Chen, W. et al. (Eds.) (2016). Workshop Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education

Does the proximity-based network of MOOCs condition changes in universities? An agnostic approach to visualize the process of nominating and designing MOOCs

Jingjing ZHANG, PhD[a*], Kirk PERRIS, PhD[a] & Chenchen ZHOU[a]
Beijing Normal University, Beijing, China

Abstract: This paper serves to provide evidence of institutional patterns to nominate and design a MOOC. Using the concept of proximity, this paper applies economic principles centred on revealed comparative advantage to discern an institution's propensity to nominate courses of certain disciplines over others in the design of a given MOOC. The conditions to design a certain MOOC will lead to the design of MOOCs in a similar discipline, since the conditions are similar (expertise, familiarity, etc.). Prevailing trends in open education matched with unmet demand for higher education suggest that institutions will embrace online learning more and more. We predict this will lead to greater specialization, of which most universities have one or several. Data of all available MOOCs, harnessed from three distinct MOOC aggregators, is utilized in this paper.

Keywords: MOOCs, Concept of Proximity, Higher Education

1. Introduction

Might MOOCs (massive open online courses) push institutions towards greater specialization in course offerings displacing less popular courses and programs? Since their mainstream inception in 2012, MOOCs have spread rapidly and globally. There are now over 4,000 MOOCs found in cyberspace with trends showing greater growth and differentiation. Observers increasingly bemoan the lack of sustainability in MOOCs. Without legitimate credentialing, they argue, high levels of attrition will continue and ultimately lead to the demise of these courses. It is already common for MOOCs to be discontinued after being offered only once. New approaches, however, are gradually taking shape.

Over the next year, three separate organizations will offer credit-bearing MOOCs that will lead to institutional credentials. Arizona State University has launched its first freshman year program, offering an affordable alternative to its campus-based and some of its online programs. The Modern States initiative is another venture that will serve to offer self-paced online courses leading to degree bearing credentials. Finally, the OERu (Open Education Resource Universitas) will launch a full-year of courses that will lead to a first year of a bachelor in general studies. The OERu is distinct in that all courses are repurposed from pre-existing OER, and form to create new OER that can be utilized outside of its network for varying purposes, including commercial ones. Although each presents a divergent pathway, the conclusion is towards conferring legitimate credentials.

There is little doubt that Coursera and other major MOOC providers will watch these endeavours closely (edX, in fact provides the platforms for both First Year and Modern States). Coursera, in fact, is already offering fee-based programs in partnership with universities leading towards varying forms of credentials, such as certificates.

We believe that MOOCs are pushing greater sustainability and legitimacy for online learning and posit that there are early signs of patterns of MOOC development that will continue in future. An institutions subject specialization(s) is favored when selecting which courses to nominate and design as a MOOC. Over time, we predict that this will create a feedback loop whereby courses in the same or similar disciplines in a given university will follow suit. We draw lessons from revealed comparative advantage and in particular, the concept of proximity, to further this prediction.

2. Comparative Advantage and Proximity Concept

The concept of comparative advantage is central to development economics（Costinot,2009）. In the industrial age, a country's or region's ability to extract and distribute its natural resources was central to its ability to participate in the global economy and acquire greater material wealth. In the information age, a country's ability to support the creation of knowledge and innovation among its population has supplanted natural resources as the central determinant to participate in the global economy.

Higher education is central to the development of a knowledge economy. Elites, professionals, and other highly skilled workers generally pass through an institution of higher learning before making their mark on the world. OECD countries have participation rates in higher education in excess of 50 percent of the age cohort. An extension of creating a universal system of higher education is the creation of a lifelong learning society. Individuals are enrolling in continuing education programs to augment existing skills or to learn new ones.

Despite steady enrolment at multiple levels of higher education, there are growing concerns of budget shortfalls and unsustainable enrolment. Governments are receding from meeting the same levels of financial support for higher education that they have committed to in past. The most visible outcome has been a rise in tuition fees.

Amidst other financial pressures, many universities have become more differentiated and more entrepreneurial. Creating greater online programs and links with industry have proven successful. At the same time, some universities have ended programs of low enrolment. Some institutions have closed altogether. With pressures on some programs, universities have partnered to offer joint programs, giving some semblance of sustainability for smaller departments.

Most universities have particular areas of study for which they specialize. The Massachusetts Institute of Technology specializes in engineering and technology. Beijing Normal University specializes in education. These universities, like thousands of others, are also comprehensive universities offering a broad range of programs. Otherwise referred to as the multiversity (Hayhoe, 1995） they service a broad student body.

The ubiquity of technology is shifting the manner in which universities operate. The idea of comparative advantage has surfaced in policy documents on higher education（Sahni&Shankar, 2005. A major concern is if universities specializes in certain areas, then why differentiate to accommodate learners outside of these areas? Proponents of consolidation argue that learners would be better served in institutions that specialize in the subject area of their choice.

Opposition would surface from faculty unions and current and prospective students. Some may be out of a job or uprooted to work or study elsewhere. Counterarguments would suggest that interdisciplinary is an intellectual strength and worth preserving.

In returning to an understanding of comparative advantage relative to natural resources, the advantage a country yields includes its ability to extract the resource, refine it where relevant, and get it to market. Drawing an analogy to higher education, it is important to consider other elements of an institution's infrastructure the support its comparative advantage.

A lens into understanding how universities may operate in future is through the direction of using massive open online courses (MOOCs). MOOCs need little introduction. They appeared in the mainstream in 2012 after Stanford University offered a few of their courses online for anyone to participate. MOOC providers such as Coursera and Edx followed, and operate as the major players in the organization and delivery of MOOCs. They have partners from around the world and continue to experiment with varying forms of delivery, credentialing and revenue generation.

The number of MOOCs that exist in cyberspace hovers around 3,000. Hundreds of universities are involved, and many are lined up to join Coursera or Edx. Some venture out on their own. The longevity of MOOCs is uncertain. For one, no sustainable business model has taken shape. Both Coursera and Edx are propped up by wealthy institutions or venture capital. Considering poor completion rates and lack of legitimate credentialing, the longevity of these and other MOOC providers, along with innumerable independent providers, is suspect. Yet, some 20 million learners are enrolled in these courses from around the world.

These demographic patterns underscore larger issues in higher education. Outside of the industrialized world, higher education is in dismal shape. Among myriad issues, is quality, which is

highly uneven. China, for example, is home to the world's largest higher education population, yet only several dozen universities are deemed world class. The aggregated enrolment of these elite institutions measures at several hundred thousand, a miniscule number out of a total higher education enrolment of nearly 40 million （Vol.40, 2005）.

Outside of quality is capacity. UNESCO predicts that an additional 70 million individuals will be seeking enrolment opportunities over the next 10 years. That is approximately 30% above current global enrolment. Adding one-third student numbers to a given institution around the world is seemingly impossible, as is the construction of thousands of new institutions. Construction costs, land, and the individuals to run and teach in these institutions are unfathomable. MOOCs may not be the answer, but they are a reasonable response to the growing demands on higher education that are occurring and will persist in the coming decade and onwards.

If online learning, and MOOCs by association, become integral to widening access to higher learning, might there be a way to predict how effective the delivery of MOOCs may be?

Applying the concept of comparative advantage to the global higher education context, and MOOCs as a predictor of future, it is important to look outside of the course specialization and consider the broader infrastructure and acumen in designing and delivering a MOOC.

According to a report by Southwest Jiaotong University, higher education institutions tend to choose existing courses to offer as a MOOC based on the following criteria:

First, courses are selected from those that won the national top level course award for excellence. Second, is to choose a course from the leading subjects of the school in order to highlight its ability and brand. Third is to choose from foundation courses that either are taught in English or get good evaluation from students for their resilient and innovative pedagogy.

Within the context of revealed comparative advantage, is the concept of proximity. According to Hidalgo, the concept of proximity posits that, "a country's ability to produce a product depends on its ability to produce other ones." Underlying the concept is compatibility of products. If a country produces apples, for example, there is a favorable infrastructure to produce pears. The climate, soil properties, and methods of packaging are all similar. Further, it is highly probably that the knowledge required to grow apples is transferrable to acquire the knowledge to grow pears.

Mapping the concept of proximity to an institution's propensity to nominate and design MOOCs based on preferred courses is the central objective in this investigation.

3. Methods

This investigation aims to take inventory of all available MOOCs, including paced and self-paced MOOCs as found on three MOOC aggregators: MOOC list, Class Central and Guokr (China). Preference for selection will be given to institutions of higher education, thereby excluding foundations or other educational agencies. The rationale for exclusion is to include institutions only that offer a range of courses taught on campus or online to students who have gained admission to a particular institution since the premise of the paper is to predict what other institutional courses may be transferrable to a MOOC.

3.1 The Concept of Proximity: Devising an Equation for MOOCs

We modify an equation devised by Hidalgo （Hidalgo, 2007） to ascertain proximity. Formally, the proximity f between courses A and B is the minimum of the pairwise conditional probabilities of a university offering a MOOC in a particular discipline given that it offers another MOOC in another discipline.

$F = \min\{P(A \mid B), P(B \mid A)\}$

With these figures a complex network will be constructed. In this network, a node represents a given course and the line that connects two nodes represents the proximity.

4. Results

We anticipate that the results will yield data that shows high probability that an institution offering two or more MOOCs will belong to the same or a similar discipline. Based on this assumption, we started using a web crawler tool—bazhuyu---to grab online MOOCs data. So far we have got all course data from Guokr, the total course number is 4367, and the categories are course name, school, time, platform, and basic introduction. Because Guokr doesn't provide discipline information, at this initial level we may not be able to inspect our hypothesis. The following are just some preliminary results from descriptive statistics.

4.1 Descriptive statistics

4.1.1 The number of Published courses showing an increasing trend

According to picture 1, the number of launched courses has been bigger since 2012 and reached its peak at the year 2015, which is the last year we can get course data of the whole year. Although there is a dropping trend during the next three year, we can obvious know that those are incomplete data, cannot reveal real situation. That means so far as we know, the number of published courses are showing an increasing trend. If studying the fitting curve more thoroughly, we can find that the yearly growth rate also increases. All of these tell us MOOCs have been rapidly developing during the last 4 four years , it has clearly been an educational hit recently.

4.1.2 The Self-paced courses account for a certain proportion

Another conclusion we can draw from picture 1 is that self-paced courses account for a certain proportion. As we can see, the number of self-paced courses take up almost 30% . One of the great advantages of online course is its ability to free students from time and space restriction, yet a lot of MOOCs are just like traditional offline education, with specific start time, and one will not be able to join in his fond course if he missed the time. The rather amount of self-paced course can fairly avoid this problem and provide learners the possibility to enjoy the merit of online education.

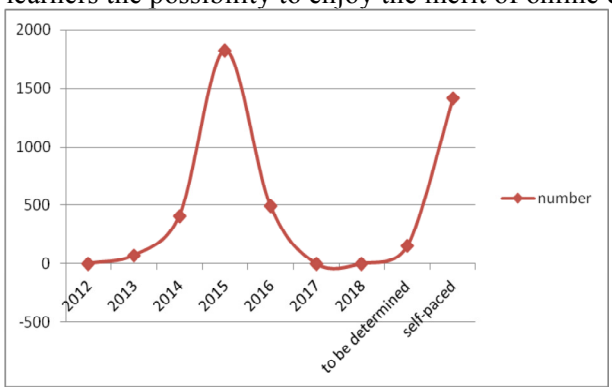

Figure 1. Time distribution of Guokr MOOCs

4.1.3 80% school have launched less than 10 courses during 2012-2018

The essence of our methodology is the concept of proximity, which will use the conditional probability to calculate. In line with the equation, the number of courses school have launched and the number of courses with similar name are indispensable data. Picture 2 is the distribution of course number a school published, over 200 school only published one course, 80% school have launched less than 10 courses the top 3 of course number a school published is 167(MIT),147(Tsinghua) and 93(Harvard).

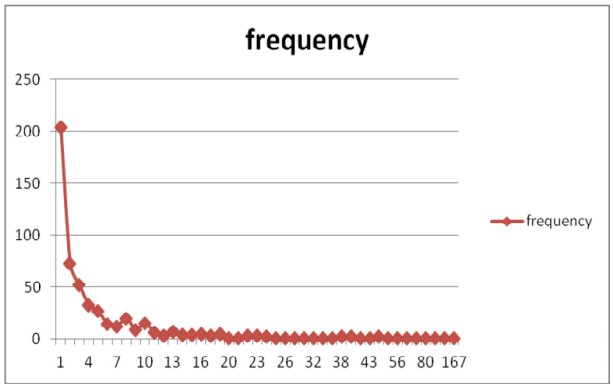

Figure 2 the distribution of course number a school published

4.1.4 Over 90% courses are with a different name

Another descriptive statistics we have done is about the course name, we want to have a general picture of how many courses are with the same name, which can show us if there are two school launched same course. Unfortunately, if we choose same course name as the indicator of connection between two separate school, we may not be capable of obtaining expected result, as 90% courses have different name.(table 1)

Table 1: the frequency of course name.

Course name	Number
1	4304
2	57
3	5
4	1

4.2 *What we planning to do next*

On account of the intial descriptive analysis, using same course name has clearly become an inferior choice, so we decide to find a way to better cluster two courses, since we have the basic introduction information of the course, text clustering is an option. Also, those data are just from Guokr, apparently there are a large amount of courses still missing, so we will keep on digging course informatin from other platform for further analysis.

5. Discussion and Conclusions

The study aims to provide evidence of institutional preference towards designing MOOCs. The hypothesis was that institutional preferences to nominate and design a MOOC will increase the likelihood of a course of the same or similar discipline to subsequently be nominated and designed as a MOOC. In a broader context, this idea was advanced using the concept of proximity located within the area of economics referred to as revealed comparative advantage. A country's propensity to extract and distribute a given resource provides conditions to extract and distribute another given resource that is similar in nature. Mapping this idea to higher education, and the prediction of greater reliance on specializations, we predicted that growth in MOOCs will reflect a similar result. A computer course that is nominated and designed as a MOOC will lead an institution to nominate and design a course in the same or similar discipline (e.g.,. mathematics), instead of nominating and designing a course from the humanities. The prediction is that this may lead to greater institutional focus towards subject specializations for which an institution has a known comparative advantage to

other institutions. The long-term result is that lesser established specializations in a given institution will be pushed to the margins, an outcome that is increasing in regularity.

Acknowledgements

The authors would like to thank the 12th five-year programme of the National Education Science Foundation (P. R. China) for financially supporting this research under Project No.ACA140009

References

Costinot, A. (2009). On the origins of comparative advantage. *Journal of International Economics, 77(2),* 255-264.

Hayhoe, R (1995). An Asian Multiversity? Comparative Reflections on the Transition to Mass Higher Education in East Asia. *Comparative Education Review, 39(3)*, 299-321.

Hidalgo, C. A., Klinger, B., Barabasi, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science, 317*, 482-487.

Sahni, R. & Shankar, V. K. (2005). GATS and Higher Education: Revealing Comparative Advantage. *Economic and Political Weekly*, 40(47), 4947-4953