Two Years of Michigan MOOCs: What Analytics Tell us About Learning in these Environments

Christopher Brooks (brooksch@umich.edu)
Stephanie Teasley (steasley@umich.edu)

Learning Education & Design Lab
School of Information and
Digital Education & Innovation
University of Michigan

Overview of MOOC Activity

- Michigan was an early member of the MOOC movement
- Offering courses on the Coursera platform
- Courses are video-based with discussion forums and assignments
- 13 different courses offered in 51 sessions
- Different institutions use MOOCs differently
- Michigan MOOCs show of diversity of programming at Michigan
- Keep in mind: MOOCs are "massive" like black holes
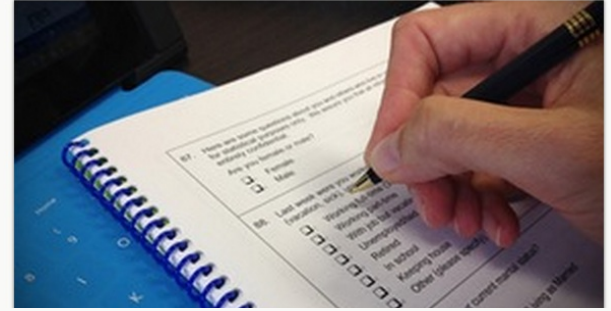
Instructional Methods in Health
Professions Education
June 1, 2015

Introduction to Cataract Surgery
October 5, 2015

AIDS: Fear and Hope
February 2, 2015

By the Numbers

Unique Learners (enrollees): 1,979,446
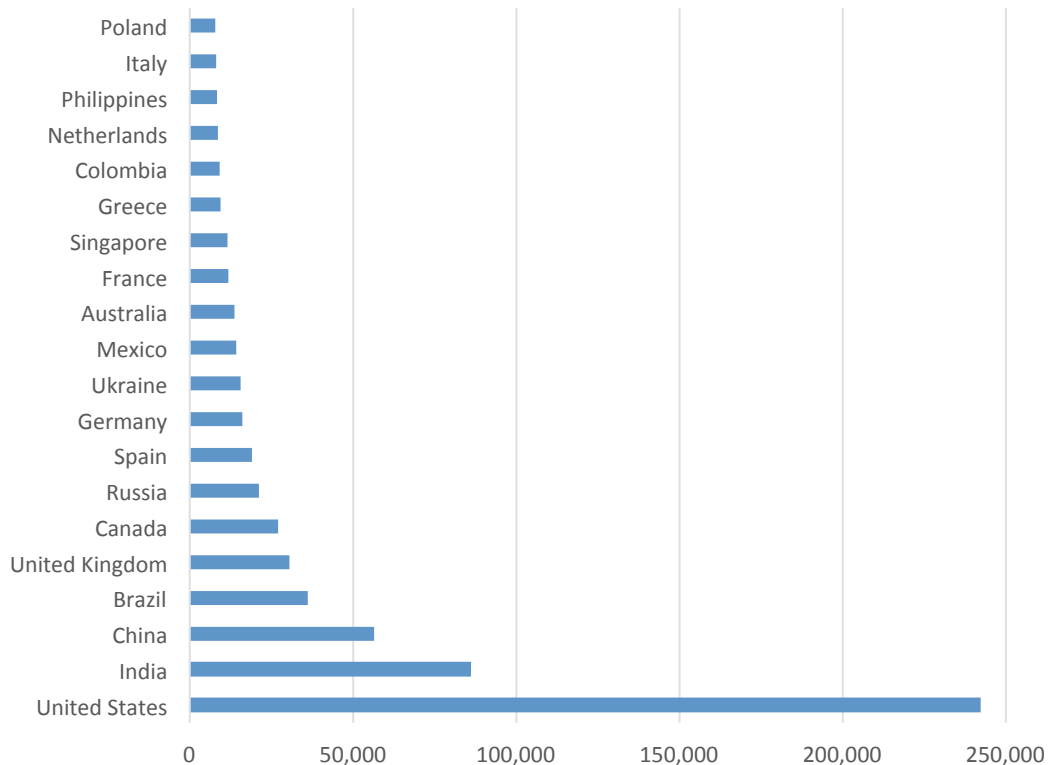
Participants: 922,163

- Graded (8.6%)
  - Certificate 48,911
  - Certificate with Distinction 30,224
- Ungraded (91.4%)
- But, roughly 300,000 participants at least tried an assessment item

Lots of data (200+ million clickstream events)

Geographic Location

Unique Learners by Country (Top 20)

- Swap between China and Brazil from Fall 2013

What do people do in MOOCs?

- Learning Analytics Fellows project
- What do "attendance patterns" look like in these kinds of courses?
- Top half is streaming, bottom is downloading
- Blue lines are Mondays, Red lines are assignment due dates

Relationship between Video Viewership and Grade?

Distribution of Course Grades
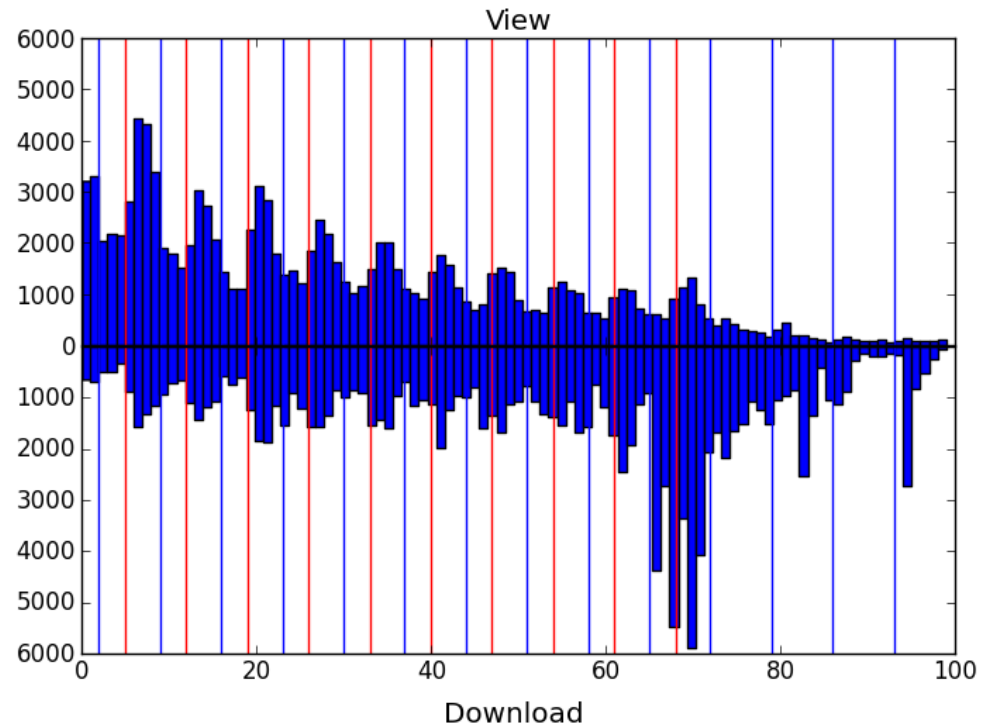
Let's back this horse up

Why are people taking MOOCs?

Coursera asks intentions:

- …mastering the course material by working through the exercises and earning a certificate. **126,312 (44%)**

- …learning the course material mainly by watching most of the lectures. **122,168 (42%)**

- None of the above. I'm just checking out the course for now. **40,650 (14%)**

But most of our participants haven't filled this survey out…

Digging Deeper

- A grounded theory analysis suggested:
  - Personal (cognitive, emotional, social understanding, pleasure, self-confidence, curiosity, fear)
  - Employment (self-employed, under employed, horizontal and vertical movement, career change)
  - Socialization (inside the mooc, and with others in the world)

Digging Even Deeper

- Knowledge Acquisition (depth and breadth of knowledge, skill development)

- Strong relationships to formal education (current students, future students, and lifelong learners)

- Versus other modalities (fast, high quality, structured, accessible, cost)

Learn with Friends

- What is the difference on learner outcomes when they sign up with friends? (χ^2 p=0.006)

	Distinction	Normal	None
Friends	34.2%	5.3%	60.5%
Alone	24.2%	4.5%	71.3%

- How does discussion forum posting habits (sociability proxy) change when learners signs up with friends? (t-test, p=0.03)

	Mean	Standard Dev.
Friends	6.96	8.83
Alone	4.86	5.06

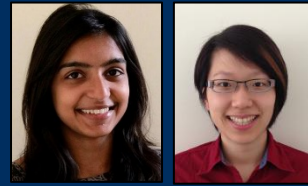# How do learners interact with friends?

- <u>Casual discussions</u> with family members
- <u>Formal study group</u> discussions with colleagues
- <u>Ad-hoc discussions</u> with colleagues who did not sign up for the course
- <u>Formal professional development activities</u>, with learners using this content directly in their teaching practice
- <u>Collaborative lecture viewing</u> activities,

What do Michigan Students think of MOOCs?

- Stratified semi-random sample of UM students ~800 responses (80 we knew enrolled in our MOOCs)
- Of students who had not enrolled in our MOOCs…
  - 40% had never heard of MOOCs
  - 18% had taken another institutions MOOC

Student Opinions

- We asked the 40% how likely they would to enroll in MOOCs (large negative skew)

Strongly disagree | Neutral | Strongly agree

| 170 | 135 | 130 | 66 | 14 |
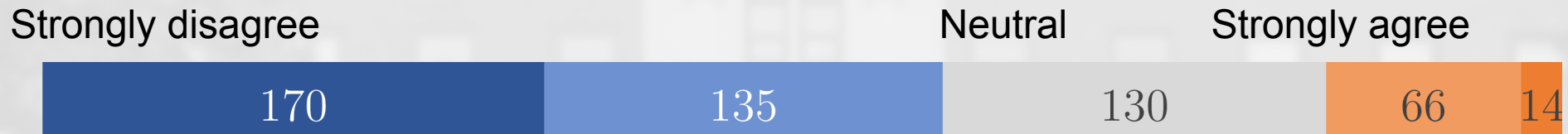
- We then asked them if the instructor recommended the MOOC for supplemental resources, how likely would they be to enroll

Strongly disagree | Neutral | Strongly agree

| 41 | 53 | 142 | 227 | 52 |

I enrolled in these MOOCs to:

- Deepen my knowledge in my area of study

- Help determine what my major should be

- Broaden my understanding of other areas

- Explore an area to see if I should take a course in that area for university credit

- Supplement a face-to-face course with additional learning resources

Beyong goals, what is impact?

- Yes, it has strengthened my knowledge in my area **(100%)**

- Yes, it has introduced me to areas I would not otherwise get exposure to **(73.1%)**

- Yes, it has encouraged me to switch institutions **(4.5%)**

- Yes, it has helped me relate my university study to career options **(22.4%)**

- Yes, it has encouraged me to consider further study (e.g. to eventually apply to a graduate degree or professional program in the area) **(37.3%)**

Digging Deeper

- Asked whether MOOCs offered by the institution are a "benefit to students enrolled in traditional programs at that institution"

- Small majority of students said yes (375)

- More than a third (34.9%, 231 students) said not sure

- Very few said no (55)

- Just as unclear as the rest of us on the value!

Why are MOOCs a good thing for traditional students?

Cost and Social Need

- – Direct and Ancillary
  - Babysitting
  - Commuting
  - Tuition
- – Institutional costs
  - Opportunity to have smaller class sizes by allowing the breadth enrollees to take the course online
  - Showcase of programs to auditing students

More Reasons

Self regulation of learning

- – Development of SRL skills
- – Showcasing of SRL ability

Learning styles and approaches

- – Acknowledgement of the diversity of learning skills

"I would be too intimidated to learn a skill like coding in a face-to-face environments because I don't have a similar skill-set or background to those studying full time...In a classroom for example, I think I'd find the gender disparity of computer science courses u n n e r v i n g   a n d   a   l i t t l e distracting...it's likely…that my confidence in my own ability to success would diminish in this kind of setting, simple because I'd be a female taking the course 'for fun' surrounded by serious students in a field that's known for being a little competitive and very male-driven."

- Brand recognition & preeminence
  - Some felt it strengthened the brand, some felt it diminished the value of the degree
  - How many of our learners who said they enrolled in a MOOC chose a UM MOOC?
    - Yes: 38%
    - No: 49%
    - Unsure: 13%

(18% of people who said they had not taken a UM MOOC actually did.)

(No record for over 50% of those who said they did take a UM MOOC.)

Quality

"...there isn't the connection with a professor or peers to facilitate learning and questioning..."

"the professors that teach these courses have an opportunity to engage with the material in a new way which might strengthen their face-to-face teaching. Additionally, many MOOCs have a wide user base which means that the professors need to teach for a multicultural and multinational population. This would most definitely benefit the students learning from this professor."
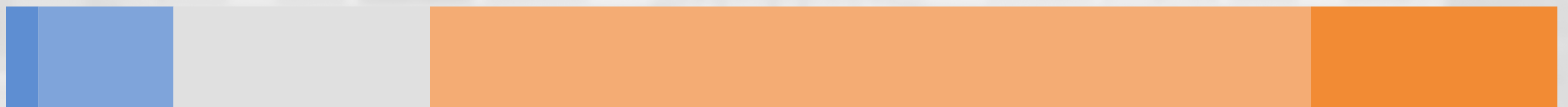
Value in volunteering in UM MOOCs?

- That this would provide a networking opportunity with faculty and graduate students?

2.6	7.6	27.3	48.4	14.1

- That this was a good way to understand your major better

3	9.5	22.9	51.8	12.8

- That interactions with MOOC students would be positive

2.6	6.2	20.2	52.1	18.9

- That over this was an opportunity worth volunteering for.

2.1	8.7	16.5	56.8	15.9

Student-led MOOC Support

- Wanted to help residential students learn through teaching

- Deepen understanding of content by helping others

- Advanced residential MBA students:

    – Created tutorial video

    – Staff live hangout "office hours"

- Made available to ~2,400 MOOC students

Online students

- Pretty small uptake from MOOC students (17% watched a video)

- Much more involvement from those in signature track (300 vs 105 watchers)

Count of learners by activity type over a rolling 7-day period

■ visited the course ■ watched a lecture ■ submitted an exercise ■ browsed the forums
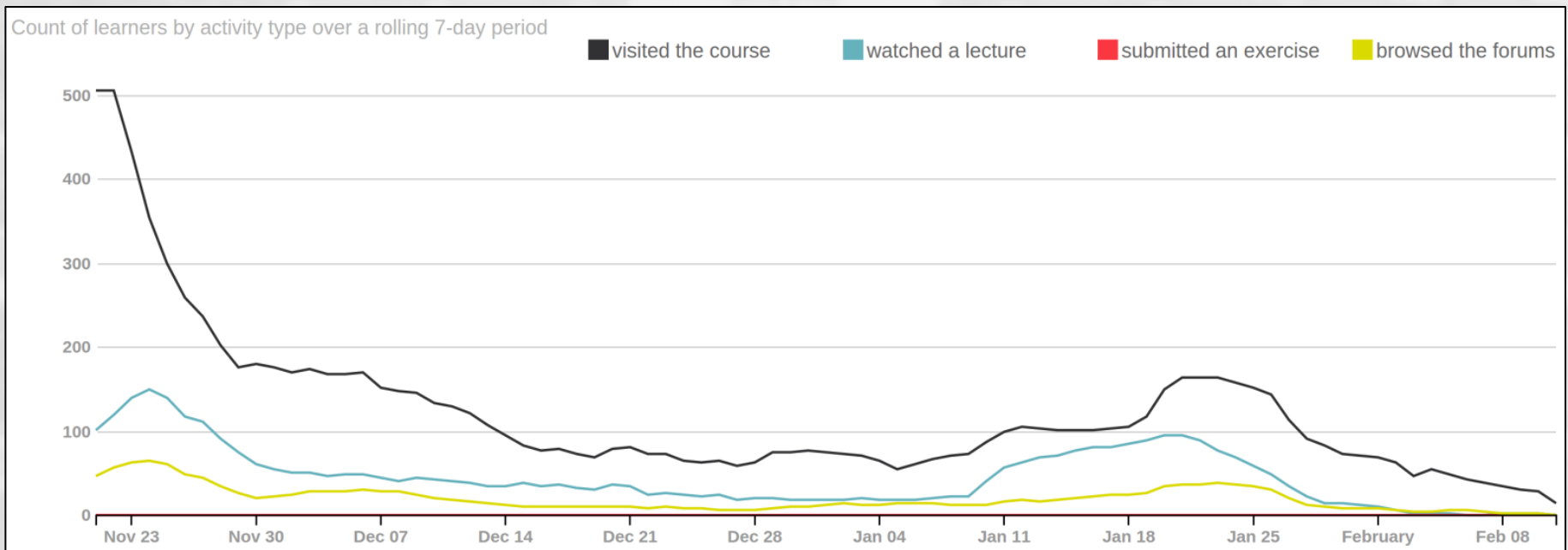
Impact on F[...] Students

"…answering liv[...] [...]way of being drilled in concepts and [...] [...]on your feet. And hearing someon[...] [...]great way of sampling the ma[...] [...]s to any given problem, and ex[...] [...]pective for [the] future…"

"The exercise of [...] [...]ewer's shoes and thinking about h[...] [...]eing explained really made me consid[...] [...]e mechanisms of finance in a mor[...]

Peng Zhang added 3 new photos.
11 hours ago near Ann Arbor, MI

Help Professor Gautam Kaul record Coursera videos! Something I dreamed of when I completed my Coursera course.

51 Likes   9 Comments

Like   Comment   Share

News Feed   Requests   Messages   Notifications   More

Other Residential MOOC Experiments

Introduction to Cataract Surgery

Understanding and Improving the US Healthcare System

Predictive Modelling

"So Harvard research confirms: people who pass courses do more stuff than people who don't. "

http://blogs.edweek.org/edweek/
edtechresearcher/2014/03/big_data
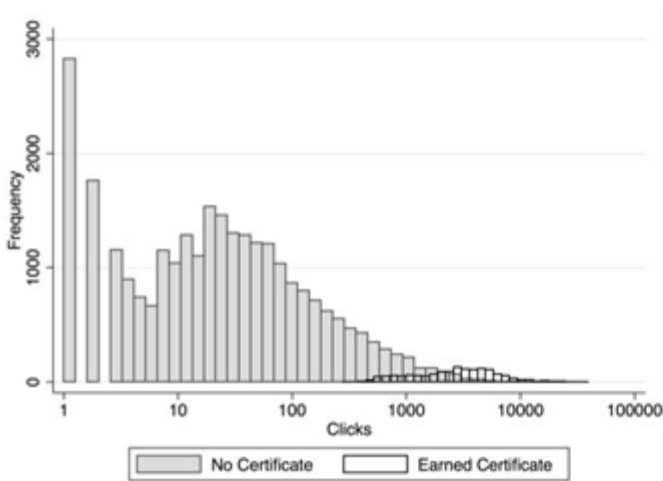_mooc_research_breakthrough_learning
_activities_lead_to_achievement.html

Figure 13: **Number of participant clicks (i.e. recorded actions) plotted on a log scale for JusticeX certificate earners (n=5,442) and non-certificate earners (n=73,345).**

- Can we do better in understanding student success in this environment?
- Could we predict who is at-risk for students who want to obtain a certificate?

It's hard to predict success in MOOCs

- MOOCs lack the diversity of data we have about residential students
  - Previous achievement (SAT/ACT, last year's course)
  - Socioeconomic status (distance from university, first in family, wealth)
  - Gender
  - Ethnicity
  - Motivation
- Building predictive models of student achievement in learning analytics is largely done on these entry-level features
- Both frustrating and refreshing
  - Want accurate models, but want causes of issues

Solution

- Built a novel feature selection algorithm inspired by work in the text-mining community
- It looks at the pattern of engagement that a student has with course resources
- Build of historical data (last year's course) to create day-by-day multilevel model's (C4.5)
- Initial work is based on student certificate achievement (pass/fail)
  - (not the only valuable outcome variable to try and predict!)

How it works

Resource	Day of Course								
	1	2	3	4	5	6	7	8	9
Video	▓		▓					▓	

Daily Accesses
Day 1: Yes
Day 2: No
Day 3: Yes
Day 4: No
Day 5: No
Day 6: No
Day 7: No
Day 8: Yes
Day 9: No

3-Day counts
Day 1-3: Yes
Day 4-6: No
Day 7-9: Yes

Weekly counts
Week 1: Yes
Week 2: Yes

Monthly counts
Month 1: Yes

For a 104 day long course, with three "resources" (videos, forums, quizzes) this gives us 408 features for the modelling activity.

Text mining often uses n-grams as features in a document

- A bigram (cat, good) is the number of pairs of these two words in a document, a trigam (cat, wasn't, good), etc.

- We build engagement n-grams up to 5 gram

Daily Accesses
Day 1: Yes
Day 2: No
Day 3: Yes
Day 4: No
Day 5: No
Day 6: No
Day 7: No
Day 8: Yes
Day 9: No

Possible bigrams
[yes, yes]: 0
[no, no]: 3
[yes, no]: 3
[no, yes]: 2

Possible trigrams:
[yes, yes, yes]: 0
[yes, yes, no]: 0
[yes, no, yes]: 1

…

For a 104 day long course, with three "resources" (videos, forums, quizzes) this gives us 717 more features for the modelling activity.

In a nutshell

We don't have <u>diverse</u> set of data, bu

And there is a lot of it (200 million+ clickstream events)

By pulling out patterns of resource access, we can use supervised machine learning (C4.5) techniques to build predictive models
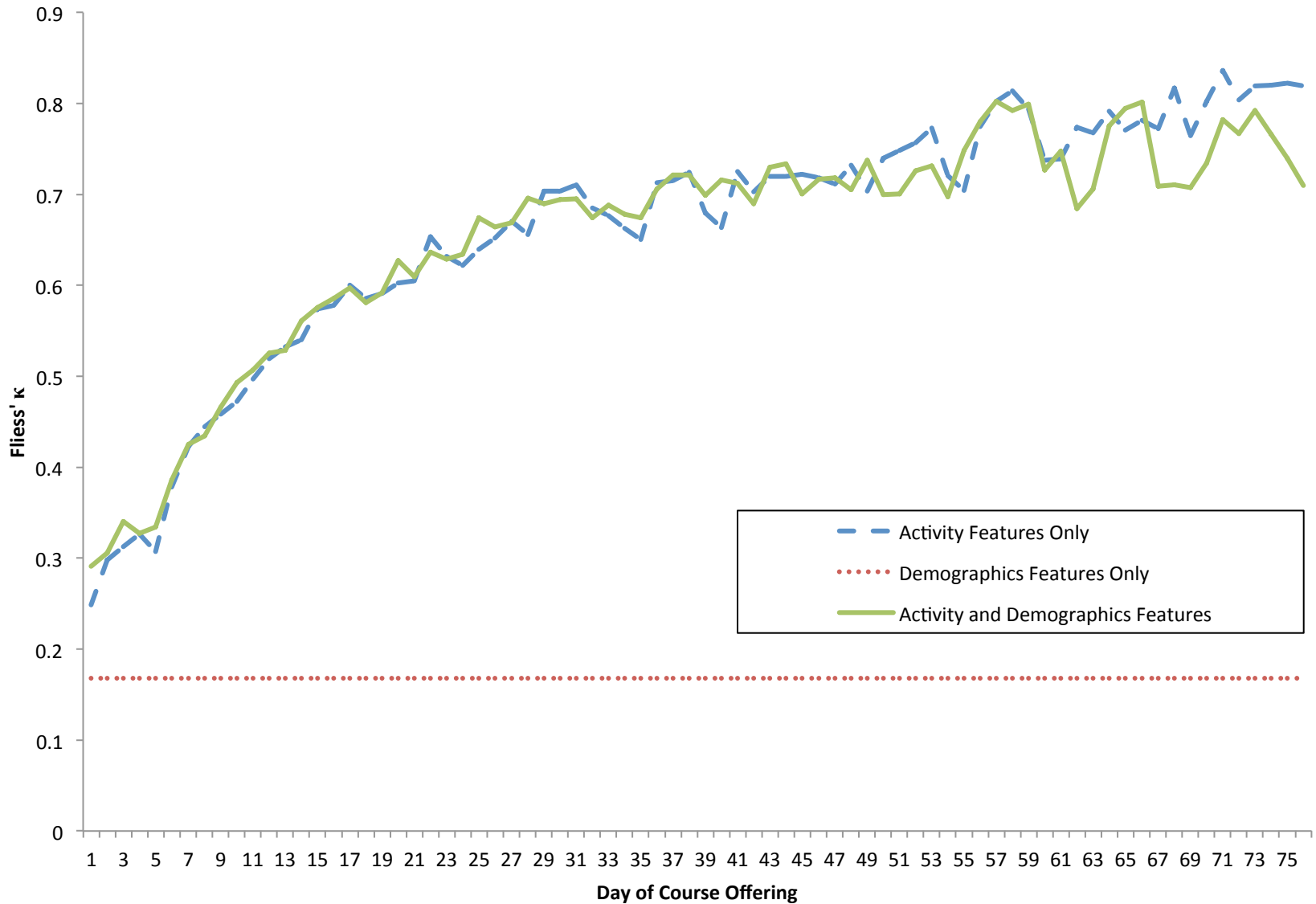
But what if we did have entry data from students?

- – Gender & Ethnicity
- – Certification status
- – Country of origin
- – Etc.

Kappa Over Time for Internet MOOC
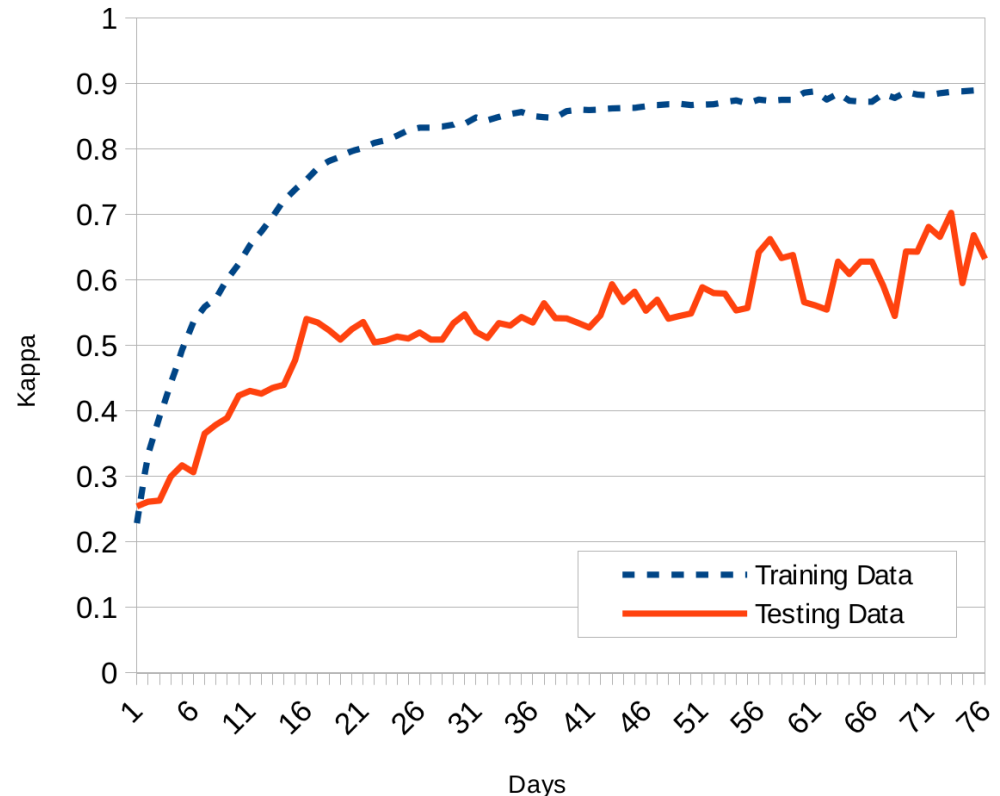
Fliess' κ versus Time in Days

Results

- It is possible to create predictive models on clickstream data for MOOCs
- 3 weeks into the MOOC seems to be an interesting point for some courses
- It is computationally intensive to create these models (daily!)
- MOOC entry/demographics information doesn't seem to add value

C. Brooks, C. Thompson, S. Teasley. (2015) A Time Series Interaction Analysis Method for Building Predictive Models of Learners using Log Data. 5th International Conference on Learning Analytics and Knowledge 2015 (LAK'15)

C. Brooks, C. Thompson, S. Teasley. (2015) Who You Are or What You Do: Comparing the Predictive Power of Demographics vs. Activity Patterns in Massive Open Online Courses (MOOCs). The second annual conference on Learning At Scale 2015 (L@S2015), Works in Progress track. Vancouver BC, March 14-15, 2015. Vancouver, BC.

Next Steps

- The what of the prediction is clear, the why is not clear
  - Prediction rules are difficult for instructors to understand, reflecting <u>latent constructs</u>
- Does this method work in blended situations?
  - Is it predictive only for MOOCs, or for blended residential education as well?
- Can we improve model accuracy through more sophisticated modeling techniques (e.g. Bayesian methods?)
- How do we integrate this with platforms like Coursera?

Conclusions

1. Every MOOC is unique (size, domain, pedagogy, student body), and so are the research questions we can ask of it
2. Student success is a nuanced concept in MOOCs
   – What do we value at UM as measures of success?
3. Residential students largely value UM MOOC offerings
   – And there are many ways to engage out students in our MOOC program and be locally impactful
4. The breadth of data available makes for novel techniques for understanding learners

Credits:
A tasty looking pie: http://www.tarasmulticulturaltable.com/wp-content/uploads/2013/03/Easter-Ricotta-Pie-5-of-6.jpg
A tastier looking cherpumple: http://beautygirlsmom.com/wp-content/uploads/2010/11/cherpumple-pie.jpg
A horses' bum: http://www.pbase.com/image/56819062