applied to MOOCs, we argue that the previously held conceptions about persistence in higher education must be adjusted for the MOOC context. For example, persistence in higher education typically focuses on persistence at the semester, year, or degree level. However, MOOCs are structured around individual, disparate courses that are offered on a rolling basis and may be combined with other courses on other platforms offered by different universities.[1] Therefore, we focused on examining within-course persistence, as this perspective leads to a more nuanced view of how students engage with higher education, whether they are piecing together a program or simply engaging in a one-off course. We applied Tinto's academic integration theory of persistence to explore engagement and persistence within individual MOOCs.

Persistence Within an Online Course: Applying Tinto's Theory to Online Learning

Tinto's classic theory of student academic and social integration and college persistence (1975, 1993, 1998) was developed to explain longitudinal student retention within a degree program within a traditional institution of higher education. However, we believe this theory can be adapted to apply to student retention within a specific online course. Throughout this section we use Tinto's original language of "integration" which we view as interchangeable to our construct of "engagement."

Tinto's theory of student persistence in higher education (1993) proposed that student background characteristics and experiences combine with institutional characteristics to affect a student's decision to voluntarily dropout. Tinto asserted that there are two major components that make up students' experiences in college: social integration and academic integration. These factors are both seen as influencing students' goals and commitment to the institution. The original model was developed for and applies best to traditional (full-time, direct from high school) students at residential colleges, but even Tinto (1998) acknowledged that the form and experience of integration varies across educational settings. Scholars have worked to adapt this framework to apply to other settings such as distance education (Sweet, 1986) community colleges (Karp, Hughes, & O'Gara, 2010–2011), and online education (Rovai, 2003; Willging & Johnson, 2009). We applied, with some adaptations adjusting for unique features of MOOCs, Tinto's theory in the new context of Massive Open Online Courses.

Applying Tinto's model to MOOCs offered some unique advantages. First, one of the fundamental tenants of Tinto's theory is that persistence is affected by institutional characteristics which affect students' integration with academics. Due to data limitations, most studies, particularly those that focus on traditional, brick and mortar higher education, have relatively few measures of course or lecture characteristics that might affect academic integration and thus persistence. Most studies in traditional higher education simply used GPA as a coarse proxy for academic integration due to data limitations. MOOC data, however, enabled the measurement of academic integration at the micro scale by observing whether each student watched every individual lecture across a wide range of courses. We were able to empirically test what was at the heart of Tinto's theory: which institutional level characteristics (in our case course and lecture level characteristics) affected student persistence.

Second, little of the prior work, including the work in online classes, has applied Tinto's model to the completion of an individual class rather than a full degree program. Although there are a few studies that examine dropout within online courses, they focus predominately on student characteristics and perceptions derived from survey data (Park & Choi, 2009; Sutton & Nora, 2008–09; Willging & Johnson, 2009). While focusing on student background and characteristics is consistent with Tinto's model, it ignores actual integration with the current academic experience.

Bernard and Amundsen (1989) argued that "[a]t the program level, individual course characteristics are likely to exert a minor influence on the decision to dropout. Within a particular course, issues like the structure and delivery of the content, and intended learning outcomes, may influence decisions to dropout as much as student characteristics and attitudes" (p. 31). Our study is one of the first to engage with this hypothesis empirically. Because we were interested in course-level decisions to dropout, we focused our analysis on course and lecture characteristics to observe which factors correlated with student persistence beyond student level characteristics.

We conceptualized academic integration within a MOOC as watching lecture videos. Because of the asynchronous learning environment, the lecture videos are the primary form of communicating content from the instructor to students, and lecture videos serve as the backbone of any course in the MOOC space. We thus used watching course lectures (our definition of engagement) as a fine grained and detailed measure of academic integration. Although Tinto argued social integration is also critical, we concentrated on academic integration in this study. We believe there is an opportunity for future research to study forum interactions as a form of social integration.

Prior Literature on MOOCs

Because MOOCs are such a new phenomenon in higher education, there is little empirical evidence on MOOCs upon which to draw. The extant literature on MOOCs generally focuses on either the demographic characteristics of MOOC users, descriptions of MOOCs and MOOC platforms, or learning analytic studies. This is changing quickly, however, and as new studies are beginning to provide more in depth analyses (see, for example, Ho et al., 2014). There is also a growing literature in computer science that employs data mining and machine learning techniques to explore MOOC behavior (see, for example, Ye et al., 2015).

Understanding who MOOC users are and where they live is a challenging endeavor. In order to keep barriers to entry as low as possible, most platforms collect virtually no information on course participants, so administrative data must be extensively supplemented with surveys and location data from internet protocol (IP) addresses.[2] Recent work demonstrates that MOOC users are concentrated in North America, India, and Europe, but that there is representation from across the globe (Ho et al., 2014; Liyanagunawardena, Williams, & Adams, 2013; Nesterko et al., 2013,). Survey data from Coursera and edX show that MOOC users tend to be employed, well educated, and young although considerable heterogeneity across courses exists (Christensen et al., 2013; Ho et al., 2014).

One of the formal analyses out of the learning analytics strand of research is an article by Kizilcec, Piech, and Schneider (2013). They used cluster analysis to determine four prototypical engagement patterns for learners in MOOCs: auditors, samplers, completers, and disengagers. They found that many MOOC users are merely exploring and have low levels of engagement early. Although useful for understanding participation patterns, their analysis only examined patterns in three computer-science MOOCs and did not examine how specific course, lecture, and student traits influenced engagement and persistence. Additionally, a few studies inspected student behaviors and characteristics related to dropout in an effort either to predict dropout as in Halawa, Greene, and Mitchell (2014) or to describe the types of students likely to dropout as in Kizilcec & Halawa (2015), but these studies did not examine course and lecture features.

The work mostly closely aligned with ours was that of Perna et al. (2014) who documented the progression of MOOC users in 16 first generation MOOCs on the Coursera platform. They demonstrated that users are sequentially driven and that a pattern of steep dropout in the initial weeks was consistent across courses. Their analysis began to identify milestones such as completing a quiz predicted course persistence, but it is not a correlational analysis examining how course, lecture, and student variables are related to persistence outcomes. In fact, they stated directly that "research to date provides few insights into how course characteristics contribute to variations in user outcomes" (p. 422). Our analysis directly answers that question.

Our contribution to the literature is two-fold. First, we provided descriptive statistics on course registration, engagement, persistence, and completion in one of the largest samples of MOOCs in the literature to date (44 courses, 2.1 million students). This sample included courses from multiple universities across a wide range of content areas and includes courses beyond their first offering. Second, we used the complete population of course registrants to examine relationships between the outcomes of student engagement and persistence and an array of course, lecture, and student level predictors. Using fixed effects models with panel data, we were able to control for a large number of unobservable characteristics to reduce bias in our estimates. Although many of the independent variables we employed were not the typical variables we see in traditional analyses in higher education, they provided important information about characteristics that are vital in online learning settings. Instructors, along with platform and course designers will be able to use these results to improve MOOC content delivery and student engagement and persistence.

Data and Methods

Coursera Data

The dataset for this descriptive and correlational analysis was comprised of administrative data from 44 MOOCs on the Coursera platform. The majority of the MOOCs were offered by Stanford University, but several courses from other American institutions of higher education are also represented. In order to take a class and watch lectures, students must generate an account on Coursera using an email address. Each time a student logged on to the Coursera platform to interact with course materials, their actions were tracked. These anonymous administrative data contained each student's course participation behavior including each time they watched or downloaded a lecture.[3] We also observed when they registered for the class, their final course grade, and whether they earned a completion certificate. In this way, these data are far more detailed and comprehensive than most educational data available.

While there are many important advantages to these data, there are significant limitations. Most MOOC platforms collect few student level variables such as demographic information and course aspirations and expectations. Although more do so now, few classes employed surveys in the initial stages of MOOC expansion to collect demographic and student intention data. Response rates from these surveys are typically very low (less than 5% in the one course for which we have data on a pre-course survey in our sample, and overall 4.3% response rate over 32 Coursera courses from the University of Pennsylvania (Christensen et al., 2013)), so using such data results in significant sample restrictions. We also did not have the ability to track students across multiple courses; in our data students received a unique identifier for each class for which they register. The dataset was from one MOOC platform and was comprised mostly of Stanford courses in STEM fields. We do not believe that students sort in any meaningful ways across platforms; however, students who enroll in STEM MOOCs may be different from students who take social sciences or humanities MOOCs.

In one of the STEM courses, we had access to the student level responses to a pre-course survey offered by the instructor. The survey asked students their goals for taking the course in addition to asking them to select a track of the course to follow. This particular course explicitly offered students the opportunity to follow one of three tracks; students could audit by simply watching the lecture videos, complete the qualitative track by taking the post lecture quizzes, or follow the quantitative track by completing problem sets. Both the qualitative and quantitative tracks were eligible for completion certificates. We used these data for our student level regression analysis discussed below with the caveat that the survey was designed by the instructor, not the researchers, and thus may be prone to concerns of reliability and validity, although this concern is minimized due to the straightforward nature of the survey questions.

Coursera captured IP (internet protocol) addresses from participants that logged onto the website to interact with course material. In order to describe the lecture watching patterns, we used IP address mapping data from Maxmind GeoIP and geographic information systems software to identify the location (latitude and longitude) of each student. Although the accuracy of the geo-location data varies by country (see http://www.maxmind.com/en/city_accuracy), we used these data to determine whether students were domestic or international with high accuracy. Coursera tracked all course participation in Unix time so we observed whether a student watched a lecture video within a specific time frame around a course email message from the instructor.

In addition to student level data, we also leveraged data on each course and each lecture within each course. These data included the time and message of all emails and announcements sent to students during the course, the length of the course in weeks, the title and length of each lecture, when the lecture was released, and whether the course was being offered for the first time.

Regression Analysis

To assess which observable characteristics best predict course persistence, we conducted regression analyses on engagement, persistence, and completion at three levels: course, lecture, and student. Each level of analysis lends itself to different persistence outcomes; hence we discuss both outcomes and predictors at each level below. Summary statistics on predictors at each of the three levels of analysis are presented in Table 1, and summary statistics on outcomes at each level are presented in Table 2.

Course-Level Measures of Persistence. At the course level, we predicted the average persistence of students using several course level predictors. There are multiple potential measures of persistence within MOOCs related to the different metrics of course participation (e.g. watching all videos, watching most of the videos, and/or earning a certificate) and who should be included in the analysis (e.g., all students who sign up for a course, all students who watch any video, etc.).[4] To account for the various possibilities, we defined course level persistence and completion in four ways.

The first two measures are the percent of students who registered for the course who watched at least 20% and 80% of the lecture videos for the course, respectively. The first metric provides a sense of whether students are exhibiting engaged and sustained interest in the course beyond watching only the first few lectures. Given the average length of courses in our sample is over 11 weeks, 20% into the course provides over two weeks for enrollment to stabilize due to late registrants and early dropouts. The 80% marker serves as a measure of the students who are engaged throughout the entire course but may not earn a certificate. We believe this level of engagement is valuable even in the absence of earning a certificate.[5] Our third and fourth measures, meant to assess completion, match measures used in studies of traditional education more closely. The third outcome measures the percentage of registered students who earn a completion certificate (e.g. completing all assignments with a minimum level of competence or completing the final assignment or quiz). The final measure of course level completion is the percentage of students who earned a certificate conditional on watching at least 20% of the lecture videos. This measure essentially excludes students who registered but never engaged with the course (a substantial number in each course).

We predicted each of these four outcomes using the following model:

$$Y_j = \alpha + X_j\beta + \varepsilon_j \qquad (1)$$

where j indexes courses and X_j is a vector of course-level characteristics including the number of students in the course, number of lectures, when the course was released relative to the earliest courses in

TABLE 1

Summary Statistics

Course-Level Regressions

# of students in class (in thousands)	48.430
	(27.765)
# of lectures	65.614
	(33.377)
When course was released (in weeks since first course release)	33.474
	(18.092)
Course requires prerequisite skills	0.205
	(0.408)
Second or higher offering of class	0.409
	(0.497)
Stanford class	0.864
	(0.347)
Avg. length of videos (in minutes)	13.591
	(6.502)
Length of course (in batches of videos)	11.65
	(11.373)
N	44

Lecture-Level Regressions

Within 8 hours of an email from the instructor	0.170
	(0.376)
Length of video (in minutes)	12.71
	(7.117)
First video of batch	0.144
	(0.351)
Lecture title includes word:	
"intro"	0.032
	(0.177)
"overview"	0.021
	(0.142)
"basic"	0.009
	(0.093)
"welcome"	0.003
	(0.055)
"summary"	0.012
	(0.107)
"review"	0.013
	(0.115)
"conclusion"	0.001
	(0.026)
"assignment"	0.008
	(0.092)
"problem set"	0.005
	(0.073)
"exercise"	0.003
	(0.055)
"optional"	0.043
	(0.203)
"example"	0.039
	(0.193)

Student-Level Regressions

	All 44 Classes	STEM Course	Pre-course Survey
Domestic student	0.191	0.193	0.224
	(0.393)	(0.394)	(0.417)
International student	0.455	0.483	0.519
	(0.498)	(0.500)	(0.500)
Missing country	0.355	0.324	0.258
	(0.479)	(0.468)	(0.437)
Student registered:			
> 4 weeks before course started	0.262	0.202	0.251
	(0.440)	(0.401)	(0.434)
3−4 weeks before course started	0.032	0.072	0.126
	(0.175)	(0.258)	(0.332)
2−3 weeks before course started	0.036	0.073	0.159
	(0.186)	(0.260)	(0.366)
1−2 weeks before course started	0.066	0.126	0.267
	(0.248)	(0.332)	(0.443)
In week before course started	0.156	0.128	0.193
	(0.363)	(0.334)	(0.395)
In first week of course	0.153	0.101	N/A
	(0.360)	(0.301)	
1−2 weeks after course started	0.067	0.075	N/A
	(0.249)	(0.263)	
2−3 weeks after course started	0.031	0.023	N/A
	(0.174)	(0.150)	
3−4 weeks after course started	0.022	0.022	
	(0.147)	(0.147)	N/A
4−5 weeks after course started	0.017	0.026	
	(0.129)	(0.159)	N/A
> 5 weeks after course started	0.159	0.152	
	(0.366)	(0.359)	N/A
gmail.com email address			0.500
			(0.500)

TABLE 1 *(continued)*
Summary Statistics

Course-Level Regressions		Lecture-Level Regressions		Student-Level Regressions			
					All 44 Classes	STEM Course	Pre-course Survey
		"advanced"	0.019	edu email address			0.028
			(0.138)				(0.164)
		"practice"	0.007	Completed pre-course survey			0.0768
			(0.084)				(0.266)
		N	2,935	*Of those that completed pre-course survey:*			
				Chose auditing track			0.226
							(0.418)
				Chose qualitative track			0.408
							(0.492)
				Chose quantitative track			0.369
							(0.483)
				Cited following reason as very or quite important:			
				Fascination with subject			0.666
							(0.472)
				Academic			0.085
							(0.279)
				Prestigious university			0.104
							(0.305)
				Fun			0.578
							(0.494)
				Curious about online classes			0.095
							(0.294)
				Job			0.076
							(0.266)
				Credential			0.045
							(0.208)
				N	2,130,907	~30,000	1,386

Note. Means for each variable are reported with standard deviations reported below in parentheses. Data in the last two columns of the right panel come from one introductory level science course. For this course, the research team had access to student emails for a subset of students (those who registered before the course began) and responses to the pre-course survey for those students that answered. We rounded the number of students in this course to hide the course's identity.

TABLE 2
Descriptive Statistics for Outcome Variables

	Means (*SD*)	*N*
Course-Level Outcomes		
Proportion of registrants who watched at least 20% of the lectures	0.216 (0.086)	42
Proportion of registrants who watch at least 80% of the lectures	0.102 (0.046)	42
Proportion of registrants who earn a certificate	0.055 (0.048)	40
Proportion of registrants who earn a certificate | watching > = 20% of the lectures	0.213 (0.106)	40
Lecture-Level Outcomes		
Percent of registrants who watched the lecture	0.160 (0.098)	2,935
Student-Level Outcomes (All Courses)		
Earned a certificate	0.053 (0.224)	2,130,907
# of videos watched	10.306 (21.158)	1,905,289
Student-Level Outcomes (STEM Course)		
Earned a certificate	0.064 (0.245)	~30,000
# of videos watched	16.488 (29.670)	~30,000
Student-Level Outcomes (Pre-course Survey)		
Earned a certificate	0.222 (0.415)	1,386
# of videos watched	35.025 (37.400)	1,386

Note. The number of students in the STEM course is rounded to hide the course's identity.

the dataset, an indicator for whether the course required any prerequisite skills, an indicator for whether the course had been offered before, an indicator if the course was offered through Stanford, the average length of the video lectures in the course, and the length of the course as measured by the number of "batches" of videos released, which were typically, but not universally, released weekly.[6]

Lecture-Level Measures of Student Behavior. We then analyzed engagement patterns at the lecture level by predicting the percentage of students who watched (either streamed or downloaded) each lecture in the course using lecture characteristics. We graphed the number of times each video was watched to demonstrate changing patterns of access over the course. We also conducted multivariate regression analyses with several fixed effects models (course fixed effects α_j and batch fixed effects δ_t) to predict the proportion of registrants in course *j* who watch each lecture *l:*

$$Y_{ljt} = \alpha_j + t_{lj} + X_{ljt}\,\boldsymbol{\beta} + e_{ljt} \tag{2}$$

$$Y_{ljt} = \alpha_j + \delta_t + X_{ljt}\,\boldsymbol{\beta} + e_{ljt} \tag{3}$$

The vector of lecture covariates, X_{ljt}, included an indicator for whether the lecture was released within 8 hours of an email sent to the class, the length of the video, an indicator for whether the lecture was the first video released in a batch, and fourteen indicator variables for whether the video title included a particular word identifying the content of the video. These indicators determined whether the title of the video lecture included words suggesting whether the video was, for example, introductory, an overview, a summary, related to assignments, optional, provided examples, or was advanced material.[7] These indicators were determined using a text search on each lecture video's title. In each model, α_j is a vector of course fixed effects.

To test whether and how temporal patterns are related to lecture watching, we modeled time differently across models. In equation (2), we included a linear time trend, t_{lj}, to account for how far into the class each lecture was. In a variant of equation (2), we accounted for the evident nonlinear pattern of decay by modeling the length into the course as an exponential decay function. In equation (3), we ran a more fully unrestrictive model in which we included "batch" fixed effects to control for the unobserved effects related to a particular time. In our most flexible model, we included both batch fixed effects as well as fixed effects for a lecture's sequence within a batch. We ran all four models for all lectures in all 44 courses.

Student-Level Measures of Student Behavior. We conducted our student level analysis on two samples. The first is the full sample of 44 courses, which is very large but for which we have very few predictors. For this full sample, we regressed two outcomes (earning a certificate and number of videos watched) on two student level observables (when students registered for the course, using a vector of week fixed effects (δ_t), and whether the student accessed the class from an international IP address). We index students with *i*.

$$Y_{ijt} = \alpha_j + \delta_t + X_{ijt}\,\boldsymbol{\beta} + e_{ijt} \tag{4}$$

To maximize the number of student level characteristics available to predict behavior at the student level, we also ran a focused analysis on one STEM course for which we have many additional variables, including email addresses and pre-course survey data. Gmail is the modal email address with over 10,000 of the 18,000 students for whom we have emails. It may also serve as a loose proxy for internet savvy. We chose to use .edu addresses because they identify a firm link with an institution of higher education. Although they may not all be currently enrolled students, having such a link suggests a high education level in the absence of demographic characteristics. We limited some of these analyses to students who registered before the course began (because we have email addresses only for these students) and only to students who completed the pre-course survey. For students that did respond to the pre-course survey, we included whether they intended to follow the auditing, qualitative, or quantitative track for the course. Survey respondents also indicated the importance of a variety of reasons for which they took the course, and we coded them as indicators for responding whether each reason was "very important" or "quite important."[8]

Fixed Effects. A significant advantage of fixed effects is that it controls for fixed characteristics even if they are unobserved by the researcher. Our lecture and student level regressions employed course fixed effects to control for all variables that are constant throughout the course across lectures and students. This includes all instructor characteristics, course structure, availability of forums, grading policies, and countless other course level variables. Batch fixed effects control for time in a similar manner; they control for all of the unobserved variables that are constant within each set of videos released together. Most importantly, this includes whether it is the first week of the course, second week of the course, etc. It also accounts for the fact that the instructor may have released more videos in one batch than another.

Fixed effects models identify the relationship between predictors and outcomes within each group as opposed to exploring the variation across groups. The results from course fixed effects used identifying variation within each course as opposed to the variation across courses. The same was true for batch fixed effects; results were identified off of variation within week instead of variation across weeks.

Our analysis was primarily a correlation analysis, and we were careful not to imply causation in our findings. Although fixed effects control for a host of fixed variables, there still could exist student unobserved variables that vary within course or that vary over time for which we could not account. However, we viewed student-level unobservables as an unlikely source of bias in our application. It is not clear one could reasonably worry that unobserved student traits that influence their persistence outcomes systematically varied within courses with the timing of lecture traits like longer lectures. This is particularly so once we controlled for "batch" fixed effects. We viewed the most likely source of omitted variables bias as coming from other lecture-specific traits. For example, if the lecture traits we observed (e.g. sequencing within a week, title wording, and length) are correlated with other persistence-relevant lecture traits, our estimates could be biased.

We may also be concerned with simultaneity issues in which students or instructors receive feedback and alter their behavior. Our reduced form estimates captured the overall effect of these simultaneities without distinguishing their direction. We acknowledge the potential for the existence of a dynamic pattern in which a characteristic such as lecture length or title may increase knowledge and engagement in a manner that influences persistence in a subsequent period, and we believe such an analysis might prove fruitful as an area of future research.

Another approach to analyzing these data would be to use a form of multilevel modeling. We viewed our fixed effects approach as one form of multilevel modeling in which the effects associated with courses and batches of lectures are fixed as opposed to random. Furthermore, multilevel modeling excels at parsing the variance between and within groups accounting for multiple levels, but our main research questions were not focused on dividing the variance. Instead, we were interested in examining the predictors of course engagement and persistence within course, which is exactly what fixed effects enable as they control for all fixed variables.

Results & Discussion

Persistence Patterns—Courses

In an online appendix, we provided details on registration, participation, and completion outcomes for each course (see Online Appendix text and Table A1 at the *JHE* Knowledge Bank for supplemental material: http://hdl.handle.net/1811/75345). We begin in Table 3, by presenting findings of equation (1): predicting four persistence outcomes using course level predictors. We lost two of the courses because they were self-study and had no defined length and another two that lacked the certificate outcome (four courses did not offer certificates, including the two self-study courses).

TABLE 3
Course Level Persistence Analysis

	(1)		(2)		(3)		(4)	
Outcome:	Prop. of registrants who watched at least 20% of the lectures		Prop. of registrants who watched at least 80% of the lectures		Prop. of registrants who earned a certificate		Prop. of registrants who earned a certificate | watching > = 20%	
# of students in class (in thousands)	0.000 (0.001)		0.000 (0.000)		0.000 (0.000)		0.001 (0.001)	
# of lectures	0.000 (0.001)		0.000 (0.000)		0.000 (0.000)		0.000 (0.001)	
When course was released (in weeks)	0.000 (0.001)		0.000 (0.000)		0.0008 (0.001)		0.001 (0.001)	
Course requires prerequisite skills	−0.010 (0.020)		−0.006 (0.010)		−0.025 (0.013)	+	−0.073 (0.030)	*
Second or higher offering of class	−0.093 (0.028)	**	−0.053 (0.015)	**	−0.051 (0.021)	*	−0.067 (0.041)	
Stanford class	0.076 (0.052)		0.034 (0.033)		0.047 (0.029)		0.054 (0.052)	
Avg. length of videos (in minutes)	−0.002 (0.002)		−0.001 (0.001)		−0.002 (0.001)		−0.003 (0.002)	
Length of course (in batches of videos)	−0.008 (0.002)	**	−0.004 (0.001)	***	−0.004 (0.001)	**	−0.006 (0.004)	
Intercept	0.579 (1.909)		0.563 (1.083)		−1.628 (1.513)		−1.984 (2.368)	
N	42		42		40		40	
Adjusted R^2	0.258		0.231		0.229		0.078	

Note. Robust standard errors in parentheses. In Models 1 and 2 our sample is 42 classes because two classes were self-paced and thus the length of the course in weeks is not a meaningful statistic. In Models 3 and 4 we include only the 40 classes in which students could earn a certificate.
$+p < 0.10$. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.

The two most consistently significant findings were that repeated courses, those not offered for the first time on Coursera, and longer courses (more batches of videos released) had notably lower levels of engagement, persistence, and completion. Nearly 10 percentage points fewer students watched at least one fifth of the videos in subsequent offerings of a class and approximately five percentage points fewer registrants earned certificates relative to courses offered the first time. Every extra batch of videos released was associated with nearly one percentage point fewer students watching at least 20% of videos and about one half percentage point fewer students earning certificates. While these point estimates were quite small, they were large relative to the outcomes means; about 22% of students watched at least 20% of lectures and under 6% earned a certificate.

Our analysis suggests that courses requiring prerequisite skills experienced rates of certificate earning three to seven percentage points lower than courses without prerequisites, controlling for other variables. Prerequisite skills were also associated with negative overall engagement, although neither result was statistically significant.

Several insignificant findings in this table are interesting. The number of students in the class, number of lectures, and average length of lectures in minutes had no statistically or practically significant relationship with any of the persistence or completion measures when controlling for other variables in the model.

These null findings were somewhat surprising. We expected that a larger concentration of students might promote a more active discussion forum that would lead to greater course engagement and increased persistence. We also expected that video length would be related to course engagement because one of the principles upon which MOOCs operate is that more concise videos teaching a shorter concept facilitates student learning. While we lacked a learning measure, we did not find that courses with shorter videos have increased rates of engagement, persistence, and completion. Although we could not measure whether students watched the entire video, we did capture whether students started streaming or downloading the video, and as the length is usually prominently stated in the video title, we assumed students would be sensitive to this characteristic and that longer videos may reduce engagement. We observed no evidence of this behavior.

Discussion & Design Implications—Courses

The finding that longer classes (as measured in the number of batches of videos released, a proxy for weeks) have lower rates of persistence and completion suggests changes in course design. As the average length of videos does not have a significant effect on student persistence and completion, this might imply that instructors should release fewer, longer lecture videos. However, as we do not have measures of within lecture attention (we can only measure if a student starts to watch or downloads a lecture) we cannot say whether students are getting all the content within a lecture. Still, these findings have implications for how instructors structure and release their lectures to optimize student persistence.

Being aware that subsequent offerings of a course have lower completion rates may prove useful to set expectations for instructors, institutions, and platforms, but it does not suggest any specific changes in course design. However, that prerequisites might deter engaged students from earning a certificate does have implications for course structure.

It is possible students without prerequisites are watching the lecture videos and declining to complete assignments because of their lack of preparation, thereby leading to reductions in certificate rates. A few courses offer multiple tracks within a single MOOC enabling students to choose their level of engagement (i.e. auditing track, qualitative track, and quantitative track). Courses with prerequisites may find it beneficial to explicitly implement such tracks to facilitate continued engagement with the material and learning even if the assignments in the full track are challenging due to the necessary prerequisite skills.

Additionally, professors of courses with prerequisites should provide advice on how students can fulfill those prerequisites. Ideally, they could refer students to other MOOCs that could be taken prior to enrolling in the course. Coursera has developed Specializations that provide such a sequence of courses, and Udacity's nanodegree program is similar, but they are currently limited in number. Further developing these sequences could lead to enhanced and continued student learning.

Persistence Patterns—Lectures

We now turn to describing the persistence patterns in more detail by examining the lecture level factors that predict students watching an individual lecture. In order to examine the drop-off of participation, we graphed the number of times each lecture video was streamed or downloaded in each course. We provide six examples in Figure 1; the y-axis

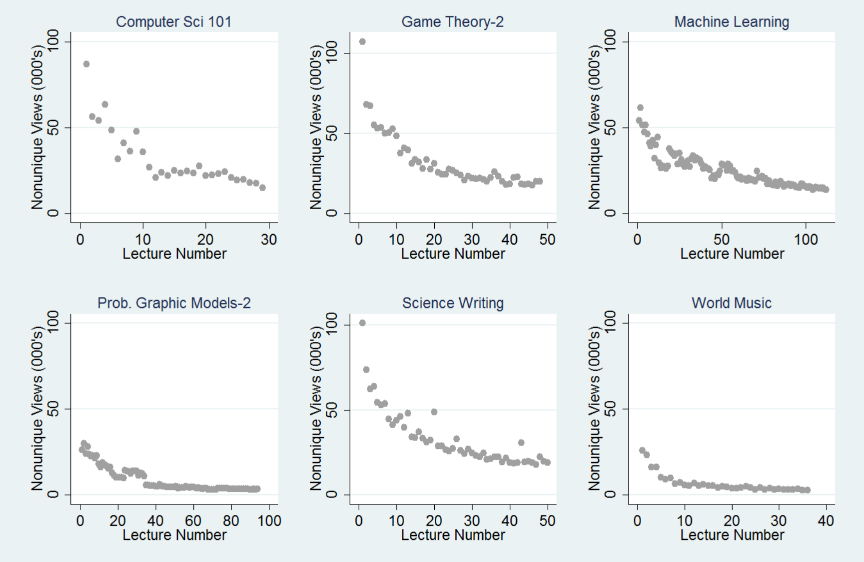


Figure 1. Nonunique Lecture Views Across 6 Courses

displays the number of times a lecture was watched, and the x-axis is the lecture's temporal position in the course. The pattern of lecture watching across courses is quite similar: high initial engagement that falls off rapidly and, in most instances, stabilizes at a low level. The rate of decline varies, but in all cases the greatest decline in participation occurs during the first ten lectures.

Despite this clear trend of rapid decline across all courses, there are noticeable outliers and discontinuities. Computer Science 101, Game Theory, and Science Writing have enormous drops between the first and second videos. Science Writing has two outliers in the middle and end of the course, and Machine Learning and Probabilistic Graphical Models have noticeable discontinuities.

We accounted for many of these outliers by examining unique views of each lecture. In Figure 2, a student who watched the same video multiple times is only counted as a single view. We have reduced the y-scale in Figure 2 to better show how removing repeat watches affects the watching patterns and to highlight design features. Removing the repeat watches makes the curve more smooth in all of the courses by eliminating outliers, most noticeably in Computer Science and Science Writing. To investigate why students watched several videos repeatedly, we examined the video content of the Science Writing course and discovered outliers were likely related to videos which discussed course assignments to which students referred back multiple times.

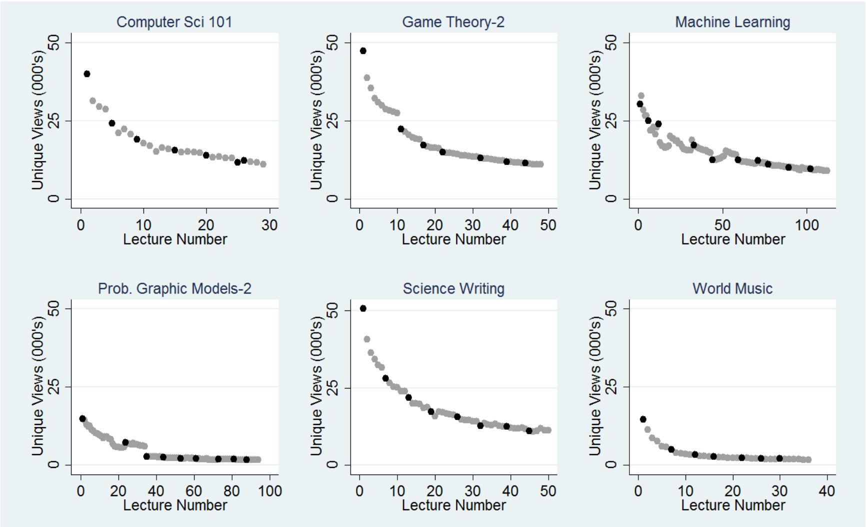


Figure 2. Unique Lecture Views Across 6 Courses 1st Video of Week Highlighted

Figure 2 also highlights the first lecture video in each week of the course in black. The first lecture of the week explains most of the discontinuities in lecture watching behavior, particularly in Machine Learning and Probabilistic Graphic Models, where a visible drop-off in course participation occurred at the transition between weeks.

Our regression analysis formalized these graphical findings. Table 4 presents regressions that predict the proportion of registrants in a given course that watched each lecture video at least once. Model (1) corresponds to equation (2) in which we model the drop-off of students over lectures linearly (% of way into course). Model (2) replicates Model (1) but employs an exponential decay function to model how far the lecture is into the course. Model (3) uses batch fixed effects to relax the functional parameterizations of this drop-off (equation (3), and model (4) adds video-within-batch fixed effects for a fully nonparametric model. Results are mostly consistent across models, although the exponential decay and nonparametric models explain more of the variance in lecture watching.

As we observed in Figures 1 and 2, how far a lecture is into the course was highly related to how many students watch the lecture. In the linear model, a video at the end of the course was viewed, on average, by almost 21 percentage points fewer students than a video at the beginning of the course, controlling for course and lecture characteristics. Given that about 44% of registrants watched the first video, this represents a 48% decline.

The regression analysis also confirmed what is obvious from Figure 2, that the highest percentage of students watches the first video within

TABLE 4

Lecture Level Analysis: Predicting the Percent of Registrants Who Watch the Lecture (Within a Class)

	Model (1)		Model (2)		Model (3)		Model (4)	
Prop. of way into course	−0.216	***						
	(0.010)							
e^(- Prop. of way into course)			0.368	***				
			(0.017)					
Within 8 hours of an email from the instructor	−0.004		−0.008		−0.013		−0.013	
	(0.010)		(0.008)		(0.011)		(0.010)	
Length of video (in minutes)	0.0008	*	0.0007	*	0.0005	*	0.0004	+
	(0.0003)		(0.0003)		(0.0002)		(0.0002)	
First video of batch	0.007	*	0.006	*	0.023	***	0.035	***
	(0.003)		(0.003)		(0.003)		(0.004)	

TABLE 4 *(continued)*

Lecture Level Analysis: Predicting the Percent of Registrants Who Watch the Lecture (Within a Class)

	Model (1)		Model (2)		Model (3)		Model (4)	
Video Title Includes the Word:								
"intro"	0.060	**	0.055	**	0.050	**	0.047	**
	(0.017)		(0.017)		(0.015)		(0.015)	
"overview"	0.043	***	0.037	***	0.029	**	0.023	**
	(0.011)		(0.010)		(0.009)		(0.008)	
"basic"	−0.001		–0.003		−0.007		−0.007	
	(0.006)		(0.006)		(0.008)		(0.008)	
"welcome"	0.191	***	0.166	***	0.124	***	0.116	***
	(0.023)		(0.023)		(0.020)		(0.019)	
"summary"	0.008		0.007		0.002		0.000	
	(0.011)		(0.009)		(0.004)		(0.005)	
"review"	−0.012		−0.017		−0.03	*	−0.028	*
	(0.013)		(0.013)		(0.012)		(0.011)	
"conclusion"	0.008		−0.002		−0.056	***	−0.053	***
	(0.014)		(0.015)		(0.011)		(0.011)	
"assignment"	–0.013		−0.012		0.000		−0.003	
	(0.012)		(0.011)		(0.014)		(0.013)	
"problem set"	−0.017	+	−0.017	+	−0.010		−0.016	
	(0.009)		(0.009)		(0.015)		(0.012)	
"exercise"	−0.026	+	−0.035	*	−0.058	**	−0.058	**
	(0.013)		(0.016)		(0.017)		(0.020)	
"optional"	−0.022	***	−0.021	***	−0.022	***	−0.020	***
	(0.005)		(0.004)		(0.005)		(0.005)	
"example"	0.002		0.001		0.003		0.003	
	(0.007)		(0.005)		(0.003)		(0.003)	
"advanced"	–0.007		−0.005		−0.008		−0.009	+
	(0.006)		(0.005)		(0.005)		(0.005)	
"practice"	0.012		0.010		0.004		0.008	
	(0.008)		(0.007)		(0.008)		(0.008)	
Course fixed effects	X		X		X		X	
Batch fixed effects					X		X	
Video-within-batch fixed effects							X	
Intercept	0.257		−0.081		0.309		0.302	
	(0.007)		(0.011)		(0.008)		(0.008)	
N	2935		2935		2935		2935	
Adjusted R^2	0.763		0.809		0.824		0.833	

Note. Standard errors clustered at the course level in parentheses. The dependent variable is the percent of all registrants who watch a lecture. Batch fixed effects are dummy variables that indicate the batch in which a video was released. They frequently but do not always align to calendar weeks. For example, if an instructor released videos on Monday and Thursday of the same week, the videos released on Monday would belong to one batch and the videos released on Thursday would be in another batch. "Video-within-batch" fixed effects indicate a video's position within a batch. In model 4 we created six dummy variables: dummies for each of the first through fifth videos of the week and a dummy to indicate sixth or higher.
$+p < 0.10$. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.

each batch. This is clearest when using batch fixed effects in Models (3) and (4). Within a batch, the first video lecture posted receives a highly significant 2–3.5 percentage points more viewers than other videos in the same week.

Surprisingly, the length of the video is statistically significantly related to an increase in the percent of students watching the video; however, the effect is extremely small. An increase of video length of 10 minutes is associated with less than a one percentage point increase in the proportion of students watching the video. This result may not indicate students are actively attracted to longer videos, but it suggests that video length, as typically displayed in the title, is not a deterrent to students beginning to watch or download it.

Instructors can place important signaling information in the title of videos, and our analysis demonstrated that specific words in the video titles are associated with different rates of watching. Videos labeled as introductory with words such as “intro,” “overview,” and “welcome” had much higher rates of watching. For example, videos labeled “intro” experience about a 5 to 6 percentage point increase in the number of registrants who watch the video. This is true even after controlling for batch and video-in-batch fixed effects, so these findings are not driven by introductory videos being watched more at the beginning of the course; it is true throughout the course.[9]

Students appear sensitive to other words as well. Videos labeled with summative words such as “review” and “conclusion” are watched by fewer students, on average, even after controlling for timing within the course. Three to five percentage points fewer registrants watch videos with these labels. As might be expected, “optional” videos were skipped by about two percentage points more students. Videos labeled “exercise” had the largest negative association with being watched, perhaps because of two groups of students: those who are fully engaged but do not need additional practice and those who are auditing and therefore not completing assignments. However, the lack of significant coefficients for “practice,” “assignment,” and “problem set” videos suggests both of those groups might be small. Finally, videos with “advanced” in their title were watched by about one percentage point fewer students than other videos.

Discussion & Design Implications—Lectures

Collectively, these findings suggest at least a subset of students pay attention to lecture titles and target specific videos to watch or ignore based on title information. Several findings are consistent with a group of students who are sporadically engaged or auditing the course.

Auditors may have been more likely to skip supplementary materials labeled "optional," skip videos about exercises, and focus on introductory materials.

As with the course level regressions, there is an interesting null finding. The proximity of the video released to an instructor sending out an email (to make an announcement or potentially remind students of lectures being posted) did not induce more students to watch the recently released videos.

Several suggestions for course design and implementation arise from these results. Most notably, many students dropped off after the first lecture of the course and never return. An instructor's best opportunity to encourage course engagement was in the very first video, which was consistently watched more than any other in the course. Furthermore, because the first video of each batch, most commonly the first of the week, was watched more frequently than subsequent videos in the same week, instructors should wisely organize their weekly content. Including important information in the first video of the week ensures that the most students will receive that information. By releasing videos in two batches per week, instructors may induce more students to watch the first video in each batch.

MOOC instructors commonly agree that dividing lectures into many shorter videos is best practice for the field. However, our results suggest that students are not deterred in their initial decision to watch a lecture by its length. While there could easily be nonlinearities in this pattern at higher lengths, videos in the five to twenty minute range are prevalent in our data, and we do not find adverse effects of video length on students' watching. To the contrary, students stream or download longer videos at slightly higher rates. Instructors should not feel obligated to divide lectures on a lengthier concept into shorter videos in order to encourage more students to watch.

Because students appeared sensitive to video titles, instructors might not wish to include critical content in lectures that include terms such as "optional," "conclusion," and "exercise" in the title knowing. Core concepts of the course should instead be presented in videos labeled "overview" or "intro." Although we cannot conclusively determine that the video titles and percent of students watching them is a causal relationship, it is hard to envision what other covariates omitted from the model could be causing bias. We have controlled for multiple fixed effects such that these estimates are accounting for course, week, and order within week in addition to lecture length.

TABLE 5

Student Level Analysis: Predicting Student Engagement

	44 Coursera MOOCs		STEM Course					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched
International student	−0.004	0.389	−0.001	−1.928 ***	0.003	−2.934 ***	0.042	−0.287
	(0.003)	(0.240)	(0.004)	(0.411)	(0.006)	(0.606)	(0.027)	(2.285)
Missing country	−0.088 ***	−14.116 ***	−0.115 ***	−26.78 ***	−0.141 ***	−31.033 ***	−0.263 ***	−38.674 ***
	(0.009)	(1.329)	(0.004)	(0.507)	(0.006)	(0.668)	(0.039)	(3.313)
<u>Student registered:</u>								
> 4 weeks before course started	−0.004	1.395 *	0.013 *	6.919 ***	−0.021	2.102	0.327 +	18.49
	(0.004)	(0.616)	(0.006)	(0.677)	(0.020)	(2.108)	(0.177)	(15.030)
3−4 weeks before course started	0.011 **	2.321 ***	0.024 ***	6.11 ***	−0.027	−0.831	0.375 *	20.077
	(0.004)	(0.537)	(0.006)	(0.752)	(0.020)	(2.115)	(0.176)	(14.961)
2−3 weeks before course started	0.015 ***	2.831 ***	0.023 ***	5.809 ***	−0.03	−1.41	0.317 +	18.094
	(0.004)	(0.572)	(0.006)	(0.746)	(0.020)	(2.113)	(0.176)	(14.932)
1−2 weeks before course started	0.018 ***	2.779 ***	0.035 ***	6.36 ***	−0.02	−0.939	0.34 +	16.618
	(0.004)	(0.487)	(0.006)	(0.649)	(0.020)	(2.078)	(0.175)	(14.862)
In week before course started	0.019 ***	2.536 ***	0.036 ***	4.846 ***	−0.018	−2.463	0.351 *	16.909
	(0.002)	(0.304)	(0.006)	(0.643)	(0.020)	(2.074)	(0.175)	(14.905)
1−2 weeks after course started	−0.028 ***	−2.161 ***	−0.046 ***	−4.151 ***	−0.069	−9.105		
	(0.004)	(0.598)	(0.006)	(0.735)	(0.081)	(8.372)		
2−3 weeks after course started	−0.049 ***	−2.258 ***	−0.061 ***	−6.834 ***	0.128	−7.139		
	(0.007)	(0.481)	(0.010)	(1.114)	(0.088)	(9.126)		

TABLE 5 (*continued*)

Student Level Analysis: Predicting Student Engagement

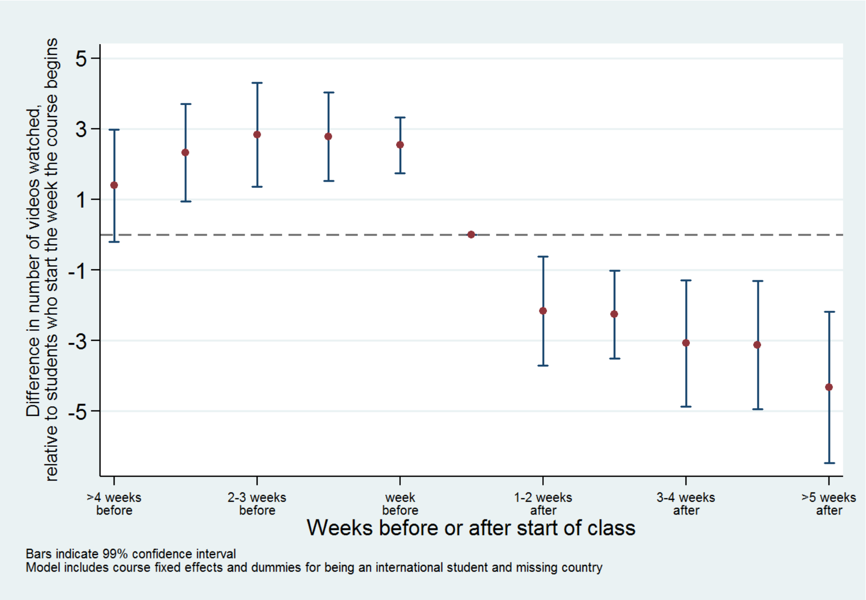
	44 Coursera MOOCs		STEM Course					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched
3−4 weeks after course started	−0.064 ***	−3.078 ***	−0.088 ***	−7.677 ***	−0.193	1.096		
	(0.008)	(0.694)	(0.010)	(1.132)	(0.138)	(14.214)		
4−5 weeks after course started	−0.072 ***	−3.125 ***	−0.087 ***	−5.648 ***	0.04	1.569		
	(0.008)	(0.705)	(0.009)	(1.057)	(0.123)	(12.744)		
> 5 weeks after course started	−0.075 ***	−4.323 ***	−0.089 ***	−9.693 ***	−0.078	−6.062		
	(0.008)	(0.831)	(0.005)	(0.619)	(0.084)	(8.722)		
gmail.com email address					−0.039 ***	−4.691 ***	−0.042 *	−4.789 **
					(0.004)	(0.430)	(0.021)	(1.796)
.edu email address					−0.042 **	−8.314 ***	−0.056	−6.493
					(0.013)	(1.310)	(0.068)	(5.820)
Completed pre-course survey					0.118 ***	12.228 ***		
					(0.008)	(0.792)		
Chose qualitative track							0.094 ***	0.772
							(0.028)	(2.362)
Chose quantitative track							0.198 ***	10.818 ***
							(0.028)	(2.385)
Cited following reason as very or quite important								
Fascination with subject							−0.011	−3.703 +
							(0.023)	(1.979)

TABLE 5 (*continued*)

Student Level Analysis: Predicting Student Engagement

	44 Coursera MOOCs		STEM Course					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched	Earned a Certificate	# of Videos Watched
Academic							−0.049	−3.736
							(0.041)	(3.516)
Prestigious university							0.054	7.556 *
							(0.037)	(3.116)
Fun							0.038 +	2.715
							(0.022)	(1.904)
Curious about online classes							−0.074 *	−2.746
							(0.036)	(3.097)
Job							−0.046	−10.058 **
							(0.046)	(3.883)
Credential							−0.025	−7.077
							(0.055)	(4.685)
Course Fixed Effects	X	X						
Intercept	0.108	15.107	0.109	24.682	0.181	35.393	−0.164	27.406
	(0.006)	(0.649)	(0.005)	(0.560)	(0.020)	(2.095)	(0.177)	(15.083)
N	1,905,289	2,130,907	~30,000	~30,000	~15,000	~15,000	~1,400	~1,400
Adjusted R^2	0.079	0.169	0.067	0.145	0.084	0.208	0.134	0.227

Note. Standard errors clustered at the course level in parentheses. Data in the left panel come from the 40 MOOCs offering certificates (column 1) and all 44 Coursera MOOCs (column 2). Data in the right panel come from one STEM course. For this course, the research team had access to student emails for a subset of students (those who registered before the course began) and responses to the pre-course survey for those students that answered. We have rounded the number of students in this course to hide the identity of the course. Students who are missing the country are students who do not have IP addresses in the data. Reference groups are: domestic students, chose to audit the course, and registered the week after the course began (first day of class through 6 days after course has begun).
$+p < 0.10$. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.



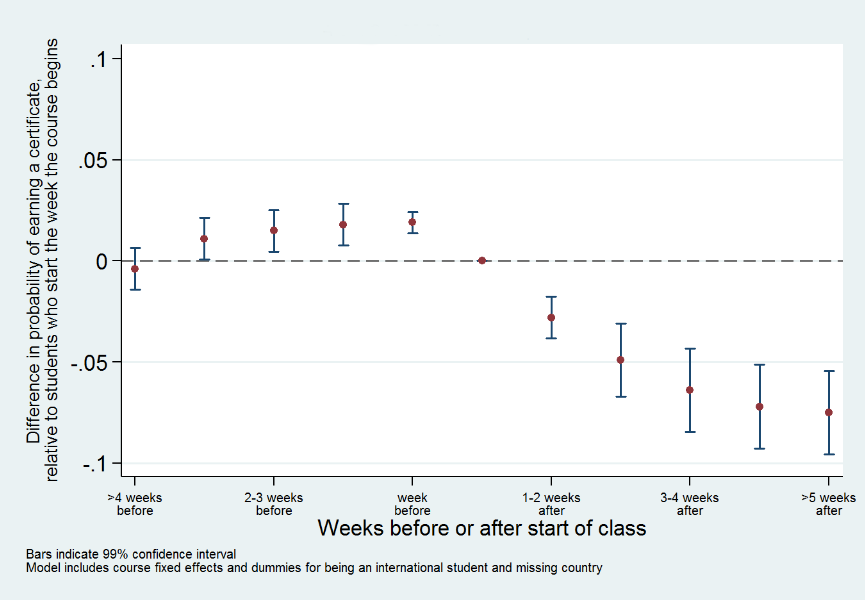


Figure 3. Relationship Between Week of Registration and Probability of Earning a Certification 44 Coursera MOOCs, 2012−2013

Persistence Patterns—Students

We now turn to using student level variables to predict persistence across all courses and within a single STEM course. The analyses across all 44 courses used an unrestricted sample of students, although four courses were excluded for the certificate outcome because they did not offer certificates. The analyses on the single course were limited in some models to students who registered for the course before the official start (because we have email addresses only for these students) and in some models to students who responded to the pre-course survey. Table 5 reports results of running equation (4) on our student level data. We examined two persistence outcomes: whether students earned a certificate and the number of videos each student watched in the course.

We first addressed registration time by including indicators for the number of weeks students registered for the class before and after the course officially began. Registering for the course within one week after it began is the reference category. For the full sample of 44 courses, we observed that, within course, students who enrolled just before the course starts had increased persistence and completion rates compared to students who registered well before or well after the course starts, as can be seen in Figure 3. Students who registered well before the course began (more than four weeks before the course launched) were statistically indistinguishable from students who registered the week after the course began. Students who registered in the week before the course launched out to four weeks early watch more videos (about 2 to 3 videos more) and are more likely to earn a certificate (about 1.5 to 2 percentage points) relative to students who registered the week after the course began.

Students who registered long after the course started were substantially less likely to earn a certificate (three to eight percentage points) and watched two to four fewer lectures compared to students who registered less than a week after the course began. As there were typically more than three MOOC lectures per week, late registrants were catching up on missed lectures, but only partially. On average, they never fully caught up to the engagement and completion levels of their peers who registered on time or early.

The only other predictor available for all students was derived from students' IP addresses and indicates whether they are domestic or international. While there appeared to be little or no difference between international students and the omitted category (domestic), students missing IP addresses (and therefore missing country of origin) appeared to watch much fewer videos and were substantially less likely to earn a certificate. Although it is unclear who these students represented, they made up approximately one-third of the sample. Future work should attempt to identify these students and to better understand their course persistence behavior relative to students with IP addresses.

The STEM course offers a more interesting analysis. The first two models replicated the analysis for all 44 courses. Students who registered very early (4 or more weeks before the class starts) were again less likely to earn a certificate, but they did watch more videos than students who registered the week that the course began. Students who registered after the course began were less likely to earn a certificate and watch fewer videos than students who registered the week the course began. Unlike across all other courses, international students watched fewer videos in this course than domestic students.

Subsequent models build by adding additional predictors drawn from email and pre-course survey data. Students with Gmail email addresses showed worse persistence than students with other email addresses, although it was hard to know exactly whom this group of students represented. Students with ".edu" email addresses also had lower persistence and completion outcomes relative to "gmail" and other email addresses. This could represent a diverse group of MOOC students: college students, college faculty, or others affiliated with a college or university.

By far the largest predictor of course completion among students who registered before the course began was whether students completed the pre-course survey. Survey respondents were 12 percentage points more likely to complete a certificate and watch 12 more lectures than non-survey responders. Completing the survey likely signals substantial interest in the course and could serve as a marker to instructors for the group of students likely to be committed to the course.

For the much smaller subset of survey respondents, the survey offered two interesting components for analysis. The first was that students were asked which track they intended to follow: auditing, qualitative, or quantitative. The quantitative track asked students to complete weekly quizzes and math based problem sets while students in the qualitative track completed weekly quizzes and a final project. Auditors were welcomed to watch the videos but were not expected to complete assignments or earn a certificate. We observed students' initial selection, but students could change their track at any time throughout the course; hence, many students intending to audit the course completed the assignments and earned a certificate. Not surprisingly, both qualitative and quantitative track students were much more likely to earn a certificate than auditing students. Quantitative track students were substantially more likely to earn a certificate than auditors (20 percentage points), and qualitative track students were nearly 10 percentage points more likely to earn a certificate relative to auditors. Quantitative track students were also much more likely to watch additional videos, almost 11 more videos than the auditing students. There was no observable difference in the number of videos watched by qualitative students and auditors.

The pre-course survey also asked students to rate the importance of seven factors in taking the course. The final two columns of Table 5 show the relationship between students who said each of the reasons was "very important" or "quite important" and persistence outcomes. These responses served as a proxy for motivation for taking the course.[10] Controlling for all of the previous factors, the strongest results appear for students who were motivated by relevance to their job. These students watched significantly fewer lecture videos (10 fewer) and had lower certificate rates (4 percentage points less), although the certificate completion finding is not statistically significant. To a lesser extent, the same is true for students who were fascinated by the subject matter. They watched fewer lecture videos, and fewer earned certificates. The largest positive relationship between reasons for taking the course and persistence outcomes was being motivated by its affiliation with a prestigious university.

Discussion & Design Implications—Students

We first considered the implications of the email address findings. To the extent that the .edu group represents current college students, it is possible that these students have traditional higher education course demands that lure them away from MOOC completion. An alternative explanation is that a subset of MOOC users might be college students using material from the MOOC to supplement their collegiate studies with little intention of completing the course. Blended learning designs are becoming common, and there are models in traditional higher education that fully incorporate lectures from MOOCs (Bruff, Fisher, McEwen, & Smith, 2013). Investigating the interplay between students in MOOCs and traditional higher education requires further study.

The motivations results suggested that there was a subset of students who pursued MOOCs for professional reasons, but they tended not to persist. This particular course likely offered little professionally relevant content. Instructors could explicitly discuss the course's application to specific jobs either in the first lecture or throughout the course in an effort to mitigate the dropout of students motivated by professional development. The fact that students who took the course for their interest in the subject were less likely to watch videos suggests students expectations differed from their experience resulting in a decision to stop watching videos. Students who were motivated by their curiosity of online courses had significantly lower rates of certificates, perhaps because they sufficiently tested the MOOC medium and then stopped participating. Finally, students who rated the prestigious university as a main factor might be motivated to complete the course to put an earned certification from the university on their resume because it would help them in the labor market. We thus far have no evidence of MOOCs' impact on labor market outcomes.

The enormous differences between students who completed the pre-course survey and those who did not and between quantitative track students and others suggest instructors can better target specific information. Selecting into the qualitative track likely signals a desire to earn a certificate but a level of discomfort with math and science. These students might have discovered quickly that the course was beyond their level of preparation, hence their lower persistence relative to the quantitative track. This finding may indicate students' preference to earn a certificate via weekly problem sets relative to an end of course project, although this requires further exploration.

It is possible to identify students most likely to be engaged even before the course begins through the pre-course survey. Those students could be grouped either homogenously or heterogeneously, depending on goals and pedagogical practice, for group work, discussion forums, or peer grading activities. This strategy also suggested benefits to enabling multiple formal tracks in the course.

There are also design implications for the registration results. Registering more than five weeks before the course starts was not related to positive completion results, whereas registering closer to the official start date was. While this may indicate certain types of students register at different times, it could be the early period limits student success. Establishing a shorter preregistration window of two to three weeks may promote persistence. Because many students register late, some consideration for enabling students to catch up would likely increase persistence and completion. Perhaps instructors could provide an avenue for late registrants to catch up by prioritizing videos every week or providing opportunities for late assignment submission. Moving towards more self-paced courses would also resolve the lower completion rates for late registrants.

Conclusion

Combining big data with regression analysis at several distinct levels of analysis, we found that the pattern of persistence across MOOCs was fairly similar across courses with an initial steep drop-off that flattened out in the later weeks. Student level MOOC persistence was related to pre-course survey completion, registering early but not too early, and desiring to take the course because of its affiliation with a prestigious university. At the lecture level, introductory and overview lectures and the first lecture of the week experienced higher viewing rates. At the course level, increased rates of participation and persistence were seen among courses that were being offered for the first time, and the number of students, number of lectures, and length of lecture videos were not predictive of persistence or completion.

We applied Tinto's model of academic integration to course engagement, persistence, and completion. We found contact from the professor in the form of an email, potentially one of the most powerful forms of academic engagement, seems to have no effect on whether students watched a lecture video released within a short time of the email. Some lecture characteristics such as lecture length were not related to engagement, but others, such as lecture titles and being the first lecture of the week were. These results support Tinto's main conjecture that institutional characteristics have important ramifications for student persistence, even in the online space.

The findings in this article illuminate certain design features of the course that instructors can put to immediate use. Because students watch the first lecture video of the week, professors should include vital information in the first release each week. The same holds true for lecture videos labeled with introductory words. Additionally, establishing more formalized tracks within a course may provide an opportunity to engage different sets of students with different expectations in positive ways. Most of our design suggestions, such as shortening the preregistration window and renaming videos, are costless, yet they could have a substantial effect on students especially given that more than one hundred thousand students can enroll in a single course. The analyses in this article also suggest more formal experimental studies could prove fruitful. Many platforms and instructors are experimenting with formal A/B testing, and several of the design suggestions we outline could easily be randomly tested to determine whether they work.

Notes

We would like to thank the Stanford Lytics Lab for their data assistance and helpful comments on our work. We also thank our three anonymous reviewers as well as participants at ASHE, the CEPA seminar at Stanford, and the Peabody College LPO Colloquium at Vanderbilt for providing thoughtful feedback on our work. Rachel Baker was supported on this project by Institution of Education Science Grant #R305B090016.

[1] Coursera launched "specializations" and Udacity launched "nanodegrees" which are both comprised of a series of courses that appear more like the traditional model in higher education of taking a sequence of courses. Gathering data across those courses may enable the application of standard theories of degree or certificate completion across multiple courses.

[2] edX is a recent exception. Ho et al. (2014) provide an analysis of edX courses that provides some basic demographic information collected from students. Hansen & Reich (2015) use mailing addresses collected from edX to locate students and to analyze the socioeconomic status of their communities.

[3] In our analysis, watching a lecture is operationalized as either downloading or beginning to stream the lecture video. We do not have access to the clickstream data to assess whether students finish watching the video.

[4] We used linear probability models for all of our binary outcomes for ease of interpretation. We tested whether they produced any out of bounds predications and found a very low frequency. We also tested logit and probit models and found consistent results.

[5] We chose 20% and 80% of videos as reasonable measures of engagement beyond the first few lectures and sustained engagement, respectively. We tested whether these measures of persistence were sensitive to our selections by testing a range of measures (10%, 30%, 70% and 90%). The results for the 10/90% and 30/70% cutoffs were substantively similar to our findings using 20% and 80%, maintaining sign, significance, and general magnitude.

[6] Course instructors typically released videos in groups once a week. We refer to each group of videos as a "batch." To account for the fact that some instructors released groups of videos more than once during a week, we conduct analyses at the "batch" rather than calendar week level.

[7] See online appendix for further details on the selection of the words in this analysis.

[8] Students were asked to indicate whether the following reasons were Very Important, Quite Important, Moderately Important, Slightly Important or Not Important:

1. The subject sounds fascinating!
2. The subject is relevant to my academic field of study
3. I want to earn some sort of credential that I can use to enhance my CV/resume
4. Because this course is offered by a prestigious university
5. I think taking this course will be fun and enjoyable
6. I am curious about what it's like to take an online course
7. This class teaches knowledge and/or skills that will help my job/career

[9] One might be concerned with temporal relationships between certain words in the lecture titles and time of release in the course. There is variation across batches for all of the words, and all but two words are distributed fairly evenly over the course. The words "welcome" and "conclusion" cluster at the beginning and end of the course, respectively. The coefficients and standard errors on these two words should be interpreted with caution due to potential issues of multicollinearity.

[10] Subsequent to the offering of this course, improved scales of learner motivations and intentions have been developed (Kizilcec & Schneider, 2015).

References

Bernard, R. M., & Amundsen C. L. (1989). Antecedents to dropout in distance education: Does one model fit all? *Journal of Distance Education, 4*, 25–46.

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), 187–199.

Chafkin, M. (2013). Udacity's Sebastian Thrun, godfather of free online education, changes course. *Fast Company*. Retrieved from http://www.fastcompany.com/3021473/udacity-sebastian-thrun-uphill-climb

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. J. (2013). The MOOC phenomenon: Who takes massive open online courses and why? (University of Pennsylvania Working Paper). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2350964

DeBoer, J., Ho, A. D., Stump, G. S., and Breslow, L. (2014). Changing "course": Reconceptualizing educational variables for Massive Open Online Courses. *Educational Researcher, 43*, 74–84.

Friedman, T. L. (2012, May 15). Come the revolution. *The New York Times*. Retrieved from http://www.nytimes.com/2012/05/16/opinion/friedman-come-the-revolution.html?_r=0

Halawa, S., Greene, D. & Mitchell, J. (2014). Dropout prediction in MOOCs using learner activity features. *eLearning Papers*, *37*, 7−16. Retrieved from http://www.moocsandco.com/sites/default/files/elearning%2037.pdf#page=7

Hansen, J. D., & Reich, J. (2015). Democratizing education? Examining access and usage patterns in massive open online courses. *Science*, *350*, 1245–1248.

Ho, A. D., Reich, J., Nesterko, S., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITX: The first year of open online courses. HarvardX and MITx (Working Paper N.1). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2381263

Johanes, P., & Lagerstrom, L. (2014). Work-in-progress: Developing online graduate courses in electrical engineering. Paper ID # 9993. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN. Retrieved from http://www.asee.org/public/conferences/32/papers/9993/view

Karp, M. M., Hughes, K. L., & O'Gara, L. (2010–2011). An exploration of Tinto's integration framework for community college students. *Journal of College Student Retention, 21*, 69–86.

Kim, J. (2012, May 21). Playing the role of MOOC skeptic: 7 concerns. *Inside Higher Ed*. Retrieved from http://www.insidehighered.com/blogs/technology-and-learning/playing-role-mooc-skeptic-7-concerns

Kizilcec, R. F., & Halawa, S. (2015). Attrition and achievement gaps in online learning. Proceedings of the Second ACM Conference on Learning @ Scale (pp. 57−66). Retrieved from http://dl.acm.org/citation.cfm?id=2724680

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. Association for Computer Machinery

Transactions on Computer-Human Interaction, *22*(6). Retrieved from http://dl.acm.org /citation.cfm?doid=2744768.2699735

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in Massive Open Online Courses. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). New York: Association for Computing Machinery.

Kolowich, S. (2013, July 19). San Jose State U. puts MOOC project with Udacity on hold. *The Chronicle of Higher Education.* Retrieved from http://chronicle.com/article/San-Jose-State-U-Puts-MOOC/140459/

Lederman, D. (2013, February 7). Expanding pathways to MOOC credit. *Inside Higher Ed.* Retrieved from http://www.insidehighered.com/news/2013/02/07 /ace-deems-5-massive-open-courses-worthy-credit

Lewin, T. (2013, August 17). Master's degree is new frontier of study online. *The New York Times.* Retrieved from http://www.nytimes.com/2013/08/18/education/masters -degree-is-new-frontier-of-study-online.html?_r=0

Liyanagunawardena, T., Williams, S., & Adams, A. (2013). The impact and reach of MOOCs: A developing countries' perspective. *eLearning Papers,* 33. Retrieved from http://centaur.reading.ac.uk/32452/1/In-depth_33_1.pdf

Mangan, K. (2012, September 6). A first for Udacity: A U.S. university will accept transfer credit for one its courses. *The Chronicle of Higher Education.* Retrieved from https://chronicle.com/article/A-First-for-Udacity-Transfer/134162/

Nesterko, S. O., Dotsenko, S., Hu, Q., Seaton, D., Chuang, I., & Ho, A. (2013, December). *Evaluating geographic data in MOOCs*. Paper presented at Neural Information Processing Systems Foundation Workshop on Data Driven Education. Lake Tahoe, NV.

Open Letter from the San Jose State Department of Philosophy. (2013). An open letter to Professor Michael Sandel from the philosophy department at San Jose State University. San Jose State University: San Jose, CA. Retrieved from http://chronicle.com/article /The-Document-Open-Letter-From/138937/

Park J., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Education Technology & Society, 12*, 207–217.

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., and Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in Massive Open Online Courses. *Educational Researcher, 43*, 421–432.

Roavai, A. P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education, 6*, 1–16.

Sutton, S. C., & Nora, A. (2008–09). An exploration of college persistence for students enrolled in web-enhanced courses: A multivariate analytic approach. *Journal of Student Retention, 10*, 21–37.

Sweet, R. (1986). Student dropout in distance education: An application of Tinto's model. *Distance Education, 7*, 201–213.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*, 89–125.

Tinto, V. (1993). *Leaving college: Rethinking the causes and curses of student attrition.* Chicago: University of Chicago Press.

Tinto, V. (1998). Colleges as communities: Taking research on student persistence seriously. *Review of Higher Education, 21*, 167–177.

Webley, K. (2012, September 4). MOOC brigade: Will massive, open online courses revolutionize higher education? *Time.* Retrieved from http://nation.time.com/2012/09/04/mooc-brigade-will-massive-open-online-courses-revolutionize-higher-education/

Willging, P. A., & Johnson, S. D. (2009). Factors that influence students' decisions to dropout of online courses. *Journal of Asynchronous Learning Networks, 13*, 115–127.

Ye, C., Kinnebrew, J. S., Biswas, G., Evans, B. J., Fisher, D. H., Narasimham, G., & Brady, K. A. (2015). Behavior prediction in MOOCs using higher granularity temporal information. *Learning @ Scale Annual Conference Proceedings*, 335–338. Vancouver, BC.