Exploring the relationship between course structure and etext usage in blended and open online courses

Daniel T. Seaton
Office of Digital Learning
Mass. Institute of Technology
Cambridge, MA 02139
dseaton@mit.edu

Yoav Bergner
Physics and RLE
Mass. Institute of Technology
Cambridge, MA 02139
bergner@mit.edu

David E. Pritchard
Physics and RLE
Mass. Institute of Technology
Cambridge, MA 02139
dpritch@mit.edu

ABSTRACT
We use a two-parameter family of bounded distribution functions (Kumaraswamy) to fit electronic textbook (etext) usage in 20 blended and online courses from Michigan State University, MIT, and edX. We observe clusters of courses in the parameter space that correlate with course structural features such as frequency of exams.

Keywords
Course structure, etexts, MOOCs, usage mining

1. INTRODUCTION
When etexts are integrated into Learning Management Systems (LMS), one can extract the number, frequency and duration of page views from the tracking logs. These data have fed back into etext interface design [2] and personalized approaches aimed at understanding student reading habits and comprehension [3], but an incomplete picture remains in terms of guiding instructors on how to integrate etexts into courses. This is particularly salient given studies which point to low use of traditional textbooks and poor correlation of use with performance [4].

We consider the fraction of etext pages accessed by students in courses of varying structure. Aspects of course structure include the primary/supplementary role of the etext, its integration with graded assessment, and the frequency of exams. Our data come from blended and distance learning courses from Michigan State University (MSU) and from open online courses from the RELATE group at the Massachusetts Institute of Technology (MIT) and edX. The MSU populations are typical for introductory science at a large state university. In contrast, the student populations in open online courses are highly variable in age and preparation [1].

Courses in this study use either LON-CAPA, with e-texts as modularized html pages, or edX, which uses digital versions of traditional textbooks within simple navigation. We fit the distribution of unique page views in each course using a two-parameter family of distribution functions with support on the interval [0,1]. The complimentary cumulative distribution function (CCDF) of the Kumaraswamy distribution is given by $F(x; a, b) = (1 - x^a)^b$. The (a, b) parameters which determine the shape of each distribution may not be familiar. We highlight four relevant regions: bimodal $(a, b < 1)$, low usage $(a < 1, b > 1)$, high usage $(a > 1, b < 1)$, and unimodal $(a, b > 1)$. Note the probability distributions (PDF, not CCDF) associated with $a = b = 0.5$ (bimodal) and $a = b = 1.5$ (unimodal) have the same "average" use.
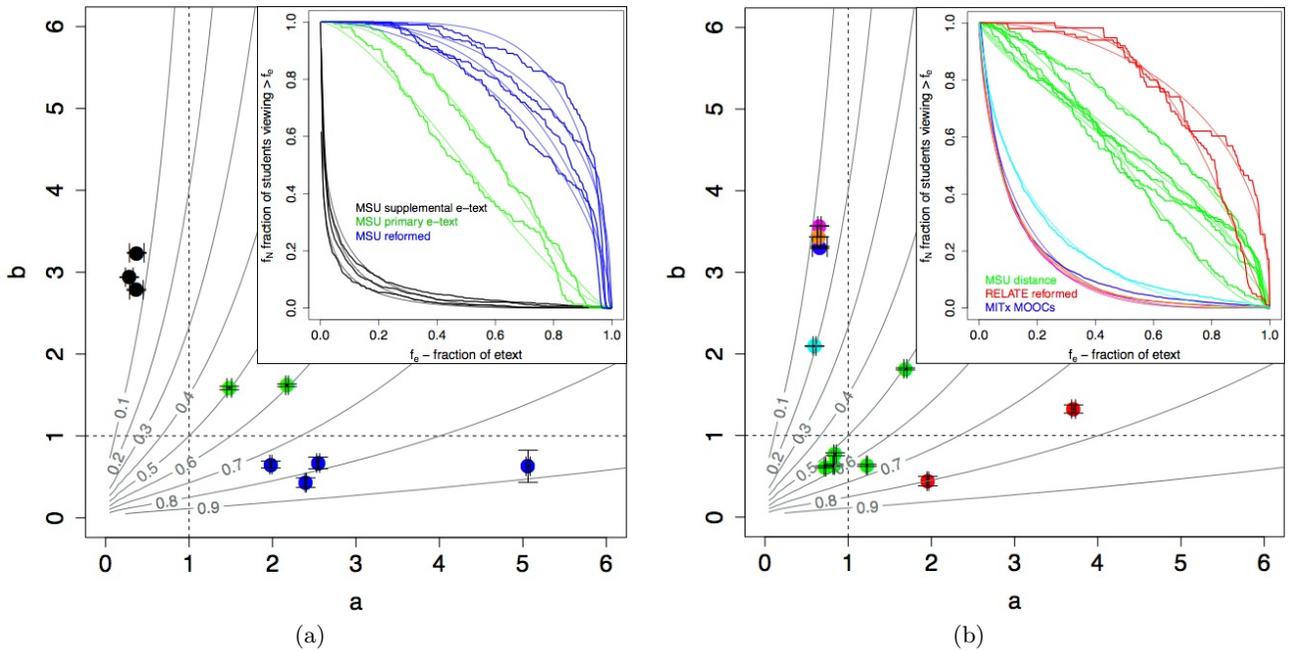
2. BLENDED COURSES
We first consider nine blended courses from MSU (same instructor), all Fall semester Mechanics using LON-CAPA for weekly homework and readings from the MultiMedia Physics[1] etext. We group the courses by structure differences as follows: MSU supplemental $(N = 898, 911, 808)$, in which the etext was available alongside a physical textbook; MSU primary $(N = 159, 190)$, in which the etext was the only learning text; and MSU reformed $(N = 211, 209, 197, 254)$, which used a primary etext along with reading quizzes and a switch to frequent exams (every two weeks).

Fig. 1(a) shows both the CCDFs (inlay) and the clustering of a, b values in the fit-parameter space, where we also indicate contours of constant fractional usage. MSU supplemental courses are in the low-usage region, average view fraction (avf) < 0.1, while MSU primary courses are in the unimodal region (avf ≈ 0.55). The change of role from supplementary to primary assigned textbook appears to have a significant impact. Proximity to the $(a = b = 1)$ saddle point indicates that the MSU primary view distributions are almost flat. MSU reformed courses are in the high-usage region (avf ≈ 0.8). Unfortunately we cannot separate the effect of frequent exams from reading questions, as both changes were implemented together.

3. ONLINE COURSES
Our second analysis involves three types of purely online courses: MSU distance $(N = 155, 231, 165, 187, 163)$, online intro physics courses; RELATE reformed $(N = 58, 100)$, open online intro physics courses from the RELATE group at MIT; MITx $(N = 7158, 4829, 2686, 1935)$, four MOOCs from MITx. Student populations for open online courses represent only certificate earners. MSU courses used the

---

[1]http://www.pa.msu.edu/bauer/mmp/

Figure 1: Etext usage via Kumaraswamy a, b**-parameters for (a) blended courses and (b) online courses. Contour lines represent the average fraction of the etext viewed. CCDFs (bold) with curve fits (smooth) are displayed as an inlay.**

same etext and LMS as the previously discussed blended courses. RELATE reformed courses used LON-CAPA with a Mechanics text developed by the RELATE group [5]. MITx MOOCs were disseminated through edX.

The online course data in Fig. 1(b) display similar, if slightly more nuanced, clustering of usage by course structure. MSU distance courses cluster near the saddle point ($a = b = 1$), but end up within three of the four usage regions. Since the average view fraction remains similar to the blended courses ($0.5 - 0.6$), the points on either side of the saddle point lend themselves to the following interpretation: while students in MSU primary (blended) courses typically accessed 55% of the etext (unimodal), students in distance courses viewed either more or less than this (bimodal).

The two RELATE courses in Fig. 1(b) both have similar average view fraction (≈ 0.78), but one (summer) appears in the unimodal region and the other (spring) in the high-usage region. The spring instance required students to complete all 14 weekly units; in summer, the last three units were optional. Only a small fraction of students completed the optional assignments, explaining the shift. The three MITx MOOCs in a tight cluster are computer science offerings (avf ≈ 0.1). A fourth, Solid-State Chemistry, lies just outside (avf ≈ 0.2). The MOOCs fall in the same region of the parameter space as the MSU supplemental (blended) courses, which is consistent with their similar course structure: all provide their etext as a supplement text.

4. DISCUSSION AND CONCLUSIONS

Course structure appears to have a dominant and predictable effect on etext usage distributions. Our framework, based on clustering in the two-parameter space of Kumaraswamy distributions, is easy to apply and generalizes to any finite resource type tracked within an LMS (e.g. lecture videos, homework). Instructors may exploit the dramatic correlation of course structure and etext usage to encourage particular student usage of their online materials.

Acknowledgments

5. REFERENCES

[1] S. Kolowich. Who Takes MOOCs?. *Inside Higher Ed*, 2012.

[2] J. Pearson, G. Buchanan, and H. Thimbleby. The reading desk: Applying physical interactions to digital documents. In *Proceedings of the 2011 annual conference on Human factors in computing systems*. ACM, 2011.

[3] T. Peckham and G. McCalla. Mining Student Behavior Patterns in Reading Comprehension Tasks. In *5th International Conference on Educational Data Mining*, 2012.

[4] T. Stelzer, G. Gladding, J. P. Mestre, and D. T. Brookes. Comparing the efficacy of multimedia modules with traditional textbooks for learning introductory physics content. *American Journal of Physics*, 77(2):184, 2009.

[5] R. E. Teodorescu, A. Pawl, S. Rayyan, A. Barrantes, and D. E. Pritchard. Toward an Integrated Online Learning Environment. In *Proceedings of the Physics Education Research Conference*, pages 321–324, 2010.