



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

Michael Yeomans¹ · Brandon M. Stewart² ·
Kimia Mavon¹ · Alex Kindel² · Dustin Tingley¹ ·
Justin Reich³

Published online: 29 November 2017

© International Artificial Intelligence in Education Society 2017

Abstract Massive open online courses (MOOCs) attract diverse student bodies, and course forums could potentially be an opportunity for students with different political beliefs to engage with one another. We test whether this engagement actually takes place in two politically-themed MOOCs, on education policy and American government. We collect measures of students' political ideology, and then observe student behavior in the course discussion boards. Contrary to the common expectation that online spaces often become echo chambers or ideological silos, we find that students in these two political courses 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.

Keywords MOOCs · Civic education · Discourse · Text analysis · Political ideology · Structural topic model

Introduction

Political theorists have long argued that exposure to diverse perspectives is vital to a robust civil society and to the development of citizens (Kahne et al. 2012). Democratic discourse requires engaging with people who hold different perspectives, and while

✉ Michael Yeomans
yeomans@fas.harvard.edu

¹ Harvard University, Cambridge, MA, USA

² Princeton University, Princeton, NJ, USA

³ Massachusetts Institute of Technology, Cambridge, MA, USA

networked technologies hold great promise for bringing people together, new technologies also have frightening capacity to separate people into ideologically segregated online spaces.

Massive Open Online Courses (MOOCs) are potential sites for students to engage with peers who hold differing beliefs, but the scope and scale of discussions among thousands of students makes tracking these interactions in detail a challenging task for faculty and course teams. In this work, we have prototyped a series of computational measures of engagement across difference: measures that could be deployed in near real-time in courses that give faculty some insight into the nature of participant interactions within and across political affiliations.

Background and Context

Internet researchers have posed two competing theories for how people confront differences on the Web (Gardner and Davis 2013). One theory holds that the Internet is a series of “silos” where individuals seek out media and communities that conform to their established beliefs (Pariser 2012). Another theory holds that the Internet contains many interest-driven spaces that serve as ideological “bridges” (Rheingold 2000), where people attracted to these interest-driven spaces can be diverse across many dimensions. Previous research has examined how technology-enabled platforms - social networks, web browsing, news aggregation - affect consumption of political content (Garrett 2009; Gentzkow and Shapiro 2011; Athey and Mobius 2012; Flaxman et al. 2016; Quattrocioni et al. 2016). This prior work generally suggests a picture of technology at odds with healthy civic discourse, in which users seek and find ideologically homogenous news sources, rather than exploring the diversity of available viewpoints (Sunstein 2017; though see Boxell et al. 2017). The media environment surrounding the 2016 U.S. Presidential Election shows a stark example of these divides (Faris et al. 2017).

The rise of open online education has potentially offered a new pathway for students to join communities of diverse learners and actively participate in political discourse with others (Reich et al. 2014). 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 (Chuang and Ho 2016). MOOCs have the potential to bridge the geographic patterns that divide students in brick-and-mortar schools along ideological lines (Orfield et al. 2012). But this optimism is tempered by several important questions that have yet to be answered empirically. Does the demographic diversity in MOOCs translate into ideological diversity? Do students use these learning communities to encounter and consider different perspectives, or are student interactions limited to communicating primarily with other students who share pre-existing ideological perspectives?

This work extends a budding literature on the discourse within online learning environments and beyond. Peer interactions are central to the pedagogical designs of online education (Siemens 2005). However, most of the previous research on discourse in MOOCs from edX and Coursera has focused on how in-course language relates to student persistence and dropout (Koutropoulos et al. 2012; Wen et al. 2014; Yang et al. 2015).

In our work, rather than focusing on how forum activity and language use predicts student performance, we are interested in peer discussion as an educational end itself.

Effective citizenship education and a healthy civic sphere require opportunities for public deliberation (Della Carpini et al. 2004). One prerequisite for health public deliberation in MOOCs is that people with differing beliefs should engage one another directly, which we can evaluate by measuring the political beliefs of students and examining whether students with different beliefs respond to one another in forums. We can also investigate the quality of deliberation by examining language use directly among political partisans. Political psychologists have observed that political partisans often shape debates through the use of competing “framings” for issues (Lakoff 2014), such as defining estate taxes as death taxes, or referring to tax cuts as tax relief (or, in the education policy space, the recent shift from away from “vouchers” and towards “funding that follows the student”). We hypothesize that extensive partisan framing would lead to divergent language whereas a convergent language would indicate that these kinds of rhetorical moves—often aimed at winning arguments rather than understanding others—were not excessively shaping forum discourse. Other interpretations of language convergence are certainly possible, Welbers and de Nooy (2014) use conversational accommodation theory, which observes that people adjust to their conversation partner’s verbal and non-verbal behaviors, to evaluate convergence in an ethnic group discussion forum. Conversational alignment is an important line of inquiry in linguistics, and new computational methods are advancing the field (Doyle and Frank 2016). Our work extends a normative perspective on this work, identifying the “inadequacies of existing discourse relative to an ideal model of democratic deliberation” (Gastil 1992). Our data allow us to test whether MOOCs breed a discussion that aligns ideologically distant students in their forum behavior and in their language use. These questions of how discourse in online forums promotes democratic deliberation have been addressed previously through hand-coding of forum interactions (Loveland and Popescu 2011), and we build on this literature by developing fully automated approaches.

In this paper we tackle this broad agenda by providing novel evidence for four research questions. First, do MOOCs attract enrollment from an ideologically diverse student body? Second, is that diversity reflected among students’ participation? Third, do students selectively interact in the forums (via comment or up-vote) based on partisanship? Finally, do students converge on a shared language, or does their discourse remain divided along partisan lines? Each of these questions identifies a unique and necessary prerequisite for engagement across differences in MOOCs.

The rest of this paper proceeds in three parts. We first introduce the two courses, their students, and their political beliefs. Then, we then describe the student interactions in course forums. Finally, we analyze the text of the posts, to evaluate the degree to which students with diverse political beliefs converge on a shared language and interact constructively with one another.

Methods

Our data are collected from two online courses run by HarvardX on the edX platform. Each course was taught by a Harvard professor and modeled on a real campus course. Course material was released in weekly chapters over 3–4 months. Students were also asked to complete a pre-course survey at the beginning of the course, which included

measures of political ideology. Both classes also had a discussion forum, in which students were asked to post regularly as part of the requirements for completing the course. However, the course administrators did not specifically check the students' posts to confirm their completion credit. Additionally, while this directive might affect the volume of posting, the nature of those posts is still determined by the students - including the contents of the posts, and the placement of posts within the forum discussion.

Courses

Saving Schools was 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 was 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 on 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, which they confirmed through an honor-based self-assessment.

The politics of U.S. education reform do not perfectly align with conservative/liberal distinctions, but the education policy preferences of the professor - Paul Peterson - 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. Prof. Peterson is a proponent of free market reforms, school and teacher accountability, charter schools, and standardized testing; and he has been critical of policies advocated by labor unions and schools boards. Our informal assessment of Saving Schools is that Prof. Peterson provides multiple perspectives on issues and gives each side a fair hearing, though he also makes clear his own, generally conservative, policy preferences.

American Government was a course about the institutions of American politics that ran from September 2015 to January 2016, taught by Harvard Kennedy School of Government faculty member Thomas Patterson. Patterson is an expert in media and public opinion. The course topics ranged from constitutional structures, political parties, the role of the media, and other elements of the national political system. The course contained 24 modules released over four months. Each module included

discussion questions; for instance, in the first unit on dynamics of American power, students were asked to discuss the prompt: “‘Money is power’ in the American system. Explain some of the ways that money is used to exert influence and who benefits as a result.” As with Saving Schools, students confirmed their participation in forums through an honor-based self-assessment. The course also used mechanisms to divide students into discussion forum cohorts by last name. Our assessment is that the selection of course topics conveys some bias or emphasis on center-left interests—concerns with income inequality and money in politics for example—in the context of a largely non-partisan explication of how American government functions. If Saving Schools tilts right, while largely providing a balanced perspective on issues, American Government tilts left, while largely remaining non-partisan.

Population of Interest

While total enrollment for these two online courses was 30,006, most of these enrollees do not actually engage with the course content - only 16,169 did anything more than enrolling through the course website. Of those students only 7,204 started the precourse survey, as one indicator of introductory activity. In Table 1, we use data collected from outside the survey to compare the demographic composition of the survey respondents, relative to non-respondents.

Among this population of survey-takers, 45.8% reported being from the United States. While these courses drew diverse international interest, we focus on these U.S. students in all our analyses below, for three reasons. Most practically, the survey in Saving Schools did not display the ideology questions to non-U.S. students (by instructor design). Second, these U.S. students form a clear plurality of the student body, and the subject matter mostly focuses on U.S. politics. Finally, there may be cultural differences in the ideological foundations of partisanship, and our measures may not capture the true diversity in our international participants’ points of view. Therefore, our theoretical interpretation of ideological differences will be most clear among the subset of students that are from the United States.

Table 1 Demographics of course participants

	Saving Schools				American Government			
	Course enrollees	Survey responses	USA sample	Any forum activity	Course enrollees	Survey responses	USA sample	Any forum activity
Observations	5,409	2,009	1,340	569	10,760	5,195	1,963	395
% Female	49.8%	57.7%	61.0%	59.4%	35.3%	39.3%	46.5%	44.6%
College Degree	79.9%	83.3%	86.2%	88.1%	58.3%	62.8%	58.7%	67.6%
Median Age (SD)	32 (14.2)	36 (14.5)	38 (14.9)	40 (15.1)	28 (13.4)	29 (14.5)	32 (17.2)	40 (17.8)
% Teachers		52.9%	57.7%	63.4%		17.9%	19.2%	30.1%
Intended Course Completion		57.2%	56.0%	75.4%		41.3%	38.4%	65.6%
Actual Course Completion	14.3%	31.8%	32.7%	61.0%	5.6%	10.0%	8.3%	28.6%

Political Beliefs

Students' ideology was measured during the pre-course survey in both courses. Almost all U.S. students who took the survey completed these items. However, the measures were different in each course, as we describe below. Broadly, to analyze American Government ($n = 1,258$) we used a single measure of generic political leaning, while in analyzing Saving Schools ($n = 1,315$) we used a four-item measure of political leaning within specific topics in education policy, which was transformed into a single ideology dimension. In all our analyses we use the continuous ideology measure. However we also divide the ideology dimension into a tripartite categorization (i.e. "liberal", "moderate", "conservative") for graphical representation.

Saving Schools Students were given a set of questions used previously in a nationally representative poll on education policy (Peterson et al. 2014; *Education Next* 2015). Many education policy issues do not map on to typical left/right divides in American politics, so we chose the four questions that were most strongly correlated with broader measures of political partisanship: questions about school taxes, school vouchers, unions, and teacher tenure. In Table 2 we show the responses in our target sample (i.e. US-based survey respondents who posted in the forum), along with the responses from the original, nationally-representative survey. In general, our sample covered a wide range of political viewpoints across the US population.

These responses generally mapped onto a single dimension of partisanship, as shown in Table 3 (note that the tenure question has an opposite sign from the others). That is, people who were generally on the "left" in terms of education policy were

Table 2 Responses by Saving Schools U.S. forum posters to education policy questions. Answers from the original nationally representative poll in italics

Question	Multiple choice response				
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 the same	Decrease	Greatly decrease
	6.6%	28.2%	17.8%	30.7%	16.7%
	<i>14%</i>	<i>44%</i>	<i>34%</i>	<i>6%</i>	<i>2%</i>
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	Somewhat oppose	Strongly oppose
	18.8%	42.9%	32.8%	3.8%	1.8%
	<i>14%</i>	<i>27%</i>	<i>19%</i>	<i>23%</i>	<i>17%</i>
Do you favor or oppose giving tenure to teachers?	Strongly favor	Somewhat favor	Neither	Somewhat oppose	Strongly oppose
	18.3%	30.7%	6.3%	22.9%	21.8%
	<i>6%</i>	<i>22%</i>	<i>30%</i>	<i>21%</i>	<i>21%</i>
Do you think teacher unions have a generally positive or negative effect on schools?	Strongly positive	Somewhat positive	Neither	Somewhat negative	Strongly negative
	9.5%	20.6%	7.7%	36.4%	25.8%
	<i>7%</i>	<i>23%</i>	<i>23%</i>	<i>16%</i>	<i>30%</i>

Table 3 Correlation matrix of ideology measures in Saving Schools

	Vouchers	Tenure	Unions
Taxes	+0.251	-.245	+0.534
Vouchers		-.158	+0.119
Tenure			-0.212

more likely to support higher taxes, tenure, and unions, while people who were generally on the “right” were more likely to support vouchers. This aligned with our understanding of the general terms of debate in current policy, so we condensed these four measures into a single-dimensional measure of ideology. We standardized the responses to each question (by subtracting the mean and dividing the standard deviation), reverse-scored the tenure question, and averaged the four measures into a single ideology index. Additionally, we confirm that our results are robust across alternate mappings to a one-dimensional ideology scale (such as the first principal component from a principal component analysis). To create tripartite categories, we divided this continuous measure into equally-sized terciles - that is, with a third of all students in each bucket. Due to the liberal skew of the student population, this meant that the liberal third was more strongly liberal than the conservative third was conservative.

American Government Students were asked a single item which targeted a more general measure of ideology, taken from the World Values Survey (2009). Participants answered the following question:

“In political matters, people talk of “the left” and “the right.” How would you place your views on this scale, generally speaking?”

They responded on a ten-point scale, and the distribution of responses is shown in Table 4, along with the results from the most recent US sample of the World Values Survey. In general, our sample provides good coverage of the ideological spectrum. The central tendency of the distribution is, on average, more left-leaning than the population at large. However, our analyses are more concerned with the range of views, and these results suggest that the student body does contain a diversity of viewpoints. To create tripartite categories, we divided the discrete scale points into three approximately equal groups: 1–3 (“liberal”), 4–5 (“moderate”), and 6–10 (“conservative”).

These results address the first of our four research questions concerning who participates in politically-themed MOOCs. That is, both classes managed to attract students that represent a diverse range of political beliefs, across the ideological spectrum. Though the demographic diversity of MOOCs has been well-established, to our knowledge, this is the first direct evidence for ideological diversity. And this is a necessary (but not sufficient) condition for engagement across differences. In the

Table 4 Ideology responses by American Government U.S. forum posters. Answers from the original nationally representative poll in italics

	Left	2	3	4	5	6	7	8	9	Right
Course Sample	7.4%	10.2%	14.2%	12.4%	25.1%	9.7%	7.5%	6.2%	3.4%	3.9%
<i>Original Poll</i>	<i>2.0%</i>	<i>3.0%</i>	<i>6.5%</i>	<i>7.2%</i>	<i>34.6%</i>	<i>17.0%</i>	<i>9.2%</i>	<i>9.1%</i>	<i>5.8%</i>	<i>5.6%</i>

remainder of the paper, we merge these data with the forum activity logs to test our three remaining research questions, which concern ideologically-driven participation and interaction in the course discussion forums.

Forum Results

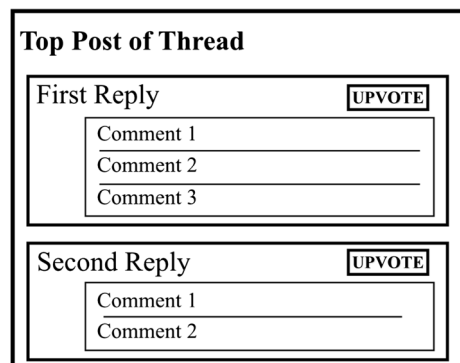
The structure and function of the discussion forums in both classes were identical, and this structure is displayed in Fig. 1. The top level of the forums included “threads”, and the forum structure allowed for up to two other levels of posting below each thread. Specifically, students could post “replies” to an original post, and each reply could be followed by an arbitrarily long list of “comments” in chronological order. Thus, the discussion threads had three levels of responses: initial posts, replies to initial posts, and comments on replies. Additionally, students were allowed to “upvote” threads and replies (but not comments), which promoted posts and replies to a higher position in the thread.

In general, each thread was self-contained, with no interaction across threads. Likewise, the replies within each thread were also self-contained, in that replies almost never responded to other replies. Thus, almost all student-to-student interactions were nested as comments within each reply, since as we shall see the vast majority of initial posts with active discussion were those started by course staff. After a student posted a reply, other students could interact with that reply by giving an upvote or adding a comment.

In our analyses, we assume that every post is directed towards the post above it in the thread - all replies are directed to posts, and all comments are directed to replies. This interpretation is in line with the intent of the commenting platform design. But in practice this was not always the case. For instance, some commenters direct their comments towards each other, rather than towards the post or reply above. In addition, some posts address a number of posters in a thread simultaneously (e.g. “let me try to synthesize four perspectives in this thread”) rather than referring to one person in particular. Still other posts were off-topic, and directed to no one at all. In these cases, the raw trace data from the forum might not match the posters’ intent.

To evaluate the extent of this mismatch - and to perform other basic qualitative coding tasks, like removing off-topic posts - we developed a new software tool: Discourse (Kindel et al. 2017). This tool handles many common formats for forum

Fig. 1 Thread Structure



data, allows individual posts to be read and rated within the context of the other posts in the thread that preceded it. We had a team of coders (at least two per post) read and rate 24,556 posts in the course-focused threads (see below) from both courses. We asked coders to classify posts according to the writers' intent in the context of the conversation. In general, their results confirmed that the metadata structure of the forums was close to accurate, with over 90% agreement between coder consensus and trace data. In fact, the trace data was as consistent with coders as the coders were with one another. So throughout we report analyses using the raw trace data. But we confirm that are substantively robust across other analytical strategies - either by assuming instead that the qualitative codes are ground truth, or else by focusing only on posts where humans and trace data agree.

Our first analytical strategy involves testing the “assortativity” of these reply-comment and reply-upvote interactions - namely, whether forum activity is self-sorted into ideologically consistent groups. To perform this test, we compare whether the ideology of a reply poster is at all predictive of the comments and upvotes they eventually receive (collapsed across the order of all threads and replies). Our second analytical strategy involves testing the partisan distinctiveness of individual posts. That is, we strip the forum context from replies and comments and treat each substantive forum post as an independent document, so that we can compare the language used by students with opposing ideologies.

Thread Types

Across both courses, we observe 2,125 threads, containing 16,522 replies, 9,889 comments, and 2,566 upvotes. But forum threads serve many different purposes in MOOCs (Stump et al. 2013; Wen et al. 2014), and forum actions were not distributed evenly across threads. Accordingly, we first consider two top-level categories of threads, based on who created the thread (student vs. course team) and the contents of the thread (administrative vs. low partisan salience vs. high partisan salience). The results of this categorization are shown in Table 5, and described below. For later analyses, we focus on course-generated content threads, and exclude both student-generated and administrative threads.

Table 5 Categorization of Thread Types. Cells indicate counts across both courses. Average forum actions per thread (comments, replies, upvotes) in parentheses

Thread Source		Coded Thread Category		
		Administrative	Low Partisan Salience	High Partisan Salience
Saving Schools	Course	34 (28.4)	43 (134)	11 (87)
	Student	71 (1.8)	185 (1.6)	153 (1.6)
American Government	Course	7 (328)	115 (137)	25 (148)
	Student	310 (0.9)	1074 (0.2)	97 (0.3)

Course Vs. Student Threads The most salient distinction is between “course threads” that were generated by a member of the teaching team, and “student threads” that were generated by the students themselves. For each chapter, the course team created a top-level “thread” in the discussion forum, and the class was usually broken into thirds (based on username or birthdate) to create smaller communities, which is a common MOOC practice (Baek and Shore 2016). The top post in these course threads provided a question about the most recent chapter, and students were encouraged to participate. Students started threads for two reasons. The first was to generate new topics of conversation. The second was that students sometimes tried to reply to a teaching-team thread, but accidentally created a new thread. The distributions in forum activity across these thread types were non-overlapping. Compared to student threads, course threads on average garnered far more replies (76.1 vs. 0.3), comments (40.7 vs. 0.2) and upvotes (9.9 vs. 0.1). The counts for student-generated threads were not much higher when we exclude all the orphaned threads that received zero comments or replies - specifically, even the subset of threads that had any activity still received only 1.5 replies, 0.9 comments, and 0.3 upvotes per thread, on average.

Thread Contents The forum threads could also be categorized in terms of their function and content. Some serve an administrative role, such as having students introduce themselves, making announcements or providing feedback from the course staff, or gathering logistical or technical questions about the course platform. The vast majority of threads focused on the content of the course itself. But even then, some threads touched on topics that were more political in nature, while others focused on less controversial topics. Though the raw activity levels were similar across thread types, we were particularly interested in how political ideology affects more- and less-politically-salient threads.

All threads (student- and course-generated) were grouped based on content into three categories. First, administrative threads were partitioned by the authors, and set aside. Second, content threads were divided into high and low political salience, according to the presence of a salient issue mentioned in the question posted by the course team to start the thread. These were coded by a research assistant, and confirmed by the authors’ own readings of the post contents. In *Saving Schools*, this coding captured the presence of themes in U.S. educational policy that were described in the course as controversial or worthy of further policy discussion, such as high stakes testing, the No Child Left Behind Act (NCLB), the Common Core, teachers’ unions, and charter schools, including the policies used in the ideology questions above. In *American Government*, this coding captured the presence of any of the twelve politically controversial issues identified as controversial in the Cooperative Congressional Election Study (Schaffner and Ansolabehere 2015), such as abortion, gun control, or tax rates. In both courses, the remaining threads were focused on conceptual or comprehension questions about educational and political institutions, and did not fall into an issue category. Accordingly, all threads that received an issue code were identified as having high partisan salience, while all threads that did not receive an issue code (and were not previously labeled as administrative) were identified as low partisan salience.

Ideology and Forum Participation

Of all the students who posted to the forums at least once, 39% of them were (i) from the U.S. and (ii) answered the ideology scale in the pre-course survey. These students accounted for 42% of forum activity in the focal (i.e. non-logistic course-created) threads. We did not try to infer the ideology of students who did not report their ideology. So the students that met all of these criteria formed the effective sample for all the following analyses of how partisanship affects forum participation.

Our second research question was whether students across the ideological spectrum were participating at similar rates. We tested this by comparing the rates of the three main forum activities (replying, commenting, and upvoting) across the entire course, among the U.S. students who reported their ideology. This analysis found essentially no relationship between political ideology and total forum activity (rank order correlation: $r_{\tau} = -.019$), and this result was consistent when we looked separately at replies ($r_{\tau} = -.024$), comments ($r_{\tau} = -.015$), and upvotes ($r_{\tau} = .007$). These relationships between ideology and forum activities are also plotted in Fig. 2 at the person level, separated by activity type.

It is possible that this diversity of activity at the course level might not translate to the thread level. That is, it is possible that liberals and conservatives participated equally overall, but did so within distinct sets of ideologically segregated threads. To investigate this, we calculated the average ideology of the posters (comments and

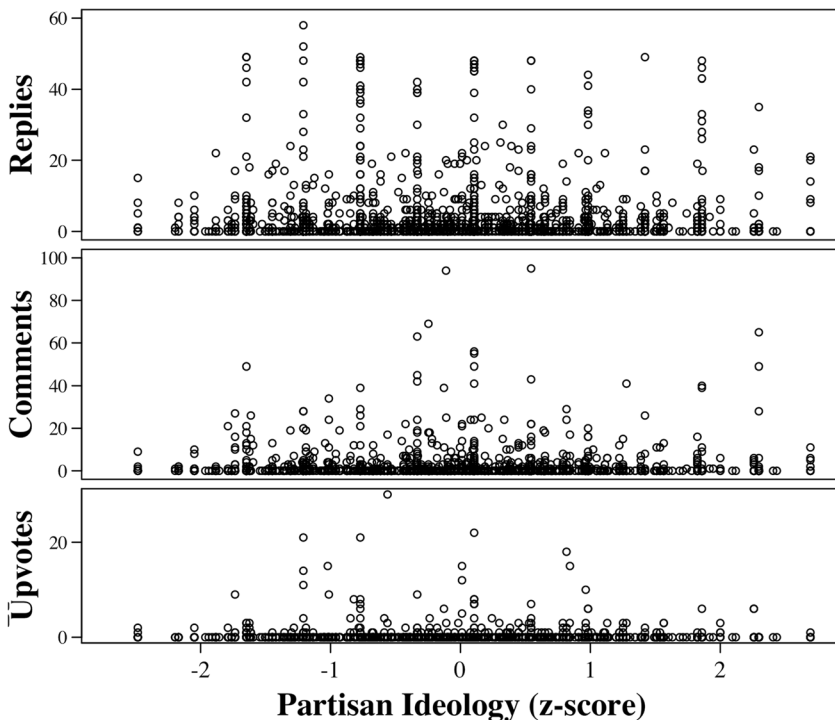


Fig. 2 Ideology and forum activity in substantive course threads

replies) in every substantive course thread. These averages are plotted in Fig. 3 - the x axis represents the average ideology of the posters (with zero being equal balance between liberals and conservatives), and the y axis represents the effective sample size of each thread (i.e. the number of posts for whom the poster's ideology is known). The results show that almost all the threads in both classes had a balanced ideological contribution, relative to the distribution of individual posters. There are some conservative-leaning threads, but they are small (under fifty posts) and limited to a small number of partisan topics. This result provides further evidence that these courses were rich in opportunities for students of different ideologies to interact with one another. For students to build “bridges” in online courses, it is first necessary that forum threads include a range of political perspectives, and this condition seems to hold in our data.

Ideology and Forum Interactions

These classes contain a diversity of viewpoints. But do students actually engage with their ideological opposites? We investigated the role of students' ideology in student-to-student forum interactions, specifically how other students respond to the replies at the top of each thread. There are two primary kinds of interaction - students can either upvote the replies directly, or they can write comments to the thread. Across both courses, we observe 400 reply-upvote pairs and 2,914 reply-comment pairs in which (a) both students were from the U.S. (b) both students answered the ideology scale, and (c) the pair occurred in a course-team-generated thread focused on course content. The analyses below focus on this sample of forum interactions.

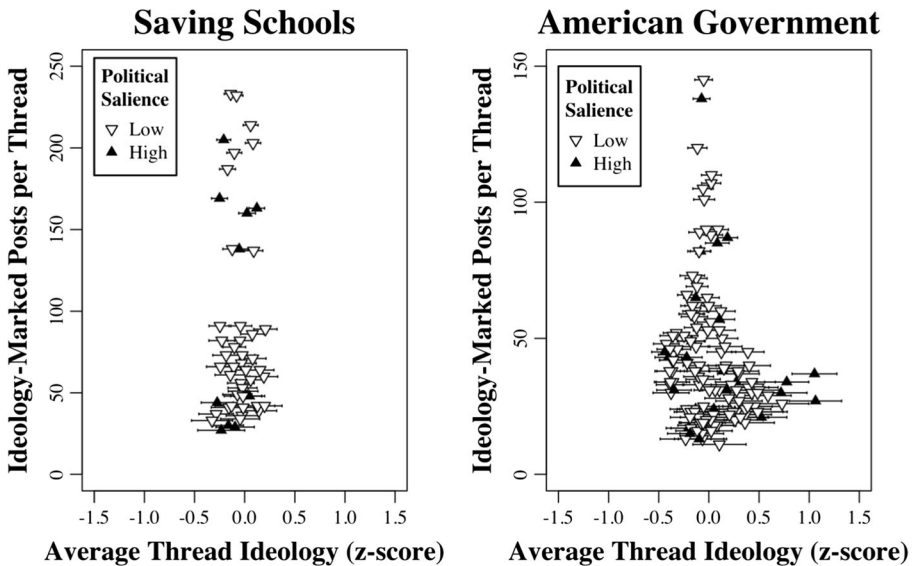


Fig. 3 Ideological distribution of posters in substantive course threads. The y axis shows the effective sample size (i.e. post for which ideology is known) and the x axis shows the average ideology among those posts. Error bars ± 1 SE

Upvotes Upvoting was uncommon, and only 10% of the replies in the focal sample received an upvote, for an average of 8.3 per thread. In all our analyses we remove all upvotes that students gave to themselves, since this could not reflect engagement with differing perspectives. We also dropped all upvote events that were immediately followed by an “unvote” event, since that was most likely the result of a mis-click rather than a true endorsement of the post.

If ideology had no impact on upvoting behavior, we would expect to see no correlation between the upvoters’ ideology and the original poster’s ideology. However, we do see some evidence of ideology-based assortativity in upvoting. In particular, American Government posters were more likely to gather upvotes from people with shared ideology than not (rank-order correlation: $r_{\tau} = .121, z(213) = 2.0, p < .05$), though this relationship was not apparent in Saving Schools ($r_{\tau} = -.019, z(155) = 0.43, p = .669$). In both classes we did not find moderation by thread type - that is, the assortativity of upvotes was constant within each class across substantive and partisan threads. We visualize these relationships in Fig. 4. For all focal threads in each course, we divide all identified upvote-reply pairs according to the tripartite classification (i.e. liberal, moderate, and conservative) of both the upvoter, and the original poster. This produces nine ideological pairings (3×3) for each course. These results suggest that the ideological sorting in American Government is stronger among conservatives, who upvote other conservative posts more than liberal posts, unlike liberals, who more evenly spread their upvotes across the ideological spectrum. We are reluctant to interpret the differences between classes, or between liberals and conservatives, though there are perhaps many class design choices that could affect these distributions.

Comments Commenting was more frequent than upvoting. The average thread had 45 comments, and 17% of all replies received at least one comment. Among these, we can exclude students who comment on their own replies, and students without ideology scores, to focus on the 2,587 reply-comment pairs that allow us to measure ideological influences on forum behavior.

Comments were not evenly distributed across replies - 83% of replies received zero comments, and most of the other threads only received one single comment. However,

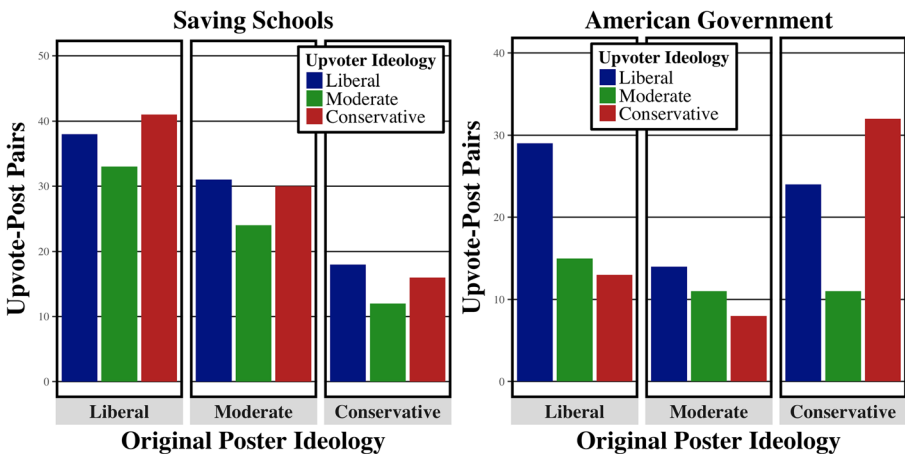


Fig. 4 Distribution of upvotes per reply in substantive course threads

some replies spawned comment sections that were much longer. Much of this heterogeneity was simply due to timing. Longer threads mostly sprung from the first or earliest replies, while replies posted later on were typically ignored (and this was most likely due to the design of the interface, which displayed earlier posts more prominently). Furthermore, these longer threads often included commenters interacting with one another. These complexities could affect our ability to measure ideology-driven connectivity between the main reply and individual commenters, especially later commenters. Accordingly, we also analyze ideological assortativity among only the first comments for each reply, as a robustness check.

The distribution of comment length was also uneven. Specifically, we noticed that many comments were disproportionately short, and this was especially true in the American Government threads - 34% of all comments in that course were under four words long, while that was true of only 2% of comments in Saving Schools (and less than 1% of other posts). Upon closer inspection, almost all of these short comments were simple content-free agreement (e.g. “well said”, “I agree”, “good points”, and so on). We suspect that many of these were written so that students could claim participation credit. This credit-seeking might also explain why a small fraction of posters were unaware of the topic of discussion (and in some cases, obviously plagiarized).

To filter these out, we recruited a team of six human coders to go through the course-created threads and, using Discourse, manually label each reply and comment as either (a) substantive, (b) a short yes, or (c) off-topic. Each thread was assigned to at least two independent coders, who each labelled every post in the thread, following the same order in which the posts were originally written. Any disagreements in the labels given by the first two coders were resolved by an independent third coder. Using the final labels, we removed the off-topic posts (0.5% of all posts), and separately counted the substantive comments (64% of on-topic posts) and short yes comments (36% of on-topic comments) as distinct forms of interaction between reply poster and comment poster.

In Table 6, we report rank-order correlations (with 95% confidence intervals) of ideology-based sorting among different subsets of reply-comment pairs culled from the

Table 6 Ideological sorting in reply-comment pairs, by comment type. Values reflect spearman correlations between poster's and reply-writer's ideology scores, with 95% confidence intervals

	Thread Type	Comments		
		Substance	Short Yes	First Comments
Saving Schools	Low Partisan	.034 [-.025,.093] (<i>n</i> = 1094)	.007 [-.327,.339] (<i>n</i> = 35)	.099 [-.020,.216] (<i>n</i> = 270)
	High Partisan	-.034 [-.152,.084] (<i>n</i> = 274)	-.301 [-.693,.229] (<i>n</i> = 16)	.027 [-.189,.240] (<i>n</i> = 84)
American Government	Low Partisan	.067 [-.029,.162] (<i>n</i> = 418)	.074 [-.075,.221] (<i>n</i> = 174)	.022 [-.135,.178] (<i>n</i> = 157)
	High Partisan	-.059 [-.227,.111] (<i>n</i> = 133)	-.112 [-.449,.252] (<i>n</i> = 31)	-.038 [-.300,.230] (<i>n</i> = 55)

course-created content threads. Following the analyses of reply-upvote pairs, these tests are all non-parametric correlations between the original posters’ ideology and their commenter’s ideology. Here again, a positive correlation between poster and commenter would indicate more siloing, on average, while a zero or negative correlation would indicate engagement across difference. In general, we find little evidence for ideology-based sorting in the forum comments. That is, most of these correlations are not significantly different from zero, which indicates that partisanship is having no aggregate effect on how commenters sort themselves among various replies. In particular, the high partisan salience threads in both classes contain a range of post-comment pairings, while the low partisan salience threads may have some modest ideological sorting among their substantive comments.

Partisan assortativity can also be represented graphically, as in Fig. 5. Here we display the percentage of substantive comments (i.e. the first column of Table 6) as divided by the tripartite ideology of the poster and the commenter. We divide these results by the partisanship of the thread in which each poster-commenter appears, and in general they agree with the analyses presented in Table 6. Posters often receive comments from students with divergent ideology. In particular, threads on subjects with high partisan salience seem to induce more interactions between students with different views of the world. These results are consistent with our hypothesis that MOOCs might provide a meaningful space for interactions between people with opposing views, rather than providing yet another echo chamber on the Internet.

Forum Language

In this section, we consider our final research question: how does ideology influence the language students use? The results so far have relied on tracking data, to simply show that partisan opponents interact in the discussion forums. However, tracking data cannot tell us the nature of that interaction. Do they diverge along partisan lines and talk past one another? Or do they converge on a shared language? To answer these questions we turn to the contents of the forum posts.

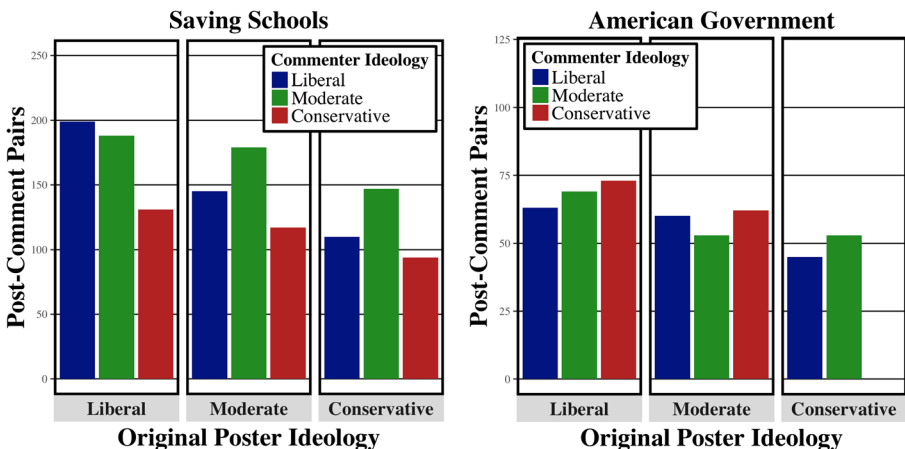


Fig. 5 Distribution of comments per reply in substantive course threads

Descriptive Statistics

After removing students who were not from the US and students who did not answer the ideology questions, the course-generated non-administrative threads in the two course forums included 4,516 and 5,452 posts (i.e. replies and substantive comments). The posts in this dataset provided us with the opportunity to model the language that partisans use in discussion with one another. The distribution of the length of these posts is given in Fig. 6. The average post across both courses was 125 words long ($SD = 116$). And main replies were somewhat longer, on average ($m = 142$, $SD = 125$) than the substantive comments they received ($m = 91$, $SD = 89$). In our analyses of partisan language to follow, we collapse the thread structure to consider each post as an independent event. This removes from analysis all information about the order of posts, and whether a post was a comment or a reply, and to which reply each comment was directed. We do, however, account for some thread-level information by considering how content systematically varies across course chapters. Each thread used here was posted for students to specifically discuss one chapter from their course, and included a chapter-related content question from the course administrators as the top-level post. There were 25 chapters in Saving Schools, and 24 chapters in American Government, and each chapter received 2–6 threads (depending on whether the threads were subdivided based on username, or whether the entire class was funneled into the same thread). We expected (and confirmed) that there would be differences in word use between the threads of each chapter, that was orthogonal to the partisanship of individual posters, and might cloud our ability to detect generic markers of partisanship. In both the classification results and topic modeling results below, we pre-process the text features to remove those that were particularly concentrated in only one chapter of a course (e.g. “cognitive skills” and “merit pay” in Saving Schools, or “equal protection clause” and “invisible primary” in American Government) so that the algorithm can prioritize the detection of partisan-leaning features that generalize across many threads and discussions.

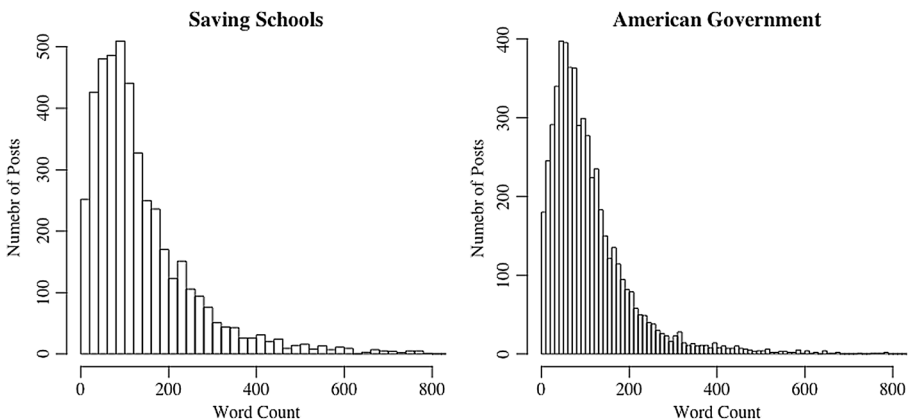


Fig. 6 Distribution of word count per post in the focal set of course threads

Partisan Distinctiveness

To construct an initial test for the existence of a linguistic partisan divide, we treated the forum data from each course as a standard supervised learning problem (Jurafsky and Martin 2009; Grimmer and Stewart 2013). Specifically, we first extracted a wide set of features from the text of the forum posts (as above, the data from the two classes were kept separate). We then used those feature counts as the inputs into penalized linear regression algorithms, which each estimated a model that could best predict each posters' self-reported ideology (Groseclose and Milyo 2005; Gentzkow and Shapiro 2010). If opposed students were simply talking past one another, we would expect the language they use to reliably reveal their ideology. That is, we could expect it to be relatively easy to detect a person's ideology, because they would process the course material through a biased partisan filter. On the other hand, if students were converging on a collaborative discourse, we would expect few linguistic markers of partisanship in the students' posts.

Feature Extraction We followed a typical “bag of words” approach, in which we simply counted the most common words and phrases from each post, removing all information about the order in which those words and phrases appear. The text from each post was parsed according to the following steps (Feinerer et al. 2008; Benoit and Paul 2017). In order, the text was converted to lowercase; then contractions were expanded; then punctuation was removed. Common stop words (“and”, “the”, and so on) were also dropped. The remaining words were stemmed using the standard Porter stemmer, and then grouped into “ngrams” - groups of one, two, or three sequential word stems. For example, “state and local government” would be parsed into six stemmed ngrams [“state”, “local”, “govern”, “state local”, “local govern”, and “state local govern”].

To focus on the most prevalent features, ngrams which appeared in less than 1% of all posts were excluded. Additionally, any ngram which was concentrated only in one particular chapter (i.e. 80% of all occurrences are in a single chapter) was also dropped - this was true of 24 ngrams in *Saving Schools*, and 90 words in *American Government*. The end result of this process was a “feature count matrix”, in which each post was assigned a row, while each ngram feature was assigned a column, and the value of each cell represented the number of times that ngram appeared in that post. Both classes provided rich vocabularies, with 1192 ngrams in *Saving Schools* and 983 ngrams in *American Government*. However, this dataset is sparse – specifically, 96% of the cells are zero, since most posts only included a few of the full set of ngrams.

Model Estimation These steps processed the unstructured text into a high-dimensional set of features, and we needed to determine how those features could be chosen and weighted to best distinguish the writers' ideology. We followed a bottom-up approach, using a common method, the LASSO, implemented in the *glmnet* package in R (Tibshirani 1996; Friedman et al. 2010). This algorithm estimates a linear regression, and shrinks the effective feature space by imposing a constraint on the total absolute size of the coefficients across all features. The size of that constraint is determined empirically, by minimizing out-of-sample error via cross-validation within the training set. This process reduces many coefficients in

the regression to exactly zero, leaving a smaller set with non-zero coefficients in the model.

To estimate the algorithm's accuracy, we used a nested cross-validation procedure (Stone 1974; Varma and Simon 2006). The entire dataset was randomly divided into ten folds of equal size. To produce out-of-sample predictions for each fold, a classification model was trained and tuned on the other nine folds, and applied directly to the held-out data to predict the ideology of those posts. To smooth out the random fluctuation across folds, we performed this whole procedure five times, and averaged across all five predictions to determine a final predicted partisanship for each post. The predicted partisanship of all posts in both classes are plotted against the actual partisanship of the author in Fig. 7. We also fit a loess regression, which is shown in the figure (with 95% confidence intervals). These results suggest that there is indeed some predictive distinction that can be made between ideologically opposed students. However the relationship between predicted and actual partisanship is not especially strong. Additionally, it seems to be somewhat stronger in American Government ($r_\tau = .179$, $z(5452) = 19$, $p < .001$), while it is weaker in Saving Schools ($r_\tau = .077$, $z(4516) = 7.7$, $p < .001$). These results provide evidence for the existence of some modest ideological divisions in the language of forum posters.

To highlight the features that distinguished partisan posts in this algorithm, we conducted a separate analysis that considered the partisan distinctiveness of each feature individually. Specifically, we calculated two commonly used statistics that capture strength of association - variance-weighted long-odds ratio and mutual information - to evaluate the relationship between feature frequency and ideology in each class (Monroe et al. 2008). In Fig. 8 we plot these two metrics against one another for every word in both classes, which gives a visual representation of the words and phrases that were the most distinctively partisan in our data. The words towards the upper corners of these plots are among the most useful. Some of these distinctive words do carry partisan connotations - for example, in American Government liberals were more likely to discuss "corporations", while conservatives were more likely to discuss the "constitution". However, the linguistic differences represented here are modest, and

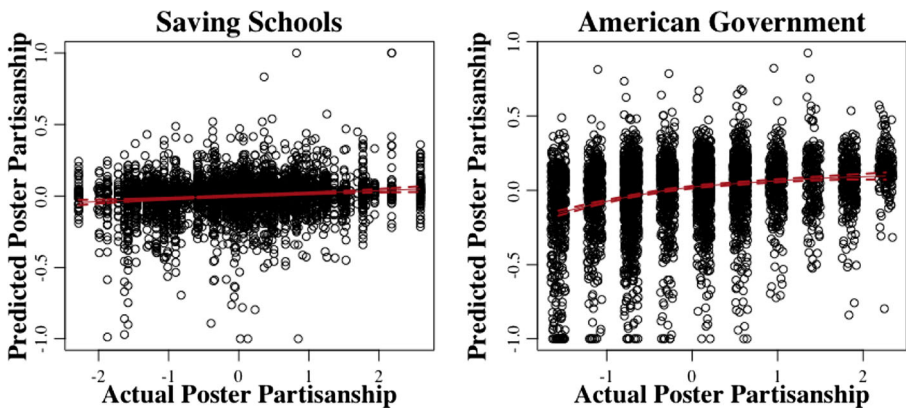


Fig. 7 Distribution of predicted versus actual partisanship from language classifier. A loess regression line is also plotted, with 95% CI. All units are in standardized ideology scores, with higher values indicating more conservative posters, and lower values indicating liberal posters

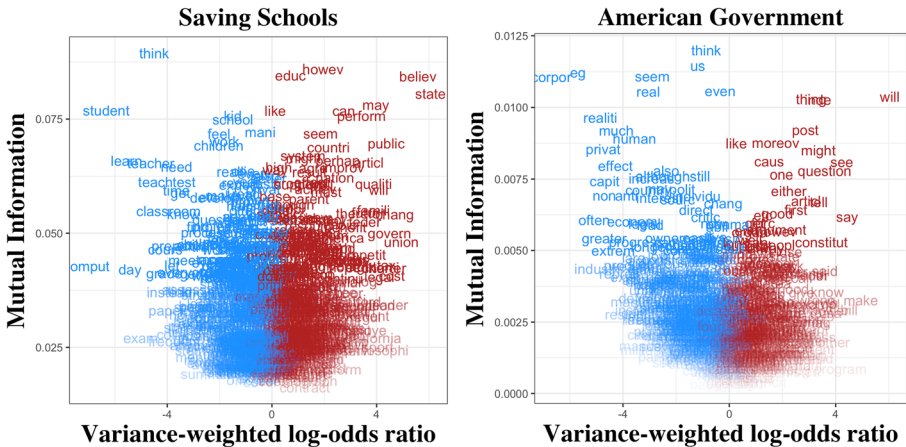


Fig. 8 Partisan distinctiveness of words in both classes. The variance-weighted log odds ratio of each word is given on the x axis, capturing the direction of partisan distinction in each word. The mutual information of each word is on the y axis, an absolute measure of distinctiveness

do not cleave along familiar partisan lines, for the most part. This provides additional evidence that the weak results of the classifier above represent a property of our data, and not just the limitations of our algorithm. That is, the language of posters does not diverge sharply along partisan lines.

Partisan Topics.

One limitation of the ngram-level analysis above is that the analysis is too granular. That is, the number of ideas and themes in the forum is much smaller than the number of unique ngrams. Thus, it is possible our ngram-based analysis could miss the effect of partisanship broader trends in topic use. To examine this we use a form of unsupervised text analysis from the topic modeling tradition (Blei et al. 2003) called the Structural Topic Model (Roberts et al. 2014; Reich et al. 2015; Roberts et al. 2016a). Topic models are designed to identify sets of words, “topics,” that tend to occur together. This reduces the high-dimensional space of “all words used in the forums” to a more manageable space of re-occurring common themes, which we can then map onto partisan ideology. The STM estimates the relationship between metadata whether it was written by a liberal, moderate, or conservative, and the proportion of the post belonging to a particular topic. From each course we estimated a separate model 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.

For each class, we processed the text of the posts using the same steps as above, with the exception that we only use unigrams (and not bigrams and trigrams) in the topic model. We estimated a separate 30-topic model for each class, using a spectral initialization procedure (Arora et al. 2013; Roberts et al. 2016b). After the model was estimated we extracted seven representative words from each topic, using the FREX metric (Bischof and Airoidi 2012; Roberts et al. 2014). The resulting topic word lists

are given in Tables 7 and Table 8. Additionally, we also estimated the partisan lean of each topic, by comparing the difference in prevalence of each topic among liberals and conservatives (as defined by the tripartite metric described above). These estimates are plotted in Fig. 9, and suggest that all partisan differences in topic use amount to less than 2% of all posts. Consistent with the classifier results, the topic model estimates suggest that for the most part, students with conflicting ideology still converge around similar topics.

For example, in Saving Schools even topics that one would expect to be politically charged, such as racial achievement gaps (Saving Schools # 29), or common core (Saving Schools # 26) seem to be evenly distributed across the ideological spectrum. The results holds in American Government, as well. Topics that are politically divisive, such as the Supreme Court (American Government #10) and national security

Table 7 Most distinctive words from topics in Saving Schools

#1	day, year, three, secondari, work, per, plan, lesson, summer, primari
#2	comput, prepar, content, stem, human, model, program, collabor, couss, key
#3	teacher, profess, profession, tenur, certif, recruit, teach, master, train, evalua
#4	measur, valu, easi, ad, find, accur, effectieess, advantag, judg, pictur
#5	quision, interest, cous, answer, post, mention, comment, debat, topic, first
#6	look, colleg, anoth, attend, high, compani, document, unies, readi, graduat
#7	well, fie, person, far, traie, care, enough, truli, bottom, young
#8	court, case, advoc, total, yet, forc, suprem, decis, freedom, v
#9	special, languag, egish, disabl, appropri, address, govern, adequ, need, feder
#10	cultur, gap, contribut, problem, academ, poverti, solv, success, diesiti, social
#11	politician, power, polit, vote, collectiebargain, bad, elect, parti, money, voter
#12	class, sie, smaller, classroom, reduc, larger, larg, watch, small, number
#13	test, score, standardi, compar, profici, math, rank, intern, th, assess
#14	good, think, thing, great, realli, point, just, lot, get, make
#15	right, opinion, law, limit, word, express, activ, third, action, act
#16	agre, said, matter, sens, push, etc., individu, true, thus, pass
#17	econom, growth, economi, correl, strong, articl, link, connect, countri, wonder
#18	local, board, district, member, superintend, communiti, run, control, elect, town
#19	policia, maker, eviec, dispar, data, bias, experiet, suspend, zero, reli
#20	account, system, must, fail, held, respons, hold, ensur, progress, hybrid
#21	research, suggest, studi, possibl, approach, perhap, thank, kind, found, outcom
#22	merit, pay, market, salari, battl, compens, paid, reward, determin, employe
#23	kid, child, children, parent, go, everi, get, want, feel, goe
#24	learn, skill, knowleg, world, centuri, method, today, math, critic, basic
#25	voucher, issu, vie, support, lectur, argument, origin, cut, discuson, still
#26	common, standard, core, state, nation, curriculum, adopt, implement, creativ, texa
#27	will, technolog, reform, digit, cost, chang, continu, union, futur, replac
#28	charter, privat, choie, public, school, option, competit, tradit, fund, famili
#29	segreg, american, equal, integr, tax, impact, america, black, immigr, de
#30	come, improv, effecti, mani, also, order, can, way, rather, much

Table 8 Most distinctive words from topics in American Government

#1	cultur, america, american, freedom, self, democraci, belief, lieti, conflict, western
#2	well, said, agre, though, noth, believ, even, word, speak, clear
#3	public, offic, opinion, polici, measur, improv, action, politician, factor, influenc
#4	becom, etc., human, societ, natur, men, man, fact, women, struggl
#5	contribut, non, lobbi, interest, group, econom, rie, resourc, grous, approach
#6	sure, never, total, make, eg, everyon, everyth, complet, rememb, learn
#7	bank, fed, money, rate, profit, reserv, economi, monetari, tool, spend
#8	insur, health, care, tax, mandat, penalti, aca, regul, afford, claus
#9	voter, vote, minor, europ, poor, poll, mayb, lower, elig, citi
#10	suprem, court, justi, judg, judici, decis, interpret, restraint, correct, rule
#11	agenc, knowleg, budget, view, point, bureaucraci, implement, field, sourc, depart
#12	thing, seem, good, get, done, alway, might, lectur, least, last
#13	movement, contrast, religi, religion, abort, street, republican, parti, democrat, church
#14	feder, grant, govern, state, aid, author, local, expand, fund, patterson
#15	issu, partisan, dietion, media, inform, control, hear, watch, report, stori
#16	presiet, branch, power, check, separ, institut, executiebranch, balanc, share, congress
#17	trade, demand, incom, manufactur, product, market, worker, labor, wealthi, wage
#18	primari, campaign, name, candid, recognit, advantag, senat, constitu, re, reelect
#19	way, term, long, year, work, still, new, thought, everi, need
#20	exist, establish, chang, base, togeth, law, effect, ousd, prove, exampl
#21	amend, protect, right, th, process, due, eviecc, search, claus, fourteenth
#22	program, assist, post, secur, social, limi, abus, secret, thegvern, threat
#23	school, educ, concept, children, opportun, equal, child, qualiti, area, parent
#24	elector, popular, singl, elect, win, district, system, represent, proport, colleg
#25	war, bush, world, militari, becam, privat, weapon, mass, rest, us
#26	legisl, committe, broad, easier, pass, congress, polar, major, ielog, chamber
#27	concentr, cost, diffus, benefit, bill, eneg, spread, pay, taxpay, lawmak
#28	uni, great, countri, structur, target, professor, system, mani, complex, key
#29	see, want, think, peopl, know, just, will, someone, take, now
#30	involv, concern, number, explain, small, much, relatiec, similar, larger, differ

(American Government #25) are discussed at essentially equal rates by both liberals and conservatives. These results suggest that for many of these politically sensitive topics, students are converging around shared language and concepts.

These topic models also suggest that there were in fact some topics where language diverged across partisan lines. In *Saving Schools*, one large ideological division was among the topics that focused on the teachers. Left-leaning writers focused more often on the teachers' certifications (*Saving Schools* #3), and time commitments (*Saving Schools* #1). Right-leaning writers instead focused more often on teachers' compensation (*Saving Schools* # 22) and school board governance (*Saving Schools* #18). In *American Government*, the most distinctive topics for liberals were economic issues, such as international trade (*American Government* #17) and interest group lobbying (*American Government* #5). Interestingly, many of the most conservative topics were

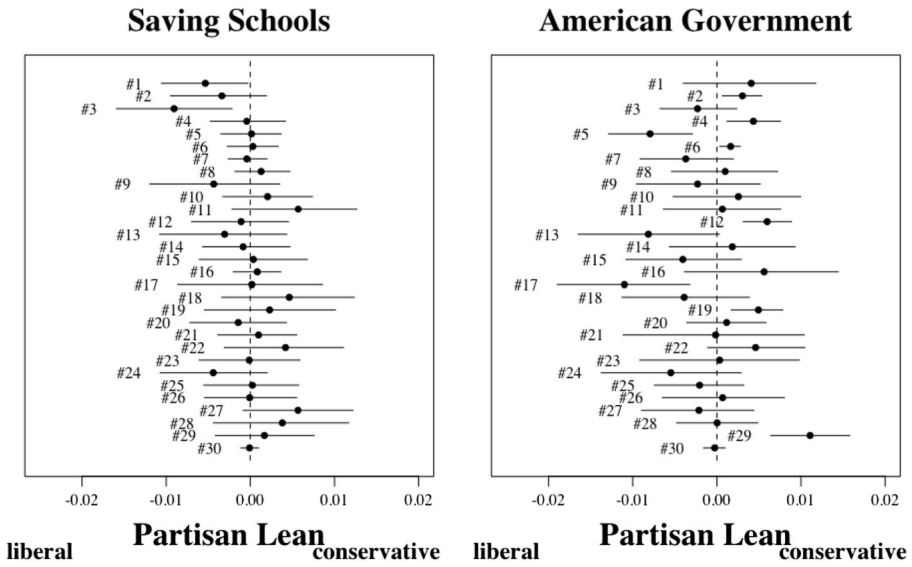


Fig. 9 Distribution of estimated change in prevalence of topics between liberal- and conservative-authored posts. Point estimates and 95% confidence intervals are shown

not substantive, but simply captured the syntactical structures of abstract principles (American Government #4) and delineating disagreement (American Government #29), perhaps because they recognize themselves to be a minority. There was also some modest partisan taunting (Grimmer and King 2011). For example, some right-leaning students in Saving Schools did explicitly complain about left-leaning policy makers in education (Saving Schools #11). Additionally, the left-leaning students in American Government did levy their complaints about religious conservatives (American Government #13).

Overall, though, these partisan-leaning topics were rare, and not especially heated based on our own reading of the posts. Instead the general pattern seemed to reflect an open and diverse conversation that welcomed views from across the political spectrum, and brought opposed students together around common language in much of the discussion forums. For transparency, we selected some representative posts from the topics mentioned in the main text here, and included them in Appendix.

Linguistic Interaction Style

So far, our analyses have considered the language of each post in isolation, evaluating the contents of each post with respect to the poster’s own ideology. However, this does not reflect the context in which many posts are generated. As Fig. 5 shows, many posts are direct comments on other students’ posts, and these comments come from students across the ideological spectrum. That is, liberals leave comments on posts from both conservatives and liberals, and conservatives’ comments are similarly distributed. These post-comment pairs, then, can span a range of ideological distance - some comments are ideologically close to their parent post (“intra-party” comments) while other comments are ideologically distant from their parent post (“cross-party”

comments). Does this ideological distance between comment and parent affect the contents of the comments themselves?

Pair-Level Data Though commenters were not randomly assigned to parents, our dataset can provide some insight into these interactions. However, our effective sample size is limited because this analysis requires that both the parent post and comment be written by someone in our target sample (from the U.S., answered the ideology survey question, etc.). We also decided to remove pairs that involved a moderate poster to highlight the contrast in ideological distance, leaving only intra-party pairs (i.e. liberal-liberal or conservative-conservative) and cross-party pairs (i.e. liberal-conservative or conservative-liberal). All in all, we found 569 such parent-comment pairs in the substantive course-created threads in *Saving Schools*, and 297 pairs in *American Government*. To increase our power, we pooled the courses together, for a total dataset of 866 parent-comment pairs.

Stylistic Measures Our sample was not large enough to conduct a bottom-up analysis of the words that distinguished ideological distance in the same way that our earlier analyses distinguished liberals and conservatives. Instead we draw on the literature to test whether established markers of linguistic style are more common in intra-party or cross-party pairs. In particular, we focus on three kinds of linguistic styles that might relate to engagement across difference. First, we considered the emotional content of the posts (often called “sentiment analysis”) by tallying the use of positively- and negatively-valenced words in the comments, as defined by a commonly used dictionary (Pennebaker et al. 2007). Second, we considered several markers of linguistic complexity, including: the average word count; the Flesch-Kincaid readability score, a measure of syllable- and sentence-level complexity (Kincaid et al. 1975); and vocabulary depth, as measured by the (reverse-scored) average frequency of the words used in the comment (Brysbaert and New 2009). Third, we considered markers that might reflect accommodation in the comments. One simple marker of accommodation is the presence of hedging language in the comment post (Hübler 1983; Jason 1988). We also explore more complex measures of accommodation, by measuring two kinds of matching in word use between the parent and comment post (Giles et al. 1991; Doyle and Frank 2016). One version measures stylistic matching, by similarity in function word use - broad categories of word classes such as pronouns, negations, quantifiers, and expletives (Ireland et al. 2011). Another version measures semantic matching, by similarity in use of topics, measured by the (reverse-scored) Hellinger distance between the distribution of topics in the parent and comment post (Blei et al. 2003).

Results For each of the measures listed above, we calculated a value for each parent-comment pair in our data. However, our dataset was somewhat imbalanced, in that liberal commenters were somewhat over-represented in the cross-party pairs, relative to the intra-party pairs. Conceptually we were most interested in the relationship between these linguistic style markers and each comment’s ideological distance from its parent, holding constant the raw ideology of the comment. To test this we conducted weighted regressions, with the weight on each observation inversely proportional to the prevalence of the ideological configuration of the parent-comment pair (i.e. liberal-liberal;

liberal-conservative; conservative-liberal; conservative-conservative). This put equal weight on each of the four configurations in our estimates, eliminating any mechanical correlation between ideology and ideological distance.

We summarize these estimates in Table 9, by reporting the results of a series of regressions that included the binary intra- vs. cross-party pair variable as a predictor, and each (standardized) linguistic style measure as an outcome. We do see a highly significant relationship between ideological distance and word count - that is, commenters who were ideologically opposed to the writer of the parent post wrote shorter posts, on average. There was also a modest trend in function word matching, which was more common from intra-party commenters. These results are not, by themselves, definitive evidence for engagement, because dictionary-based methods can in some cases be context dependent, and unreliable in small sample sizes. However, these results corroborate the main conclusions of the other analyses here, and provide evidence that the conversations being held in these MOOC forums provide a rich, engaging source of civic discourse for students.

General Discussion

Open online courses attract a wide diversity of students, and this diversity can potentially serve many of the institutional goals of MOOCs. In this paper we present results that potentially identify a neglected benefit of this diversity - that is, can these courses allow students to bridge political differences and interact with their ideological opponents? Contrary to the concerns of observers that the internet has become a place of echo chambers and silos (Sunstein 2017), we find evidence that, at least in these two examples, online courses are a space in which people with different political opinions can learn and engage together.

We found that the student body contained participants with diverse policy preferences. Only a subset of participants chose to engage in online forum discussions, but the subset that did so had a range of political ideologies. Within forums, we found that most threads contained a balanced proportion of liberal and conservative posters, and that liberals and conservatives directly responded to each others' posts. We argue that

Table 9 Estimated relationship between measures of linguistic style in comments and ideological distance from their parent posts. Positive coefficients indicate higher levels in comments from cross-party pairs, compared to intra-party pairs. (* = $p < .1$; ** = $p < .05$; *** = $p < .01$)

Measure	Coefficient	Standard Error
Word Count	-0.196***	0.067
Emotional Language		
Positive Valence	-0.057	0.067
Negative Valence	-0.039	0.066
Linguistic Complexity		
Readability	0.105	0.068
Vocabulary Depth	0.003	0.068
Accommodation		
Hedging	0.044	0.067
Function Word Matching	-0.133**	0.068
Topic Matching	0.059	0.069

these online courses present a case study where at least the pre-conditions of deliberative discourse appear to be met.

Additionally, text analysis of student forum posts suggests that students with different political beliefs tend to discuss many similar topics in roughly equal proportion. However, we still found evidence of some partisan division in some of the topics they discussed. In particular, discussions did seem to diverge more around issues related to teacher's contracts in *Saving Schools*, and economic issues in *American Government*. However, these divisions were modest, and combined with the tracking data, these results suggest that, generally speaking, students did not segregate themselves within rhetorical frames that inhibited meaningful discussion. Finally, we found that the linguistic style of comments was not meaningfully different between those that replied to intra-party posts, and those that replied to cross-party posts.

These results fit into a broader research framework driven by two categorizations that could be applied to any online community that attracts diverse participation. These categories form a 2×2 matrix that maps onto our latter two research questions, and summarizes the ways in which political differences might affect online discussion forums, shown in Fig. 10. 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 together, using a common vocabulary. Here we focus primarily on describing and measuring these possible categories of discussions. In future work we hope to understand what causes these different categories to emerge, and what might be done to promote deliberative discourse, at the expense of the other types of discussions.

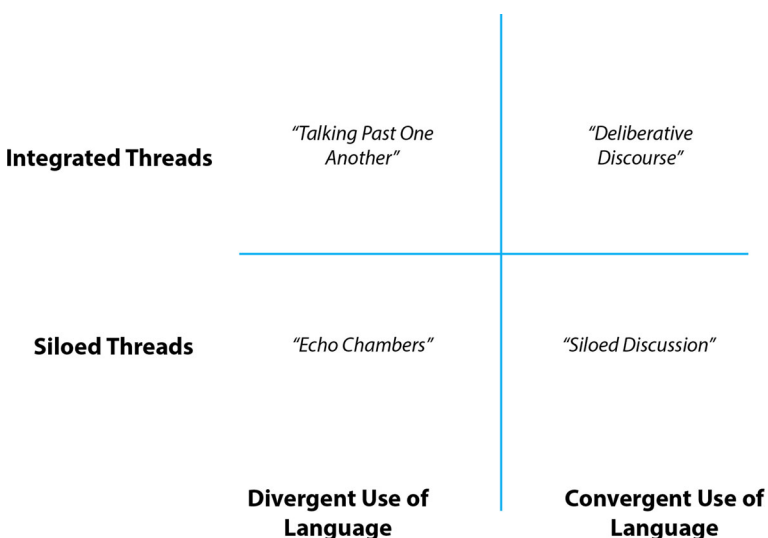


Fig. 10 Two-by-two schematic of dimensions of engagement in online discussion forums

One clear limitation of these results is that we can only analyze the content of MOOC discussions among students who endogenously choose to enroll. This means that we cannot determine whether the generally positive intellectual climate in these discussion forums is inherent to the MOOC format, or whether these MOOCs attract a particularly open-minded and reasonable student body (or some combination thereof). It is hard in our sample to detect the extent of selection effects, in part because the individual-level determinants of civil discourse may be unobservable. Furthermore, we do not at present have comparable data from some other reference population. Intuitively, one could easily imagine situations in which MOOC enrollment is positively correlated with some latent propensity to be civil to one's ideological opponents. On the other hand, one could also imagine situations in which even well-intentioned students might fail to bridge across partisan divides. We cannot conclude whether the MOOCs' student selection or the MOOCs themselves are independently necessary to produce constructive civic discourse. However, our results suggest that these two factors are jointly sufficient to produce constructive civic discourse.

Another important limitation of this current work is that we lack a "ground truth" measure of partisanship with which to calibrate our measures of engagement across political difference. For instance, our analysis of the forum text suggests some (but not total) partisan division in language use. Are these divisions large or small? Though we try to provide some context, ultimately we do not have a perfect benchmark. Additionally, we have presented analyses that draw from many established approaches to analyzing open-ended text, and all of these approaches rely on simplifications and assumptions to quantify the unstructured data from the forums. However, these are not exhaustive, and the results we present here do not preclude the possibility of other linguistic differences that might be better captured by other modeling choices or feature sets. We hope that our research is a starting point to spur new work, as other datasets and language models are adopted in the scientific community.

One important advantage of the methods we use here are that they are generalizable across contexts. That is, the same analytical framework could be applied in other settings where the structure and contents of their discussions are tracked, including settings in which other dimensions of interest (such as gender or ethnicity) are the focus of diversity efforts. In parallel with this basic research, our future efforts include building a dashboard that incorporates these analyses in a standard suite of tools. This dashboard could be used by administrators to monitor engagement across difference (or lack thereof) in their own class discussion forums, informing classroom policy in real time. We also hope that in future research, we might use experimental interventions explicitly designed to increase engagement across political differences, and to evaluate how the measures we describe here respond to those interventions.

Although many online discussion spaces tend towards partisan division, our results suggest that MOOCs stand out from that trend, and can provide a space where students' exposure to divergent perspectives can be enriched. 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.

Acknowledgements We gratefully acknowledge grant support from the Spencer Foundations New Civics initiative and the Hewlett Foundation. We also thank the course teams from Saving Schools and American Government, the Harvard VPAL-Research Group for research support, Lisa McKay for edits, and research assistance from Alyssa Napier, Joseph Schuman, Ben Schenck, Elise Lee, Jenny Sanford, Holly Howe, Jazmine Henderson & Nikayah Etienne.

Appendix: Examples of posts from partisan topics

Here we provide the reader with some example posts from both classes. We drew these from the data after reading through the forums, and tried to select posts that were representative of the topics and threads that showed the most partisan divergence in our data. We used the STM functionality to select four posts from every topic mentioned in the main text. We first selected the ten posts which had the highest estimated prevalence of each topic. We then selected the four shortest posts from that list of ten, for brevity. The resulting posts are listed below.

Time commitments (Saving Schools #1)

"I currently teach two 5th grade Reading classes. One class with 30 students and one with 22 students. These are heterogeneously grouped classes, with reading levels that span at least 2 grade levels in each class. I am expected to do a 20 minute whole group lesson, and then three small group lessons, one below grade level, one on grade level, and one above grade level. I also need to provide meaningful, leveled independent work for students while I'm meeting with the other groups. And don't forget about meeting the objectives/skills of my students who have IEP and 504s. There is too much being put on the plates of teachers."

"I once heard that being a teacher is like being a lawyer who is in court all day long. Your prep work including lesson planning, grading papers, tutoring students, contacting parents all happens outside of that time. Please tell me how it is humanly possible for a teacher to grade 150 papers and plan lessons and write assessments every day in a 60 minute conference period. Until someone can explain that, they cannot argue that teachers work shorter days. That exact comment in the second article made me cringe. That and the one about how easy it is for teachers to call in substitutes. Really? If you mean going in at ten o'clock at night or 6 in the morning to write a plan for the substitute which won't be followed anyway."

"**thank you** for pointing out the challenge of having a second career during vacation times. I am a high school English teacher. I work at least eight hours a day (the school day is 6.5 hours and I spend at least 1.5 hours working at my desk or at home, plus a couple of hours a week advising student clubs). I work a minimum of 4-6 hours every weekend. That's when I have planning to do and homework to check, but *don't* have a pile of essays waiting to be graded. I rely on the summer to take courses and earn PDPs in order to maintain licensure, improve my craft, and, yes, advance my salary. So I do not have time to get another job during the summer."

"I am paid relatively well—certainly better than national averages—so I am not complaining at all for myself. But I take issue with anyone who says teachers have the summers "“off,”" so they can just take vacations or get a second career. I put so many hours in during the school year that I rely on the summer time to catch up with other professional demands."

"The best teaching situation I ever had was at a public middle school where I taught in a team with two other teachers, we met the parents of the all of the students, we had block scheduling (we could divide the morning hours however we chose), and we had a block of planning time (2 periods) EVERY DAY when the three of us could coordinate our efforts. The class was a 7/8 mix and students stayed with us for 2 years. These were not privileged kids and it was not a wealthy community, but a new principal came in and gave us 100% freedom within our team. The test results in the years I was there were positive and significant, especially for the students who had been lagging. We all knew what would be tested, but we were free to teach however we wanted—our team used an experiential, inter-disciplinary approach. I couldn't wait to get to work, and our attendance records suggested the vast majority of our students felt the same way. That was more than twenty years ago, and I haven't heard of anything like it recently, but the freedom, money and (these days) safety to do that would make a big difference. "

Teacher certifications (Saving Schools #3)

"If you eliminate paying teachers more for a Master's degree, would you start all teachers at a higher salary to begin with? "

"I can tell you that I retired from the military (Navy, Commander) and then became a teacher. I have two master's degrees and a doctorate. The requirements to become a teacher are frustratingly bureaucratic."

"I feel that teachers are underpaid. They should be paid according too how they perform. As they show teachers are getting paid more say getting a Masters Degree but in studies it has shoven that they show no better performance than teachers without a Masters."

"Here is an overview of the Stanford Hollyhock Fellowship Program that aims to improve teacher retention given that nearly half of all teachers leave the classroom within five years: [link]"

Liberal Complaints (Saving Schools #11)

"The people who were for desegregation stopped working for it once it was implanted, as they only fought for their political ideals."

"As elected officials, they should be the voice of the people. Politics does unfortunately play a role, so our voice is not always heard."

"Really disappointed to hear that only 10-15% of people vote in school board elections. I knew it was probably low, but not that low. Seems to reflect our priorities."

"I agree mostly, and while working as a journalist, I definitely saw times where unions wanted school boards to think of absolutely everything but the children. In general though, I think unions are good or bad dependent on their individual leadership, not the concept, although I'm less supportive of unions in public than private settings."

Teacher compensation (Saving Schools # 22)

"I would think battle pay is reasonable, merit pay is controversial and market pay is not good."

"Among the three types of teacher pay, market pay and battle pay seem to be less controversial and supported by data. The implementation of merit pay needs to be supported by the teachers being subjected and a collaborative environment."

"I'd say all three, but especially battle pay and market pay, because they can be implemented in the capitalist sense in response to supply and demand. If these pay scales are in effect, then the tough jobs or the jobs that are accepted in lieu of private sector positions will have to compensate competitively in order to be filled at all. And the better the compensation, then the better the candidates that are willing to take them."

"The disparity between the two articles appears to indicate, again, that it is not what we pay teachers but how we pay them. Teachers are not paid based on merit. Systems based largely on seniority alone tend to encourage the weakest to remain. Passionately motivated individuals work hard to achieve results and receive the same compensation as an individual who merely has a pulse. As talented individuals watch everyone move ahead and get paid the same increases for longevity, they become discouraged and many leave."

School board governance (Saving Schools #18)

"I personally don't believe in school boards. I feel that school boards should function much like company boards - hire an effective CEO (superintendent) and let that person run the corporation (school system). Of course the key is hiring an "effective" CEO."

"The problems with school boards, particularly how political they are, are disheartening. I understand the impulse to eliminate them altogether. Yet, the local community needs to have a say in their local schools and the school board meetings is a way for local community to be involved and to have a voice. A more top-down approach allocating the school board decisions to a mayor, or state and federal government would lead to less buy in and more resentment. When the people they have a say, or simply the option of having a say, more collaboration ensues."

"As an advocate for innovation and school choice, I agree with [name] that local school boards should be more of a board of trustees, and less involved in the day-to-day operations of the district. But I strongly disagree that the superintendent should be a state employee under the civil-service rules. If the local school board is responsible for the overall performance of the local district (as a board of trustees would be) and the superintendent is the CEO of that district, or the person ultimately responsible for the performance, then the superintendent should be hired by and be responsible to the local board."

"I'd be very curious to see more specific examples of the difference in school districts/school boards around the States, especially between very different communities i.e. a small rural school district versus a big city school district. I do think some local control via school boards or campus based committees as mentioned above are good, but of course it can be tricky to avoid corrupt officials leading these groups. I have heard my fair share of complaints about corrupt school district officials growing up and attending school in the Philadelphia school district! In the past few years the Philly school district has gone through a ton of turmoil. Here's an interesting article on it: [link]"

Common core (Saving Schools # 26)

"Common core appears to be a step in the right direction to national standards of excellence. However, its success or failure will be determined by implementation."

"I support states adopting the common core. The standards and supporting curriculum guidelines are a significant educational move toward skills required in the 21st century."

"The Common Core standards would give people a easy way to evaluate students' proficiency between states. It would also be standards for educational staffs to know where they are and where they want to go. I support states to adopting the Common Core standards."

"Yes, Texas does have its own Core Standards. It is similar to the Common Core. I would like to see how Texas Core Standards compare with CCSS. Which is more rigorous? How are students' performance compared on Texas Core Standards vs. Common Core State Standards?"

Racial achievement gaps (Saving Schools # 29)

"Desegregation without integration is not a benefit, neither is it equal.. All need to be treated equally and receive the same education and benefits."

"i agree that other forms of institutional racism beyond schools impact the achievement of minority students in the united states. simply declaring segregation to be unlawful will not lessen the impact of other forms of discrimination and inequity. integration doesn't alleviate the economic disparity between different racial groups."

"It is often said that desegregation wasn't intended to raise student achievement. Rather, it was meant to fix a social condition of inequality. Simply mixing races within a school doesn't change root causes for low achievement, such as low-income status, parental education levels, or the actual quality of instruction. Although desegregation hasn't driven up student achievement as many had hoped, I still believe it has an important legacy. It was a deliberate attempt to improve equality amongst children, and might be more symbolic than anything else. It showed that addressing racial inequality had entered the national conscious, and that is a major step forward. Hopefully, in the future, the lingering issues that cause unequal achievement will also be addressed, furthering the legacy of equality in the United States."

"I am appreciative of desegregation and grateful to those who came before me who were willing to put their lives on the line in order to see the movement materialize. When I think about the many atrocities that African Americans endured prior to and since desegregation, I am saddened. Desegregation has had an indelible impact on generations of people, not just those of African ancestry. It is true that student achievement levels may or may not have been raised following desegregation—the playing field has yet to be leveled. There are many reasons (some deep rooted) as to why achievement levels have not been raised since desegregation was legally put in place. Nevertheless, many have benefited from desegregation and I am hopeful that our progeny will face no boundaries in the future as they live and thrive as students and as human beings. "

Abstract Principles (American Government #4)

"The first thing a person should become is student of human nature. By nature people are selfish and self centered. They can be noble on occasion and with our ability to learn might grow——— but don't bet your farm on it."

"it seems that many religious people are able to find a ""middle ground"" on same sex marriage. They still see it is a sin but see the sinners as God's children worthy of their love. I understand your comment on wedding cakes etc but couldn't I use the same or similar argument with Jews (they killed Jesus), blacks-(slave honor your master), etc? If I am in the cake business shouldn't it about cakes and not who sleeps with who? "

"Regarding your natural law argument it seems this is dependent on when you consider the fetus as a human being? If it is from the moment of conception, when it is viable outside the womb or something else? And what of the woman's natural right to her own body? What about rape and incest? On same sex marriage, what about the natural rights of the same sexers to love and dignity? While I concede the contribution of religion to morality, atheists can also have and contribute to the morality of the community. Morality is not dependent on religion."

"Yes, a relatively unspoken example of de facto discrimination is the undermining of Black families by removing black fathers from households. In slavery families were often broken up and sold separately. In this and the last century, Welfare rules denied benefits to

any home where an able bodied man lived regardless of his ability to find a job. The incarceration of many black men is another way in which the structure of black families is undermined. Although outright discrimination against black families was never written into the laws, the effect of policy and regulation manifests a discriminatory result."

Interest Group Lobbying (American Government #5)

"Corporations or economic-based special interest groups have financial resources available to them that other non-economic groups do not have. In addition, they can offer benefits such as jobs to their members. Non-economic groups have to rely on fundraising and are hit hard by those who do not contribute but may benefit."

"Economic groups have access to financial resources that allow for better organization and influencing politicians. Non-economic groups suffer due to low financial resources and lack of a strong organized effort. The free-rider problem refers to people enjoying public goods without contributing to the effort or joining the organization."

"Free riders get the benefits of whatever is passed without attaching their name to it. Lobbying groups are organized around economic interests because these groups have resources such as corporate profits. They usually are successful. In contrast, non-economic interests have to rely on voluntary donations to support their lobbying efforts. They are usually unorganized

"American lobbying groups organized around economic interests, particularly business firms because these firms are able to provide financial backing to these groups. Money generated from their business activities. Non-economic groups are forced to raise their own funds to support their cause, funds which usually come from donations. Also, these non-economic groups are faced with the issue of the "free rider" problem people who are benefiting from the group activity and policy goals do not fund or participate in the push for the non-economic group's political agenda."

Supreme Court (American Government #10)

"Justices on the Supreme Court tend to be picked for political reasons and adhere to those through their lifetimes. For example conservatives judges are picked by conservative politicians."

"The Supreme Court decisions are so important because unlike lower courts there is no appeal. The Supremes are nominated by a president and generally follow partisan politics. Studies of court decisions bear out this simple observation. Because the Supremes are the final say, deal with issues of distribution of power and generally stick to appellate decisions, politics plays a greater role in their decisions than lower courts.

"Judicial restraint holds that judges should generally defer to precedent and to decisions made by legislatures. Judicial activism holds that judges should actively interpret the

Constitution, statutes, and precedents in light of fundamental principles and should intervene when elected representatives fail to act in accord with these principles"

"Judicial restraint suggests that the court narrowly construes the law and reviews the matter for its legality without saying what the law means or how it is to be interpreted. Sebelius case being an example. Judicial activism is when the court reviews a matter and interprets what the law means and announces its interpretation suggesting what is permissible under the law as interpreted by the court. Citizens v. Fed Elect."

Conservative Complaints (American Government #13)

"Republican still have religious values as part of their political strategy as oppose to Democrats with a my liberal strategy."

"Democratic strategies tend to move towards the social issue of whatever cause it is, whereas Republicans gravitate towards the moral issue of that cause."

"Republicans hold stronger religious values than democrats and liberals. for those support republicans are often those who go to church regularly, they more likely tend to oppose same-sex marriage, gay rights, and abortions."

"The Republican Party attracts conservatives and evangelicals while the Democratic Party attracts women and liberals. Republicans balked at court rulings regarding women's rights, abortion, and the growth of the counterculture movement. Democrats focus on LGBT issues and women's rights."

International trade (American Government #17)

"Democrats and labor union tend to protect blue workers' jobs through Protectionism, while republicans favor free trade."

"Free trade advocates reduced or eliminated tariffs to encourage trade and access to products and services abroad. Protectionism proposes higher tariffs on imports to protect domestic firms. Republicans favor free trade while Democrats oppose them, since they represent labor groups that are threatened by labor moved overseas."

"Demand-side economics puts money directly in consumer's hands, while supply-side economics benefits through tax cuts to businesses and the upper-class. Democrats tend to support demand-side economics because it supports the lower incomes, while supply-side economics is usually supported by Republicans because those are the entities that support them."

"Protectionism and free trade are opposites. Protectionism is when tariffs are levied on goods made abroad to protect domestic jobs. Free trade is elimination of tariffs so that goods flow freely between borders. Specific industries where goods can be product

cheaply produced abroad favor protectionism. Multinational companies favor free trade since it allows them to sell their goods abroad more cheaply."

National security (American Government #25)

"George W Bush told the American public that Iraq had weapons of mass destruction and had that message dominate the American media with this information until the American public and Congress gave Bush support for the war. The sentiment of the American public can be a deterrent to the President initiating war."

"After the World War II, the U.S. had a advantage over other nations. Its manufacturing grown during the war and was intact at the ended. The opposite of other nations that their factories had been destroyed by the end of the war. So the U.S. manufactured goods to sell to the ones need of imported goods. But some of them like Germany and Japan rebuilt their manufacturing sectors."

"President Bush made speeches to the public where he declared an "axis of evil" which necessitated the going declaring of war in Afghanistan and later in Iraq. He falsified intelligence information and presented it to the UN as justification of American use of force. The Congress could have held its own investigation of the intelligence and withheld its needed declaration of war. Congress could have withheld the funds for the war effort before troops were in harms way."

"To build support for the War in Iraq, George Bush did a few things:

1. George Bush slowly revealed his plan to attack Iraq in many speeches beforehand. Instilling the name of Iraq in the public's mind was the first step.
2. Bush enacted a new doctrine, the Preemptive War Doctrine. This stated that the US could not wait to be attacked, as it will be too late. This gave the power of the president to go to war, based on the premise that the US will be attacked or harmed.
3. Citing intelligence reports, Bush asked Congress to declare war on Iraq because they were starting to gather weapons of mass destruction.
4. Bush's message about Iraq and weapons of mass destruction was quoted over 10× more than that of the opposition. The message he wanted the public to hear was broadcast to them whether they wanted to hear it or not."

Delineating disagreement (American Government #29)

"Extreme weather events are not hard to see. When their water runs out people more people will start to understand! "

"The controversy of "Obamacare" aside, would you eliminate the EPA and FDA? Most people want the safety of our food and drugs regulated (FDA). Most people want clean air and water (EPA)."

"How about the constitutionality of these programs? When people are too indoctrinated to know the Constitution, don't you think that the government conditions people to think the way the government wants? So of course people agree with them."

"I think as we go through the next 16 weeks (or whatever this is?) you will find we agree about much and the disagreements will be on the 'what to do about it?' part. As I've said above, persistent problems are complex, with no easy answers."

References

- Arora, S., Ge, R., Halpern, Y., Mimno, D. M., Moitra, A., Sontag, D., Wu, Y., & Zhu, M. (2013). A practical algorithm for topic modeling with provable guarantees. *International Conference on Machine Learning*, 2, 280–288.
- Athey, S., & Mobius, M. (2012). The impact of news aggregators on internet news consumption: The case of localization. *Working Paper*.
- Baek, J., & Shore, J. (2016). Promoting student engagement in MOOCs. In *Proceedings of the third ACM conference on learning@ scale*, 293–296.
- Benoit, K. (2017). Quantda: Quantitative analysis of textual data. *R package version 0.99.22*. <https://doi.org/10.5281/zenodo.1004683>.
- Bischof, J.M. & Airoldi, E.M. (2012). Summarizing topical content with word frequency and exclusivity. In *International Conference on Machine Learning (ICML)*, 201–208.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences*, 201706588.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990.
- Chuang, I. & Ho, A.D. (2016). HarvardX and MITx: Four years of open online courses — Fall 2012–summer 2016. *Working Paper*.
- Della Carpini, M. X. D., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annu. Rev. Polit. Sci.*, 7, 315–344.
- Doyle, G., & Frank, M. C. (2016). Investigating the sources of linguistic alignment in conversation. In *Proceedings of ACL*.
- Education Next. (2015). Results from the 2015 *Education Next* Poll. Retrieved October 27, 2015 from <http://educationnext.org/2015-ednext-poll-interactive>
- Faris, R. M., Roberts, H., Ething, B., Bourassa, N., Zuckerman, E., and Benkler, Y. (2017). *Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election*. Berkman Klein Center for Internet & Society Research Paper. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:33759251>.
- Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5), 1–54.
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, Echo chambers and online news consumption. *Public Opinion Quarterly*, 80, 298–320.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22.
- Gardner, H., & Davis, K. (2013). *The app generation: How today's youth navigate identity, intimacy, and imagination in a digital world*. New Haven: Yale University Press.
- Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among internet news users. *Journal of Computer-Mediated Communication*, 14(2), 265–285.
- Gastil, J. (1992). Undemocratic discourse: A review of theory and research on political discourse. *Discourse & Society*, 3(4), 469–500.
- Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35–71.

- Gentzkow, M., & Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4), 1799–1839.
- Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press.
- Grimmer, J., & King, G. (2011). General purpose computer-assisted clustering and conceptualization. *Proceedings of the National Academy of Sciences*, 108(7), 2643–2650.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 267–297.
- Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237.
- Hübler, A. (1983). Understatements and Hedges in English. *Pragmatics and Beyond*, 4(6), 1–192.
- Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, 22(1), 39–44.
- Jason, G. (1988). Hedging as a fallacy of language. *Informal Logic*, 10(3), 169–175.
- Jurafsky, D., & Martin, J. (2009). *Speech and natural language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River: Prentice Hall.
- Kahne, J., Middaugh, E., Lee, N. J., & Feezell, J. T. (2012). Youth online activity and exposure to diverse perspectives. *New Media & Society*, 14(3), 492–512. <https://doi.org/10.1177/1461444811420271>.
- Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.
- Kindel, A., Yeomans, M., Reich, J., Stewart, B., & Tingley, D. (2017, April). Discourse: MOOC Discussion Forum Analysis at Scale. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale* (pp. 141–142). ACM.
- Koutropoulos, A., Gallagher, M. S., Abajian, S. C., de Waard, I., Hogue, R. J., Keskin, N. Ö., & Rodriguez, C. O. (2012). Emotive vocabulary in MOOCs: Context & participant retention. *European Journal of Open, Distance and E-Learning*, 15(1).
- Lakoff, G. (2014). *The all new Don't think of an elephant!: Know your values and frame the debate*. Chelsea Green Publishing.
- Loveland, M. T., & Popescu, D. (2011). Democracy on the web: Assessing the deliberative qualities of internet forums. *Information, Communication & Society*, 14(5), 684–703.
- Monroe, B. L., Colaresi, M. P., & Quinn, K. M. (2008). Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4), 372–403.
- Orfield, G., Kucsera, J., & Siegel-Hawley, G. (2012). E pluribus... separation: Deepening double segregation for more students. Working Paper.
- Pariser, E. (2012). *The filter bubble: How the new personalized web is changing what we read and how we think*. New York: Penguin Books.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. Austin, TX: liwc.net.
- Peterson, P.E. (2010) Let the Charters Bloom. Retrieved August 7, 2015 from <http://www.hoover.org/research/let-charters-bloom>
- Peterson, P. E., Henderson, M., & West, M. R. (2014). *What Americans think about schools and how to fix them*. Washington, D.C.: Brookings Institution Press.
- Quattrociochi, W., Scala, A., & Sunstein, C. R. (2016). Echo chambers on facebook. *Working Paper*.
- Reich, J., Romer, A., & Barr, D. J. (2014). Dialogue across difference: A case study of Facing History and Ourselves' Digital Media Innovation Network. In B. Kirshner, E. Middaugh (Eds.), *Becoming political in a digital age*. Charlotte, NC: Information Age Publishing.
- Reich, J., Tingley, D., Leder-Luis, J., Roberts, M. E., & Stewart, B. M. (2015). Computer-assisted reading and discovery for student generated text in massive open online courses. *Journal of Learning Analytics*, 2(1), 156–184.
- Rheingold, H. (2000). *The virtual community: Homesteading on the electronic frontier*. Cambridge: MIT press.
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082.
- Roberts, M. E., Stewart, B. M., & Airolidi, E. M. (2016a). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, just-accepted, 1–49.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2016). Navigating the local modes of big data. In R. M. Alvarez (Ed.), *Computational Social Science*. Cambridge: Cambridge University Press.

- Schaffner, B., Ansolabehere, S. (2015) 2010-2014 cooperative congressional election study panel survey. <https://doi.org/10.7910/DVN/TOE811>.
- Siemens, G. (2005). Connectivism: Learning as network-creation. *ASTD Learning News*, 10(1).
- Stone, M. (1974). Cross-validated choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B: Methodological*, 111–147.
- Stump, G. S., DeBoer, J., Whittinghill, J., & Breslow, L. (2013). *Development of a framework to classify MOOC discussion forum posts: Methodology and challenges*. In: NIPS Workshop on Data Driven Education, 1–20.
- Sunstein, C. R. (2017). *# Republic: Divided democracy in the age of social media*. Princeton: Princeton University Press.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B: Methodological*, 58(1), 267–288.
- Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7(1), 91.
- Welbers, K., & de Nooy, W. (2014). Stylistic accommodation on an internet forum as bonding: Do posters adapt to the style of their peers? *American Behavioral Scientist*, 58(10), 1361–1375.
- Wen, M., Yang, D. & Rośe, C.P. (2014). Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the International Conference on Weblogs and Social Media*, 525–534.
- World Values Survey Association. (2009). *World Values Survey 1981–2008 official aggregate v. 20090901*. Madrid: ASEP/JDS.
- Yang, D., Wen, M., Howley, I., Kraut, R., & Rose, C. (2015, March). Exploring the effect of confusion in discussion forums of massive open online courses. In *Proc 2nd ACM Conference on Learning@Scale*, 121–130.