The Civic Mission of MOOCs: Measuring Engagement across Political Differences in Forums

Justin Reich
MIT
Cambridge, MA
jreich@mit.edu

Brandon Stewart
Princeton
Princeton, NJ
bms4@princeton.edu

Kimia Mavon
Harvard
Cambridge, MA
kmavon@seas.harvard.edu

Dustin Tingley
Harvard
Cambridge, MA
dtingley@gov.harvard.edu

ABSTRACT
In this study, we develop methods for computationally measuring the degree to which students engage in MOOC forums with other students holding different political beliefs. We examine a case study of a single MOOC about education policy, *Saving Schools*, where we obtain measures of student education policy preferences that correlate with political ideology. Contrary to assertions that online spaces often become echo chambers or ideological silos, we find that students in this case hold diverse political beliefs, participate equitably in forum discussions, directly engage (through replies and upvotes) with students holding opposing beliefs, and converge on a shared language rather than talking past one another. Research that focuses on the civic mission of MOOCs helps ensure that open online learning engages the same breadth of purposes that higher education aspires to serve.

Author Keywords
MOOCs; civic education; discourse; text analysis; political ideology; structural topic model

ACM Classification Keywords
K.3.1. Computer Uses in Education.

INTRODUCTION
Political theorists have long argued that exposure to diverse perspectives is vital to a robust civil society and to the development of citizens [7]. Democratic discourse requires engaging with people who hold different perspectives, and differing perspectives are often a function of different life circumstances. One serious threat to contemporary civic education, therefore, is the growing homogeneity of school populations. In the United States, residential segregation has lead to growing de facto K-12 school segregation [8], while admissions processes and self-selection has resulted in class segregation in higher education.

L@S 2016, April 25-26, 2016, Edinburgh, Scotland, UK.

ACM 978-1-4503-3726-7/16/04…$15.00 .
DOI: http://dx.doi.org/10.1145/2876034.2876045

Open online education offers one pathway for students to join communities of diverse learners beyond the bounds of geography [13]. Demographic research into massive open online courses (MOOCs) has shown that these courses are among the most diverse "classrooms" in the world, with students of different ages, levels of education, and life circumstances [5]. A globally diverse online learning community, however, does not guarantee that students encounter and consider different perspectives. Internet researchers have posed two competing theories for how people confront differences on the Web [6]. One theory holds that the Internet is a series of "silos" where individuals seek out media and communities that conform to their established beliefs [9]. Another theory holds that the Internet contains many interest-driven spaces that serve as ideological "bridges" [15], where people attracted to these interest-driven spaces can be diverse across many dimensions. At present, we know little about which of these theories best characterizes open online courses.

In this study, we investigate whether students in online course forums are building silos or bridges. We examine a case study of the HarvardX course HKS1368x, Saving Schools, a course on U.S. education policy. Using survey data, we locate student political beliefs along an education policy ideology spectrum—from left-leaning supporters of unions and school boards to right-leaning supporters of vouchers and charters—and then we analyze how political beliefs predict engagement in the course.

We present two approaches to evaluating whether students in MOOC forums engage across their political differences. First, by examining the connections between posts, replies, and upvotes in forum threads, we evaluate the degree to which students in forums respond directly to students with differing opinions. Second, we use text analysis to evaluate whether the use of language in online forums converges or diverges among students with different political beliefs. Even if students with different beliefs form integrated networks, it may still be possible for students to use siloed language to talk past one another. For instance, in a conversation about guns in American society, conservatives might address Second Amendment rights whereas liberals might discuss issues of public health. We hypothesize that in high-quality discussions in which students directly address each others' ideas, computational text analysis will reveal modest differences in the language used by people with opposing political beliefs.

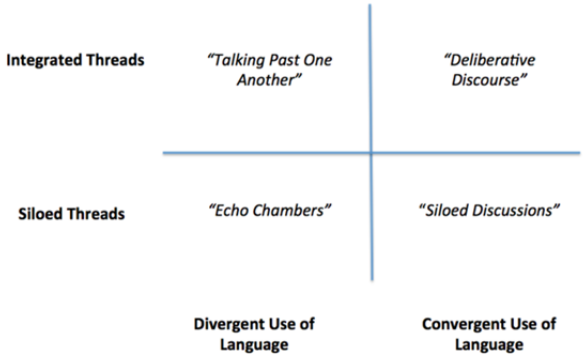

Figure 1: Two-by-two schematic of dimensions of engagement in online discussion forums.

Based on these two analyses, we propose a two-by-two matrix to summarize dimensions of engagement across political differences in online discussion forums, shown in Figure 1. The bottom left quadrant describes forums where people with different political beliefs separate into silos and use different language; these are the echo chambers of Internet discourse. The top left quadrant describes integrated threads in which partisans use different language; these are spaces where students with different beliefs talk past one another. In the bottom right quadrant, students discuss topics using a shared language, but they divide themselves into conversational silos with like-minded others. In the top right quadrant is the ideal condition of deliberative discourse, where people with diverse beliefs converse with a common vocabulary. This initial study, part of a larger project on measuring engagement across political differences in open online courses [18], presents early efforts at measuring these two dimensions.

We address four specific research questions in this study. First, do students with diverse political beliefs register for Saving Schools? Second, political belief correlate with course participation patterns? Specifically, do people with left- or right-learning beliefs on education reform participate more in the course forums? Third, do students in online discussion forums engage directly, through replies and upvotes, with students with different political beliefs? Fourth, do students with different political beliefs use different words and language to discuss topics, or do students converge on a shared language?

In addressing these questions, we respond to Siemens's [17] call to focus MOOC research not just on individual cognitive development but on improving society. As occurs in all sectors of education, there is constant tension between the many possible purposes of MOOCs: as online job and workforce skill training, as a stimulating leisure activity for lifelong learners, or as a modern Agora where citizens from a global commonweal can learn and share together. The pivot of MOOCs towards workforce development is not a foregone conclusion [3]. Research that examines the quality of engagement across political differences in MOOCs can lead the field towards more investigations of the potential of open online learning to benefit civic education and civil society.

This paper proceeds in four parts. We introduce the Saving School course, its students, and their political beliefs. We then present descriptive statistics on how political beliefs correlate with course participation. Next, we present social network analyses of the forum interactions of participants. Finally, we use Structural Topic Models, a method of text analysis, to evaluate the degree to which students with diverse political beliefs converge on a shared language [16].

SAVING SCHOOLS: THE COURSE AND PROFESSOR

Saving Schools is a course about U.S. education policy and reform offered by HarvardX on the edX platform that ran from September 2014 to March 2015. The course is taught by Paul Peterson, Director of the Program on Education Policy and Governance at Harvard University and Editor-In-Chief of *Education Next*, a journal of opinion and research. The course was designed around Peterson's (2010) book *Saving Schools* and consisted of four mini-courses based chapters of the book: "History and Politics of U.S. Education," "Teaching Policies," "Accountability and National Standards," and "School Choice."

Each mini-course was 5-6 weeks long, with content released in weekly bundles according to topic. Each week included a package of materials, such as video lectures, assigned reading, multiple choice questions, and discussion forums. For example, in the second Saving Schools module, "Teaching Policies," the weekly modules included discussions of "Teacher Compensation" and "Class Size Reduction." The "Teacher Compensation" module included three video lectures with the homework questions "are teachers paid too little?"; "are teachers paid too much?"; and "are teachers paid the wrong way?" Students were then instructed to read two opposing *Education Next* pieces on teacher pay and to respond in the forums to a discussion prompt on that topic. Some weeks, students were split into discussion cohorts by letter of last name or date of birth. Learners earning a certificate were required to post at least once in the discussion forum each week.

The politics of U.S. education reform do not perfectly align with conservative/liberal distinctions, but Paul Peterson's education policy preferences are generally associated with conservative positions. His journal, *Education Next*, is considered one of the leading publications for conservative viewpoints on education policy issues, and executive editor Martin West was an educational advisor to Mitt Romney's presidential campaign. Peterson is a proponent of free market reforms, school and teacher accountability [11], charter schools [10] and standardized testing [11]; and he has been critical of policies advocated by labor unions and schools boards. Our informal assessment of Saving Schools is that Petersen provides multiple perspectives on issues and gives each side a fair hearing, though he also makes clear his own, generally conservative, policy preferences.

	Participants	Survey Respondents	U.S. Forum Posters
Total (n)	5,408	1,982	592
Median age	32 years	36 years	40 years
US learners	76%	69%	100%
Percent female	55%	58%	64%
Bachelor's degree or above	79%	78%	88%
Average median grade	41%	67%	93%
Ever taught		55%	63%
Now teaching		26%	45%
Intent to complete course		57%	75%
Intent to audit course		31%	18%

Table 1: Demographics of Saving Schools participants.

The Students of Saving Schools

Demographics

There were 10,478 students who registered in at least one of the four modules for Saving Schools. As is typical for HarvardX MOOCs, about half of registrants did not ever actually enter the courseware, leaving 5,408 participants. Demographic data for these participants is shown in Table 1: 55% were female, 79% had a Bachelor's degree or advanced degree, 76% were participating from an IP address registered in the United States, and the median age was 32. The median grade, calculated by averaging the median grade of all four modules among students who got at least one graded problem correct, was 41%.

Of the 5,408 participants, 1,982 students completed the pre-course survey. Survey respondents are demographically similar to participants as a whole. Among these respondents, 55% report that they have worked as a teacher or instructor, and 26% report currently teaching. Respondents from the United States were asked questions, derived from a national survey conducted by *Education Next* [12], about their education policy preferences.

In this study, we focus our attention on the 592 students from the United States who both participated in the forum and completed the pre-course survey. Of the 1,008 students who posted in forums at least once, 613 were in the United States, and 592 American forum posters took the course survey, for a response rate among this group of 96%. These students differ in key dimensions from other course participants: on average they are older, more likely to be female, more likely to have a Bachelor's degree, more likely to have taught, more interested in earning a certificate, and more likely to actually do so. This group is our sample of interest for the remainder of the study.

Do you think that [government funding/taxes] for public schools in your district should increase, decrease, or stay about the same?				
Greatly increase	**Increase**	**Stay about the same**	**Decrease**	**Greatly decrease**
19%	42%	33%	4%	2%
Do you favor or oppose giving low-income families attending failing public schools the choice to attend private schools instead, with government assistance to pay the tuition?				
Strongly favor	**Somewhat favor**	**Neither favor nor oppose**	**Somewhat oppose**	**Strongly oppose**
18%	30%	5%	23%	24%
Do you favor or oppose tenure for teachers?				
11%	20%	7%	33%	27%
Do you think teacher unions have a generally positive or negative effect on schools?				
Strongly positive effect	**Somewhat positive effect**	**Neither positive nor negative effect**	**Somewhat negative effect**	**Strongly negative effect**
7%	29%	16%	34%	15%

Table 2: Responses by Saving Schools U.S. forum posters to politically salient questions derived from *Education Next* poll (*n*= 592)

Political Beliefs

In the pre-course survey, students from the United States were given a set of questions taken from the *Education Next* poll, a commissioned annual public opinion poll. (The authors had no involvement in creating the survey; these were secondary data) Of the various questions asked, we identified four questions that are most strongly correlated with measures of political partisanship [12]: questions about school taxes, school vouchers, unions, and teacher tenure. In Table 2, we show the distribution of responses to these four questions among the 592 U.S. survey respondents who were also forum posters. While the respondents generally favored increasing taxes and funding for public schools, with 61 percent in favor and only 6 percent opposed (perhaps unsurprisingly given the number of enrolled educators), beliefs varied more widely for the other three questions. Among respondents, 48 percent favored vouchers while 47 percent were opposed; 31 percent favored tenure while 60 percent were opposed; and 36 percent favored unions while 49 percent were opposed. In the 2015 *Education Next* poll, 50 percent of the public supported increased school funding, 42 percent supported vouchers, 29 percent favored tenure, and 30 percent favored unions [4].

Our assessment of these survey results is twofold. First, Saving Schools attracted students with diverse political beliefs about education policy. This alone is an important finding: MOOCs have the capacity to enroll students with diverse beliefs even when the faculty is recognized for having partisan beliefs. Second, U.S. forum posters' policy preferences are similar to those of the general U.S. public.

To generate a composite representation of a student's political ideology, we use principal components analysis (PCA) to reduce these four responses into a single representation of political ideology, drawing on all 1,295 U.S. students (out of 1,982 total pre-course survey respondents) who completed the *Education Next* poll questions (Table 3).

Factors	Political Ideology	Comp. 2	Comp. 3	Comp. 4
Eigenvalue	1.79	.92	.83	.45
Weights				
Taxes	0.3601	-0.8084	0.4286	0.1823
Vouchers	-0.4059	0.273	0.8714	-0.0366
Tenure	0.5688	0.4768	0.143	0.6547
Unions	0.6181	0.2113	0.1909	-0.7327

Table 3: Four factors derived from a Principal Components Analysis of 1,295 U.S. Saving Schools student responses to four *Education Next* questions.

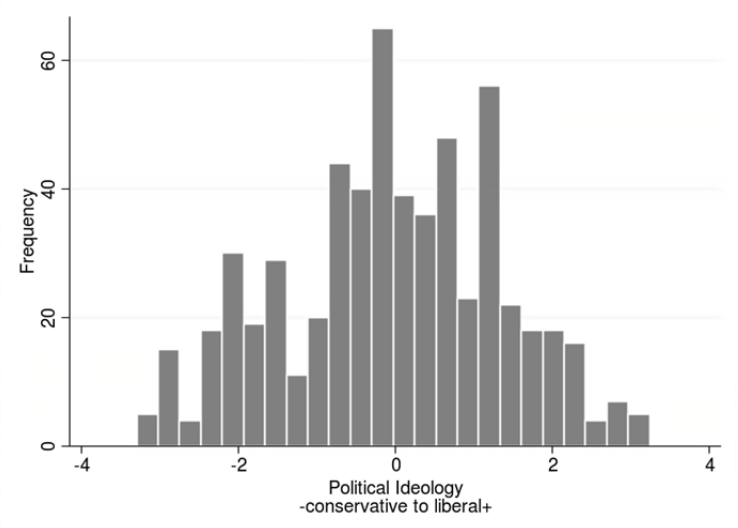

Figure 2: Distribution of political ideology scores of all U.S. Saving Schools survey respondents who posted in the forums (*n*=592).

The first principal component eigenvalue is substantially larger than the others and is the only component above one. As expected, this principal component has positive loadings for the questions related to taxes, tenure, and unions, and negative loading on the question related to vouchers. Given these favorable qualities for a summary measure, we call this new variable Political Ideology, where positive values are associated with typical liberal positions and negative values are associated with typical conservative positions. In Figure 2, we show the distribution of political ideology across all U.S. Savings Schools forum posters.

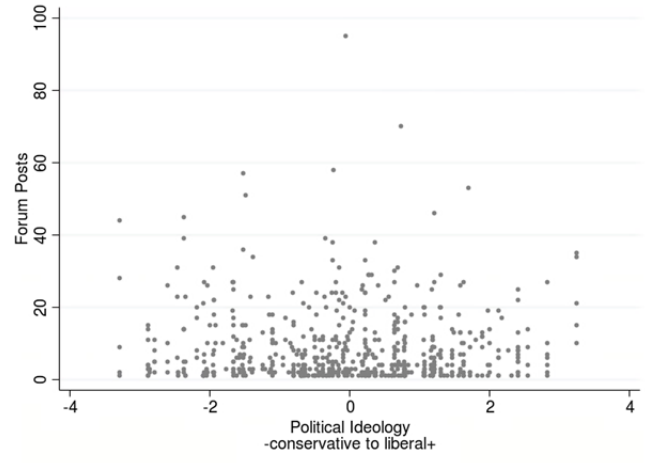

Figure 3: Scatter plot of U.S. Saving Schools forum posters by political ideology and number of forum posts (*n*=592).

FORUM ACTIVITY IN SAVINGS SCHOOLS: DOES POLITICAL BELIEF PREDICT STUDENT FORUM PARTICIPATION?

The corpus of discussions across all four modules includes 8,649 posts from 1,008 participants. The number of posts written per learner throughout the course varied from 1 to 95 among those who completed the *Education Next* poll. In Figure 3, which shows a scatterplot of the number of forum posts and political ideology, we see little relationship between political ideology and posting behavior, with posting behavior evenly distributed across conservative, neutral, and liberal political ideology scores. Moreover, choosing to post at least once is uncorrelated with political ideology. We conducted a logistic regression model where the outcome is posting at least once for the 1,166 U.S. students who completed the political survey and demographic questions. Controlling for gender, education, and age, we find no relationship between political ideology and posting at least once (β=.03, SE=.04, p=.50).

Students with different political beliefs participated in the Saving Schools forums at equitable rates. We turn to forum activity analysis and text analysis to examine whether these diverse students talk with one another.

Network Structure by Political Affiliation

For students to build "bridges" in their online course forums, it is first necessary that forum threads include a range of political perspectives. In this section we assess whether individual forum threads contain contributions from posters with a diversity of political beliefs. We then consider whether partisanship is a determinant of patterns in whom individuals post in response to and who they upvote.

Categorizing threads

Threads provide a natural mechanism for partitioning separate conversations in the forum. Different forum threads serve different purposes in MOOCs [19, 20]. Some serve an administrative function, such as having students introduce themselves, making announcements about the timing of homework, or providing feedback from the teaching staff. Others contain content of a more political nature such as discussions of collective bargaining, the common core, or vouchers. While engagement on all sides of the political spectrum may be of interest in administrative threads, we are particularly interested in whether people with different political beliefs engage in politically-salient threads.

To identify the threads of interest for this analysis, one researcher manually coded the initial topic of all 476 threads of the course into three categories: administrative, low political salience, and high political salience. Administrative threads included questions or comments about how to use the EdX platform to turn in homework assignments and a thread at the start of each mini-course inviting each student to share something about his/herself. Topics related to education policy but not demarked as controversial topics within the Saving Schools course were coded for low political salience. An example includes threads looking at the relationship between economic growth and education that began: "I certainly agree that economic growth and education are closely related. All the countries that scored low in the [international assessment] do not have an outstanding economic growth. Schools are the main source for human capital." Topics coded as high political salience included posts addressing themes in U.S. educational policy demarked as controversial or worthy of further discussion in the Saving Schools course, such as high stakes testing, the No Child Left Behind Act (NCLB), the Common Core, teachers' unions, and charter schools.

Threads vary substantially in both size and content. In order to ensure that our analysis was not dominated by low participation threads, we limit our analysis to threads that contain more than 10 posts, which leaves 76 threads containing 7677 posts. Of these 76, 15 were administrative, 11 were high political salience, and 50 were low political salience. There are 136 and 114 posts on average for high and low political salience threads compared to 64 for administrative threads.

Threads Have Ideologically Diverse Posters

We first address the question of whether posters segregate themselves across threads. In Figure 4, we show the average ideology level (as measured in the factor analysis above) for each thread, organized by our salience coding. Each point represents a thread and the intervals around them are 95% confidence intervals around the estimates of the average ideology of the posters in that thread. Almost all of the confidence intervals cover zero, which indicates

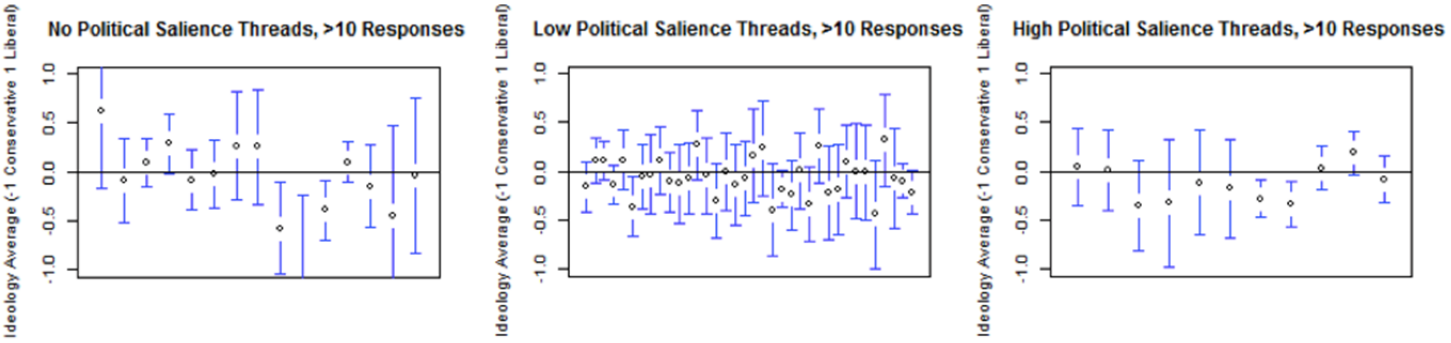

Figure 4: Ideological average in Saving Schools forum threads with more than 10 posts.

that individual threads are not being dominated by a single perspective. Even among high political salience threads the ideological average never strays far from the global mean.

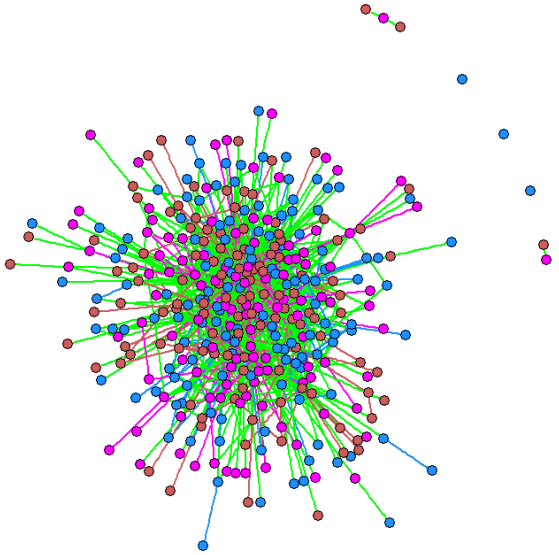

Figure 5: Exploratory social network graph of Saving Schools threads with >10 responses

We plot the network in Figure 5. We label nodes of liberal forum posters as blue, moderates as magenta, conservatives as red, and exclude non-respondents. Edges represent direct responses in threads with at least 10 posts. Edges are same color for matched nodes, and green otherwise. Nodes are laid out using the Fruchterman-Reingold layout. We attempted numerous community detection techniques including stochastic blockmodels, mixed-membership stochastic blocks and latent space models. None yielded a clustering consistent with political ideology, and the assortativity coefficient is small (0.037 using the continuous ideology measure).

One metric of engagement across partisan difference is the expected number of replies for a given poster. If a post generates only responses from authors of the same ideological type, it suggests little opportunity for bridging dialog. Figure 6 shows the patterns of posts and replies by post-type category. Along the x-axis we have the poster ideology, represented in a common U.S. tripartite division of liberal, moderate, and conservative, and the y-axis gives the expected number of replies by the replying author's ideology. Particularly in the high salience setting, an individual posting is likely to get at least one response from another forum participation with different beliefs. This is a necessary if not sufficient condition for bridging dialog.

Posting in a thread with politically diverse perspectives indicates engagement with thread participants but it does not necessarily imply acknowledgement of alternative arguments or perspectives. For this we turn to upvote behavior, which is how forum participants signal their positive response to specific posts. While there are multiple possible meanings to an upvote, we view it as an explicit action to promote the visibility of a post and thus the poster's argument. If an individual upvotes a post written

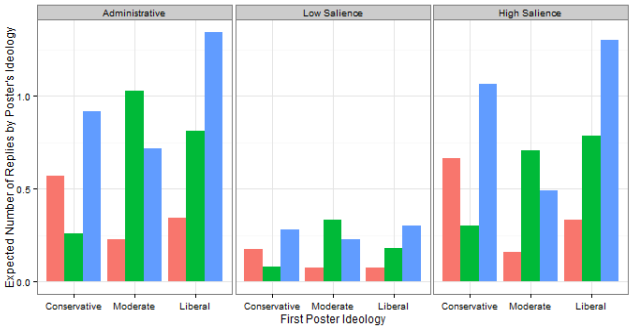

Figure 6: For each category of post, the expected number of comments per response post by response poster's ideology and commenter's ideology.

by someone of a different ideology in a politically salient thread, we take this as an indicator of a bridging action and an acknowledgment of alternative views.

Upvoting is a less frequent activity than posting within this particular class. There are 822 upvotes in the 76 threads we consider. Each thread contains 10.8 upvotes on average (coming from 5.7 unique voters) in comparison to 101 posts per thread on average. We have the voter's and the original poster's political ideology score for 285 upvotes.

To investigate whether people upvote across political differences, we look at upvotes where the voter and the post author have political ideology scores that are more than a standard deviation apart. For each thread for which we have both the original poster's and the upvoter's political scores, we calculate the percentage of upvotes that are made by a voter whose political ideology differs from that of the poster. The average thread-level percentages indicate a high rate of difference upvotes: 55% (administrative threads), 48% (low salience threads), and 49% (high salience threads). These rates are all consistent with rates we would expect if upvotes were assigned randomly across upvoted posts. This suggests that participants upvote posts that reflect an opposing ideology at the same rate as they

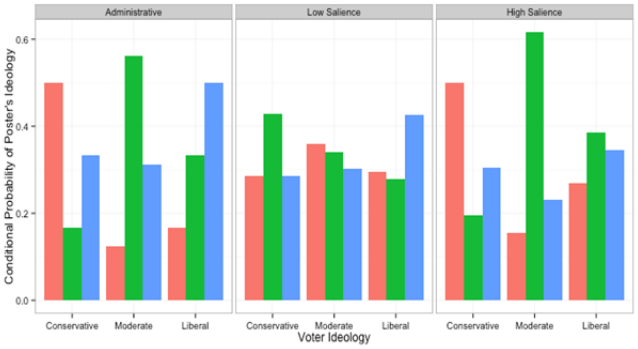

Figure 7: Conditional probability of poster ideology by upvoter ideology among upvotes in Saving Schools

upvote posts reflecting their own ideology. These results hold for metrics of assortativity on the bipartite graph of voters and posters. The assortativity coefficient on the continuous ideology score is 0.049.

We can also conceptualize upvote patterns in terms of categories of ideological affiliation. We split the voters and posters into terciles for conservative, moderate, and liberal. For each type of thread, we compute the conditional probability of the poster's ideological label given the voter's ideological label, displayed in Figure 7. We notice that the probabilities are roughly equal in the low salience group, and a chi-square test fails to reject the null distribution of independence. This suggests engagement across partisan differences. For the high salience threads, moderates show a clear preference for upvoting the posts of moderate posters. Conservatives show a preference for upvoting the posts of conservatives but also are willing to bridge, while liberals show a preference for moderates and other liberals. Interestingly, the administrative threads show a tendency for each group to vote for their own type.

The overall picture of forum activity in *Saving Schools* is one of enagement across diverse beliefs. Threads with more than 10 posts almost universally include an ideologically balanced mix of contributors. When looking at upvoting behavior, we find evidence that students indicate approval of posts by other students with substantially different political beliefs. We see some ideological preferences in upvoting behavior, but also upvoting of posts by peers with differing beliefs. Overall, the structure of forum activity suggests that in this one politically-salient MOOC, ideologically diverse students engage directly with one another over politically contentious issues.

Language Use by Political Affiliation

In this section, we investigate whether people with different beliefs use different language to discuss similar topics, or whether they converge on a shared language. We hypothesize that measurable differences in language use by partisan groups may represent situations where partisans "talk past one another," discussing the same issue in different terms that limit meaningful exchange. To examine these patterns, we a form of unsupervised text analysis from the topic modeling tradition [2] called the Structural Topic Model [14, 16]. Topic models are designed to identify sets of words, "topics," that tend to occur together.

We focused our analysis on two different forum threads that dealt with politically contentious issues: the Common Core State Standards and vouchers for private schools. For both threads we estimated a separate STM to evaluate whether particular topics are more likely to be discussed by students from one side of the political ideology scale. Differences in the distribution of topic usage by one partisan group may be evidence of fracturing discussions.

Common Core

Students were asked to respond to a prompt about the Common Core: "Based on this week's videos and readings, what is your opinion of the Common Core debate? Do you support states adopting the Common Core standards or do you oppose them doing this." From this data we have 195

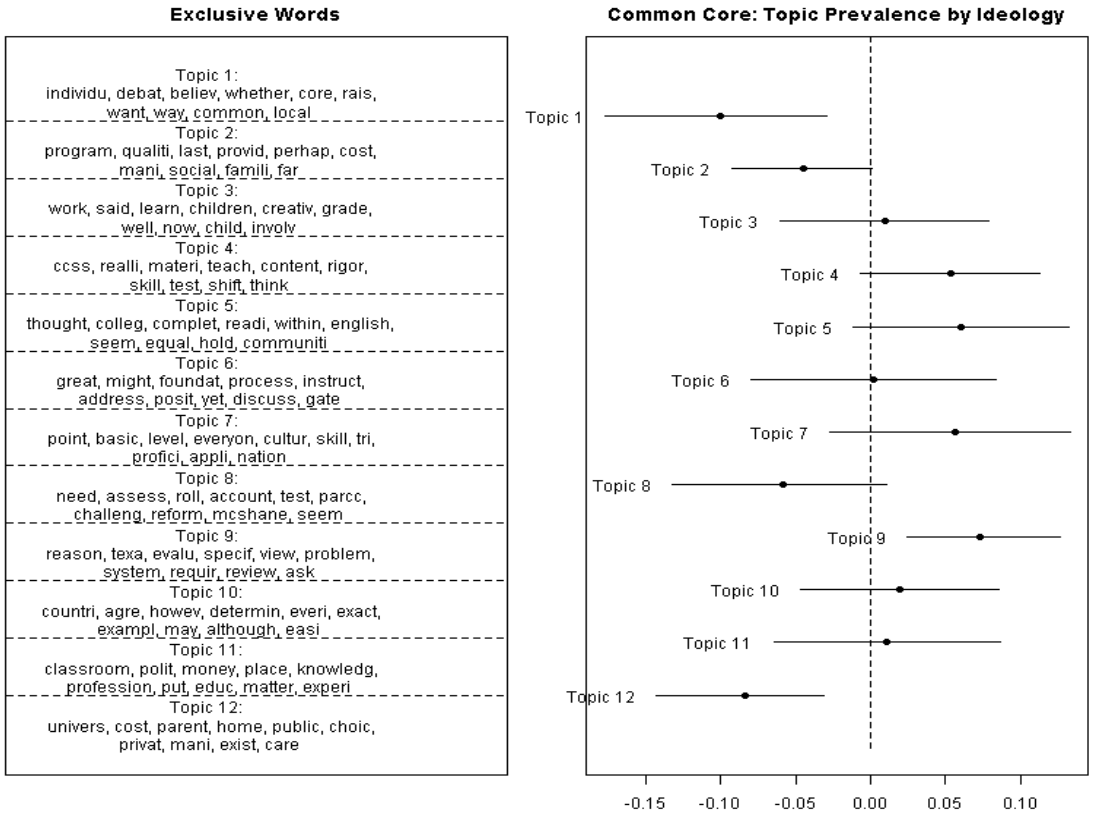

Figure 8: Results of Structural Topic Model Analysis of 195 posts in Saving Schools thread about Common Core State Standards.

posts for which we have political ideology information on the poster. Figure 8 plots two sets of results from a 10-topic model for the Common Core thread. On the left side are sets of words, known in the literature as FREX words [1, 16], that are associated with a specific topic while being less likely to appear in other topics. We use these FREX words as well as examples of posts highly representative of particular topics (as identified by the STM model) to semantically interpret these topics. For example, topic 3 deals in part with individual student needs and the need to encourage creativity. Topic 8 focuses on the role of tests and assessments in reforms based on the Common Core.

An advantage of the Structural Topic Model over other topic modeling algorithms is that we can incorporate information about the person who makes a forum post. In particular, we analyze whether liberals and conservatives write about different topics more or less frequently. The STM estimates several important quantities. First, like other topic models, it estimates a set of topics, which are defined as a set of words that tend to occur together. More specifically, it estimates a probability distribution where for each topic, each word has an associated probability of belonging to that topic. Second, and unique to the STM, it estimates the relationship between metadata that we know about a post, here whether it was written by a liberal, moderate, or conservative, and the proportion of the post belonging to a particular topic. Due to our focus on these political categories, we created a tripartite coding of the ideology data. We classified conservatives as those between 0 and the 33rd tercile, moderates between the 33rd and 66th tercile, and liberals from the 66th to 100th tercile.

We investigated representative posts from the three topics that did see a difference in frequency among liberals versus conservatives. Topic 9 focused on problems with the Common Core, and it was more likely brought up by liberals. As one student wrote in a representative post, "Actually, some [of] the greatest criticisms of the CC come from K and early elementary teachers -- saying that the standards are too high for student development (and they might be totally right, as no K-3 teachers were involved in

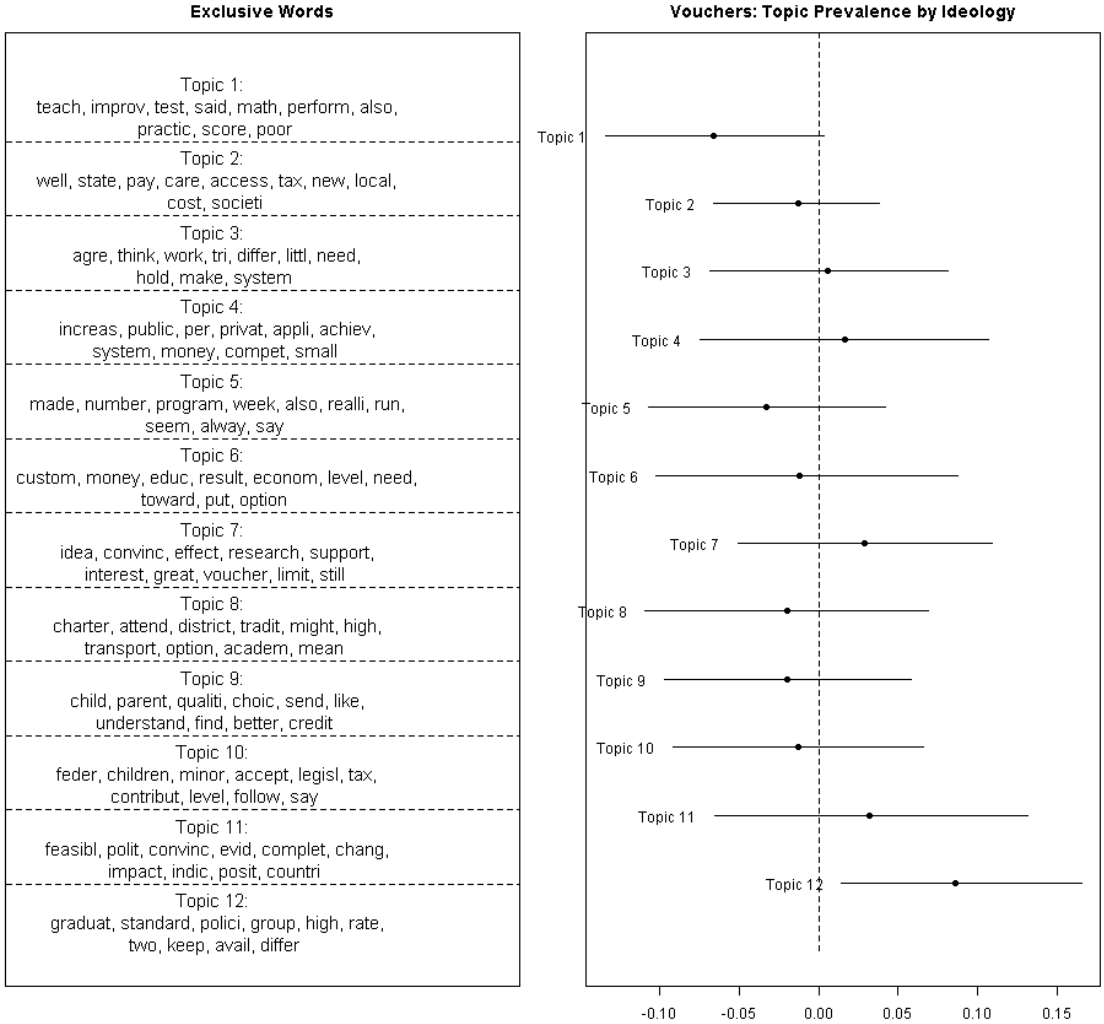

Figure 9: Results of Structural Topic Model analysis of 155 posts about vouchers.

the development of the standards)." Close inspection of these posts reveals that not every post is actually critical, however, and some address criticisms positioned by others. For example, one student responded to another, "As a public high school teacher, I totally agree with you that diversity and letting parents take the lead is absolutely important (and, clearly, that is something that you have set as a priority for your own children). However, the Common Core standards are about public schools -- cases in which parents have not chosen to take the lead." These examples show that even in topics where we see some quantitative evidence of partisan difference in topic distribution, meaningful exchange can be occurring.

Topics 1 and 12 were more associated with conservative posters. Topic 12 dealt with issues of costs, federal and state funding, and school choice. For example, one post read "NCLB is still on the books, so this will not permit some states to benefit from federal funds unfairly and without solvency while others balance budgets and invest in youth. Meanwhile, those states that CHOOSE to act by doing nothing will be identified as such. Liberty and freedom means CHOICE. But choice can be good and bad." Topic 1 dealt with how the debate about the common core had become politicized at the national level. For instance, a poster wrote, "Diversity and letting parents take the lead will always win out here against national curriculum that gets dumbed down to the lowest common denominator and politicized so as not to offend any group or to cater only to the largest sections of the population."

While these differences reveal some differences in how Liberals and Conservatives emphasized different language in the Common Core debate, overall, there were few significant differences in the topic distributions among students with different political beliefs in this thread.

Vouchers

Our second analysis focuses on a discussion thread about vouchers. Our results identified patterns of ideological engagement similar to those in the Common Core thread. For this thread, the discussion prompt was "Are you convinced by the evidence presented in the videos and readings in support of vouchers? If not, where do you think the evidence is lacking? If you are convinced, do you think vouchers are politically feasible?"

We analyzed the 155 posts for which we had a political ideology score using a Structural Topic Model with 12 topics. The left hand side of Figure 9 presents the words most exclusive to each topic. The right hand side presents estimates of whether the prevalence of a topic in a forum post varied between conservatives and liberals. Across all 12 topics we see only two significant differences. Topic 1, which dealt with opportunities for improvement in teaching practices, was more likely to be brought up by conservatives. Topic 12, on the other hand, was more likely to be brought up by liberals. This topic dealt with the need to keep graduation standards high. As with our analysis of the discussion about the Common Core, our text analyses reveal minimal differences in patterns of language used by liberals and conservatives. Tentatively, we present this as evidence that students with different political beliefs are engaging directly with one another in the forums, using a common set of themes and discussion topics.

In these threads about the Common Core and vouchers, and every other individual thread that we examined, we found little evidence of different distributions of topics among posts from liberals and conservatives. One possible interpretation of these findings is that our text analysis instruments are not sensitive enough to pick up on important differences in language use. Another interpretation is that in the forums, a meaningful exchange of ideas on controversial topics using a shared language is taking place. In the context of our other findings—the even distribution of political beliefs among students, the even forums participation of students with different beliefs, and the lack of evidence of assortative mixing in forums—we believe the latter interpretation is plausible.

DISCUSSION

Saving Schools represents a potentially surprising example of a politically-charged MOOC, taught by a professor with transparently partisan positions, that engages a politically diverse community of online students. Contrary to the concerns of observers that the internet has become a place of echo chambers and silos, we find evidence that Saving Schools is a space in the contentious debates over U.S. education reform where people with different opinions can learn and engage together. We found that the student body, at least among U.S. students, contained participants with diverse education policy preferences. Only a subset of participants chose to engage in online forum discussions, but the subset that did so was politically representative of the full set. Within forums, we found that most threads contained a balanced proportion of liberal and conservative posters, that liberals and conservatives directly responded to each others' posts, and that liberals and conservatives upvoted posts across partisan lines. Text analysis of student forum posts suggests that students with different political beliefs tend to discuss similar topics in roughly equal proportion. We find little evidence of students segregating themselves within rhetorical frames that inhibit meaningful discussion. Returning to our two-by-two schematic of online discourse from Figure 1, we argue that Saving Schools presents a case study where at least the pre-conditions of deliberative discourse appear to be met.

The most obvious limitation of this current work is that we lack any kind of "ground truth" with which to calibrate our instruments designed to measure engagement across political difference. In future studies, we hope to triangulate our forum analysis and text analysis with more robust and systematic qualitative evaluations of forum threads, conducted by coders with expertise either in online learning or in the substantive subject of the discussion. We also hope

to include self-reported data from students about their experiences and their perceptions of engagement across political differences in the course. Finally, we hope that in future case studies, we might find evidence of assortative mixing, and then use experimental interventions to increase engagement across political differences and evaluate how our measures respond to those interventions.

Ultimately, our hope is that greater research and attention to non-cognitive and civic outcomes in MOOCs can broaden the conversation about the purposes of open online learning. Historically, public education has not only served the purpose of developing young people for professions but also for their roles as citizens in civil society. MOOC research should engage with questions as broad as our hopes for higher education.

ACKNOWLEDGMENTS

We gratefully acknowledge grant support from the Spencer Foundations New Civics initiative. We also thank the Saving Schools course team, the Harvard VPAL-Research Group for research support, and Lisa McKay for edits.

REFERENCES

1. Jonathan M. Bischof and Edoardo M. Airoldi. 2012. Summarizing topical content with word frequency and exclusivity. In *International Conference on Machine Learning* (ICML), 201–208. http://www.icml.cc/2012/papers/113.pdf
2. David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research* 3: 993-1022.
3. Max Chafkin. 2014. Udacity's Sebastian Thrun, Godfather Of Free Online Education, Changes Course. Retrieved October 27, 2015 from http://www.fastcompany.com/3021473/udacity-sebastian-thrun-uphill-climb
4. *Education Next*. 2015. Results from the 2015 *Education Next* Poll. Retrieved October 27, 2015 from http://educationnext.org/2015-ednext-poll-interactive
5. Andrew Dean Ho, Isaac Chuang, Justin Reich, Cody Austun Coleman, Jacon Whitehill, Curtis G Northcutt, Joseph Jay Williams, John D Hansen, Glenn Lopez, and Rebecca Petersen. 2015. HarvardX and MITx: Two Years of Open Online Courses Fall 2012-Summer 2014. HarvardX & MITx Working Paper No. 2.
6. Howard Gardner, Katie Davis. 2013. *The app generation: How today's youth navigate identity, intimacy, and imagination in a digital world.* Yale University Press
7. Joseph Kahne, Ellen Middaugh, Nam-Jin Lee, and Jessica T. Feezell. 2012. Youth online activity and exposure to diverse perspectives. *New Media & Society* 14,3: 492-512. doi:10.1177/1461444811420271
8. Gary Orfield, John Kucsera, Genevieve Siegel-Hawley 2012. E pluribus… separation: Deepening double segregation for more students. Working Paper.
9. Eli Pariser. 2012. *The filter bubble: How the new personalized Web is changing what we read and how we think.* Penguin Books.
10. Paul E. Peterson. 2010 Let the Charters Bloom. Retrieved August 7, 2015 from http://www.hoover.org/research/let-charters-bloom
11. Paul E. Peterson. 2015. No Child Left Behind and testing help hold schools accountable. Retrieved August 7, 2015 from http://www.latimes.com/opinion/op-ed/la-oe-peterson-nclb-defense-of-testing--20150224-story.html
12. Paul E. Peterson, Michael Henderson, and Martin R. West. 2014. *What Americans Think about Schools and How to Fix Them*. Brookings Institution Press.
13. Justin Reich, Anna L. Romer, and Dennis Barr. 2014. Dialogue across difference: A case study of Facing History and Ourselves' Digital Media Innovation Network. In *#youthaction: Becoming Political in a Digital Age*, Ben Kirshner and Ellen Middaugh. (Eds.). Information Age Publishing.
14. Justin Reich, Dustin Tingley, Jetson Leder-Luis, Margaret E. Roberts, and Brandon M. Stewart. 2015. Computer-Assisted Reading and Discovery for Student Generated Text in Massive Open Online Courses. *Journal of Learning Analytics*. 2, 1:156-184.
15. Howard Rheingold. 2000. *The virtual community: Homesteading on the electronic frontier*. MIT Press.
16. Margaret E. Roberts, Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G. Rand. 2014. Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science.* 58, 4: 1064-1082.
17. George Siemens. 2015. Keynote Address at Learning with MOOCs Conference, Teachers College, Columbia University.
18. Spencer Foundation. 2015. Computational Measures of Engagement Across Difference in Online Courses. Video. (2 October 2015). Retrieved October 27, 2015 from https://www.youtube.com/watch?v=9v_vJJfdqIY
19. Glenda S. Stump, Jennifer DeBoer, Jonathan Whittinghill, Lori Breslow. 2013. Development of a framework to classify MOOC discussion forum posts: Methodology and challenges. In *NIPS Workshop on Data Driven Education*.
20. Miaomiao Wen, Diyi Yang and Carolyn Penstein Rośe. 2014. Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the International Conference on Weblogs and Social Media*, 525-534