2018

# Three essays on how social context shapes engagement online

BOSTON UNIVERSITY

QUESTROM SCHOOL OF BUSINESS

Dissertation

# THREE ESSAYS ON HOW SOCIAL CONTEXT SHAPES ENGAGEMENT ONLINE

by

**JIYE BAEK**

B.S., Ewha Womans University, 2009
M.S., Yonsei University, 2012

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2018

Approved by

First Reader ______________________________

Jesse Shore, Ph.D.
Assistant Professor of Information Systems

Second Reader ______________________________

Chrysanthos Dellarocas, Ph.D.
Associate Provost for Digital Learning & Innovation
Richard C. Shipley Professor in Management
Everett W. Lord Distinguished Faculty Scholar
Professor of Information Systems

Third Reader ______________________________

Stephanie Watts, D.B.A.
Associate Professor of Information Systems

## DEDICATION

To My Family.

## ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere appreciation for my advisor, Professor Shore. He truly has been a tremendous mentor for me, and it has been an honor to be his first Ph.D. student. He was always available for advice on both professional and personal matters, and I appreciate all of his contributions of times and ideas, encouragement, and patience. Professor Shore is a tremendous role model for how to produce careful and groundbreaking scholarship. I could not imagine having had a better advisor for my Ph.D. program.

I would also like to thank my dissertation committee, Professor Dellarocas and Professor Watts, for their insightful comments and encouragements. None of my dissertation chapters would have been possible without the support of Professor Dellarocas. I admire that he can identify important research questions and develop studies that provide clear answers to said questions. Professor Watts encouraged me to think beyond the specific topics, generalize my findings, and to speak to a broader literature and audience. It has been a great privilege to work with them, and I thank them for making my dissertation defense an enjoyable moment.

Many Professors and fellow students at Boston University have made this journey fun and exciting. They have all been there to support me since the beginning of my Ph.D. process and have been a source of friendship as well as good advice and collaboration. I would like to thank them for their kindness throughout this extensive process.

Additionally, I would like to extend a special thanks to my family. Words cannot express how grateful I am for their love and encouragement throughout this journey, particularly for my parents, who raised me with unconditional love and support. I am especially grateful for my mom's constant prayers for me; her prayer was what sustained me thus far. Most of all, I am thankful for my beloved husband and best friend, K. Ever since we first met in Boston when I was in the master's program, he has been incredibly supportive and understanding despite the emotional roller coaster involved. I could not have done this without him.

Finally, I thank my God for getting me through all of the difficulties. He is the one who let me finish my degree. Thank you, Lord.

**THREE ESSAYS ON HOW SOCIAL CONTEXT**

**SHAPES ENGAGEMENT ONLINE**


**JIYE BAEK**

Boston University Questrom School of Business, 2018

Major Professor: Jesse Shore, Ph.D., Assistant Professor of Information Systems


ABSTRACT


Understanding online user engagement is a key challenge for social platforms that support the communal creation or transfer of knowledge and information. Engagement is not only a function of individual attributes but also the result of the social context that derives from platform choices. This dissertation presents several empirical examples of how social context shapes online engagement in social platforms such as social media or online communities. In the first chapter, I investigate how the social network structure influences Twitter users' information sharing behavior. I reconcile contradictory theories of the diversity of information sharing on social media using data representative of the whole population of Twitter users. In the second chapter, I investigate how online community size impacts users' platform engagement. By conducting a randomized field experiment on edX, I show a causal influence of community size on individual user's knowledge-sharing behavior, retention and performance. In the third chapter, I examine how social learning impacts out-group users' engagement in an online learning community in terms of language and culture. I broaden the scope of my research in this last chapter by studying a context that has received little attention in the platform engagement literature. I use an interdisciplinary multi-method approach in my research that includes social network

analysis, randomized field experiment, and econometrics. This dissertation involves a combination of these methods to understand user-behavior in the social platform and introduce interventions to maximize the benefit for digital platform and users alike.

## TABLE OF CONTENTS

DEDICATION .......... iv

ACKNOWLEDGMENTS .......... v

ABSTRACT .......... vii

TABLE OF CONTENTS .......... ix

CHAPTER 1: Network structure and patterns of information diversity on Twitter .......... 1

**1.1 Introduction** .......... 1

**1.2 Theories of information diversity on social media** .......... 4

1.2.1 Echo Chambers and polarization .......... 7

1.2.2 Mainstreaming Theory .......... 9

1.2.3 Beyond average behavior: macroscopic and subnetwork analyses .......... 11

**1.3 Data and methods** .......... 16

1.3.1 The Twitter Dataset .......... 16

1.3.2 Measurement of Political Slant .......... 17

1.3.3 Variables .......... 18

1.3.4 Statistical Models .......... 20

1.3.5 Descriptive Statistics .......... 26

**1.4 Results** .......... 28

1.4.1 Average behavior .......... 28

1.4.2 Beyond average behavior: macroscopic and subnetwork analyses .......... 32

**1.5 Discussion**......42

1.5.1 Summary of empirical findings......42

1.5.2 Broader implications......43

1.5.3 Limitations......47

1.5.4 Conclusion......48

CHAPTER 2: Forum size and content contribution: a MOOC field experiment......50

**2.1 Introduction**......50

**2.2 Theoretical background**......53

2.2.1 Prior studies on group size......53

2.2.2“Collaborative engagement” type forum and “individual engagement” type forum......56

2.2.3 Hypotheses......59

**2.3 Research Methodology**......60

2.3.1 Dependent Variables......61

2.3.2 Experiment Procedure......62

2.3.3 Independent variables......64

**2.4 Results**......65

2.4.1 Hypothesis 1 – Main hypothesis......65

2.4.2 Hypothesis 2 and Hypothesis 3 – performance and retention......70

**2.5 Discussion and Conclusion**......75

2.5.1 Discussion......75

2.5.2 Future extensions ....................................................................................77

2.5.3 Conclusion ..............................................................................................79

CHAPTER 3: Heterogeneous engagement patterns between in-group and out-group users ..................................................................................................................80

**3.1 Introduction** ...................................................................................................80

**3.2 Theoretical background** ................................................................................84

3.2.1 Language in today's educational setting ....................................................85

3.2.2 Language in the organizational setting .....................................................85

3.2.3 Cultural background and learning..............................................................87

**3.3 Research methodology**....................................................................................89

3.3.1 Dependent variables ....................................................................................89

3.3.2 Independent variables..................................................................................90

3.3.3 Other independent variables .......................................................................90

**3.4 Results** .............................................................................................................91

3.4.1 Effect of cultural and linguistic background on social engagement ................92

3.4.2 Effect of cultural and linguistic background on retention...............................93

3.4.3 Effect of cultural and linguistic background on performance .........................94

3.4.4 Effect of cultural and linguistic background on the rate of perfect scores.......95

3.4.5 Result validation with a different course subject: The Art of Poetry ..............96

**3.5 Discussion and Conclusion** .............................................................................98

3.5.1 Discussion....................................................................................................99

3.5.2 Future extensions ........................................................................................ 100

3.5.3 Conclusion ................................................................................................. 101

BIBLIOGRAPHY ......................................................................................................... 102

CURRICULUM VITAE………………………………………………………………….. 115

## LIST OF TABLES


Table 1: Descriptive statistics…………………………………………………………… 26

Table 2: Correlation matrix………………………………………………………………27

Table 3: Tabulation of mean incoming and outgoing political slant by account………...28

Table 4 Relationship between incoming and outgoing political slant…………………... 33

Table 5: Estimated parameter for incoming slant in regression of outgoing slant on incoming slant for news-centric core, with different definitions of which nodes belong to the core, considering only communication within the core…………... 35

Table 6: Estimated parameter for incoming slant in regression of outgoing slant on incoming slant for news-centric core, with different definitions of which nodes belong to the core, considering all tweets……………………………………… 35

Table 7: Number of accounts in news-centric core, with different definitions of which nodes belong to the core……………………………………………………… 36

Table 8: likelihood of *Z* under various permutations of nodes of the core……………… 37

Table 9: likelihood of Z under various permutations of nodes of the moderate users subgraph……………………………………………………………………………….38

Table 10 Summary of results of hypothesis tests……………………………………….. 39

Table 11: Field experiment design……………………………………………………….64

Table 12: Effect of treatments on discussion forum contributions……………………… 67

Table 13: Concentration of posts per person……………………………………………. 69

Table 14: Effect of treatments on performance and retention………………...………… 72

Table 15. Number of users for each user category-1…………………………………….92

Table 16. Number of users for each user category-2…………………………………….92

Table 17: Effect of IVs on discussion forum contribution……………………………… 93

Table 18: Effect of IVs on retention…………………………………………………….. 94

Table 19: Effect of IVs on performance………………………………………………… 95

Table 20: Effect of IVs on # assignment submit, # perfect score, and #perfect/#submit.. 96

Table 21: Effect of IVs on #Posts & Grade………………………………………………... 97

## LIST OF FIGURES

**Figure 1**: The adjacency matrix of the news-centric core, permuted by outgoing slant (top left) by incoming slant (top right) by the method of Clauset, Newman and Moore (2004) (bottom left) and by the values of the eigenvector corresponding to one of the smallest eigenvalues of the Laplacian matrix (bottom right)…….... 40

**Figure 2**: The adjacency matrix of the "moderate users" subgraph taken from the periphery, permuted by outgoing slant (top left) by incoming slant (top right) by the method of Clauset, Newman and Moore (2004) (bottom left) and by the values of the eigenvector corresponding to one of the smallest eigenvalues of the Laplacian matrix (bottom right)…………………………………………...41

**Figure 3.**: Summary diagram of connectivity patterns, distinguishing core from periphery. Overall, there is a tendency to centrism, but a majority of tweets received originate in the network core, which has a tendency to polarization. Grey circles represent accounts in the network core and periphery. Circle size is proportionate to number of accounts. Arrows indicate percentage of total connectivity within and between core and periphery. **A:** arrow size is proportional to total number of follower-followee relationships in the full data set and labeled with a percentage (e.g. 79.8% of all links are within the periphery). **B:** arrow size is proportional to (an upper bound on) the number of tweets received in the full data set, calculated as number of tweets sent multiplied by the number of followers those tweets were sent to (e.g. only 33.5% of all tweets received were both sent and received by accounts in the periphery)…………………………... 42

**Figure 4:** Cumulative distribution functions of posts per person by treatment………… 69

# CHAPTER 1: Network structure and patterns of information diversity on Twitter[1]

## 1.1 Introduction

Because anyone can post and re-share content, social media has been connected to increased participation and diversity of expression, raising hopes for a role for social media in promoting innovation, building social capital and empowering workers within firms and in society in general (An et al., 2011; Bertot, et al., 2010; Kane et. al, 2009; Woodly, 2008). Given the opportunities available in big data and the imperative to make use of them (LaValle, et al, 2013), business leaders have turned in increasing numbers to analyzing social media data in order to learn from customers (Culnan, McHugh, and Zubillaga, 2010; He, Zha and Li, 2013; Chen, Chiang, and Storey, 2012), and computer scientists have developed many tools to help achieve these ends (see e.g. Pang and Lee, 2008). Firms have also adopted internal social networking platforms in great numbers. Yammer, a popular enterprise social networking platform, claims to be used by more than 500,000 firms, including 85% of the Fortune 500 (Yammer, 2015).

It is easy to see why many see social media as potentially valuable external sources and internal conduits of diverse knowledge (Kane, Majchrzak and Ives, 2010). Innovation has long been seen as deriving from recombining diverse ideas (Schumpeter, 1934), and diverse ideas are assumed to flow through diverse networks (Hampton, Lee, and Her, 2011) like those created by connecting a diverse user base via social media. In general, diversity among individuals is thought to lead to better performance in solving

[1] This chapter is a joint work with Jesse Shore and Chrysanthos Dellarocas

problems (Hong and Page, 2004) and modern crowdsourcing approaches to innovation would seem to thrive on the fuel of diversity (Jeppesen and Lakhani, 2010).

However, information technologies, while providing historically unprecedented potential for free public expression, also provide self-regulating mechanisms that allow users to customize content feeds. In making these choices, people tend to connect with similar others (McPherson, Smith-Lovin, and Cook, 2001) and seek out information that confirms their previously-held beliefs (Nickerson, 1998). It is therefore unclear if the diverse points of view of social media users ever actually come into contact with each other, or if they cyber-balkanize themselves into “echo chambers” in which they are only exposed to ideas they already hold (Van Alstyne and Brynjolffson, 2005).

Many of the most lucid and powerful research studies on this topic to date have been in the setting of political information diversity and communication – a setting we also study in the present paper. In addition to being economically and societally consequential, political communication is an appealing setting for the study of information diversity: there is a clear left-right spectrum of opinion, which simplifies the difficult issue of how to measure diversity in a meaningful way (Page, 2010). Additionally, it is not too much of a stretch to view political communication among social media users in the United States as an example of a market for information in which two principal organizations (the Democratic and Republican political parties) are competing for attention and influence.

In prior literature, there is some evidence in favor of a tendency to echo chambers, some evidence in favor of polarization and still other evidence in favor of a tendency for people to limit themselves to expression of moderate and mainstream ideas on social media. We believe that a likely reason for the conflicted nature of the literature is that earlier work has generally focused too narrowly on unrepresentative or incomplete data sets. Social networks and online communities often have a core-periphery structure consisting of a highly interconnected core of important and active nodes, surrounded by a larger, less densely connected periphery (Borgatti and Everett, 2000; Dahlander and Fredriksen, 2012, Wu, et al., 2011).

By focusing on highly active users, prior research on the phenomena of echo chambers and polarization has arguably only emphasized the study of the network core, whose behavior is not representative of the average user of the platform (Adamic and Glance, 2005; Conover, et al, 2011; Bakshy, Messing and Adamic, 2015). Moreover, it could even be argued that by constructing their data sets by including only those individuals with clear partisan affiliation (Adamic and Glance, 2005; Bakshy, Messing and Adamic, 2015), or those who posted about politically divisive topics (Conover, et al., 2011; Barbera et al., 2015), prior research studied only users prone to political division and therefore sheds little light on the nature of social media in general. Due to its traditional survey methodology, Hampton, et al. (2014) does not have this limitation but on the other hand it also cannot answer those questions which would require large-scale network data as evidence.

Here, we seek to reconcile the differing perspectives on patterns of diversity in social media with a study of a complete cross-section of Twitter posts ("tweets") of hyperlinks, together with the associated follower network data. Our data set includes 15 million unique URLs posted by 2.7 million users based on a 300 hour data set, representing a complete record of all such activity on Twitter during the collection period. We test hypotheses implied by prior research as well as characterize the overall structure of the Twitter follower network with respect to ideological diversity. Rather than echo chambers or cross-sectional evidence of polarization, we find that, on average, Twitter accounts post links to more politically moderate (but not necessarily centrist) news sources than the links they receive in their own feed. Members of a tiny but highly followed network core behave differently from the typical user, however, and post links to sources that are more politically extreme than what they receive in their own newsfeeds. While our empirical setting is political slant, we believe that the implications go beyond this narrow application and provide a basis for understanding the structure of self-organization in social media more generally.

## 1.2 Theories of information diversity on social media

No one can read every article or interact with every user on the internet; instead, internet users must make choices about where to direct their attention. Given the human tendency toward homophily (McPherson, Smith-Lovin and Cook, 2001) and confirmation bias (Nickerson, 1998), social media users are likely to follow other users whose opinions are similar to their own. At the extreme, this could lead to fragmentation of users into ideologically narrow groups, in which people are only exposed to information that confirms their previously-held opinions (Van Alstyne and Brynjolffson, 2005; Burt, 2004). We refer to this as the "echo chambers" theory of social media. Empirical studies have confirmed some of these fears: "there is a tendency for blogs with the same political and ideological inclination to link to each other" (Adamic and Glance, 2005; Conover, et al, 2011; Hargittai, Gallo and Kane, 2008) and a tendency of readers to engage with content aligned with their ideological preferences (Lawrence, Sides and Farrell, 2010). Homophilous behavior is then magnified by algorithmic information filters on certain social media sites such as Facebook (Bakshy, Messing and Adamic, 2015; Lazer, 2015).

A related view says that homophily may not lead people to be disconnected and ignorant of opposing views, as echo chambers theory would have it. Instead, the relationship between groups of connected individuals may be mutually aware and antagonistic. Sunstein (2002, 2008) argues that when like-minded individuals discuss a controversial topic, there is a tendency for them to adopt an even more extreme position on that topic than they initially held. Barbera et al. (2015) document this process unfolding over time in partisan debate of controversial issues on Twitter. Conover et al. (2011) show that while people follow and retweet[2] like-minded others on Twitter, they mention[3] users they disagree with in the context of argument and other negative commentary, illustrating that separate groups are antagonistic, not ignorant of each other. We refer to this as the "polarization" theory of social media.

---

[2] To "retweet" is to re-share a message one has received with one's own followers
[3] On Twitter a "mention" is to include another user's handle in a tweet (post), prepended with an @ sign, which uniquely identifies the specific individual, creates a clickable hyperlink to their profile page and notifies the target individual about the mention

Despite all of this evidence, however, the idea that social media users segregate themselves into homogeneous or polarized communities is far from an established fact. Some have theorized that while social network ties may tend to be formed among similar others, there are many dimensions along which that similarity may be manifest (Watts, Dodds, and Newman, 2002) and social media users may be connected not only to people with whom they agree politically, but also to people with whom they share other similarities, such as workplace, alma mater and so on. This phenomenon of simultaneous contact with people from different contexts has been called "context collapse" and can lead users to limit their expression of potentially controversial beliefs (Marwick and boyd, 2010; see also Bernstein (2012) for a similar finding in an organizational context). Centola and Macy (2007) argue that certain phenomena – including potentially controversial expressions such as political beliefs – are most likely to occur and exert influence in the context of a highly clustered network such that there is the possibility of receiving multiple reinforcing signals from one's network neighbors. Finally, these recent theories echo the pre-internet theory of public opinion that people tend to articulate what they perceive to be the mainstream point of view or withhold their voice entirely, creating a "spiral of silence" for minority viewpoints (Noelle-Neumann, 1974). Collectively, we refer to these ideas as the "mainstreaming" theory of social media.

There is empirical support for the mainstreaming narrative of social media use as well. On average, it has been found that people are much less likely to discuss controversial topics on social media than in private (Hampton et al. 2014). For political hashtags on Twitter, repeated exposures are important precursors to an individual's adoption of those hashtags in their own posts (Romero, Meeder and Kleinberg, 2011), which could be interpreted as seeking repeated confirmation from their community before sharing something potentially controversial.

Our primary goal is to consider evidence for and against the theories of echo chambers, polarization, and mainstreaming. In this and the following sections, we therefore ask what we would expect to find in a complete cross-section of Twitter posts if the above theories were in fact true. We articulate a number of detailed hypotheses to test on this basis, but our overarching questions are simply whether Twitter shows evidence of (1) echo chambers (2) polarization and (3) mainstreaming.

### *1.2.1 Echo Chambers and polarization*

For our purposes, what cross-sectional observations would be consistent with echo chambers and polarization? First, we expect to find homophily. In other words, we would expect the typical Twitter user to tweet links to news sources with similar political slant to the slant of the content they receive from the people they follow: we expect followers and followees to tweet at a similar level of political slant (we define how we measure slant below).

> *Hypothesis 1a: The mean political slant of news sources in tweets by individuals is significantly correlated to the mean political slant of the tweets that they receive from their followees.*

If the homophily of Hypothesis 1a is strong enough to create echo chambers, we would expect not just correlation between political slants, but indeed for people to tweet at the same mean level of political slant as those that they follow.

> *Hypothesis 1b: The mean political slant of news sources in tweets by individuals is statistically indistinguishable from the mean political slant of the tweets that they receive from the people they follow.*

Sunstein (2002, 2008) argues that when like-minded individuals are connected, the views they express can be more extreme than what they would have expressed prior to deliberation, in part because of social pressure toward conformity. He refers to this phenomenon as polarization. If Twitter accounts are not just homophilous but also polarized, it would suggest that they tweet at more extreme levels of slant than the information they receive in their news feeds.

> *Hypothesis 1c: The mean political slant of news sources in tweets by individuals is more extreme than the mean political slant of the tweets that they receive from the people they follow.*

Alternatively, rather than treating the individual as the unit of analysis, we could treat network ties as the unit of analysis. In this case, we expect to see ties (follower-followee relationships) between people who tweet links to content with similar political slant. In other words, in social network terminology, we expect "assortativity" – a correlation between the presence of network ties and similarity on some attribute (Newman, 2003) – based on political slant.

> *Hypothesis 2: The level of network assortativity on the mean political slant of news sources in individuals' tweets is significantly higher than could be explained by random chance.*

Additionally, if Twitter contains echo chambers, we would expect people to follow other individuals who follow each other, amounting to political feedback loops. We expect to see network clustering. Clustering is the degree to which the people with whom a person is connected are themselves connected to each other. For example, if Julie follows Romy on Twitter because they share ideology, and Romy likewise follows John, then it is also likely that Julie follows John.

> *Hypothesis 3: The level of clustering in the follower-followee network is significantly higher than could be explained by random chance.*

Individuals in such dense clusters accrue shared, mutual knowledge as a consequence of communicating with each other (Granovetter, 1973; Hansen, 1999; Burt, 2004). Moreover, people are more likely to strongly influence one another within, rather than between, clusters of ties (Centola, 2010). As a result, we would expect people within dense clusters to be more politically similar to each other than people who are not in highly clustered network positions.

> *Hypothesis 4: The greater the clustering around an individual, the stronger the correlation between the political slant in their own tweets and the political slant in the tweets they receive from the people they follow.*

### *1.2.2* Mainstreaming Theory

What observations would constitute evidence of mainstreaming behavior? The spiral of silence (Noelle-Neumann, 1974) and context collapse (Marwick and boyd, 2010) describe a tendency for people to withhold opinions that they think are not in accordance with the mainstream or potentially offensive. Therefore, most basically, we would expect to observe more "silent reading" at less centrist levels of slant. That is, if people exhibit mainstreaming behavior, more people would choose to read tweets, but not post tweets themselves, the further away they were from the political center.

We test for silent reading in two ways. First, we simply ask if more centrist individuals post more tweets than less centrist individuals.

*Hypothesis 5a: The number of tweets sent per person is highest at the center of the political spectrum and lowest at the extreme left and right of the political spectrum.*

On the other hand, if less centrist individuals do post fewer tweets, it could simply be that they use Twitter less overall, rather than actively using it for silent reading. We therefore also test the following.

*Hypothesis 5b: The ratio of tweets sent to tweets received per person is highest at the center of the political spectrum and lowest at the extreme left and right of the political spectrum.*

The component theories that make up our "mainstreaming" theory (especially context collapse and the spiral of silence) do not explicitly make predictions about the choices individuals make about what information to consume – only the information they choose to put into the public domain themselves. Whatever information is consumed, however, we would expect individuals to Tweet material that is more politically neutral (centrist) than what they receive.

*Hypothesis 6a: On average, the mean political slant of news sources linked to in an individual's own tweets is more politically centrist than the mean political slant in the tweets they receive from the people they follow.*

A very strong version of this hypothesis that takes the original idea of a spiral of silence very literally is that whatever they read, we would find people tweeting only politically centrist content themselves.

> *Hypothesis 6b: The political slant of news sources linked to in the average Twitter user's tweets is statistically indistinguishable from the mean political slant of the population.*

### *1.2.3 Beyond average behavior: macroscopic and subnetwork analyses*

The above hypotheses are specified in microscopic terms, in that we treat individuals and network dyads as the units of analysis, and will be tested on the entirety of link-posting behavior on Twitter during the study period. These hypothesis tests serve our theoretical questions and provide the foundation of our empirical analysis.

To paint a fuller picture, however, and to better connect our work with prior research on social media, we also include a series of analyses that take other perspectives on the data. In particular, scholars of online communities have been concerned with their macroscopic core-periphery structure (Dahlander and Fredriksen, 2012; Collier and Kraut, 2012; Wasko, Teigland and Faraj, 2009), which Wu, et al. (2011) have demonstrated also describes Twitter networks. In a classic core-periphery structure, the network core is a set of nodes (individuals) that tend to be connected to each other; the periphery is a (typically larger) set of nodes that tend to be connected to nodes in the core, but not to each other (Borgatti and Everett, 1999). In the setting of Twitter, this is to say that there is a set of highly-followed accounts (the core) that tend to follow each other; more typical users (the periphery) follow members of the core, but are less likely to follow other typical users.

1.2.3.1 Analyses of macroscopic structure

Members of the news-sharing core differ from other users with respect to their network position; it is possible that they also differ from other users in terms of the correlation between incoming slant and outgoing slant. To check for this possibility, we repeat a basic analysis of the correlation between incoming and outgoing slant on two subgraphs[4] of the Twitter news-sharing network. In particular, we distinguish those accounts that are highly followed and active in posting many links to news items from those that are not.

We expect to find a higher correlation between incoming and outgoing slant in the 'news-centric core' — the subgraph of individuals who are both highly followed and post many news items — than in subgraphs defined by the other three combinations of those two variables. People in the news-centric core may be maintaining a public identity centered around news, and so may connect with fewer people for reasons other than discussion of news. They may also engage in self-conscious management (Marwick and boyd, 2010) of their list of followees — to demonstrate party loyalty, for example — which would result in a higher correlation between incoming and outgoing slant.

In contrast to members of the news-centric core, those who are highly followed but do not post many links to news items are probably highly followed for other reasons, such as celebrity, and may not pay as much attention to the variable of political slant when choosing whom to follow. Those who post many links to news items but are not

[4] A subgraph consists of a certain subset of nodes of a larger network, along with all of the links between them.

highly followed may well demonstrate homophily on political slant, but because they are less likely to be public figures, they may not be curating their followee list as self-consciously as members of the news-centric core. Finally those who are neither highly followed (among the individuals in our data comprising people who posted hyperlinks and their followers and followees) nor highly active posters of news may be less active users of Twitter, or actively using Twitter for other purposes, and thus are not expected to demonstrate less homophily on political slant than those in the news-centric core.

> *Hypothesis 7a: The mean political slant of news sources in tweets by individuals in the news-centric core of Twitter users is more highly correlated to the mean political slant of the tweets that they receive from their followees than that of people outside of the news-centric core.*

Following the logic above and Wu, et al.'s (2011) observation that 'coreness' is not a binary but rather a continuous variable, we would expect that the higher the thresholds we use to separate individuals who are "highly followed" and "post many links to news" from everybody else (i.e. the stricter the definition of what constitutions the news-centric core), the higher the correlation will be among members of that core.

> *Hypothesis 7b: The stricter the definition of what constitutes the news-centric core, the greater is the effect*[5] *of incoming slant on outgoing slant.*

1.2.3.2 Correspondence of macroscopic network structure and political slant

Earlier studies of political division on social media have shown a clear correspondence between the macroscopic structure of a network and the political slant of

---

[5] NB: here and below, "effect" is intended only in the sense of "statistical effect" and not in the sense of "causal effect."

its nodes (Adamic and Glance, 2005; Conover et al, 2011). In particular, this work shows that the network is starkly divided into two (one liberal and one conservative) modular clusters of nodes, such that nodes tend to be connected within each cluster, but only sparsely connected between clusters. By showing how cleanly political slant corresponds to network structure, these excellent studies lend strong support to the cyber-balkanization theory (Adamic and Glance, 2005) and polarization theory (Conover et al., 2011), discussed above.

These studies are nevertheless limited in two important ways. First, both studies use only a binary, liberal v. conservative representation of slant, preventing more nuanced examination of homophily. Second, both studies only consider the behavior of elites and self-identified partisans (i.e., members of the news-centric core, whom we have just argued are not representative of the typical user) and thus shed no light on how social media works as a platform for discourse for the vast majority of users.

Because our data includes a continuous representation of political slant and includes all Twitter users rather than only elites, we are able to address these two limitations. First, we are able to analyze the core separately from other users. Second, rather than consider *classifications* of nodes into two categories (liberal v. conservative), we focus instead on *permutations* of nodes defined either by political slant or other means.

A permutation is simply an ordering of the nodes of a network such each node is assigned an ordinal number from 1 to N (where N is the number of nodes in the network). Permutations can be defined by any number of means, but in the present context we will be particularly interested in permutations derived from the political slant variables: those in which the nodes are ordered from most liberal (and thus given the number 1) to the most conservative (and thus given the number N). A “good” permutation is one in which nodes that are close together in the network are close together in the ordering of nodes. In the following hypotheses (8a – 8f), we compare the quality of permutations in this sense. Section 3.4.3, below, provides more concrete details on measurement of permutation quality.

Essentially, just as prior work showed that classifying nodes into liberal v. conservative was a good fit to the macroscopic division of the network into two distinct communities, we will ask if a continuous measure of political slant is a good one-dimensional description of network structure. We wish to ask this question for both incoming and outgoing slant and for the network core and network periphery. This involves making a number of comparisons among nodal permutations. First, to establish whether permutations based on incoming or outgoing slant are “good” descriptions of macroscopic network structure, we compare them to permutations derived from standard community discovery algorithms (see below). Second, since it is not a given that incoming slant and outgoing slant are equally closely related to network structure, we compare these two slant permutations to each other. Third, we repeat this process separately for the core and the periphery.

Accordingly, we test the following hypotheses comparing the quality of nodal permutations. First we compare incoming and outgoing slant for core and periphery.

*Hypothesis 8a: An ordering of nodes in the news-centric* ***core*** *based on* ***outgoing political slant*** *is of equivalent quality to an ordering of those nodes based on* ***incoming political slant***

*Hypothesis 8b: An ordering of nodes in the* ***periphery*** *based on* ***outgoing political slant*** *is of equivalent quality to an ordering of those nodes based on* ***incoming political slant***

Then, we compare incoming and outgoing slant to community discovery algorithms for the core.

*Hypothesis 8c: An ordering of nodes in the news-centric* ***core*** *based on* ***outgoing political slant*** *is of equivalent quality to orderings of those nodes derived from* ***community-discovery algorithms****.*

*Hypothesis 8d: An ordering of nodes in the news-centric* ***core*** *based on* ***incoming political slant*** *is of equivalent quality to orderings of those nodes derived from* ***community-discovery algorithm****s.*

Finally, we compare incoming and outgoing slant to community discovery algorithms for the periphery.

*Hypothesis 8e: An ordering of nodes in the* ***periphery*** *based on outgoing political slant is of equivalent quality to orderings of those nodes derived from community-discovery algorithms.*

*Hypothesis 8f: An ordering of nodes in the* ***periphery*** *based on* ***incoming political slant*** *is of equivalent quality to orderings of those nodes derived from* ***community-discovery algorithms****.*

## 1.3 Data and methods

### *1.3.1 The Twitter Dataset*

Our Twitter data comes from Galuba et al. (2010), and contains 15 million unique URLs, tweeted by 2.7 million users. For 300 continuous hours, starting on Thursday, September 10th, 2009, 19:56:47 GMT, the Twitter Search API was continuously queried for the search string “http”. The text of each tweet returned by the query was parsed for any URLs and user names it contained. Each URL mentioned in the tweets was stored. If the URL was created by one of the popular URL shortening services (e.g. bit.ly), HTTP redirects were recursively followed to expand the URL to its original form. All the URLs were also URL-decoded to ensure uniform representation under the percent-encoding (%xx) notation. For each tweet, the Twitter API was queried for the metadata about the tweet’s author as well as all the users that the author follows.

### *1.3.2 Measurement of Political Slant*

Gentzkow and Shapiro (2011) published measurements of the political slant of the 119 most widely visited sources of online news in the United States, building on data from comScore Plan Metrix with 12 months data in 2009 (the same year as the Twitter data). Plan Metrix data come from a survey distributed electronically to approximately 12,000 comScore panelists. The survey asks panelists the question “In terms of your political outlook, do you think of yourself as. . .? [very conservative / somewhat conservative/ middle-of-the-road/ somewhat liberal / very liberal]”. The average number of daily unique visitors in each category is reported by comScore for each site for each month.

Using this data, they posit the model of utility of a visit to a website in equation 1. The utility is that of user $i$ going to site $j$ on visit $k$ on a given day, given the site quality $\alpha$, political slant $\gamma$, and dummy variable $c$ set to 1 if visitor $i$ is conservative and -1 if they are liberal (they omit data from individuals who answered “middle of the road”).

$$u_{ijk} = a_j + 2(c_i - 1)\gamma_j + \varepsilon_{ijk} \tag{1}$$

They fit a Generalized Mixed Model to the visit data, under the discrete choice modeling assumption that the visit would be made if and only if $u_{ijk} \geq u_{irk}\ \forall r \neq j$. We use the estimated parameter $\gamma$ as our measure of political slant. We also use $\alpha$ as a control variable indicating site quality. For the analysis, we use all tweets that contain any of the 119 domain URLs from Gentzkow and Shapiro (2011).

Although Plan Metrix data are only available for relatively large sites, visits to news sites are highly concentrated. The 119 sites in the sample represent over 95% of all visits to news sites via independent browsing online (Gentzkow and Shapiro, 2011), and given the greater expected concentration of exposure on social media than independent browsing (Hong, 2012), the sample is expected to be even more completely representative for the setting of Twitter.

### *1.3.3 Variables*

#### 1.3.3.1 Individual level variables

For each user, we calculated mean incoming political slant (incoming slant) and mean outgoing source slant (outgoing slant). For the outgoing slant, we average the political slant of each URL source that the user tweeted. Incoming slant is the averaged slant score of every URL tweeted by the individuals whom a user follows (his/her followees). For both incoming and outgoing slant, if a news source was tweeted more than once, its slant score would be counted more than once in the average. Similarly, we calculate mean incoming (outgoing) quality for each individual from Gentzkow and Shapiro's $\alpha$. Finally we tabulate the count (number) of incoming and outgoing tweets for each user. Note that we fit models on data for users that both sent and received tweets containing links to news sources; to calculate incoming slant, however, we consider tweets from all users, including those who did not receive any news links in their own timelines. The output of those twitter users who are widely followed but do not follow other accounts (typically public figures) is therefore still accounted for in the data.

To test Hypothesis 5, we estimate the empirical frequency distribution and probability density function of tweets across the domain of political slant present in our data, using a kernel density estimator, for both incoming and outgoing tweets.

1.3.3.2 Network variables

Using the Twitter data, we construct a follower-followee network. A directed network tie exists from user *i* to user *j* if user *j* is a follower of user *i*. From this data, we calculated aggregate clustering (Watts and Strogatz, 1998), and assortativity on political slant (Newman, 2003). The clustering coefficient captures the degree to which one's followers and followees also follow each other. More specifically, it is a measure of how many links there are among a node's neighbors, divided by the number of links that could exist among a node's neighbors. Assortativity is analogous to a measure of correlation between two nodes having a link and having similar values on an attribute (political slant, in this case).

### *1.3.4 Statistical Models*

We fit ordinary least squares (OLS) models to test Hypotheses 1a, 1b, 1c, 4, and 6. Hypotheses 1a, 1b, 1c, and 6a and 6b are concerned with the relationship between the mean incoming slant and the mean outgoing slant, while Hypothesis 4 is concerned with the mediating influence of network clustering.

The significance of a difference between observed values and what would be expected by chance of clustering and assortativity on slant (Hypotheses 2 and 3) must be established by comparison to null distributions. To calculate such null distributions for clustering and assortativity, we use the 'configuration model' to generate random graphs (Newman, Strogatz and Watts, 2001) that preserve both the degree distribution and the joint distribution of outgoing slant and degree over individuals. We then calculate clustering and assortativity on these random graphs to form distributions of these values that would be found under the null hypothesis that there was no true tendency toward clustering or slant-based assortativity.

To test Hypothesis 5, we need to assess whether there are systematic differences in the ratio of tweets read to tweets received across the spectrum of political slant. To do this, we consider two regressions using the logarithm of the count of tweets sent divided by the logarithm of the count of tweets received as the outcome variable. In one regression, we use the mean outgoing slant as the predictor variable, to see if more politically central tweeters are more active. In the other, we use the difference between mean incoming slant and mean outgoing slant as the predictor variable to test whether people tweeting more centrist material tweet more, even if they are not centrist in absolute terms.

1.3.4.2 Core-Periphery structure

Hypotheses 7a and 7b concern the difference between the behavior of people who are highly followed and post many news articles and other individuals. To test these, we select nodes that are greater than or equal to some threshold quantiles of outdegree and number of news stories posted, for $s, t \in \{0.75, 0.80, 0.85, 0.95\}$. We then consider the induced subgraph containing only those nodes that have outdegree greater than *s%* of nodes and have posted more news stories than *t%* of nodes in the full network. We then regress outgoing slant on incoming slant for only those tweets coming from within this subgraph and separately, for all tweets coming from any source using OLS and report the estimated parameter.

1.3.4.3 Concordance of community structure and slant

Hypotheses 8a-8f stipulate a concordance between the macroscopic community structure of the network and the political slant of the nodes. For these hypotheses, we are asking if permutations based on slant are good in the sense that nodes that are closely connected in the network are also close together in the permutation ordering. However, it is unclear a priori how to measure such correspondence between slant and structure, and then, how to determine if a given level of correspondence between slant and structure is a lot of correspondence or only a little. In other words, how good is good? .In order to test these hypotheses, we therefore (1) define a measure of permutation quality (2) use standard community-discovery algorithms from the literature to define permutations that represent network structure well and measure their quality, (3) measure the quality of permutations based on slant, and (4) define a significance test to determine if the quality of the slant-based permutations are significantly worse than the community-discovery algorithmic permutations.

Hypotheses 8a, 8c, and 8d concern the news-centric core, and hypotheses 8b, 8e and 8f concern those outside of the news-centric core. The "core" subgraph is defined as above, using nodes greater than or equal to some quantiles *s*, *t* of outdegree and news posting activity such that a regression of outgoing slant on incoming slant yields the highest estimated parameter. Because of the computational expense of conducting these analyses on all ~213,000 nodes outside of the news-centric core using our methods, we test the latter hypotheses on a subgraph consisting only of moderate users. We define this subgraph as giant component of those accounts between the 25$^{th}$ and 75$^{th}$ percentiles for outdegree and less than the 75$^{th}$ percentile for number of news items posted. This results in a subgraph of 75,640 Twitter accounts (a little more than one third of all news-active accounts), which omits the large number of least active and least followed accounts.

To measure quality of permutations, we start with the intuitions that connected nodes (those that follow each other) should be close together in a "good" permutation and that ties (matrix entries equal to 1) between nodes whose indices are close together in a given permutation will be close to the diagonal of the permuted adjacency matrix. Conversely, of course, ties between nodes that are far apart in the permutation will be far from the diagonal in the permuted adjacency matrix. We use these intuitions to define an idealized model against which to compare the observed permuted data such that we can evaluate them quantitatively.

We define the idealized model as a probability matrix, Z, such that matrix entries (network ties) closest to the diagonal are modeled as having probability 1, with linearly declining probability further from the diagonal. Note that this is not a fitted model, so the "probabilities" are not estimated from the data; as an idealized model, the matrix of probabilities functions more as a "scoring matrix:" we calculate the likelihood of the idealized model, *Z*, under the observed permuted data in question.

Concretely,

$Z_{i,j} = 0$ (No self loops)

$Z_{i,j} = 1$ , $j \in \{i+1, i-1\}$ (Highest probability closest to diagonal)

$Z_{i,j} = Z_{i,j-1} - 1/n$ , $j \geq i+2$

(Decreasing probability with distance from diagonal)

$Z_{i,j} = Z_{i,j+1} - 1/n$ , $j \leq i-2$

where *i* and *j* are row and column indices, respectively, and *n* is the number of nodes. To calculate the likelihood $L_p$ of *Z* under some permutated adjacency matrix, *P*, we simply take the Hadamard (pointwise) product of *Z* and *P* and take the sum of all the entries in the resulting matrix.

$$L_p = \sum_{ij} Z_{ij} P_{ij} \qquad (2)$$

The higher the likelihood, the more closely the permuted observed matrix adheres to the idealized model.

In addition to the permutations implicit in sorting the nodes according to their outgoing and incoming slants, we also consider two algorithmically-defined permutations deriving from the network structure (pattern of follower/followee ties) alone, rather than taking into account political slant or any other nodal attribute. Algorithmically defined permutations attempt to place nodes close together in the permutation ordering if they are close together in the network. In the first of these, we follow the usual procedure of spectral clustering: we calculate the eigenvectors of the Laplacian matrix (a transformation of the adjacency matrix representation of the network) and then rank nodes according to the values in the eigenvector corresponding to one of the smallest eigenvalues not equal to zero[6] (see e.g. (Dhillon, 2001; Von Luxburg, 2007) for more detail). To find the smallest eigenvalues and corresponding eigenvectors of the "moderate user" subgraph, we use ARPACK numerical methods (Lehoucq, Sorensen, and Yang, 1998), which are therefore approximate. In the second algorithmically-defined permutation, we use the method of Clauset, Newman and Moore (2004), as implemented in the igraph analytical software package (Csardi and Nepusz, 2006), which produces a full hierarchical dendrogram as a side effect of finding a smaller number of communities. We simply take the ordering of nodes at the bottom level of that dendrogram as our permutation.[7]

We visualize the difference between these four permutations of nodes (two by

---

[6] Usually the best description of the macroscopic structure of a network is found in values of the eigenvector corresponding to the smallest non-zero eigenvalue, but not always. We therefore consider the eigenvectors corresponding to the 5 smallest eigenvalues and take the best, where best is defined as yielding the highest likelihood of $Z$ (see below).

[7] We grant that this permutation is based on a partial, rather than full ordering of nodes, since the first pair of nodes that are grouped together in a given branch of the dendrogram could appear in either order in the final permutation. However, since the number of such interchangeable pairs is small, and the distance that each node in these pairs could move in the permutation is at maximum 1 spot, we take the partial ordering output from the R function to be representative of the quality of all such possible permutations.

slant and two by community discovery algorithm) by plotting the adjacency matrix, with the rows and columns in permutation order for the core and moderate users subgraphs (Figures 1 and 2). On these plots, if rows are indexed by $i \in \{1 \ldots N\}$, and columns are indexed by $j \in \{1 \ldots N\}$, then a point at location $(i,j)$ on the visualization indicates that there exists a tie between node $i$ and node $j$ (account $j$ follows account $i$ on Twitter). The closer a permutation is to the idealized model, $Z$, the more the points in these plots will be concentrated toward the matrix diagonal.

It remains to determine how much higher a likelihood has to be to be considered significantly better than the likelihood of an alternative permutation of the observed matrix. Typically, likelihoods are compared via likelihood ratio tests. Strictly speaking, however, likelihoods calculated from a matrix probability model on two different permutations of the same data are not nested, and thus the chi-squared limiting distribution on the traditional likelihood ratio test cannot be assumed. Instead, we calculate a critical value for distinguishing between the likelihoods of this model under these two permutations computationally.

We calculate the worst-case reduction of likelihood due to incorrect ordering for each of 5% of the total number of nodes and tabulate the reduction of likelihood that would occur if the edges incident to those nodes were moved as far away from the diagonal as possible. We repeat this procedure 1000 times and take the 95% percentile of the resulting distribution to be the critical value, greater than which we would consider two likelihoods different assuming a 5% type one error rate.

### *1.3.5 Descriptive Statistics*

In our dataset, after processing and selecting those who both received and sent tweets containing links to the sources we covered, we were left with a group of 215,174 Twitter accounts that posted 27,127,798 tweets, of which 908,565 contained a hyperlink to one of the 119 news sources for which Gentzkow and Shapiro provide an estimated political slant. There were 14,870,199 follower-followee relationships among these accounts, and only 7177 accounts did not follow and were not followed by any of the other accounts that posted links to the 119 domains.

There were 165,624 accounts that had outgoing slant less than zero (liberal) and 49,550 accounts that had outgoing slant greater than zero (conservative). This is consistent with Pew's survey results, which indicate that liberals significantly more active on social media (Pew Research Center, 2012). Descriptive statistics are in Table 1 and correlations are in Table 2. We also tabulated the counts of users by mean incoming slant and mean outgoing slant. As Table 3 shows, we find that some people read tweets from the opposite side of the political spectrum from the side they tweet on themselves.

Table 1: Descriptive statistics

| | Min | Mean | Max | Sd | # NAs |
|---|---|---|---|---|---|
| Mean outgoing slant | -1.5568 | -0.233 | 2.263 | 0.393 | 0 |
| Count of outgoing tweets | 1 | 4.156 | 3321 | 22.495 | 0 |
| Quality of outgoing tweets | 0 | 5.802 | 8.630 | 1.251 | 0 |
| Mean incoming slant | -1.557 | -.226 | 1.879 | 0.232 | 6708 |
| Count of incoming tweets | 0 | 715.8 | 100984 | 2238.6 | 0 |
| Quality of incoming tweets | 0 | 5.847 | 8.630 | 0.768 | 6708 |
| Count of outgoing retweets | 0 | 0.7631 | 867 | 3.96 | 0 |
| Outdegree | 0 | 69.11 | 52731 | 348.7 | 0 |
| Indegree | 0 | 69.11 | 32511 | 252.66 | 0 |
| Clustering Coefficient | 0 | 0.108 | 1 | 0.116 | 6110 |

Table 2: Correlation matrix

| | Mean outgoing Slant | Count of outgoing tweets | Quality of outgoing tweets | Mean incoming Slant | Count of incoming tweets | Quality of incoming tweets | outdegree | indegree | Clustering Coefficient | Count of retweets | ln(outdegree+2)/ ln(indegree+2) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| outSlant | 1 | 0.017 | -0.1342 | 0.395 | 0.063 | -0.087 | 0.025 | 0.0352 | 0.0163 | 0.0315 | 0.0276 |
| outCount | 0.017 | 1 | -0.005 | 0.034 | 0.086 | -0.015 | 0.0554 | 0.0463 | -0.0419 | 0.2044 | 0.1243 |
| outQuality | -0.134 | -0.005 | 1 | -0.084 | -0.023 | 0.243 | -0.0095 | -0.0154 | -0.0084 | -0.0221 | -0.0167 |
| inSlant | 0.395 | 0.034 | -0.0836 | 1 | 0.09 | -0.216 | 0.0413 | 0.0586 | 0.0351 | 0.0423 | 0.0278 |
| inCount | 0.063 | 0.086 | -0.023 | 0.09 | 1 | -0.033 | 0.6449 | 0.8956 | -0.1116 | 0.0893 | -0.0269 |
| inQual | -0.087 | -0.015 | 0.243 | -0.216 | -0.033 | 1 | -0.021 | -0.0345 | 0.033 | -0.0256 | -0.0818 |
| outdegree | 0.025 | 0.055 | -0.0095 | 0.041 | 0.645 | -0.021 | 1 | 0.7685 | -0.0992 | 0.0413 | 0.0943 |
| indegree | 0.035 | 0.046 | -0.0154 | 0.059 | 0.896 | -0.035 | 0.7685 | 1 | -0.1125 | 0.0537 | -0.0074 |
| Clustering Coef | 0.016 | -0.042 | -0.0084 | 0.035 | -0.112 | 0.033 | -0.0992 | -0.1125 | 1 | -0.0159 | -0.1747 |
| rtCount | 0.031 | 0.204 | -0.0221 | 0.042 | 0.089 | -0.026 | 0.0413 | 0.0537 | -0.0159 | 1 | -0.0011 |
| ln(OD+2)/ ln(ID+2) | 0.028 | 0.124 | -0.0167 | 0.028 | -0.027 | -0.082 | 0.0943 | -0.0074 | -0.1747 | -0.0011 | 1 |

Table 3: Tabulation of mean incoming and outgoing political slant by account

| Mean out. slant | Mean incoming slant (-1.75,-1.25] | (-1.25,-0.75] | (-0.75,-0.25] | (-0.25,0.25] | (0.25,0.75] | (0.75,1.25] | (1.25,1.75] |
|---|---|---|---|---|---|---|---|
| (1.75,2.25] | 0 | 0 | 9 | 38 | 18 | 2 | 0 |
| (1.25,1.75] | 0 | 2 | 29 | 100 | 71 | 6 | 0 |
| (0.75,1.25] | 0 | 25 | 1073 | 2847 | 1153 | 100 | 3 |
| (0.25,0.75] | 1 | 112 | 4959 | 11640 | 3335 | 123 | 4 |
| (-0.25,0.25] | 6 | 536 | 27949 | 47400 | 2432 | 125 | 7 |
| (-0.75,-0.25] | 4 | 519 | 43881 | 34305 | 973 | 67 | 1 |
| (-1.25,-0.75] | 2 | 1840 | 14856 | 7241 | 281 | 19 | 0 |
| (-1.75,-1.25] | 0 | 3 | 222 | 122 | 1 | 0 | 0 |

## 1.4 Results

### *1.4.1 Average behavior*

We begin with results for average behavior of all individuals who tweeted a link to one of the sites covered by the Genztkow and Shapiro data. We report regression coefficients from OLS models in Table 4. Most notably, the estimated parameter for the mean political slant of sites linked-to in incoming tweets was very stable at 0.6568 to 0.6720 in all models. Additional statistical results are mentioned in line with the text, below.

1.4.1.1 Echo Chambers

Hypotheses 1a and 1b are statements about homophily, operationalized as the correlation between the political slant in incoming versus outgoing tweets. Hypothesis 1a stipulates that there is a significant correlation between those quantities. As just mentioned above, we estimated a positive and significant regression parameter for this relationship across all models. We do find homophily, and the hypothesis is therefore supported.

Hypothesis 1b makes a stronger statement about the relationship between incoming and outgoing slant, claiming that they are equal. If this hypothesis were true, we would expect the regression parameter to be equal to 1.0 (meaning the outgoing slant is equal to 1.0 times the incoming slant, and therefore equal). Given the standard errors of the estimated coefficients, this hypothesis is rejected: the outgoing slant is not equal to the incoming slant. The slant at which people tweet is correlated with but not equal to the slant of the material they receive. Hypothesis 1c says that the estimated parameter for incoming slant's effect on outgoing slant should be greater than one. Hypothesis 1c is therefore likewise rejected.

Hypothesis 2 is similar to Hypothesis 1a in its focus on homophily, but considers network ties to be the unit of analysis. Specifically, it states that individuals are more likely to follow and be followed by people with similar politics, as measured by political-slant based assortativity. We found an observed assortativity of 0.1624 and calculated a mean assortativity of 1000 null models (described above) of $-2.767 \times 10^{-6}$, with a standard deviation of 0.0003. The frequentist probability of the observed assortativity being drawn from the null distribution is less than one tenth of one percent, and thus we reject the null and support hypothesis 2. There is a significant assortativity based on political slant.

Hypothesis 3 states that there is a statistically significant tendency toward network clustering, that is, that the people whom an individual follows and is followed by are likely to also follow each other. We found an empirical aggregate level of clustering equal to 0.1083 and calculated a mean clustering of 1000 configuration models equal to

0.0863 with a standard deviation of 0.0002. The observed level of clustering is greater than that expected by chance, with the probability of the observed value being drawn from the null distribution being less than one tenth of one percent. We thus fail to reject Hypothesis 3. Like most social networks, there is a statistically significant tendency to clustering over and above what we would expect by chance when we hold the degree distribution constant.

Hypothesis 4 speaks to the notion that clustering is associated with "echo chambers" in social media. It is intended to represent the notion that people in clustered positions (those whose followers and followees also follow each other), may be even more likely than those in unclustered positions to tweet at a similar political slant to their network neighbors. Table 4, model 4 reports a regression coefficient of 0.0585 for clustering. Therefore we fail to reject Hypothesis 4, as we do find evidence that people in positions of high clustering tweet more similarly to the people they follow than people in positions of low clustering. However, we must note that the effect of clustering, while statistically significant, is not of great magnitude. A one-standard deviation increase in clustering coefficient would only result in a predicted increase in outgoing slant from 0.6721 times the incoming slant to 0.6789 times the incoming slant.

1.4.1.2 Mainstreaming

Our mainstreaming theory states that people are less likely to voice opinions that they perceive are not widely held. Hypothesis 5a operationalizes this theory as a claim that individuals on either end of the political spectrum (far away from the political center) will tweet fewer times per person than those in the political center. Hypothesis 5b says that people on either end of the spectrum will tweet less for each tweet they receive; those at the political center will tweet at the highest rate per tweet they receive. Our evidence on these hypotheses is mixed. For Hypothesis 5a, we do not find evidence of more tweets by centrist accounts. Rather, we find a slightly higher rate of tweeting by more conservative accounts. For Hypothesis 5b, we do find that accounts that tweet at an outgoing slant of -0.15 (between the sample mean of -0.23 and the political center, 0.0) tweet slightly more on average per tweet that they receive than more politically distal Twitter accounts. We also find that accounts that tweet more centrally than the mean slant of their followees tweet more times per tweet received. However, despite the vanishingly small p-values for the estimated parameters in these regressions, the effect magnitudes and the $R^2$s are also tiny for tests of both hypotheses. Therefore, we reject Hypothesis 5a and accept Hypothesis 5b but only trivially, and we do not report the parameter estimates here.

Hypothesis 6a returns to the relationship between incoming and outgoing political slant and stipulates that people tend to tweet more centrist material than the material they read in their own newsfeeds. Hypothesis 6b is much stronger and stipulates that the average Twitter user tweets at the same level of slant as the mean of the whole population. As already stated, the estimated coefficient from Table 4 was 0.67, which is statistically significantly less than 1, and greater than 0. In other words, hypothesis 6a is supported and 6b is rejected. Overall, Twitter accounts do tend to tweet more centrist material than the material posted by the accounts they follow, but not necessarily at or near the political mean of the population.

### *1.4.2 Beyond average behavior: macroscopic and subnetwork analyses*

#### 1.4.2.1 Core-periphery structure

Hypotheses 7a and 7b concern a core of highly followed users who are active in posting links to news stories. Hypothesis 7a states that the correlation between incoming slant and outgoing slant is stronger within the core, and hypothesis 7b states that the higher the standards used to define the core, the more similar outgoing slant will be to incoming slant. Tables 5, 6, and 7 summarize evidence relevant to these hypotheses: the estimated parameter for the effect of incoming slant on outgoing slant is reported for different definitions of the core with respect to both outdegree and news posting activity. When we consider only those tweets from inside the core, the maximum parameter estimate that we find is 1.0863; when we consider all tweets from all sources, we find an even higher parameter: 1.1723. Given our previous results on the centrist tendencies of the majority of users, this difference in parameters is expected.

All specifications we tested for the core yielded a higher parameter for the effect of incoming slant on outgoing slant than the one we found in our study of the whole population (Table 4); we thus fail to reject Hypothesis 7a.

As for Hypothesis 7b, there is a clear pattern evident in Tables 5 and 6: the more restrictive the definition of the core, the higher the estimated effect of incoming slant on outgoing slant. In both tables, the higher the quantile of degree used as a threshold for core membership, the greater the estimated parameter. The magnitude of the effect of raising the quantile threshold of news posting activity is smaller than that for degree and in Table 5 is generally highest at the $90^{th}$ quantile of news posting in each column, except the last column, corresponding to the strictest definition of the core. In this right-most column of both tables, the maximum parameter estimate is found when we define the core as consisting only of those individuals who are above the 95th percentile for both outdegree and number of news items posted. We therefore cannot reject hypothesis 7b: the stricter the definition of what constitutes the news-centric core, the greater is the effect of incoming slant on outgoing slant.

Table 4 Relationship between incoming and outgoing political slant

| | DV: Mean slant of sites in outgoing tweets | | | |
|---|---|---|---|---|
| | I | II | III | IV |
| Mean slant, sites in incoming tweets | 0.6674 ***<br>0.0034 | 0.6575 ***<br>0.0035 | 0.6568 ***<br>0.0035 | 0.6721 ***<br>0.0036 |
| ln(Count of incoming tweets) | | 0.0082 ***<br>0.0004 | 0.0142 ***<br>0.0007 | 0.0153 ***<br>0.0007 |
| ln(Count of outgoing tweets) | | 0.0044 ***<br>0.0009 | 0.0023 *<br>0.0010 | 0.0019<br>0.0010 |
| Mean quality of sites in incoming tweets | | -0.0018<br>0.0011 | -0.0022 *<br>0.0011 | -0.0023 *<br>0.0011 |
| *ln*(# followers+2) | | | -0.0102 ***<br>0.0010 | -0.0102 ***<br>0.0011 |
| *ln*(# followers+2) ÷ *ln*(#followees+2)† | | | 0.0554 ***<br>0.0035 | 0.0618 ***<br>0.0035 |
| Clustering coefficient | | | | 0.0585 ***<br>0.0075 |
| Intercept | -0.0831 ***<br>0.0011 | -0.1167 ***<br>0.0063 | -0.1634 ***<br>0.0066 | -0.1766 ***<br>0.0073 |
| # of Twitter accounts | 208,463 | 208,460 | 208,458 | 204,465 |
| Adjusted $R^2$ | 0.156 | 0.158 | 0.159 | 0.164 |

Notes: standard errors are printed below parameter estimates.
***: p< 0.001; **: p< 0.01; *: p < 0.05
† the logarithm of the number of followers/ees plus 2 is taken to avoid dividing by zero for those with no followees.

1.4.2.2 Polarization in the core

Tables 5 and 6 show that the more restrictive the definition of the network core, the higher the parameter estimate for the estimate of the relationship between incoming slant and outgoing slant. For moderately restrictive definitions of the network core (for example, those accounts with greater than the 85$^{th}$ percentile of outdegree and 90$^{th}$ percentile of news items posted in Table 5) the parameter estimates for incoming slant are not significantly different from 1.0. We therefore would not be able to reject Hypothesis 1b – that the mean political slant of news sources in tweets by individuals is statistically indistinguishable from the mean political slant of the tweets that they receive from the people they follow – for the news-centric core thus defined. However, for the most restrictive definitions of the core the average outgoing slant is in fact more extreme than the average incoming slant, indicating not so much echo chambers, in which we would expect people to be reading and tweeting at the same political slant, but rather a tendency to polarization, in which we see people reading more centrist material on average than what they tweet themselves.

What emerges is a more nuanced picture of the whole. The vast majority of Twitter accounts that post news items do not post many of them, have a moderate number of followers among other news-posting accounts, and tend to post news items from more centrist sources than what they read themselves. On the other hand, a small minority of Twitter accounts constituting the network core post relatively many news from more politically polarized news sources than those in their own news feeds.

This is not to say that the core only posts material from the political extremes or that the periphery only post centrist material, simply that on average the core posts more extreme material and the periphery posts more centrist material *than the accounts they follow*. Our results also do not support the extrapolation that the centrist tendency of accounts in the periphery is due to a tendency of following more extreme accounts in the core. We regressed outgoing slant on incoming slant after excluding core accounts and the tweets originating from those accounts (in the manner of Table 5, but for the periphery rather than for the core). After thus removing the effects of the core from the periphery, the estimated parameter for incoming slant's effect on outgoing slant was 0.7030, only slightly higher than the estimate for the complete data.

Table 5: Estimated parameter for incoming slant in regression of outgoing slant on incoming slant for news-centric core, with different definitions of which nodes belong to the core, considering only communication within the core

| | | Quantile of outdegree | | | | |
|---|---|---|---|---|---|---|
| | | 75th | 80th | 85th | 90th | 95th |
| Quantile of news posting | 95th | *0.9522 | 0.9731 | 0.9929 | 1.0099 | *1.0863 |
| | 90th | *0.9630 | 0.9804 | 1.0029 | 1.0281 | *1.0802 |
| | 85th | *0.9500 | *0.9689 | 0.9867 | 1.0196 | *1.0623 |
| | 80th | *0.9451 | *0.9622 | 0.9789 | 1.0162 | *1.0659 |
| | 75th | *0.9236 | *0.9433 | *0.9663 | 0.9988 | *1.0447 |

Note: * indicates 95% confidence interval for the mean does not contain 1.0

Table 6: Estimated parameter for incoming slant in regression of outgoing slant on incoming slant for news-centric core, with different definitions of which nodes belong to the core, considering all tweets.

| | | Quantile of outdegree | | | | |
|---|---|---|---|---|---|---|
| | | 75th | 80th | 85th | 90th | 95th |
| Quantile of news posting | 95th | *1.0362 | *1.0594 | *1.0878 | *1.1078 | *1.1723 |
| | 90th | *1.0294 | *1.0452 | *1.0714 | *1.1016 | *1.1477 |
| | 85th | 1.0072 | *1.0284 | *1.0505 | *1.0874 | *1.1270 |
| | 80th | 0.9944 | 1.0142 | *1.0370 | *1.0811 | *1.1315 |
| | 75th | *0.9613 | 0.9848 | 1.0093 | *1.0573 | *1.1093 |

Note: * indicates 95% confidence interval for the mean does not contain 1.0

Table 7: Number of accounts in news-centric core, with different definitions of which nodes belong to the core.

| | | Quantile of degree | | | | |
|---|---|---|---|---|---|---|
| | | 75th | 80th | 85th | 90th | 95th |
| Quantile of news posting | 95th | 6332 | 5489 | 4553 | 3478 | 1956 |
| | 90th | 11003 | 9427 | 7708 | 5705 | 3157 |
| | 85th | 15929 | 13480 | 10799 | 7866 | 4303 |
| | 80th | 20882 | 17473 | 13805 | 9879 | 5347 |
| | 75th | 30435 | 24973 | 19392 | 13513 | 7034 |

1.4.2.2 Correspondence of community structure and political slant

Is the Twitter follower network organized according to the political slant of its nodes? Here we make several comparisons between permutations of nodes based on slant to those deriving from the patterns of ties alone using community discovery algorithms. Figures 1 and 2 visualize the adjacency matrix of the core and typical users subgraph according to the spectral and slant permutations of the nodes. The following paragraphs quantify these comparisons.

1.4.2.3 Core permutations

Section 3.4.3, above, describes the matrix probability model of which we calculate the likelihood on the spectral and slant partitions. In short, the likelihood of this model is calculated by pointwise multiplication of $Z$ (see above) with some permuted adjacency matrix $P$ and will be high to the extent that a given permutation concentrates tie weight toward the diagonal of a matrix. Critical values for differences in likelihoods were determined computationally.

The likelihoods of the diagonal gradient model under the four permutations of the core subgraph are presented in Table 8. Critical values are also given, which represent the 95th percentile of likelihood reduction expected when 5% of nodes are removed from their proper place in the permutation and placed in a worst-fit location in the permutation.

Table 8: likelihood of $Z$ under various permutations of nodes of the core

| | Likelihood of probability model $Z$ | Critical value to be considered worse than this permutation |
|---|---|---|
| Clauset, Newman and Moore | 165047.4 | 8864.9 |
| Laplacian Eigenvector | 164627.8 | 8917.3 |
| Incoming Slant | 163685.5 | 8806.4 |
| Outgoing slant | 149002.7 | |

Strikingly, the outgoing slant permutation is a much poorer fit to the diagonal gradient model than any of the other three permutations, and indeed the likelihood of Z given the outgoing slant permutation is significantly less than the other three according to our critical values, leading us to reject Hypothesis 8c. The incoming slant permutation is a better representation of the whole network than the outgoing slant permutation in the sense that nodes that are close together in the incoming slant permutation tend to be more closely connected in the network than nodes that are close together in the outgoing slant permutation are. We therefore reject Hypothesis 8a. Additionally, for the core, the likelihood of the incoming slant permutation is not less than the likelihoods of the Clauset, Newman and Moore (2004) and Laplacian eigenvector-based permutations, minus their critical values. We fail to reject Hypothesis 8d and find incoming slant to be an equivalently good description of network structure as standard community discovery algorithms.

1.4.2.4 Typical users subgraph permutation and classification

Table 9: likelihood of Z under various permutations of nodes of the moderate users subgraph

| | Likelihood of probability model *Z* | Critical value to be considered worse than this permutation |
|---|---|---|
| Clauset, Newman and Moore | 209335.8 | 10718.6 |
| Laplacian Eigenvector | 199606.2 | 14282.2 |
| Incoming Slant | 170276.5 | 10733.2 |
| Outgoing slant | 163009.4 | |

The likelihood of the diagonal gradient model under the four permutations of the "moderate users" subgraph is presented in Table 8. Like the results for the core, this subgraph yields the highest likelihood under the Clauset, Newman and Moore (2004) permutation, followed by the spectral permutation, the incoming slant permutation and the outgoing slant permutation. However, although the order of results is the same, we draw different conclusions as follows. The outgoing slant permutation is not significantly worse than the incoming slant permutation given our definition of significance based on the critical value. We therefore fail to reject hypothesis 8b: incoming and outgoing slant are equivalently good descriptions of network structure for moderate users in the periphery. Additionally, both of the slant permutations are worse than the likelihood minus the critical value for both of the community discovery algorithms. We therefore reject Hypotheses 8e and 8f: community discovery algorithms produce better descriptions of network structure than either incoming or outgoing political slant.

A summary of hypothesis test outcomes is presented in Table 10.

Table 10 Summary of results of hypothesis tests

| Hypothesis | Result |
|---|---|
| Hypothesis 1: relationship between inslant and outslant | |
| 1a: outslant correlated with inslant | * |
| 1b: outslant = inslant | |
| 1c: outslant more extreme than inslant | *(core only) |
| Hypothesis 2: Assortativity based on political slant | |
| 2: slant assortativity > random | * |
| Hypothesis 3: Transitivity | |
| 3: transitivity > random | * |
| Hypothesis 4: clustering and slant | |
| 4: clustering increases effect of inslant on outslant | * (trivially) |
| Hypothesis 5: Higher rate of tweeting at political center | |
| 5a: #sent highest at political center | |
| 5b: #sent ÷ #received highest at political center | * (trivially) |
| Hypothesis 6: Tendency to centrism | |
| 6a: outslant more centrist than inslant | * |
| 6b: individual outslant = population mean outslant | |
| Hypothesis 7: Members of core are less centrist | |
| 7a: effect of inslant higher for members of the core | * |
| 7b: stricter definition of core → higher effect of inslant on outslant | * |
| Hypothesis 8: political slant is a good summary of network structure | |
| 8a: outgoing slant permutation ~ incoming slant permutation (core) | (incoming better) |
| 8b: outgoing slant permutation ~ incoming slant permutation (non-core) | * |
| 8c: outgoing slant permutation ~ community discovery alg. Permutations (core) | (algo better) |
| 8d: incoming slant permutation ~ community discovery alg. permutations (core) | * |
| 8e: outgoing slant permutation ~ community discovery alg. Permutations (non-core) | (algo better) |
| 8f: incoming slant permutation ~ community discovery alg. Permutations (non-core) | (algo better) |

*Notes: "*" indicates the null hypothesis was rejected, and evidence was found for the stated alternative hypothesis. "~" indicates that the quality of one permutation is equivalent to the quality of the other permutation*

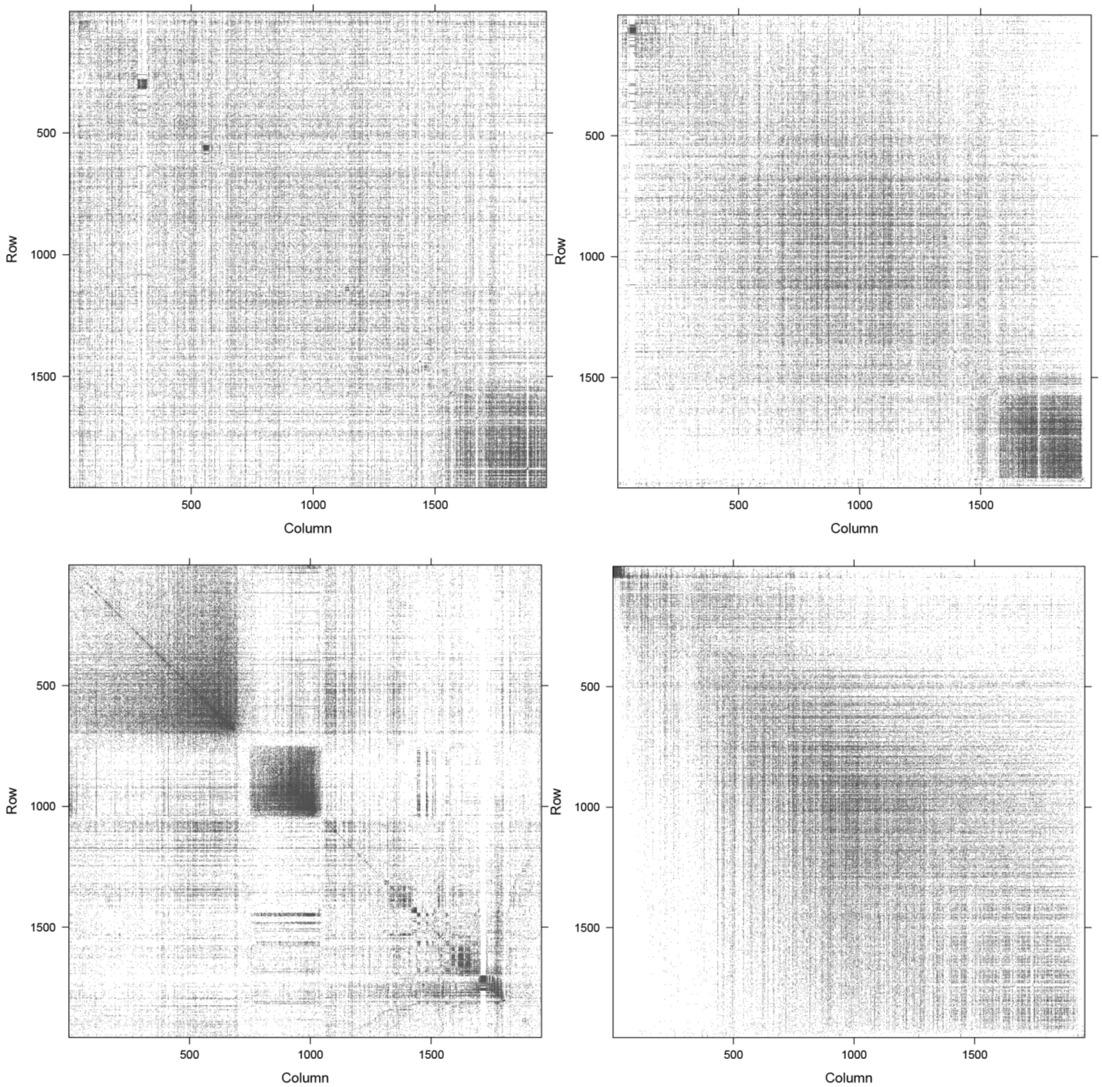


**Figure 1**: The adjacency matrix of the news-centric core, permuted by outgoing slant (top left) by incoming slant (top right) by the method of Clauset, Newman and Moore (2004) (bottom left) and by the values of the eigenvector corresponding to one of the smallest eigenvalues of the Laplacian matrix (bottom right).

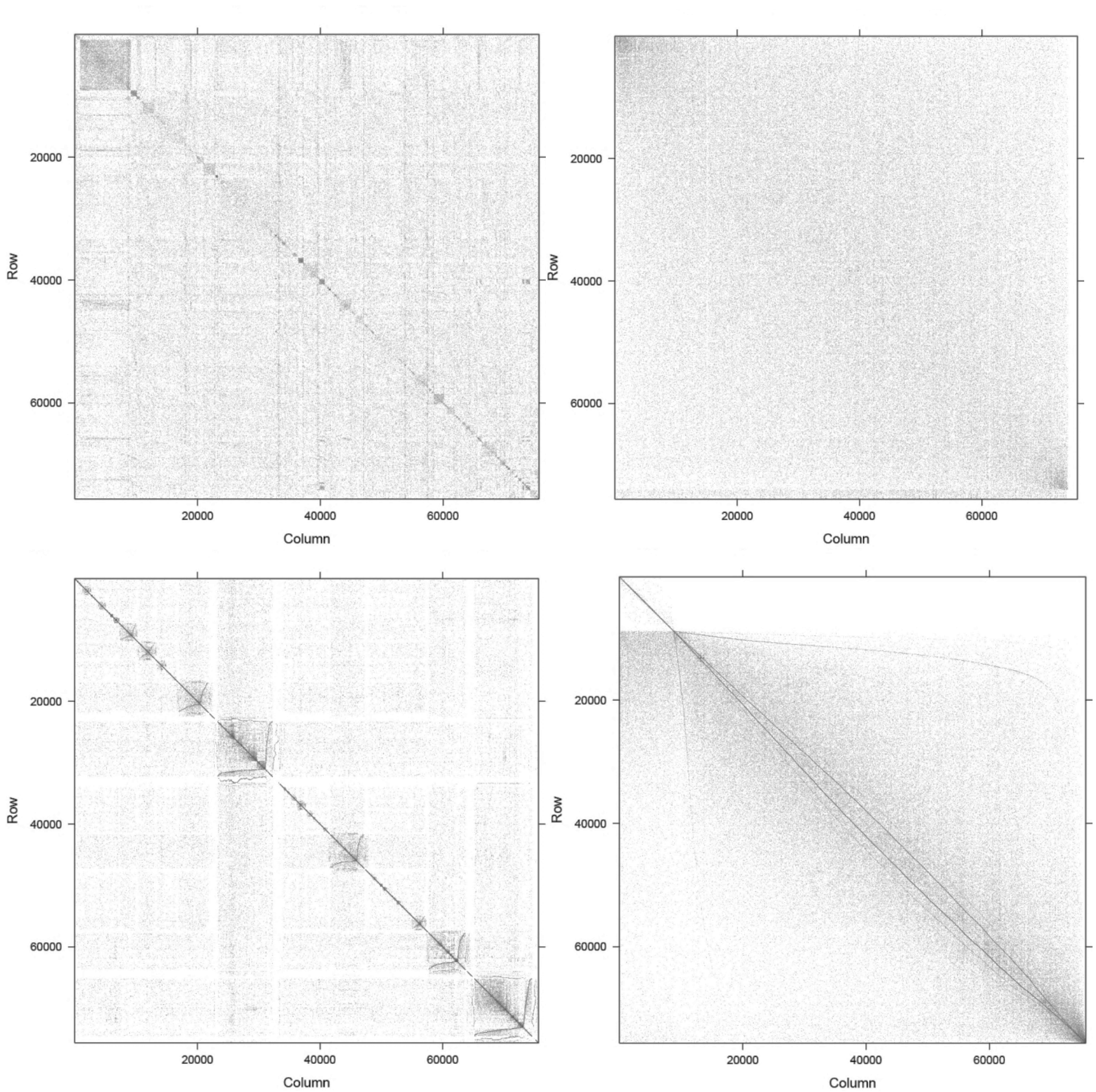


**Figure 2**: The adjacency matrix of the "moderate users" subgraph taken from the periphery, permuted by outgoing slant (top left) by incoming slant (top right) by the method of Clauset, Newman and Moore (2004) (bottom left) and by the values of the eigenvector corresponding to one of the smallest eigenvalues of the Laplacian matrix (bottom right).

## 1.5 Discussion

### *1.5.1 Summary of empirical findings*

Overall, our results are only partially consistent with theories of echo chambers, polarization and mainstreaming. Although small echo chambers may exist, we do not see clear evidence for them in the aggregate. We do find evidence of homophily (outgoing slant is correlated with incoming slant), but also an average tendency to moderation and many points of contact among different points on the political spectrum (see slant-permuted matrices in Figures 1 and 2). We do see a polarized and active core in which network structure closely corresponds to political slant, but we also see a much larger (albeit much less active) generally moderating majority for which network structure is more weakly related to slant.

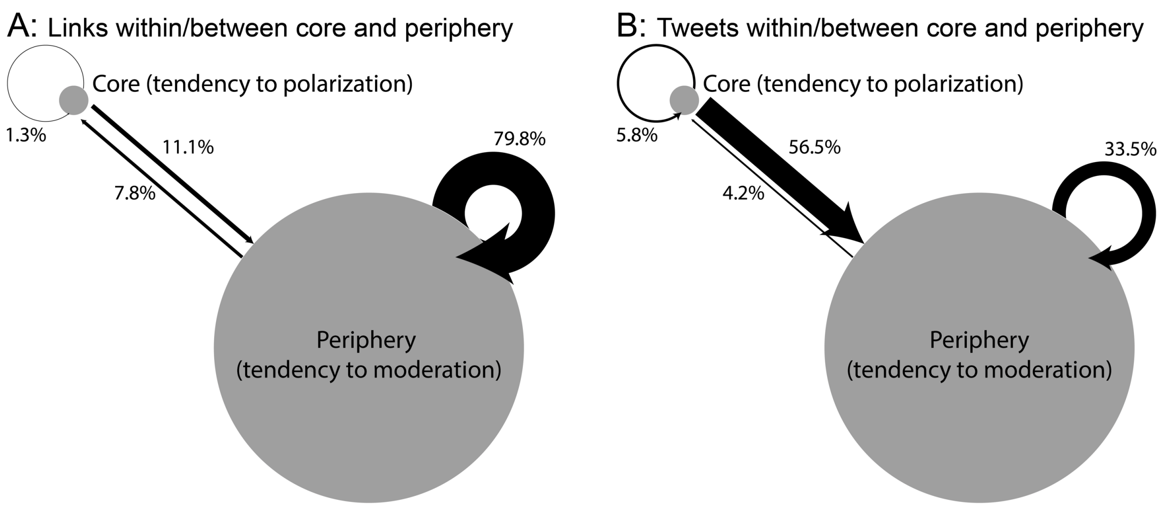


**Figure 3.**: Summary diagram of connectivity patterns, distinguishing core from periphery. Overall, there is a tendency to centrism, but a majority of tweets received originate in the network core, which has a tendency to polarization. Grey circles represent accounts in the network core and periphery. Circle size is proportionate to number of accounts. Arrows indicate percentage of total connectivity within and between core and periphery. **A:** arrow size is proportional to total number of follower-followee relationships in the full data set and labeled with a percentage (e.g. 79.8% of all links are within the periphery). **B:** arrow size is proportional to (an upper bound on) the number of tweets received in the full data set, calculated as number of tweets sent multiplied

by the number of followers those tweets were sent to (e.g. only 33.5% of all tweets received were both sent and received by accounts in the periphery).

A diagrammatic summary of the overall communication structure is in Figure 3. The widespread concern over polarization may be due to the over-representation of tweets originating in the core, constituting a sort of network paradox (Feld, 1991). As for mainstreaming, we do not find an absolute, but rather a relative tendency to political centrism. We also note that accounts outside of the core are tweeting across the political spectrum, which undermines a literal theory of a spiral of silence (Noelle-Neumann, 1974).

### *1.5.2 Broader implications*

#### 1.5.2.1 What is read versus what is said

In cross section, we find that communication patterns look very different when one looks at what is read (incoming information) instead of what is said (outgoing information). Because incoming slant is more closely related to network structure than outgoing slant, in one limited sense we can conclude that what is read is the more meaningful measure. This may have substantial consequences for our understanding of influence in social networks, which typically only looks at expressed behavior (analogous to what is said in the context of this study). For example, in a network study of influence in the spread of a product, an individual's social media posts about that product could be interpreted as an expression of interest in the product or as the outward expression of desire to conform without any true interest in the product.

Additionally, we find that the relationship between what is read and what is said is strikingly different in the network core from outside of it: core accounts tend to position themselves in a more extreme position that what they are exposed to, while the typical account positions itself in a more moderate position. In the setting of influence in networks, it could well turn out that there is a similar regularity such that those within a core systematically express their preferences in an extreme manner, while those outside of the core systematically express their preferences in a moderate and dampened manner.

Because of this marked heterogeneity between core and periphery, it is necessary to study communicating systems as a whole as we seek to understand the technologically mediated crowd that is of increasing importance in our evolving economy and society.

5.2.2 The core versus the periphery in online communities: the "multiplex public"

Like other social networks, online communities have a core-periphery structure (Dahlander and Fredriksen, 2012; Collier and Kraut, 2012; Wasko, Teigland and Faraj, 2009) and are composed of individuals with shared goals and interests that communicate over the internet (Preece, 2000), in a self-organized manner consisting of voluntary participation and without formal organization (Dahlander and O'Mahoney, 2011). Our data could therefore be considered an online community of political discussion with liberal and conservative sub-communities, or alternatively, two overlapping communities in conflict with each other.

In general, prior research has treated membership in the core versus the periphery as essentially an issue of the *level of engagement* in the community.  Some attention has been paid to how individuals end up in the core (Collier and Kraut, 2012; Dahlander and

O'Mahoney, 2011; Johnson, Safadi and Faraj, 2015), and the sources of motivation for "heavy weight" participants in the core compared to "light weight" participants in the periphery of an online community (Haythornthwaite, 2009). Our results, however, reveal that those in the periphery are not only different from those in the core in terms of the amount of participation or reason for participation in the community, but indeed also in terms of the very nature of their information sharing behavior. Again, we find that on average, core members share links to more politically extreme news sources than the links they receive in their own timelines. Periphery members, on the other hand, are the opposite.

People tend to express themselves freely to the extent that the topic of conversation is consistent with their public or professional identity, and that their audience is homogenous (Marwick and boyd, 2010). For most people these conditions do not apply, since they use a personal (rather than professional or other narrowly constructed public identity) social media account to connect to multiple contexts and identities (Rainie and Wellman, 2012; Hampton, Lee and Her, 2011; Marwick and Boyd, 2010). In other words, most people cannot assume that their followers also follow each other, which accords with the fact that the periphery of a social network is not highly interconnected within itself by definition (Borgatti and Everett, 2000). For people who both have a clear public identity and surround themselves with others with shared interests and goals – in other words, for members of the core – Marwick and boyd's conditions for free expression are met. This free expression could then be amplified by social influence (Centola and Macy, 2007; Shore, Bernstein and Lazer, 2015) and made more extreme by group polarization processes (Sunstein, 2002).

If individuals in the core and the periphery have different characteristic behaviors and social environments, then lumping them together under the single term “community” is insufficient. Instead, a new term is needed to describe this social structure that is most pervasive in our data. We offer the term “multiplex public” to describe the social structure that such typical users of social networking services inhabit. “Multiplex” refers to the multiple network layers (a work network, a school network, a friend network and so on) that come together to form the overall follower-followee network, and “public” emphasizes the environment that is neither a single cohesive community nor a disconnected crowd, but in which individuals are still visible to sparsely-connected others.

We suggest that this multiplex public has received less attention in the past in part because it has not been an obvious source of peer production. Because of their economic consequence, online communities and crowds have been obvious and important to researchers in and around the disciplines of management. Now, as data science uses digital traces for all manner of social scientific and business intelligence purposes, we should also acknowledge the significance of this prominent social structure and identify they ways it diverges from cohesive groups and network cores in future research.

1.5.2.3 Research methods

Network research nearly always faces a boundary definition problem (Laumann, Marsden and Prensky, 1989): the researcher must define who is in and who is out of the research data. As a matter of convenience, this often means selecting nodes on the basis of their activity; in the case of political slant, prior work has sampled people to study on the basis of their obvious political partisanship (Adamic and Glance, 2005; Conover, et al, 2011; Bakshy, Messing and Adamic, 2015; Barbera et al., 2015). While all of these studies go to some lengths to account for their data collection strategy, at a certain level they cannot fully escape the fundamental limitations that come with sampling on the dependent variable. That partisans are polarized does not imply that social media users in general are polarized.

The implications for future research on social media are clear: the behavior of members of the core is not representative of people outside of the core. Networks constructed by choosing obviously relevant individuals (because they post a lot about the research topic, for example) are likely to consist only of the network core and leave out the more representative (in terms of ordinary users) periphery.

### *1.5.3 Limitations*

Although our data are broadly representative in terms of their inclusion of typical Twitter users, our coverage consists of a cross-section in a non-election year. This means that we cannot speak to issues of influence or other dynamic processes on or of networks – only the cross-sectional organization of Twitter. More importantly, however, is the fact that our data was collected from a relatively "typical" period of time: 2009 was not an election year and September 10th-23rd (the data collection window) did not contain any major news stories[8] that might spark an increase in partisan conflict. If the data were

[8]See http://www.infoplease.com/year/2009.html#us

collected at an atypically polarized time, we may have observed different results. Finally, Twitter was 3 years old when the data were collected, so while it no longer was only the home of early adopters, it had not yet gained the reach and user base that it has today. It is impossible to say for certain how this might affect results if this study could be repeated with current data.

A second set of limitations comes with our use of Gentzkow and Shapiro's slant scores. Although they cover over 95% of all direct news browsing and an even higher percentage of exposure to news on social media, we do not cover all sources of news. We cannot rule out the possibility that there are echo chambers built around the sharing of news from sites representing a tiny minority of news exposures, including those from hate sites. Indeed, if there were a total absence of such phenomena at the fringe, it would be surprising. However, this doesn't affect our results, which characterize the vast majority of news exposures on Twitter. Finally, we study sharing and receiving links to news sites, so our data do not cover other types of speech; it is possible, for example, that free text tweets follow different patterns than those we observe here in shares of news content.

### *1.5.4 Conclusion*

By using data representative of the whole population of Twitter users, we were able to reconcile apparently contradictory theories of diversity of information sharing on Twitter. The aggregate picture cannot be described as just a collection of echo chambers on the one hand, or a clear pattern of mainstreaming on the other. Rather, with elements of both tendencies, we instead see a whole system comprising a vast moderating majority

– a multiplex public – with a polarized two-part community at its core. Predicted behavior depends on which part of the system you are looking at, but on average, Twitter accounts post more centrist information than they receive in their own timelines, undercutting the prevailing narrative of the social media echo chamber. Instead, the widespread perception of such polarization may be the result of a network paradox, in which the behavior of nodes with a high degree is mistaken to be typical (Feld, 1991).

# CHAPTER 2: Forum size and content contribution: a MOOC field experiment[9]

## 2.1 Introduction

Digital communication technologies make it easy to bring together large groups of people with shared interests in online discussion communities or forums to share and learn from each other. A critical factor for the discussion communities' success is the active contributions of each individual because the provided knowledge is a key resource which attracts other users to the community (Butler 2001). However, a common challenge is that only a small minority of users actively contribute to discussion by posting content or asking or answering questions. Promoting greater engagement and active contribution is thus a key challenge for online sites that support the communal creation or transfer of knowledge. In this paper, we study the challenge of promoting a higher level of active contribution of posts per person in the context of a discussion forum on a Massive Open Online Course (MOOC). MOOCs can have over 100,000 students learning from a single professor posting digital content, and from each other on discussion forums (Breslow et al. 2013, Hood et al. 2015). However, MOOCs – like other online platforms – suffer from low level of engagement as only 3-5 percent of users interact in the forum (Breslow 2013, Rosé et al. 2014) and more than 90 percent of users drop out of course.

A substantial body of research on online engagement has studied individual motivations and antecedents of active contribution of content (Butler, et al., 2002; Wasko

---

[9] This chapter is a joint work with Jesse Shore

and Faraj, 2005; Kuk, 2006; Ma and Agarwal, 2007; Bateman, et al., 2011; Ren, et al., 2012), the influence of social networks and norms (Ransbotham, et al, 2012; Huang, et al., 2017; Burtch, et al., 2017), explicit calls to action (Zalmanson and Oestreicher-Singer, 2015), as well as the development of user engagement over time (Preece and Schneiderman, 2009; Dahlander and O'Mahony, 2011; Oestreicher-Singer and Zalmanson, 2013, Butler, et al., 2014; Kokkodis and Lappas, 2016). Here we consider another variable that we believe is relevant to the online communities: the number of people in discussion forums. Long literatures on digital collaboration and communication, group dynamics, and education (e.g. Chidambaram and Tung 2005; Latane et al. 1979; Mao et al., 2016; Kim, 2013) shows that the level of contribution per person is negatively related to the number of people interacting with each other: the more people, the less the contribution we should expect per person. In other words, it may be the very fact that discussion forums can attract large numbers of participants that depresses contribution on a per-person basis.

However, a key limitation of these studies on the effect of size (hereafter "size refers to the number of people in a single group, cohort or other set) on participation per person is that prior research overwhelmingly analyzes settings in which people work together to collectively solve a single problem or work on a shared project. Such "collaborative engagement" is the norm in formal organizations and is also common in digitally-enabled collaborative work settings such as Wikipedia editing or open source software programming. In collaborative engagement settings, group or community members exert individual effort, while output (and thus often incentives for completion or performance) is at the collective level.

Other digital platforms do not work by collaborative engagement and instead support a pattern of "individual engagement" in which both efforts and outcomes are at the individual level. In individual engagement settings, such MOOCs, as well as Q & A sites like Quora or Stack Overflow, people are motivated to pursue diverse individual outcomes (e.g. to receive information or gain status by providing it) and there is no hierarchical direction of collective output. Less is known about the effects of size on contribution in such individual engagement settings.

In this paper, we conduct a field experiment on an edX MOOC to study the effect of cohort size on the per-person level of contribution to discussion forums. We focus on testing a three-way treatment, randomizing 6000 pre-registered users into discussion forums containing 125, 500 or 2000 people. We also tested a 2-way treatment, in which users were either required or encouraged to participate in the forum for full course credit. Contrary to prior research on size, we find that the contribution per user *increases* with the discussion forum size. This increased contribution was primarily in the form of comments on existing posts: while the number of threads initiated per person was not significantly different between forums of different size, the number of comments per person on other user's threads was substantially and significantly higher in larger forums. We also found that much of the increase in participation in larger cohorts was in the form of greater participation at the highest percentiles of the distribution of posts per person: the greatest contributors contributed a larger share of all posts (including new threads and comments) in larger cohorts.

We attribute the difference between our results and those in prior literature to a difference in research setting, theorizing that the important difference is between collaborative engagement settings and individual engagement settings. In the former, larger cohorts of participants lead to less contribution per person, while in the latter, larger cohorts lead to greater contribution per person.

## 2.2 Theoretical background

The number interacting people in one online space is one of the key factors mediating individual engagement. Different literatures use different terms to refer to the number of interacting people. For example, organizational research refers to "*group size,"* educational research refers to "*class size,"* and information systems research may use different terms depending on the specific context, including "*community size*" or "*number of users*." In this paper, when reviewing existing literature, we mirror the vocabulary used by earlier authors, which is most often "group size" or "class size." However, there may be different social processes in effect in groups, classes or communities that are not present in our research setting. Therefore, to avoid connotations from these terms, we use the term *cohort size* in our research context to refer to the number of students in a single MOOC discussion forum.

### *2.2.1 Prior studies on group size*

Much prior work argues that greater size creates challenges for both online and in-person groups, and that the amount of interaction per person goes down as group size increases because there is more free-riding: obtaining value from a group without contributing to that group (Albanese and van Fleet 1985). Similar to free-riding, prior research has studied "social loafing" (Ingham, et al., 1974; Latane et al., 1979) and the "bystander effect" (Latane and Darley, 1968), showing that a person is less likely to take action or assume responsibility when there are others in a group or otherwise co-present, and that the probability of contribution of effort or helping others is inversely related to the number of people present (Barron and Yechiam, 2002; Bray, et al., 1978; Lowry et al., 2006). Thus, in smaller groups, individuals tend to contribute more time and energy to interact and share information with others because they feel responsible to the group.

It has long been known that, in traditional offline organizations, people in larger groups feel that they matter less, make less of a difference, and that others may not recognize their contributions (Gooding and Wagner, 1985; Kerr 1989; Kerr and Bruun, 1983). Large group size also leads to higher communication and coordination costs (Pendharkar and Rodger 2009) and it can be harder to tell how much and how each individual has contributed (Jones 1984; Kerr and Bruun 1981). Therefore, in larger groups, individual contribution is likely to be lower (Bales and Borgatta 1966, Diehl and Strobe 1987; Wheelan 2009).

Large group size is less of an obstacle to a high level of individual contribution online than it is offline, because digital collaborators do not suffer from the "production blocking" effect (Gallupe et al. 1992) in which group members must wait for each other to finish before initiating their own engagement; however, larger groups still have lower participation per person than smaller groups in digital collaboration. For example, empirical evidence on problem-solving within technology-supported groups shows that size is correlated with a decrease in participation per person, idea quantity per person, decision quality and group cohesiveness because of more free riding (Chidambaram & Tung 2005, Alnuaimi et al. 2010; Valacich et al. 1995, Yap and Bock, 2006).

Prior research has examined the antecedents of free riding in digital collaboration, with various proposed mechanisms. Increased group size makes interaction between group members more difficult and complex (Riopelle et al., 2003). Additionally, moral disengagement by individuals can reduce effort per person; this can take the form of feeling less responsible for ensuring a good outcome, increased ability to blame others for poor collective outcomes, and increased feelings of dehumanization that come from being among too many others (Alnuaimi et al., 2010). Members of larger groups also have more difficulty establishing relationships with others and can be overwhelmed by the high volume of communication, which can lead to reduced contribution or attrition from online groups (Jones, et al., 2004; Wang et al., 2013).

There do exist some studies that show a positive effect of size on collective output (e.g. number of posts on a forum), but only for *total* output, and not contribution per person (Koh et al. 2007; Carillo and Okoli 2011). Indeed, Mao and collaborators (2016) show experimentally that even when the total group output goes up with size, the level of contribution per person still goes down, even in complex tasks requiring division of labor. While larger brainstorming groups generate more ideas than smaller ones (e.g., Valacich et al. 1992, Fellers 1989), the contribution per person is either not evaluated in these studies or displayed no significant differences (Chidambaram and Tung 2005, Valacich et al. 1992, Gallupe et al. 1992).

An important outlier in the literature on size is reported in a paper by Zhang and Zhu (2011). Zhang and Zhu document the effect of the Chinese government unexpectedly blocking Chinese-language Wikipedia from access by users and editors within mainland China. Immediately after the block, editors who were not blocked reduced their level of contribution to the site, especially those who interacted heavily with editors who disappeared suddenly due to the block. As Zhang and Zhu write, this reduction in effort by individuals who lost their collaborators is due to "social effects" and not to a pure effect of size.

Results in the context of online education have been equivocal. We are not aware of any research results on the effect of size in MOOCs per se, but several studies present evidence from semester-long online university courses. One field experiment shows that students in smaller online discussion forums read more posts and interact more with other students (Kim 2013). This study employed real-time class discussions, however, raising the possibility that production blocking or other coordination losses may have been responsible for the connection between size and depressed performance. A retrospective study on observational data showed the opposite result: students in larger classes contributed more to class discussions (Qiu et al. 2012), but size and student contribution are endogenously correlated, and it is unclear whether there is a true effect of size.

### *2.2.2 "Collaborative engagement" type forum and "individual engagement" type forum*

Existing literature consistently shows that group, class or community size is negatively related to user engagement per person. However, this literature overwhelmingly studies “collaborative engagement” contexts in which people participate collectively to create one shared group-level outcome (e.g. content, solution, etc.). In such collaboration engagement settings, the logic of social loafing is a natural consequence of a mismatch between costs borne by the individual (each person must contribute effort individually) and benefits accruing to the group or collective (outcomes and thus incentives for success are attributed to the group). Thus, there is a natural tendency for individuals to free-ride off of the efforts and engagement of others. For example, in Mao and collaborators’ (2016) experiment, paid workers from Amazon Mechanical Turk collaborated to complete a collective task: creating a single disaster map by aggregating their individual contributions.

Collaborative engagement is not the only type of online platform, however (Budhathoki and Haythornthwaite 2013). Many individuals are engaged voluntarily in diverse online discussion forums such as Quora or StackOverFlow, and they are motivated to participate by individual-level motivations such as learning from others, fun, gaining social status, or promoting a sense of community (Brabham, 2010; Wasko and Faraj, 2005). In such individual engagement settings, there is no hierarchical direction of collective output, and people participate or not according to their own individual motivations.

In individual engagement settings, each individual’s comments become available as a resource for future users; thus, larger communities have more resources and are more valuable and attractive to new users (Butler 2001). For example, individuals will have a greater chance of finding interesting information or the answer to a question in a bigger discussion forum where many participants have already engaged. By this logic, the larger the community, the more new users will be attracted to the resources provided by that community. In other words, online communities display “network effects” or “network externalities.” The theory of network effects says that the value of an interaction technology or platform increases with the number of users on that platform (Katz and Shapiro 1994), especially users one might want to interact with (Lin & Lu, 2011) and this concept is often applied in research to explain behavioral intentions toward and engagement in interaction-based platform such as online forum or social network services (*e.g.* Kang and Namkung, 2016). Individual engagement settings often display network effects, where people are not locked into the participation as a group member, individuals are more likely to join a larger community because they find higher value in it.

What is not clear from the “resource availability” or “network effects” perspectives is whether an increased tendency to *join* larger communities translates into an increased tendency to *participate* in them. It seems reasonable that a more attractive community might motivate more engagement per person, but it could also be that the presence of more resources (usually in the form of archived existing discussions) means that there is less need to interact: if the answer to my question is already archived, I can just read the existing answer without actively asking again.

We are left with a puzzle. Size depresses active contribution in collaborative engagement settings. Size promotes joining a community or platform in individual engagement settings. However, we do not know the effect of size on active contribution in individual engagement settings.

### *2.2.3 Hypotheses*

What relationship between size and contribution should we expect in a MOOC discussion forum? Overwhelmingly, prior research points to a negative relationship between size and engagement, but we have argued that this prior research has studied collaborative engagement settings. In contrast, MOOC discussion forums are individual engagement settings in which users are motivated by their individual goals, rather than collaborating to create a single collective output or outcome. For example, their goals and motivations include learning, achievement, differentiating themselves in the job market, interacting with other users and so on (Breslow, et al., 2013).

In MOOC discussion forums, the cohort size provides a measure of resource availability to engage individuals (Butler 2001) because each individual's posts and potential posts are available to the rest of the forum participants. For example, users of the forum provide feedback on each other's posts. The more other users there are, the more feedback a given user is likely to get on his or her posts. Similarly, the more existing posts there are, the more likely a user will be to find something that catches their interest. These available resources could increase the probability of user engagement and may, in turn, imply a virtuous cycle, bringing more unengaged users into the discussion. Therefore, our main hypothesis is the following.

*H1: Users in larger cohorts contribute more per person to MOOC discussion forums.*

Because of the strong positive links between engagement and student success (Kuh 2009), we also examine grades as an outcome variable. If forum users get more answers to their questions or find more informative comments from peers in larger discussion forums, they may be more likely to understand the course materials (Marzano et al., 2001) and thus perform better on quizzes or exams. Therefore, our second hypothesis is,

*H2: Users in larger cohorts achieve higher grades.*

Finally, we consider course completion. Not only is the course completion rate one of the traditional outcomes in educational research, but it is also one of the major concerns of current MOOCs; more generally, user attrition is a common issue for platform managers. Greater engagement has been repeatedly been found to be correlated with greater course completion in prior research (Reich, 2015). It is possible that if larger cohorts increase engagement, this effect could translate into greater course completion. In larger cohorts, users may get greater benefit from interacting with others than they do in smaller cohorts because there is more information generated by peers. When they get greater benefit from the course, they are more likely to return to the course for more. Therefore, our last hypothesis is,

*H3: Users in larger cohorts complete more of the course.*

## 2.3 Research Methodology

We conducted a field experiment in Boston University's 'Sabermetrics 101'[10]

[10] "Sabermetrics" denotes the statistical analysis of baseball player and team records; the word derives from the acronym SABR, for the Society for American Baseball Research. According to the course website, the Sabermetrics 101 course primarily covers the "basics of data science and how it applies to the study of baseball" and the "fundamentals of the R and SQL programming languages."

course offered on edX, which began on July 7th, 2015 and ran for 10 weeks. The edX platform was launched through a partnership between Harvard University and the Massachusetts Institute of Technology in May 2012 and has become one of the largest MOOC providers, along with Coursera and Udacity. Participating institutions have expanded to include 48 core "charter universities" (including Boston University, which provided access to the course we studied), along with many additional partner schools and organizations. It hosts online university-level courses to internet users all around the world and provides an online discussion forum for user interaction.

### *2.3.1 Dependent Variables*

We measure forum contributions several different ways. We consider all posts made by a user (*# posts*), as well as distinguishing between posts that are new threads (*# threads*) on the forum from comments made on an existing thread (*# comments*). We test the effects of the treatment on the number of posts for each variable, as well as on a dichotomous indicator of whether the variable is greater than zero (*any posts*, *any threads,* and *any comments*).

We measure *grade* as an average grade of all assignments, quizzes and the final exam, including any 0s for any that were not completed. We also use whether the user passed the course by earning 60% of available points as a binary dependent variable (*passed course*). Course completion is highly correlated with grade but distinct from it, and we measure it with the number of assignments – lecture questions, quizzes and final exam – that had a non-zero score (*# assignments*) as well as a binary version, indicating whether a user completed at least one assignment (*any assignments*).

Additionally, although most students are auditors in the sense that they take the course for free and do not receive any official credit, edX provided the option to work toward an edX *ID Verified Certificate* by paying a $10 fee and achieving a passing grade. For the present course, a user could opt-in to the ID-verified track up through August 11th, 2015, which was approximately halfway through the course. Whether or not the student passed and thus actually received the certificate was determined at the end of the course. We treat the decision to pay the fee (*paid verification*) as a binary dependent variable that indicates a high intention to complete the course.

### *2.3.2 Experiment Procedure*

Our experimental design focuses on manipulating the number of students interacting with each other in the discussion forum portion of the edX platform. In a typical edX course, all students use the same discussion forum to communicate with each other; in our experiment, we replaced the single catch-all forum with multiple forums serving cohorts of different numbers of students. Each student in our study population was assigned uniformly at random to one and only one cohort, and student forum posts were only visible to other members of their own cohort. Students were aware that other cohorts existed, but they could not see or participate in discussions in those cohorts.

Randomization into separate cohorts was performed one week before the course's official start date, just prior to the time that the course website was made accessible to registered students. At this point in time, there were just over 6000 pre-registered students, and we randomized 6000 of these pre-registered students into experimental treatments. Of the 6000 randomized individuals, 4104 logged in to the course website and were exposed to the experimental treatment. Exposure to the treatment was independent of treatment assignment per Fischer's exact test, and thus we consider our sample size to be 4104 in statistical analyses.

We picked our cohort sizes assuming a high rate of initial attrition. Prior research on MOOCs had shown that 60-70% of initially enrolled students do not end up participating in the course (Clow, 2013). We therefore created three conditions for cohort size of 125 people per cohort (treatment "S"), 500 people per cohort (treatment "M"), and 2000 people per cohort (treatment "L"), with the expectation that the effective sizes would be on the order of 40, 160, and 640 people per cohort after initial attrition. Because the overall number of people randomized to each treatment group is the same, there are different numbers of cohorts within each treatment (see Table 11).

Within the smaller two treatments, we also ran a two-way treatment in which we informed half of the cohorts that forum participation was required to get full course credit. In traditional learning environments, instructors often grade students on their participation the discussion to motivate students' engagement and we adapted this to our study. We implemented the participation requirement by notifying students by email and by posting the requirement on the course webpages (treatment "R"). Students in the non-treated groups received messages that participation in the forums was encouraged (treatment "E").

Table 11: Field experiment design

| | Treatment S<br>125-person cohorts | Treatment M<br>500-person cohorts | Treatment L<br>2000-person cohort |
|---|---|---|---|
| Treatment E<br>(posting encouraged) | SE<br>8 cohorts | ME<br>2 cohorts | LE<br>1 cohort |
| Treatment R<br>(posting required) | SR<br>8 cohorts | MR<br>2 cohorts | (empty) |

Other than the forums, all cohorts had access to the same course materials. Throughout the course, participants received prompts reminding them of their discussion cohort size and participation requirement condition via email and message on the courseware pages. Many students continued to register for the course after the beginning of the course, but they were not included in the experiment and instead assigned to a "default" cohort, which had a varying number of enrolled students. Finally, students remained in a single cohort for the entire duration of the MOOC.

### *2.3.3 Independent variables*

Users provide self-reported data to edX at the point of sign-up for the platform, which for our experimental subjects was strictly prior to the course beginning. We use self-reported gender, self-reported level of education, self-reported age, number of weeks since signing up for edX as additional control variables.

Additionally, we use a free text self-reported "goals" for the course as a source for additional controls as follows. Users that included any of the words "career", "work" or "job" were coded with a 1 on the variable *goals: work*. Users that included the word "baseball" were coded with a 1 on the variable *goals: baseball*. Users that included variants of the word "sabermetrics" were coded with a 1 on the variable *goals: sabermetrics*. Users that included the phrase "data science", variants of the word "statistics," variants of the word "analytics" or "SQL" or "R" were coded with a 1 on the variable *goals: statistics*. Additionally, users that left that field blank were coded with a 1 on the variable *no goals given*.

## 2.4 Results

### *2.4.1 Hypothesis 1 – Main hypothesis*

#### 2.4.1.1 Model choice

The variables *# posts*, *# threads* and *# comments* show evidence of over-dispersion: the variances of these forum engagement variables are substantially larger than their means. When the data are over-dispersed, the standard errors in Poisson models are biased downward and, therefore, negative binomial models are preferred (Cameron and Trivedi 2013). Another characteristic of our data is that very few students in our sample posted questions and comments. Of the total sample of 4104 students, only 1063 posted; most participants have a count of zero for the forum engagement. To counter the effects of excessive zeros in our model we used zero inflated negative binomial (ZINB) regression. ZINB models account for overdispersion and model the presence of excessive zeroes. In addition to being preferred for our count data on theoretical grounds, the ZINB model outperforms other models in terms of overall model Akaike Information Criterion

(AIC), as well as Vuong's test. For models of the dichotomous variables *any posts*, *any threads,* and *any comments*, we used logistic regression.

We tested models with interaction effects between size and the participation requirement, as well as between size and control variables. Interaction terms were insignificant in analysis of variance tests, and models that contained them had worse AIC scores. We therefore omitted interaction terms from the models reported here.

Since students are nested within cohorts, we also fit random effects versions of models. Surprisingly, we found no within-cohort covariance in the dependent variable after conditioning on the covariates. The more complex random effects specifications were thus discarded in favor of the simpler models reported below.

2.4.1.2 Level of contribution per person

Overall, the evidence supports rejecting the null for Hypothesis 1 and concluding that users in larger cohorts contribute more per person to MOOC discussion forums. We find that cohort size has a positive and statistically significant association with total number of post per person in the discussion forum (Table 12: column "all posts"). Specifically, the expected number of posts per person in large cohorts is 1.51 times that of small cohorts (effect magnitudes are obtained by exponentiating the estimated coefficients). There was no statistically significant difference between medium-sized cohorts and small-sized cohorts, but the estimated coefficient for medium size was positive and between the coefficients for small size and large size, consistent with Hypothesis 1.

We also examined the cohort size's effects on different types of posts: posts that initiate a new thread (table 12: "# *threads*") and posts that comment on existing threads (table 12: "# *comments*"). Results show that larger cohorts stimulate more contribution per person especially in the form of comments on existing threads rather than posts of new threads. Specifically, the expected number of comments per person in the large cohort was 1.77 times that of small cohorts.

Table 12: Effect of treatments on discussion forum contributions

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | # posts | # threads | # comments | any posts | any threads | any comments |
| | *ZINB* | *ZINB* | *ZINB* | *logit* | *logit* | *logit* |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| size = L | $0.410^{**}$ | 0.180 | $0.573^{***}$ | 0.127 | $0.264^{*}$ | 0.121 |
| size = M | 0.125 | 0.069 | 0.185 | 0.127 | 0.164 | 0.052 |
| Participation req. | $1.475^{***}$ | $1.596^{***}$ | $1.359^{***}$ | $0.699^{***}$ | $0.970^{***}$ | $0.678^{***}$ |
| Weeks on edX | $-0.009^{***}$ | $-0.009^{***}$ | $-0.010^{***}$ | $-0.009^{***}$ | $-0.007^{***}$ | $-0.009^{***}$ |
| Age < 20 | 0.410 | 0.574 | 0.141 | −0.091 | 0.030 | −0.005 |
| 30 <= Age < 40 | $0.388^{**}$ | $0.341^{*}$ | $0.409^{**}$ | 0.197 | 0.197 | 0.124 |
| 40 <= Age < 50 | $0.648^{***}$ | $0.729^{***}$ | $0.601^{***}$ | $0.362^{**}$ | $0.387^{**}$ | $0.310^{*}$ |
| 50 <= Age < 60 | $1.440^{***}$ | $1.530^{***}$ | $1.329^{***}$ | $0.663^{***}$ | $0.904^{***}$ | $0.509^{**}$ |
| 60 <= Age < 70 | $1.377^{***}$ | $1.506^{***}$ | $1.234^{***}$ | $0.954^{***}$ | $1.153^{***}$ | $0.801^{***}$ |
| Age >= 70 | 0.746 | 0.639 | 0.775 | 0.459 | 0.645 | 0.188 |
| no Age given | 0.379 | 0.441 | 0.351 | 0.217 | 0.292 | 0.037 |
| no goals given | $-0.305^{**}$ | $-0.297^{*}$ | $-0.291^{*}$ | $-0.290^{**}$ | $-0.292^{**}$ | $-0.287^{**}$ |
| goals: work | 0.327 | 0.474 | 0.162 | 0.140 | $0.478^{*}$ | −0.005 |
| goals: baseball | $0.468^{*}$ | $0.533^{*}$ | 0.412 | 0.295 | $0.351^{*}$ | 0.263 |
| goals: statistics | −0.113 | −0.079 | −0.096 | 0.225 | 0.214 | 0.214 |
| goals: sabermetrics | 0.185 | 0.026 | 0.293 | $0.420^{**}$ | 0.125 | $0.504^{**}$ |
| edu: elementary | −0.473 | 0.282 | −1.970 | 0.876 | 1.397 | 0.014 |
| edu: jr. high | 0.271 | 0.100 | 0.517 | 0.701 | $1.075^{**}$ | 0.381 |
| edu: high school | −0.028 | 0.171 | −0.208 | 0.106 | 0.120 | 0.007 |
| edu: associate | 0.116 | 0.199 | 0.015 | 0.255 | 0.192 | 0.078 |
| edu: masters | $-0.236^{*}$ | −0.246 | −0.239 | −0.060 | −0.160 | −0.059 |
| edu: doctorate | −0.397 | −0.288 | $-0.510^{*}$ | $-0.446^{*}$ | −0.275 | −0.359 |
| edu: other | −0.095 | 0.103 | −0.276 | −0.309 | −0.104 | −0.334 |
| edu: none | $-1.656^{**}$ | −1.370 | $-1.951^{**}$ | −1.056 | −1.386 | −1.068 |
| no edu. given | 0.073 | 0.396 | −0.168 | −0.027 | −0.038 | 0.038 |
| female gender | −0.072 | −0.089 | −0.026 | 0.109 | −0.043 | 0.229 |
| no gender given | −0.133 | −0.431 | 0.060 | −0.203 | −0.247 | −0.048 |
| Constant | $-0.384^{*}$ | $-1.315^{***}$ | $-0.893^{***}$ | $-1.136^{***}$ | $-2.037^{***}$ | $-1.303^{***}$ |
| Observations | 4,104 | 4,104 | 4,104 | 4,104 | 4,104 | 4,104 |
| Log Likelihood | -4,831.702 | -3,153.763 | -3,790.191 | -2,162.586 | -1,679.743 | -1,963.243 |
| AIC | 9723.404 | 6367.527 | 7640.382 | 4,381.172 | 3,415.485 | 3,982.486 |

*Note:* $^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001
Reference categories are size = S, 20<=age<30, bachelor's degree and male
ZINB = zero-inflated negative binomial regression

2.4.1.3 Lurkers v. contributors

The logistic regressions on binary variables model not the average number of posts, but whether a user posted at all. For these models, the effect sizes are smaller than for models of the number of posts, although the directions and magnitudes are consistent with Hypothesis 1. One significant finding was that the odds of a user in the large forum posting at least one new thread was 1.3 times the odds of a user posting at least one thread in a small forum. Converting from odds to probabilities, users in small cohorts had a conditional probability of 0.115 of posting at least one new thread, while users in the large cohort had a conditional probability of 0.145 of posting at least one new thread.

2.4.1.4 Required participation treatment

Unsurprisingly, requiring participation had a strong positive effect on both the expected number of posts per person and the probability that an individual user posted at all. Users in the participation required treatment made 4.37 times as many posts as those in the participation encouraged treatment. This does not mean that the majority of users actually posted, however. Overall, the conditional probability of posting at least once in the participation required treatment was 0.392, compared to 0.243 in the participation encouraged treatment. For both the number of posts and whether a user posted at all, effect sizes were larger for initiating new threads than for commenting on existing threads.

2.4.1.4 Distribution of posts per person

For further context, we visualized the empirical cumulative distribution functions of # posts, # threads and # comments per person (Figure 4). Additionally, we used two standard measures of concentration – the Herfindahl-Hirschman Index of concentration (HHI) and the Gini coefficient – for each treatment condition (Table 13). Higher values of both HHI and Gini coefficient indicate more concentrated distributions of posts per person.

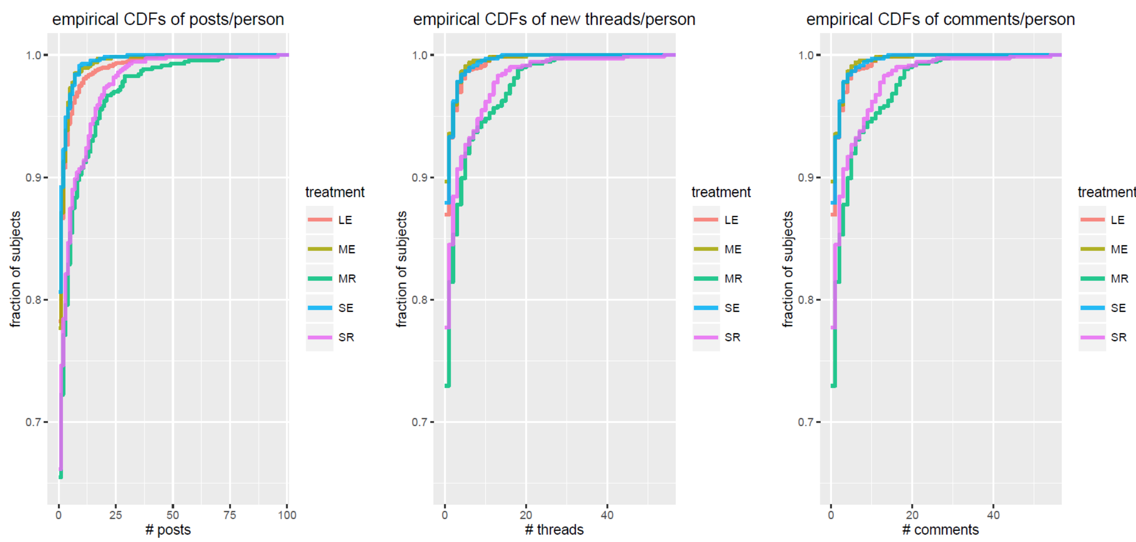


**Figure 4:** cumulative distribution functions of posts per person by treatment.

Table 13: Concentration of posts per person

| Treatment | HHI | Gini |
|---|---|---|
| Large, partic. not req. | 0.015 | 0.922 |
| Medium, partic. not req. | 0.020 | 0.896 |
| Small, partic. not req. | 0.020 | 0.911 |
| Medium, partic. required | 0.016 | 0.889 |
| Small, partic. required | 0.010 | 0.851 |

The greatest apparent differences in the distributions of posts per person are between the participation required treatment and the participation encouraged treatment.

Both participation required treatment groups appear less concentrated in the visualization and have the lowest values for both HHI and Gini coefficient in Table 13. However, beyond that observation, differences in concentration between our treatment groups are not clear from inspection of the raw data. Across all three panels of Figure 4, it would appear visually that the top contributors in the MR (medium, participation required) treatment post more than the top contributors in the SR (small, participation required) treatment. Additionally, top posters in the large cohort appear to comment more than individuals in the small and medium treatments with participation not required.

Using the method of Machado and Santos Silva (2005; Geraci, 2016) for quantile regression on count data, we estimated conditional quantiles of the distribution of # *posts*. As expected, the 95$^{th}$ quantile of the participation required treatment was significantly larger than that of the participation encouraged treatment (coefficient = 1.574, $p<0.0001$; coefficients are on the log scale). We also found that the conditional 95$^{th}$ quantile of # *comments* was higher in the large cohort than the small cohorts (coefficient = 0.573, $p=0.041$). Overall, the most marked increases in posting activity due to size and requiring participation appear at the top of the distribution of posts per person.

### *2.4.2 Hypothesis 2 and Hypothesis 3 – performance and retention*

Similar to the models in Table 12, we fit a zero inflated negative binomial regression when modeling # *assignments.* For the binary variables *paid verification*, *any assignments*, and *passed course*, we fit logistic regressions. For *grade*, we fit an ordinary least squares (OLS) model. As above, we rejected random effects models in favor of the simpler models presented in Table 14 due to a lack of intra-class correlation.

2.4.2.1 Cohort size effect on performance and retention.

Unlike the results for forum contribution, above, our results for the effect of size on performance and retention are mostly statistically indistinguishable from zero; we thus cannot reject the null for Hypotheses 2 and 3. The one positive result we found was for the *paid verification* variable: by half-way through the course, users in the large forums opted in to paying the fee to obtain an ID-verified certificate of completion, indicating that more of them had a greater intention to complete the course. Converting coefficients to conditional probabilities, the probability of users in the small cohorts paying the fee was 0.105, while the probability of users in the large cohort paying the fee was 0.143, keeping other covariates at baseline values.

Table 14: Effect of treatments on performance and retention

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | paid verification | # assignments | any assignments | grade | passed course |
| | *logit* | *ZINB* | *logit* | *OLS* | *logit* |
| | (1) | (2) | (3) | (4) | (5) |
| size = L | 0.351* | 0.044 | 0.056 | 0.007 | 0.022 |
| size = M | 0.176 | 0.014 | 0.068 | 0.007 | 0.065 |
| Participation req. | 0.272 | 0.156* | 0.110 | 0.019* | 0.255 |
| Weeks on edX | −0.010*** | −0.003*** | −0.009*** | −0.001*** | −0.007*** |
| Age < 20 | −0.168 | 0.358 | −0.133 | 0.016 | 0.336 |
| 30 <= Age < 40 | −0.148 | 0.120 | −0.090 | 0.011 | 0.254 |
| 40 <= Age < 50 | 0.193 | 0.272* | −0.150 | 0.025 | 0.502** |
| 50 <= Age < 60 | −0.055 | 0.366** | 0.208 | 0.048** | 0.555* |
| 60 <= Age < 70 | −0.327 | 0.513*** | 0.560** | 0.101*** | 0.967*** |
| Age >= 70 | −13.767 | 0.200 | 0.319 | 0.004 | −0.147 |
| no Age given | −0.221 | 0.016 | 0.094 | 0.012 | 0.051 |
| no goals given | −0.323* | 0.027 | 0.056 | 0.006 | 0.113 |
| goals: work | 0.130 | 0.223 | 0.208 | 0.037 | 0.484* |
| goals: baseball | 0.333 | −0.191 | 0.277 | −0.005 | −0.234 |
| goals: statistics | 0.301 | 0.233 | 0.248 | 0.049** | 0.497* |
| goals: sabermetrics | 0.619** | 0.034 | 0.314* | 0.028 | 0.193 |
| edu: elementary | −13.840 | −2.049* | 0.247 | −0.105 | −11.303 |
| edu: jr. high | 0.371 | −0.483 | 0.448 | −0.024 | −0.870 |
| edu: high school | 0.374* | −0.161 | 0.025 | −0.020 | −0.190 |
| edu: associate | −0.283 | −0.334 | −0.159 | −0.047* | −0.818* |
| edu: masters | −0.248 | −0.097 | −0.137 | −0.018 | −0.151 |
| edu: doctorate | −0.603 | −0.229 | −0.146 | −0.033 | −0.403 |
| edu: other | −0.593 | 0.238 | −0.731 | −0.022 | 0.022 |
| edu: none | −0.475 | 0.222 | −0.071 | 0.024 | 0.574 |
| no edu. given | 0.058 | −0.105 | −0.084 | −0.012 | −0.089 |
| female gender | −0.286 | −0.098 | −0.231* | −0.021 | −0.323 |
| no gender given | −0.744 | 0.027 | −0.279 | −0.013 | −0.186 |
| Constant | −2.140*** | 2.253*** | −0.209 | 0.112*** | −2.386*** |
| Observations | 4,103 | 4,104 | 4,104 | 4,104 | 4,104 |
| Log Likelihood | -1,024.560 | -7,911.161 | -2,575.279 | 1.660 | -1,170.282 |
| Akaike Inf. Crit. | 2,105.119 | 15882.32 | 5,206.558 | 52.681 | 2,396.564 |
| *Note:* | *p<0.05; **p<0.01; ***p<0.001 | | | | |
| | Reference categories are size = S, 20<=age<30, bachelor's degree and male | | | | |
| | ZINB = zero-inflated negative binomial regression | | | | |

2.4.2.2 Effect of requiring participation on performance and retention.

Interestingly, requiring participation did have a small but significant effect on both the number of assignments completed and grade. Students in the required participation treatment completed 16.8% more assignments and received approximately 17% higher grades than those outside that treatment group. It should be noted that although these are large relative improvements in performance, they are based on small absolute gains over low baselines: requiring participation caused students to complete 11.12 assignments instead of the baseline 9.52 (out of a total of 34), and an expected grade of 13.1 instead of the baseline of 11.2 (out of 100).

Still, the positive effect of this treatment suggests that greater engagement and contribution to discussion does spill over to an increase in performance and retention. It is possible that the increased engagement due to cohort size was not large enough to be detectable in our data. With even larger cohorts, it is possible that an effect may detectable in future research, though that remains an open question.

2.4.2.3 Survival analysis of login data.

In addition to the data on course completion modeled above, we also have the date of each learners' last login to the edX platform, which is a weaker and noisier signal of engagement than the completion of course assignments and quizzes. It is weaker, because a student can log in without actually going further and engaging with the course content, and it is noisier, because students can stay logged in for extended periods of time, engaging with the course, but without triggering a login event. Analyses of these data were less clear, but broadly consistent with our analysis of assignment completion data.

We fit Cox proportional hazard models to conduct a Survival analysis on users' last recorded logins ("login survival" for short). The course's end date is not a good end date for the survival analysis. By the time of the course's end, all users stop logging in by design, so we must choose points in time prior to the course's completion to assess differences in login survival. Additionally, end dates for the analysis that are too close to the course end are likely to provide especially noisy data, because students may have stayed logged-in to complete final assignments and exams. Moreover, especially diligent students might complete their final work as rapidly as possible and stop logging in before some that have procrastinated. Therefore, for the purposes of avoiding these problems, we choose to right-censor the data by stipulating artificial end dates, such that users who log in after the end dates are treated as having "survived" (stayed engaged) for the whole course. For these end dates, we use the release of the 5th and 6th (final) course modules on 8/20/15 and 8/27/2015, respectively, which are approximately 70% and 80% through the course duration.

As for the results, learners in the medium-sized cohorts – especially those in the MR cohorts – left the course at a lower rate than those in other cohorts. Surprisingly, given the strength of the intervention on other measures, the "participation required" treatment by itself had no effect on login survival. Specifically, learners in medium-sized cohorts left the course at a lower rate than learners in small cohorts before the release of the 5th module. When analyzed by cohort type, we found that learners in the MR cohorts left the course at a lower rate than learners in the LE and SR cohorts (marginally significant) before the release of the final module, and they left the course at a lower rate than learners in SE (marginally significant) and SR ($p$=0.015) before the release of the 5th module. These results are broadly consistent with our other findings, in that the MR cohort had the most posts per person and the greatest login survival. Nevertheless, due to the limitations of login data noted above, we put more weight on the results of analyzing the quiz data than the survival analysis of the login data.

## 2.5 Discussion and Conclusion

This study provides understanding of how people are engaged in online discussion forum especially in the learning context. Contrary to prior research, we find that students in larger cohorts contributed more per person to MOOC discussion forums. Our strongest specific result was that larger cohort size results more comments made per person on existing posts. We also found an increase in the probability than any given individual would post at least one new thread in the large cohort, relative to the small cohorts. The bulk of the increase in participation in larger cohorts, however, occurred at the top of the engagement distribution.

We did not find clear evidence that size had an effect on grade or retention, but we did see a significantly larger number of users pay for the ID verified enrollment track, indicating a greater investment in the course among larger cohorts. Additionally, given that all estimated coefficients were positive, it is possible that with a larger data set or larger differences in cohort size, future research may detect a statistically significant result. Echoing prior research (e.g. Zalmanson and Oestreicher-Singer, 2015), simply stating that participation was required led to large and statistically significant increases in forum contribution across all variables. It also caused a small but significant increase in average grade and number of assignments completed.

### *2.5.1 Discussion*

We think our contradicting findings to the existing studies are due to the different characteristics of discussion forum. Most of prior organizational studies on group size were conducted on the collaborative engagement type forum while our experimental setting – MOOCs discussion forum - was individual engagement type where users pursue individual goals, individuals' contribution is purely voluntary and collaborative interaction is not required. In MOOCs discussion forum, the number of users (cohort size) reflects amount of available resource because each individual brings in potentially useful information or knowledge to the forum. And this resource increases perceived value of the forum thus, larger cohort increase engagement from both active and passive users because of the network effect. Since network effect increase exponentially with the size, there also be substantially more people who stimulate others engagement in larger cohort, thus increase discussion contribution per person as a result.

Another nature of the discussion forum in this learning platform is that people come to ask and learn about what they do not know well. This means that people might feel embarrassed if their questions, which they think silly, get much attention from strangers. Therefore, in a larger forum where they can easily observe many others doing the same thing, people might feel more comfortable to write on the forum, in turns, participate more. In fact, students in Large size cohort answered in the survey that they feel more comfortable to post questions or replies on the forum than those in Medium size cohort ($b=0.063$, $p=0.06$; survey Q; How comfortable are you posting questions or replies on the forum?). However, we need to note that this is only a suggestive result and interpretation since the response rate was extremely low (3.8%) and we do not have reliable evidence on this.

In addition to the cohort size's effect, students who were required to participate also engaged more per person in MOOC's discussion forum and had higher course completion rate and performance. In the survey, students in a required participation condition replied that the forum posts are more informative. This might suggest that they get more benefit from the posts than the students in the condition where participation is not a requirement and, in turns, it was reflected to their grade. Again, given the fact that the response rate was very low, we only interpret the survey results in suggestive way.

Our study has clear implications for the cohorting features currently being tested on MOOCs. Despite the appeal of "small class sizes" our results suggest that course designers should not divide learners into separate cohorts. Doing so would reduce the amount of user-generated content that stimulates further engagement with the course. Beyond MOOCs, our results provide a clear guide for how to use cohort size to promote per-person engagement: when people are interacting to collaborate, larger cohorts depress individual contribution; when people are interacting to pursue their own goals and interests, larger cohorts promote greater individual contribution. Dividing users into smaller cohorts may be justified on other grounds – for example to bring together groups of subgroups of users who have particular interests in common. Such shared interests could promote greater participation, but this would need to be weighed against the deleterious effects of smaller cohort size.

### *2.5.2 Future extensions*

Building on this work, further experiments and empirical work could be done to advance our understanding of size in online engagement. The range of cohort sizes we investigated was large for experimental studies but still did not cover the full range of discussion forum sizes observed on the internet. It is possible that at even larger scales the positive effects of size that we observed are outweighed by negative effects only seen within extremely large populations. Additionally, much work in the area of online engagement discusses how cohort composition changes over time, while we studied a short-lived discussion forum. In principle, our results should hold over time as well: the greater resources available in larger forums do stimulate further increase in community membership (Butler 2001). On the other hand, larger forums build up a larger archive of "answered questions" that no longer need to be asked. Future research should establish whether greater contribution per person persists over time.

We have examined a setting which is clearly an example of what we call an "individual engagement" setting, and much prior work has examined clear examples of "collaborative engagement" settings. However, it should be acknowledged that some settings share characteristics of both individual and collaborative engagement. For example, in open source programming communities, there is collaboration on a single collective product, but individual contributors have diverse personal incentives, such as wanting to use the software themselves, career signaling, social capital, or just having fun (Lerner and Tirole, 2002; Lakhani and Wolf, 2005). According to our reasoning, in these mixed settings increased size would have a mixed effect. The individual engagement part of the effect of size is that more users result in more user-generated content that could stimulate others into a state of active contribution. The collaborative engagement part of the effect of size is that more users mean more opportunity to free ride on the efforts of others. A further complication is that the presence of monetary incentives can dramatically alter motivations and behavior (Li and Zhang, 2016). Future work should work to shed further light on settings with mixed attributes to better understand the critical role of contextual variables for the effect of size on contribution.

### *2.5.3 Conclusion*

All prior studies on the effects of size have shown that larger cohorts lead to less engagement per person. To the best of our knowledge, our study is the first to show the opposite effect. We attribute the difference between our results and prior literature to a difference in experimental setting: whereas prior studies have focused on collaborative engagement settings, we studied an individual engagement setting, which are especially common among digital platforms. As such, our subjects were not subject to free-riding pressures but seem instead to have been attracted to the forums by the resources that other users created by their own participation. Our results provide a clear implication for using cohort size as a tool for motivating greater contribution per person: when outcomes are collective, small cohorts elicit more contribution; when outcomes are individual, large cohorts elicit more contribution per person.

# CHAPTER 3: Heterogeneous engagement patterns between in-group and out-group users

## 3.1 Introduction

Digital platforms connect a wide group of individuals, drawn from diverse backgrounds from all over the world, and offer them the opportunity to expand their knowledge through an extensive amount of information provided by each user. For example, question and answer (Q&A) sites such as Stack Overflow and Quora use the wisdom of crowds to complement search engines (Harper et al. 2009). Massive open online courses (MOOCs) allow users to expand their knowledge not only by providing them high-quality lectures, but also by encouraging them to learn collaboratively with diverse individuals in a discussion forum, where they can ask and answer questions regarding the course contents.

Despite its benefit of bringing together new information, this diversity has its pitfalls. For example, interacting with others from different backgrounds takes additional effort from individual users, because they have to manage different behaviors and ways of thinking. Even more energy is required of the minorities in a group. In particular, linguistic and cultural backgrounds are known as the most critical factors to be considered in individuals' knowledge sharing and acquiring process because they are the frameworks through which humans communicate and understand reality (Vygotsky, 1968). In this paper, I study how individuals from different linguistic and cultural backgrounds vary in their approach to platform engagement, especially in terms of knowledge sharing and acquiring in a digital platform.

Lakoff and Johnson (1999) argue that soci0-cultural phenomena are embodied in our minds based on the environment we live in, and this is culturally and linguistically consistent. In other words, a person's conceptual system or frame of reference (e.g., how he or she perceives and reacts to the situation) is developed via his or her cultural experiences and is widespread across languages and cultures. In the offline world, two types of frame of reference exist. For example, assuming the situation of an international student studying in the United States, one frame of reference is the way this student looks at the situation based on his or her life in the United States. The other is based on his or her experience or embodied cultural cognition "back home" (Ogbu 1998; Leung et al. 2011). Conversely, in a digital platform, international users' frames of reference are predominantly from the latter case because the majority of users access digital platforms from their home country. Hence, understanding users' platform engagement through consideration of their hometown situation in terms of language and culture is critical on digital platforms.

These days, the majority of multinational digital platforms use English as a shared language for more efficient knowledge transfer (Altbach 2014, Barak et al. 2016). This means that 75% of the internet population (i.e., non-native English Speakers) use a different frame of reference from native English Speakers, which suggests that they also differ in their attitude and behavior (Ogbu 1998). Research reveals that using English as a common language could lead to miscommunication or misunderstanding among users who are not native English speakers, thus hinder their motivation to engage in the knowledge-sharing process and decrease the performance of them (UNESCO, 2008; Slavin, 1987; Vygotsky, 1978). Nevertheless, we still lack knowledge about how users react to this reality, because English has become a universal language and we consider this a given.

An individual's cultural background – or cultural relevance to the subject – is another overlooked factor in today's digital platforms, although it is a well-known fact that people from different cultures think and behave differently (Lambert 1973). For example, let's assume a case in which users take a course, "War for the Greater Middle East", through an edX platform. This course covers the history of Islamic war and American conflict in the Middle East. There is no doubt that users taking the course from Middle Eastern countries have different frames of reference towards the course contents than those taking the course from the United States. The current platform system pays little attention to understanding how cultural relevance influences users' platform engagement, because the cultural differences are not as visible as they are in offline interaction settings.

In reviewing the literature on the linguistic and cultural diversity of individuals in social interaction, I find that the key limitation is that most of the evidence is assuming face-to-face communication, thus cannot particularly explain virtual communications (Tenzer and Pudelko 2016). In addition, researchers who study user engagement in digital platforms primarily focus on the users' average behavior, rather than considering their background. In this paper, to address this gap, I raise following question: how do individuals' cultural background, i.e., cultural relevance to the subject, and linguistic background, i.e., whether English or non-English, influence their approaches to engaging in a digital learning platform?

In this paper, I empirically investigate how individuals from different cultural and linguistic backgrounds behave differently in the Sabermetrics course on edX, an MOOC platform. For cultural background, I measured the popularity of baseball (cultural relevance to the baseball) in each country using the country ranking and membership in the International Baseball Federation (IBF), because users learn baseball analytics in the edX Sabermetrics course. For linguistic background, since the shared language in the platform is English, I classified users into two groups: users from English-speaking countries and users from non-English-speaking countries. Results indicate that, on average, users from English-speaking countries contribute more to the forum, stay longer, and perform better. However, among those users who put at least some effort into the course materials, users from non-English-speaking countries significantly perform better, stay longer, and demonstrate no difference in terms of forum contribution. The effect of cultural background shows that users from the culture in which baseball is popular contribute more to the forum, stay longer, and perform better than others in the platform. However, the effect is insignificant among users who put at least some effort into the course materials. I then validate these results with another course subject: the Art of Poetry.

## 3.2 Theoretical background

Explosive growth in information technology facilitates communication among people using different languages and from different cultures. These diverse users have unique frames of reference, thus perceive and react to situations differently. A person's frame of reference is developed via his or her cultural experiences and is widespread across languages and cultures (Lakoff and Johnson 1999). Language has a profound effect when individuals with different language backgrounds communicate to each other, because it is central to all aspects of life (Chomsky, 1992; Klitmoller and Lauring 2016). Cultural background is also critical in communication, because people are not 'blank slates'; they interpret the situation within their frame of reference, which comes from their cultural background (Lemke, 2001; Palincsar, 1998; Vygotsky, 1978).

In the offline world, two frames of reference exist. One is the way a person views a situation based on life in the foreign country in which he or she lives, and the other is based on his or her experience or embodied cultural cognition "back home" (Ogbu 1998; Kovecses, 2000; Leung et al. 2011). However, in a digital learning platform, international users' frame of reference largely comes from the latter, because the majority of users access digital platforms from their home country. Hence, understanding users' platform engagement by considering their hometown situation, in terms of language and culture, is critical. This calls attention to the effect of linguistic and cultural background in participants' online engagement with the social learning platform edX, a MOOC.

### *3.2.1 Language in today's educational setting*

Learners use language to manifest their thoughts and knowledge (M. Barak et al. 2016). Proper use of language facilitate communication among learners when learners understand and interpret the meaning correctly (Lemke, 2001; Palincsar, 1998). On the other hand, when the language use is ineffective, it may lead to misunderstanding of the contents or miscommunication among users, thus hinder learning outcomes (Slavin, 1987; Vygotsky, 1978). These days, learners who do not share the same native language use English as a communication medium for effective interactions especially in online contexts (Altbach, 2014). For example, in learning platforms such as edX or Coursera, many courses – even for courses from non-English speaking universities – are produced and delivered in English to provide common communication medium for all learners (Altbach, 2014). Using English as a shared language increases the efficiency of knowledge transfer. However, for the learners who are not native English speakers, the learning motivation might be impeded and the learning process might get slower, even if they can speak and understand the language. This is a common problem in today's learning platform, but we still lack knowledge about how learners react to this reality.

### *3.2.2 Language in the organizational setting*

Language differences are also a major concern for global organizational settings. Prior studies indicate that language differences influence communication dynamics and knowledge transfer (Tenzer and Pudelko 2017; Harzing and Pudelko 2014; Peltokorpi 2017). For example, employees prefer to interact with colleagues who share their native language, rather than with people who speak a foreign language (Tenzer and Pudelko 2017). Some studies have revealed “language-based shadow structures”, which are “communication networks functioning independently from official organizational structures” (Harzing and Pudelko 2014). Related research notes that proficiency in the official language allows conversation participants to take key intermediary roles, which function as informal “language nodes” or “gatekeepers” (Tenzer and Pudelko 2017). Some qualitative studies show that language homogeneity fosters knowledge flow throughout global networks (Peltokorpi 2017), and lack of a shared language delays knowledge transfer and increases transfer costs.

In multinational virtual communication, using a shared language is critical for efficient and effective knowledge transfer. However, in this context, individuals’ ability to use the shared language is critical, because it not only affects their own communication experience but others’ as well. Since English is official corporate language in most real-life global organizations, native English speakers are more likely to achieve better positions of power in the process of knowledge transfer because they have high ability with the shared language.

Though there is much evidence in both organizational and educational settings, most of it assumes communication to happen face to face, thus cannot directly explain virtual communications (Tenzer and Pudelko 2016). This does not match with global platforms’ reality, where much knowledge transfer is conducted virtually, as in social learning platforms such as Coursera or edX.

### *3.2.3 Cultural background and learning*

Learners also acquire knowledge through their belief systems or frames of reference, which have been adopted from their cultural in-groups. Learners who have different cultural backgrounds have different communication style and learning motivation. In today's educational setting, students who study abroad often "experience a culture shock when the organization, behaviors, and expectations of the host university are different from those of the students' culture" (Zepke & Leach, 2005; Zhou, Jindal-Snape, Topping, & Todman, 2008). Naturally, great number of studies on global education have focused on international students' learning styles (De Vita, 2001), stress, and anxiety (Rienties, et al., 2012; Ward, Okura, Kennedy, & Kojima, 1998), and how they engage with the host university.

In the context of the social learning platform, because most of them are open and free learning environments, there exists a large number of users with particularly diverse cultural backgrounds. However, since the cultural differences are not as visible as in offline learning settings, we have limited understanding of how cultural background influences learners' platform engagement. Hence, in this study, I investigate the effect of linguistic and cultural backgrounds on learners' engagement with the social learning platform.

For users to gain knowledge from a digital platform and eventually become committed, engaged members, they must first decide to join the platform, then stick around long enough to learn. When linguistic and cultural out-group users consider joining the platform, they face barriers that cause them to suffer, because they experience two different frames of reference that are psychologically inconsistent (e.g., English vs. native language or platform culture related to the subject vs. individual's hometown culture) (Kraut et al. 2011). Kraut et al. (2011) argue that those who overcome the entry barrier to join the platform eventually contribute and commit more by finding a way to reconcile the cognitive dissonance of two different references. The cognitive dissonance theory states that people come to like things and contribute more in order to lessen their mental stress because this is the only way to reconcile their references or views (Aronson 1997). Hence, the linguistic and cultural out-group users in MOOC platforms who overcome the entry barriers to participate are expected to show higher platform engagement. The hypotheses are:

H1: The out-group users who pass the entry barriers will show higher forum engagement than the majority users;

H2: The out-group users who pass the entry barriers will show higher performance than the majority users;

H3: The out-group users who pass the entry barriers will remain in the platform longer than the majority users.

## 3.3 Research methodology[11]

I tested the impact of cultural and language background on platform engagement by investigating user activities on the 'Sabermetrics 101' course offered on edX, which began on July 7th, 2015 and ran for 10 weeks. Harvard University and the Massachusetts Institute of Technology launched the edX platform in May 2012, and it has become one of the largest MOOC providers, alongside Coursera and Udacity. It offers online university-level courses to internet users around the world and provides an online discussion forum to support user interaction. Participating partners have expanded to more than 110, including universities, nonprofit organizations, and corporations.

### *3.3.1 Dependent variables*

I measured platform engagement in three ways: *social engagement, retention,* and *performance*. For *social engagement*, I considered all posts a user made on a discussion forum, as well as distinguished between new thread posts and comments made on an existing thread. I measured *retention* by recording the number of assignments – including lecture questions, quizzes, and a final exam – that had a non-zero score, as well as a binary version, indicating whether a user completed at least one assignment (any assignments). *Performance* was measured by an average grade of all assignments, quizzes, and the final exam, including any 0s for any that were not completed.

[11] The platform and data description in this method section is overlapped with those in the method section (section 3) in chapter 2 since I use the same dataset.

### *3.3.2 Independent variables*

To construct the independent variable, *cultural background*, I measured the *popularity of baseball* (cultural relevance to the baseball) in each country, because the country's baseball culture would impact participants' subject familiarity when they learned baseball analytics in the edX Sabermetrics course. I collected data from the logs preserved by the edX platform and extracted the country name using the IP address from which a user had logged in. I categorized each country into one of two groups based on whether it is a member of the International Baseball Federation (IBAF; 71 countries), and classified the top 10 countries in IBAF world ranking. In addition to the culture variable, I also measured users' *linguistic background* using the official language of each country. Since English is the shared language in the platform, I classified users into those from an *English-speaking country* and a *non-English-speaking country*. For example, Australian users are categorized into "Top 10 IBAF member" from "English-speaking country", while Spanish users are "Non-Top 10 IBAF member" from "Non-English-speaking country".

### *3.3.3 Other independent variables*

Users provide self-reported data to edX at the point of sign-up for the platform, which was strictly prior to the beginning of the course. I used self-reported gender and self-reported level of education as additional control variables. I also used free-text self-reported "goals" for the course as a source of additional controls; e.g., users that included the word "baseball" in their response were coded with a 1 on the variable "goals: baseball".

Users were selected equally from three different sizes of cohorts: small cohort (125 people per cohort), medium cohort (500 people per cohort), and large cohort (2,000 people per cohort). All users have access to the same course materials but have unique, cohort-specific forums. I used cohort size as another control variable. Further, one-third of users in the data were required to participate in a forum to receive full course credit, while the other two-thirds of users were not. I used participation requirement as a control variable.

Additionally, although most students are auditors in the sense that they take the course for free and do not receive any official credit, edX provided the option to work toward an edX ID-Verified Certificate by paying a $10 fee and achieving a passing grade. I treated the decision to pay the fee (paid verification) as a binary control variable that indicates high intention to complete the course.

## 3.4 Results

Table 15 and Table 16 display the number of users in different categories. They indicate that Sabermetrics is a male-dominant course and most of them are from an English-speaking country, which is one of the top 10 members of IBAF. The out-group users are female, from a non-English speaking country, which is either ranked below top 10 for members of IBAF or is a non-IBAF member. Tables 17, 18, 19, and 20 contain the results of empirical analysis for each dependent variable. For the analysis, I compared generalized linear models with Poisson, Quasi-Poisson, and Gaussian response, checking Akaike Information Criterion (AIC) criteria. The results in Table 17, Table 18, Table 19, and Table 20 are the generalized linear models with Poisson response.

Table 15. Number of users for each user category-1

| | |
|---|---|
| **IBAF Member** | **IBAF Non-Member** |
| 3826 | 278 |
| **IBAF Top10** | **IBAF Below11 + Non Member** |
| 3088 | 1016 |
| **English speaking country** | **Non-English speaking country** |
| 3251 | 853 |
| **Male** | **Female** |
| 3711 | 393 |
| **Goal_baseball** | **Goal_not baseball** |
| 255 | 3849 |

Table 16. Number of users for each user category-2

| | IBAF Top10 Member | Below11+Non Member |
|---|---|---|
| English Speaking Country | 2903 | 348 |
| Non-English Speaking Country | 185 | 668 |
| Male | 2412 | 797 |
| Female | 300 | 93 |
| Goal_baseball | 241 | 14 |
| Goal_not baseball | 2847 | 1002 |

### *3.4.1 Effect of cultural and linguistic background on social engagement*

Model 1 in Table 17 demonstrates that, on average, users from English-speaking countries and members of IBAF contribute significantly more to the forum than others. However, Model 2 shows that the users from English-speaking countries no longer contribute more than others when the top 10 IBAF membership factor is controlled for, which means that the effect of linguistic background disappears. Furthermore, Model 3 and Model 4 show that the effect of English-speaking background and IBAF membership factors is insignificant for users whose grade is higher than 0 or who complete at least one assignment. This indicates that for users who put at least some effort into the course materials, language and cultural background are no longer a barrier for social interaction.

Table 17: Effect of IVs on discussion forum contribution

| | Dependent Variable: Number of Posts | | | |
|---|---|---|---|---|
| | | | grade>0 | retention>0 |
| | (1) | (2) | (3) | (4) |
| | Estimate | Estimate | Estimate | Estimate |
| (Intercept) | 0.32 | 0.06 | 1.34 | 1.66 |
| EngSpeak | 0.79*** | 0.19 | 0.92 | 0.92 |
| ibaf_member | 0.61* | | 0.29 | 0.28 |
| ibaf_top10 | | 1.05*** | | |
| gender_male | -0.02 | 0.00 | -0.27 | -2.27 |
| gender_na | -0.34 | -0.92 | -0.44 | -0.44 |
| CohortSize_M | -0.17 | -0.14 | -0.47 | -0.47 |
| CohortSize_S | -0.51* | -0.49* | -1.14* | -1.14* |
| Participation Required | 2.18*** | 2.17*** | 5.44*** | 5.43*** |
| collegeGrad | 0.11 | 0.11 | 0.52 | 0.52 |
| model | glm | glm | glm | glm |
| Obervations | 4104 | 4104 | 1527 | 1527 |
| AIC | 25406 | 25390 | 10483 | 10681 |

*3.4.2 Effect of cultural and linguistic background on retention*

Model 1 and Model 4 in Table 18 show that users from an English-speaking country stay significantly longer when IBAF membership is controlled for (the result is also the same without controlling for the IBAF membership). Interestingly, for users whose grade is higher than 0 or who complete at least one assignment, the effect of the English-speaking factor becomes significantly negative (seen in Model 2 and Model 3).

Users from non-English-speaking countries stay longer in relation to all users who put some effort into the course materials. Also, when IBAF membership moderates the effect, male users stay significantly longer than female users.

Table 18: Effect of IVs on retention

| | Dependent Variable: Retention | | | |
|---|---|---|---|---|
| | | grade>0 | retention>0 | |
| | (1) | (2) | (3) | (4) |
| | Estimate | Estimate | Estimate | Estimate |
| (Intercept) | 0.02 | 0.33*** | 0.34*** | 0.17** |
| EngSpeak | 0.03** | -0.07* | -0.07* | 0.03** |
| ibaf_member | 0.06** | | | -0.10 |
| ibaf_member_top10 | | 0.03 | 0.03 | |
| gender_male | 0.02* | 0.02 | 0.02 | -0.14* |
| gender_na | 0.01 | 0.03 | 0.03 | -0.14* |
| CohortSize_M | 0.00 | -0.01 | -0.01 | 0.00 |
| CohortSize_S | -0.01 | -0.02 | -0.02 | -0.01 |
| Participation Required | 0.02* | 0.04* | 0.05* | 0.02* |
| collegeGrad | 0.02 | 0.05* | 0.05* | 0.02* |
| gender_male : ibaf_member | | | | 0.17** |
| gender_na : ibaf_member | | | | 0.16* |
| model | glm | glm | glm | glm |
| Obervations | 4104 | 1527 | 1527 | 4104 |
| AIC | 873 | 1023 | 1043 | 870 |

*3.4.3 Effect of cultural and linguistic background on performance*

Results on user performance show similar patterns. Model 1 and Model 4 in Table 19 indicate that users from English-speaking countries achieve significantly higher grades. For users whose grade is higher than 0 or who complete at least one assignment, the effect of the English-speaking factor on performance becomes significantly negative (seen in Model 2 and Model 3). Users from non-English-speaking countries rather achieve higher grades among those who put some effort into the course materials. Also, when IBAF membership moderates the effect, male users achieve significantly higher grades than female users.

Table 19: Effect of IVs on performance

| | Dependent Variable: Grade | | | |
|---|---|---|---|---|
| | | grade>0 | retention>0 | |
| | (1) | (2) | (3) | (4) |
| | Estimate | Estimate | Estimate | Estimate |
| (Intercept) | 0.01 | 0.24*** | 0.26*** | 0.15** |
| EngSpeak | 0.02* | -0.05* | -0.06* | 0.02* |
| ibaf_member | 0.04** | 0.01 | -0.01 | -0.10* |
| ibaf_member_top10 | | | | |
| gender_male | 0.02 | 0.02 | 0.02 | -0.12* |
| gender_na | 0.01 | 0.04 | 0.05 | 0.13* |
| CohortSize_M | 0.00 | 0.00 | -0.01 | 0.00 |
| CohortSize_S | -0.01 | -0.01 | -0.01 | 0.01 |
| Participation Required | 0.02* | 0.04* | 0.04* | 0.02* |
| collegeGrad | 0.02* | 0.05* | 0.05** | 0.02* |
| gender_male : ibaf_member | | | | 0.15* |
| gender_na : ibaf_member | | | | 0.15* |
| model | glm | glm | glm | glm |
| Obervations | 4104 | 1527 | 1527 | 4014 |
| AIC | 93 | 977 | 989 | 90 |

### *3.4.4 Effect of cultural and linguistic background on the rate of perfect scores*

In this section I analyze how well a user performs on each assignment he or she submits. Table 20 displays the results for users whose grade is higher than 0 on (1) how many assignments the user submits (2) how many assignments on which the user achieves a perfect score and (3) the rate of assignments that a user achieves a perfect score on among those he or she submits. In Model 1, as seen in the previous sections, non-English-speaking users complete significantly more assignments when controlling for IBAF membership. Model 2 shows that non-English-speaking users achieve more perfect scores than users from English-speaking countries when controlling for IBAF membership. Interestingly, in Model 3, English-speaking users achieve a higher rate of perfect scores among those they submit. Also, unlike in any other model, participation requirement has a negative effect in Model 3.

Table 20: Effect of IVs on # assignment submit, # perfect score, and #perfect/#submit

| | Dependent Variable | | |
|---|---|---|---|
| | # assignment submit | # perfect score | #perfect/#submit |
| | (1) | (2) | (3) |
| | Estimate | Estimate | Estimate |
| (Intercept) | 11.31*** | 5.78*** | 0.42*** |
| EngSpeak | -1.86* | -0.76** | 0.06* |
| ibaf_member | 0.00 | 0.31 | 0.01 |
| gender_male | 0.71 | 0.49 | 0.01 |
| gender_na | 1.17 | 1.35 | 0.06 |
| CohortSize_M | -0.36 | -0.15 | 0.01 |
| CohortSize_S | -0.71 | -0.21 | 0.00 |
| Participation Required | 1.51* | 0.06 | -0.72*** |
| collegeGrad | 1.71* | 1.27** | 0.06** |
| model | glm | glm | glm |
| Obervations | 1527 | 1527 | 1527 |
| AIC | 11723 | 10432 | 613 |

### *3.4.5 Result validation with a different course subject: The Art of Poetry*

In addition to seeing the effect of cultural and linguistic background in the Sabermetrics course, I validate these results using a different course subject which requires an unrelated set of skills or talent: “The Art of Poetry.” In general, learning poetry is more likely to be influenced by language as compared to learning baseball statistics. Since the lecture mostly covers American poems, including some English poems, I hypothesized that people who are familiar with American or UK culture would have some benefit when taking the course.

Table 21: Effect of IVs on #Posts & Grade

| | Dependent Variable | | | | | |
|---|---|---|---|---|---|---|
| | # Posts (1) | # Posts (2) | Grade (3) | Grade (4) | # Posts (3) | Grade (4) |
| | Estimate | Estimate | Estimate | Estimate | Estimate | Estimate |
| (Intercept) | 1.68*** | 1.61*** | -1.93*** | 0.15*** | 6.6*** | 0.15*** |
| EngSpeak | 0.20 | | -0.04 | | | |
| Literature_relevance | | 0.36** | | -0.01 | | |
| Multilingual | | | | | -3.46** | -0.03* |
| gender_male | -0.29** | -0.27* | -0.16** | -0.02** | -1.65* | -0.02* |
| collegeGrad | 0.20 | 0.08 | 0.04 | 0.01 | 0.61 | 0.00 |
| model | quasipoisson | quasipoisson | quasipoisson | quasipoisson | poisson | poisson |
| Obervations | 4601 | 4601 | 4601 | 4601 | 4601 | 4601 |

While the Art of Poetry has a balanced male to female ratio (11,422 males vs. 11,414 females), there are 250% more users from English-speaking countries than those who are not. Table 21 shows the analysis results. In terms of gender, in every culture I analyze (e.g., US, UK, American, English-speaking, Asian, etc.), females show significantly higher performance and social engagement than males. As seen in Models 1 and 3, there is no significant difference between users from English-speaking countries and non-English-speaking countries in terms of forum engagement and performance, which is consistent with the previous sections. As seen in Models 2 and 4, users familiar with American or English literature contribute more to the forum; however, they show no difference in terms of performance. Interestingly, however, users from bilingual countries in which English is an official language, but using native language is more common, show significantly lower forum engagement and performance.

### 3.5 Discussion and Conclusion

This research study examined how linguistic and cultural background influences users' platform engagement and found mixed results. On average, users from English-speaking countries contribute more to the forum, stay longer, and perform better. However, among users who pass the higher entry barrier (e.g., decide to put at least some effort into the course), users from non-English-speaking countries significantly perform better, stay longer, and demonstrate no difference in terms of forum contribution. The effect of cultural background shows that users from countries who are members of IBAF contribute more to the forum, stay longer, and get perform better than others in the platform. However, the effect becomes insignificant among users who pass the higher entry barrier.

I also validate these results with another course subject, the Art of Poetry, which has a higher entry barrier due to English proficiency and cultural relevance to the subject being more critical. The additional analysis shows that there is no significant difference between users from English-speaking countries and non-English-speaking countries in terms of forum engagement and performance, which is consistent with the previous sections. Users familiar with American or English literature contribute more to the forum; however, they show no difference in terms of performance. Users from bilingual countries in which English is an official language, but using native language is more common, show significantly lower forum engagement and performance. This indicates that English proficiency, which can lower the entry barrier for international students, rather negatively affects their forum engagement.

### *3.5.1 Discussion*

These results indicate that the common belief that social engagement is unfavorable to linguistic or cultural out-group users is not always valid in digital platforms. In general, entry barriers to engage with the platform are likely to drive away potential users. However, when users overcome the entry barrier, they are more likely to show high commitment and contribution. In this study, the linguistic and cultural out-group users who decided to take the course and interacted with the course contents overcame high entry barriers, and they showed higher performance and retention. This suggests that digital platform designers should be able to lower or raise the entry barrier depending on the situation, such as user background.

Another explanation for out-group users' engagement patterns could be their higher motivation. Users who access from foreign countries have had fewer chances to find information about the Sabermetrics course on the edX platform. For example, it is possible that Korean users consume Korean content much more than English content, not only because of the Language, but also because every person has limited resources to acquire and process information. Also, there are only 6 Korean posts regarding the Sabermetrics 101 course, which are searchable via the Korean version of Google. Hence, people who register for the course despite low exposure to the information are likely to be more highly motivated to participate, as compared to the majority users who consume contents primarily in English, thus show higher engagement.

One interesting result is that, unlike other patterns, English-speaking users achieve a higher rate of perfect scores among assignments they submit than non-English-speaking users. This suggests that the goal of taking the course might be different between them. Out-group users who overcome entry barriers tend to show higher commitment to the course, thus have a tendency to complete the course. However, users from English-speaking countries might have the goal to acquire specific knowledge from the subset of contents, therefore achieve higher grades for what they submit.

### *3.5.2 Future extensions*

This paper has limitations, and further empirical work could be done to broaden our understanding of out-group users' heterogenous engagement patterns. For example, the length of the course was relatively short compared to the lifecycle of general online communities. Future research should study how the engagement pattern changes over a longer period of time. Also, although I investigate user engagement patterns in the additional subject of the Art of Poetry, the scope is still not comprehensive. In the future, broader subjects should be covered. Another potentially critical limitation is that cultural background and linguistic background could be confounding variables because individuals' cultural habits may have side effects. For example, the reason for English-speaking users' higher social engagement in the platform might not just be due to the linguistic advantage, but also because people from Western cultures generally engage more in social settings. Another reason could be that users from different cultures might have different values of education or work ethic, thus behave differently. A future study should focus only on the cultural background effect in order to remove any confounding effect of linguistic background. The linguistic background effect should be tested separately in a course which has the least cultural component, such as mathematics or statistics. Another possible analysis could be done in the future to see how subtitles influence non-native-English speakers' platform engagement.

### *3.5.3 Conclusion*

The most notable characteristic of current digital platforms, like MOOCs or any other multinational platforms, is that they tend to bring diverse users from all over the world. For example, more than 70% the of students are international students with diverse backgrounds in MOOC platforms, which is a much higher portion than that of universities in the U.S., which hover around 20% (USNews 2018). Nevertheless, there is lack of policy to support these international users' platform engagement. This is one of a few studies investigating the effect of the cultural and language background of users in a global social learning setting. It offers several insights that extend our understanding of the under-investigated online engagement of users with diverse backgrounds in digital platforms.

## BIBLIOGRAPHY

Adamic, L.A. and Glance, N. (2005). "The political blogosphere and the 2004 U.S. election: divided they blog," Proceedings of the 3rd International Workshop on Link Discovery, Chicago, Illinois: ACM,

Albanese, R., & Van Fleet, D. D. (1985). Rational behavior in groups: The free-riding tendency. Academy of Management Review, 10(2), 244-255.

Alnuaimi, O. A., Robert, L. P., & Maruping, L. M. (2010). Team size, dispersion, and social loafing in technology-supported teams: A perspective on the theory of moral disengagement. Journal of Management Information Systems, 27(1), 203-230.

Altbach, P. G. (2014). MOOCs as neocolonialism: Who controls knowledge? International Higher Education, 75, 5e7

Bakshy, E., Messing, S., and Adamic, L. A. (2015). "Exposure to ideologically diverse news and opinion on Facebook." Science, 348(6239), 1130-1132.

Bales, R., & Borgatta, E. (1966). Size of group as a factor in the interaction profile. In A. P. Hare, E. Borgatta, & R. Bales (Eds.), Small groups: Studies in social interaction (pp. 495-512). New York: Knopf

Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. Computers & Education, 94, 49-60.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? Psychological Science, 26(10), 1531–1542.

Barron, G., & Yechiam, E. (2002). Private e-mail requests and the diffusion of responsibility. Computers in Human Behavior, 18(5), 507-520.

Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). Research note-the impact of community commitment on participation in online communities. Information Systems Research, 22(4), 841-854.

Bernstein, E. S. (2012). The transparency paradox a role for privacy in organizational learning and operational control. Administrative Science Quarterly, 57(2), 181-216.

Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. Government Information Quarterly, 27(3), 264-271.

Borgatti, S. P., & Everett, M. G. (2000). Models of core/periphery structures. Social Networks, 21(4), 375-395.

Brabham, D. C. (2010). Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. Information, Communication & Society, 13(8), 1122-1145.

Bray, R. M., Kerr, N. L., & Atkin, R. S. (1978). Effects of group size, problem difficulty, and sex on group performance and member reactions. Journal of Personality and Social Psychology, 36(11), 1224.

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. Research & Practice in Assessment, 8.

Budhathoki, N. & Haythornthwaite, C. (2013). Motivation for open collaboration: Crowd and community models and the case of OpenStreetMap. American Behavioral Scientist, 57(5), 548 – 575

Burt, R. S. (2004). Structural holes and good ideas. American Journal of Sociology, 110(2), 349-399.

Burtch, Gordon, Yili Hong, Ravi Bapna, and Vladas Griskevicius. (2017) "Stimulating online reviews by combining financial incentives and social norms." Management Science, 64(5), 2065–2082.

Butler, B. S. (2001). Membership size, communication activity, and sustainability: A resource-based model of online social structures. Information Systems Research, 12(4), 346-362.

Butler, B. S., Bateman, P. J., Gray, P. H., & Diamant, E. I. (2014). “An Attraction–Selection–Attrition Theory of Online Community Size and Resilience.” MIS Quarterly, 38(3), 699-728.

Butler, Brian, Lee Sproull, Sara Kiesler, and Robert Kraut (2008). "Community effort in online groups: Who does the work and why." In S.P. Weisband (Ed.), Leadership at a distance: Research in technologically supported work, (pp. 171-194). New York: Lawrence Erlbaum Associates

Cameron, A. C., and Trivedi, P. 2013. Regression Analysis of Count Data (Vol. 53), Cambridge, UK: Cambridge University Press

Carillo, K., & Okoli, C. (2011). "Generating quality open content: A functional group perspective based on the time, interaction, and performance theory." Information & Management, 48(6), 208-219.

Centola, D. (2010). "The spread of behavior in an online social network experiment." Science, 329(5996), 1194-1197.

Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. American Journal of Sociology, 113(3), 702-734.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly, 36(4), 1165-1188.

Chidambaram, L., & Tung, L. L. (2005). "Is out of sight, out of mind? An empirical study of social loafing in technology-supported groups." Information Systems Research, 16(2), 149-168.

Chomsky, N. (1992). Deterring democracy. Hill and Wang.

Clow, D. (2013, April). "MOOCs and the funnel of participation." In Proceedings of the Third International Conference on Learning Analytics and Knowledge (pp. 185-189). ACM.

Coetzee, D., Fox, A., Hearst, M. A., & Hartmann, B. (2014a). Should your MOOC forum use a reputation system? In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 1176-1187). ACM.

Coetzee, D., Fox, A., Hearst, M. A., & Hartmann, B. (2014b). Chatrooms in MOOCs: all talk and no action. In Proceedings of the First ACM Conference on Learning@ Scale Conference (pp. 127-136). ACM.

Collier, Benjamin and Robert Kraut, "Leading the Collective: Social Capital and the Development of Leaders in Core-Periphery Organizations," Proceedings Collective Intelligence, (2012), arXiv:1204.3682 [cs.SI]

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). "Political Polarization on Twitter." Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM) (pp. 89–96).

Culnan, M. J., McHugh, P. J., & Zubillaga, J. I. (2010). How large US companies can use Twitter and other social media to gain business value. MIS Quarterly Executive, 9(4), 243-259.

Dahlander, L. and Frederiksen, L. (2012) "The core and cosmopolitans: A relational view of innovation in user communities," Organization Science, 23, 4, 988-1007.

Dahlander, L., & O'Mahony, S. (2011). Progressing to the center: Coordinating project work. Organization Science, 22(4), 961-979.

Dahlander, Linus, and Siobhan O'Mahony (2011). "Progressing to the center: Coordinating project work." Organization Science 22, no. 4: 961-979

Diehl, M., & Strobe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. Journal of Personality and Social Psychology, 53, 497-509.

Feld, S.L. (1991) "Why your friends have more friends than you do." American Journal of Sociology: 1464-1477.

Fellers, J. W. (1989). The effect of group size and computer support on group idea generation for creativity tasks: An experimental evaluation using a repeated measures design. Unpublished Doctoral Dissertation, Indiana University, Bloomington

Ferschke, O., Yang, D., Tomar, G., & Rosé, C. P. (2015, June). Positive impact of collaborative chat participation in an EDX MOOC. In International Conference on Artificial Intelligence in Education (pp. 115-124). Springer International Publishing.

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Nunamaker, J. F. (1992). Electronic brainstorming and group size. Academy of Management Journal, 35(2), 350-369.

Galuba, W., Aberer, K., Chakraborty, D., Despotovic, Z., & Kellerer, W. (2010, June). Outtweeting the Twitterers-predicting information cascades in microblogs. Proceedings of the 3rd Conference on Online Social Networks. USENIX Association.

Gentzkow, M., & Shapiro, J. M. (2011). Ideological Segregation Online and Offline. The Quarterly Journal of Economics, 126(4), 1799-1839.

Geraci, Marco (2016). Qtools: Utilities for Quantiles. R package version 1.0. http://CRAN.R-project.org/package=Qtools

Gooding, R. Z., & Wagner III, J. A. (1985). A meta-analytic review of the relationship between size and performance: The productivity and efficiency of organizations and their subunits. Administrative Science Quarterly, 462-481

Granovetter, M. S. (1973). The strength of weak ties. American Journal of Sociology, 1360-1380.

Hampton, K, L Rainie, W Lu, M. Dwyer, I. Shin, and K. Purcell (2014) "Social Media and the 'Spiral of Silence'." Pew Internet and American Life Project Publication.

Hampton, K. N., Lee, C. J., & Her, E. J. (2011). How new media affords network diversity: Direct and mediated access to social capital through participation in local social settings. New Media & Society, 13(7), 1031–1049.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. Administrative Science Quarterly, 44(1), 82-111.

Hargittai, E., Gallo, J. and Kane M. (2008). "Cross-ideological discussions among conservative and liberal bloggers," Public Choice, 134(Jan.), pp. 67-86.

Harper, F. M., Moy, D., & Konstan, J. A. (2009, April). Facts or friends? Distinguishing informational and conversational questions in social Q&A sites. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 759-768). ACM.

Harzing, A. W., & Pudelko, M. (2014). Hablas vielleicht un peu la mia language? A comprehensive overview of the role of language differences in headquarters–subsidiary communication. The International Journal of Human Resource Management, 25(5), 696-717.

Haythornthwaite, C. (2009). Crowds and communities: Light and heavyweight models of peer production. In 42nd Hawaii International Conference on System Sciences, 2009. HICSS'09. (pp. 1-10). IEEE.

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. International Journal of Information Management, 33(3), 464-472.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proceedings of the National Academy of Sciences of the United States of America, 101(46), 16385-16389.

Hong, S. (2012) "Online news on Twitter: Newspapers' social media adoption and their online readership." Information Economics and Policy 24(1), 69-74.

Hood, N., Littlejohn, A., & Milligan, C. (2015). Context counts: How learners' contexts influence learning in a MOOC. Computers & Education, 91, 83-91.

Huang, Nina, Yili Hong, and Gordon Burtch. (2017) "Digital social visibility, anonymity and user content generation: Evidence from natural experiments." MIS Quarterly, forthcoming.

Ingham, Alan G., George Levinger, James Graves, and Vaughn Peckham. "The Ringelmann effect: Studies of group size and group performance." Journal of Experimental Social Psychology 10, no. 4 (1974): 371-384.

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. Organization Science, 21(5), 1016-1033.

Johnson, S. L., Safadi, H., & Faraj, S. (2015). The emergence of online community leadership. Information Systems Research, 26(1), 165-187.

Jones, G. R. (1984). Task visibility, free riding, and shirking: Explaining the effect of structure and technology on employee behavior. Academy of Management Review, 9(4), 684-695.

Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. Information Systems Research, 15(2), 194-210.

Joy, S., & Kolb, D. A. (2009). Are there cultural differences in learning style? International Journal of Intercultural Relations, 33(1), 69-85.

Kane, G. C., Fichman, R. G., Gallaugher, J., & Glaser, J. (2009). Community relations 2.0. Harvard Business Review, 87(11), 45-50.

Kane, G.C., Majchrzak, A. and Ives, B. (2010). “Enterprise and Industry Applications of Web 2.0” MISQ Executive, December 2010.

Kang, J. W., & Namkung, Y. (2016). Restaurant information sharing on social networking sites: do network externalities matter? Journal of Hospitality & Tourism Research, 40(6), 739-763.

Katz, M. L., & Shapiro, C. (1994). Systems competition and network effects. The Journal of Economic Perspectives, 8(2), 93-115.

Kerr, N. L. (1989). Illusions of efficacy: The effects of group size on perceived efficacy in social dilemmas. Journal of Experimental Social Psychology, 25(4), 287-313.

Kerr, N. L., & Bruun, S. E. (1981). Ringelmann revisited: Alternative explanations for the social loafing effect. Personality and Social Psychology Bulletin, 7(2), 224-231.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. Journal of Personality and Social Psychology, 44(1), 78.

Kim, J. (2013). Influence of group size on students' participation in online discussion forums. Computers & Education, 62, 123-129.

Klitmøller, A., & Lauring, J. (2016). When distance is good: a construal level perspective on perceptions of inclusive international language use. International Business Review, 25(1), 276-285.

Koh, J., Kim, Y. G., Butler, B., & Bock, G. W. (2007). Encouraging participation in virtual communities. Communications of the ACM, 50(2), 68-73.

Kokkodis, M., & Lappas, T. (2016). Realizing the Activation Potential of Online Communities. International Conference on Information Systems (ICIS), 2016

Kraut, R. E., Resnick, P., Kiesler, S., Burke, M., Chen, Y., Kittur, N., ... & Riedl, J. (2012). Building Successful Online Communities: Evidence-based Social Design. MIT Press.

Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. Journal of College Student Development, 50(6), 683-706.

Lakhani, Karim R., and Robert G. Wolf. (2005). "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects." Perspectives on Free and Open Source Software, Feller, Fitzgerald, Hissam, and Lakhani, Eds., Cambridge: MIT Press

Lakoff, G., & Johnson, M. (1999). Philosophy in the Flesh (Vol. 4). New York: Basic Books.

Lambert, W. E. (1975). Culture and language as factors in learning and education. In A. Wofgang (Ed.), Education of Immigrant Students. Toronto: Ontario Institute for Studies in Education.

Latane, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. Journal of Personality and Social Psychology, 10(3), 215.

Latane, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. Journal of Personality and Social Psychology, 37(6), 822.

Laumann, E. O., Marsden, P. V., & Prensky, D. (1989). The boundary specification problem in network analysis. Research methods in social network analysis. Linton

Freeman, Douglas White and A. Kimball Romney, eds, Chapter 3. Fairfax, VA: George Mason University Press

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2013). Big data, analytics and the path from insights to value. MIT Sloan Management Review, 21.

Lawrence, E., Sides, J. and Farrell, H. (2010). "Self-segregation or deliberation? Blog readership, participation, and polarization in American politics," Perspectives on Politics, 8, 141–157.

Lehoucq, R. B., Sorensen, D. C., & Yang, C. (1998). ARPACK users' guide: Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods (Vol. 6). SIAM.

Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. Journal of Research in Science Teaching, 38(3), 296-316.

Lerner, Josh, and Jean Tirole (2002). "Some simple economics of open source." The Journal of Industrial Economics 50(2), 197-234.

Leung, A. K., Qiu, L., Ong, L., & Tam, K. P. (2011). Embodied cultural cognition: Situating the study of embodied cognition in socio-cultural contexts. Social and Personality Psychology Compass, 5(9), 591-608.

Li, Xitong, and Jiayin Zhang (2016), "The effects of monetary incentives and social comparison on MOOC Participation: A Randomized Field Experiment", Proceedings of the International Conference on Information Systems, Dublin, Ireland, Dec, 2016.

Lin, K. Y., & Lu, H. P. (2011). Why people use social networking sites: An empirical study integrating network externalities and motivation theory. Computers in Human Behavior, 27(3), 1152-1161.

Lowry, P. B., Roberts, T. L., Romano, N. C., Cheney, P. D., & Hightower, R. T. (2006). The Impact of Group Size and Social Presence on Small-Group Communication Does Computer-Mediated Communication Make a Difference? Small Group Research, 37(6), 631-661.

Ma, M., & Agarwal, R. (2007). Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. Information Systems Research, 18(1), 42-67.

Machado JAF and Santos Silva JMC (2005). Quantiles for counts. Journal of the American Statistical Association, 100(472), 1226–1237.

Mao, A., Mason, W., Suri, S., & Watts, D. J. (2016). An experimental study of team size and performance on a complex task. PLoS One, 11(4), e0153048.

Marwick, A.E. and boyd, d. (2010) “I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience” New Media & Society first published on July 7, 2010 doi:10.1177/1461444810365313

Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). Classroom instruction that works: Research-based strategies for increasing student achievement. Alexandria, VA: ASCD.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 415-444.

Newman, M. E. (2003). Mixing patterns in networks. Physical Review E, 67(2), 026126.

Newman, M. E., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. Physical Review E, 64(2), 026118.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 2(2), 175.

Noelle-Neumann, E. (1974). "The Spiral of Silence A Theory of Public Opinion." Journal of Communication 24(2): 43-51.

Oestreicher-Singer, Gal, Lior Zalmanson. 2013. Content or community? A digital business strategy for content providers in the social age. MIS Quarterly 37, 591–616.

Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: a cultural‐ecological theory of school performance with some implications for education. Anthropology & Education Quarterly, 29(2), 155-188.

Page, Scott E. Diversity and complexity. Princeton University Press, 2010.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. Annual Review of Psychology, 49(1), 345-375.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

Peltokorpi, V. (2017). Absorptive capacity in foreign subsidiaries: The effects of language-sensitive recruitment, language training, and interunit knowledge transfer. International Business Review, 26(1), 119-129.

Pendharkar, P. C., & Rodger, J. A. (2009). The relationship between software development team size and software development cost. Communications of the ACM, 52(1), 141-144.

Pew Research Center (2012). Understanding the Participatory News Consumer, Technical report, Pew Research Center, Pew Internet and American Life Project.

Pew Research Center (2014). Social media update, Technical report, Pew Research Center, Pew Internet and American Life Project.

Preece, J. (2000). Online communities: designing usability, supporting sociability. Industrial Management & Data Systems, 100(9), 459-460.

Preece, Jennifer, and Ben Shneiderman (2009). "The reader-to-leader framework: Motivating technology-mediated social participation." AIS Transactions on Human-Computer Interaction, 1(1), 13-32.

Qiu, M., Hewitt, J., & Brett, C. (2012). Online class size, note reading, note writing and collaborative discourse. International Journal of Computer-Supported Collaborative Learning, 7(3), 423-442.

Rainie, L., & Wellman, B. (2012). Networked: The new social operating system. Cambridge: MIT Press.

Ransbotham, Sam, Gerald C. Kane, and Nicholas H. Lurie (2012). "Network characteristics and the value of collaborative user-generated content." Marketing Science 31(3), 387-405.

Reich, J. (2015). Rebooting MOOC research. Science, 347(6217), 34-35.

Ren, Y., Harper, F. M., Drenner, S., Terveen, L. G., Kiesler, S. B., Riedl, J., & Kraut, R. E. (2012). Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds. MIS Quarterly, 36(3), 841-864.

Rienties, B., Beausaert, S., Grohnert, T., Niemantsverdriet, S., & Kommers, P. (2012). Understanding academic performance of international students: The role of ethnicity, academic and social integration. Higher Education, 63(6), 685-700.

Riopelle, K., Gluesing, J. C., Alcordo, T. C., Baba, M., Britt, D., McKether, W., Monplaisir, L., Ratner, H.H., and Wagner, K. H. (2003). Context, task, and the evolution of technology use in global virtual teams. In C. B. Gibson and S. G. Cohen (Eds.), Virtual teams that work: Creating conditions for virtual team effectiveness (pp. 239-264). San Francisco: Jossey-Bass.

Romero, D. M., Meeder, B., & Kleinberg, J. (2011, March). Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. In Proceedings of the 20th International Conference on World Wide Web (pp. 695-704). ACM.

Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., & Sherer, J. (2014, March). Social factors that contribute to attrition in MOOCs. In Proceedings of the First ACM Conference on Learning@ Scale Conference (pp. 197-198). ACM.

Ross, K. (2018) https://www.usnews.com/education/best-colleges/the-short-list-college/articles/2018-02-06/10-national-universities-with-the-most-international-students

Schumpeter, J., 1934, The Theory of Economic Development, Cambridge: Harvard University Press

Shore, J., Bernstein, E., & Lazer, D. (2015). Facts and figuring: An experimental investigation of network structure and performance in information and solution spaces. Organization Science, 26(5), 1432-1446.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. Review of Educational Research, 57(3), 293-336.

Sunstein, C. (2002). The law of group polarization. Journal of Political Philosophy 10(2):175–195.

Sunstein, C. (2008). Why groups go to extremes. AEI Press.

Tenzer, H., & Pudelko, M. (2017). The influence of language differences on power dynamics in multinational teams. Journal of World Business, 52(1), 45-61.

UNESCO. (2008). Mother tongue matters: Local language as a key to effective learning. Paris: UNESCO

Valacich, J. S., Dennis, A. R., & Nunamaker, J. F. (1992). Group size and anonymity effects on computer-mediated idea generation. Small Group Research, 23(1), 49-73.

Valacich, J.S., Wheeler, B.C., Mennecke, B.E., and Wachter, R. (1995). The effects of numerical and logical group size on computer-mediated idea generation. Organizational Behavior and Human Decision Processes, 62(3), 318–329.

Van Alstyne, M., & Brynjolfsson, E. (2005). Global village or cyber-balkans? Modeling and measuring the integration of electronic communities. Management Science, 51(6), 851-868.

Vita, G. D. (2001). Learning styles, culture and inclusive instruction in the multicultural classroom: A business and management perspective. Innovations in Education and Teaching International, 38(2), 165-174.

Von Luxburg, U. (2007). A tutorial on spectral clustering. Statistics and Computing, 17(4), 395-416.

Vygotsky, L. (1978). Interaction between learning and development. Readings on the Development of Children, 23(3), 34-41.

Wang, X., Butler, B. S., & Ren, Y. (2013). The impact of membership overlap on growth: an ecological competition view of online groups. Organization Science, 24(2), 414-431.

Ward, C., Okura, Y., Kennedy, A., & Kojima, T. (1998). The U-curve on trial: A longitudinal study of psychological and sociocultural adjustment during cross-cultural transition. International Journal of Intercultural Relations, 22(3), 277-291.

Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. MIS Quarterly, 35-57.

Wasko, M. M., Teigland, R., & Faraj, S. (2009). The provision of online public goods: Examining social structure in an electronic network of practice. Decision Support Systems, 47(3), 254-265.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. Nature, 393(6684), 440-442.

Watts, D. J., Dodds, P. S., & Newman, M. E. (2002). Identity and search in social networks. Science, 296(5571), 1302-1305.

Wheelan, S. A. (2009). Group size, group development, and group productivity. Small Group Research.

Woodly, D. (2008). “New competencies in democratic communication? Blogs, agenda setting and political participation,” Public Choice, 134, 109-123.

Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011, March). Who says what to whom on Twitter. In Proceedings of the 20th International Conference on World Wide Web (pp. 705-714). ACM.

Yammer (2015) https://about.yammer.com/why-yammer/, accessed 5/1/2015

Yap, Lee-Xian, & Gilbert Bock. (2006). Investigation of the Effects of Members' Closeness Size and IT in Virtual Communities: The Social Network Perspective. AMCIS 2006 Proceedings, 526.

Zalmanson, Lior and Oestreicher-Singer, Gal, 'Your Action is Needed': The Effect of Website-Initiated Participation on User Contributions to Content Websites (May 2015). Available at SSRN: https://ssrn.com/abstract=2606236

Zeileis, Achim, Christian Kleiber, Simon Jackman (2008). Regression Models for Count Data in R. Journal of Statistical Software, 27(8). doi:10.18637/jss.v027.i08

Zepke, N., & Leach, L. (2005). Integration and adaptation: Approaches to the student retention and achievement puzzle. Active Learning in Higher Education, 6(1), 46-59.

Zhang, X., & Zhu, F. (2011). Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. The American Economic Review, 101(4), 1601-1615.

Zhou, Y., Jindal-Snape, D., Topping, K., & Todman, J. (2008). Theoretical models of culture shock and adaptation in international students in higher education. Studies in Higher Education, 33(1), 63-75.

**CURRICULUM VITAE**

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]