Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS
- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Testing temporal hypotheses about response behavior in Q&A forums using a statistical network model and meta-analysis

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
Daniels, Chelsea	New York University	
Bergner, Yoav	New York University	
Lynch, Collin	North Carolina State University	
Barnes, Tiffan	North Carolina State University	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) Mar 2018

Check type of content being submitted and complete one of the following in the box below:
- ☐ If article: Name of journal, volume, and issue number if available
- ☑ If paper: Name of conference, date of conference, and place of conference
- ☐ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

International Conference on Learning Analytics & Knowledge, March 5-9, 2018, Sydney, Australia

DOI or URL to published work (if available)

Testing temporal hypotheses about response behavior in Q&A forums using a statistical network model and meta-analysis

Chelsea Daniels   **Yoav Bergner**
New York University
yoav.bergner@nyu.edu

Collin Lynch   Tiffany Barnes
North Carolina State University

ABSTRACT: In the e-learning context, social network analysis (SNA) can be used to build understanding around the ways students participate and interact in online forums. This study contributes to the growing body of research that uses statistical methods to test hypotheses about structures in social networks. Specifically, we show how statistical analysis can be used to examine the effect of students' past online interactions on their later participatory behavior, and to synthesize results across multiple networks.

Keywords: MOOC, social network analysis, ERGM, meta-analysis

1   INTRODUCTION

With the emergence of massive open online courses (MOOCs) and increasing reliance on computer supported collaborative learning (CSCL) in traditional educational contexts there is an ongoing need to increase understanding around the technical and social interactions that define learning communities online [1, 2]. The application of statistical network models, such as exponential random graph models (ERGM), to SNA in CSCL provides a means by which to examine whether social network structure can be explained by hypothesized mechanisms [3]. This study builds upon recent applications of ERGM to CSCL [3, 5], and explores students' online behaviors in terms of the interactions that contribute to the structure of a social learning environment.

2   CONTEXT AND DATA

In this study we are interested in testing what we term the "pay-it-forward" hypothesis: whether having received a response in previous weeks will contribute to an individual responding to someone else in a given week. This paper looks at the pay-it-forward hypothesis in the context of two different learning environments: (I) a blended computer science course at NCSU with 251 registered students (DM), and (II) a Coursera MOOC with 48,000 total enrollees (BDE). A key contribution of the present work is in the treatment of each week's response network as a sample of behavior at a particular time slice. Results from multiple weeks are then combined by means of a meta-analysis (two-level model). By combining the networks in a multi-level model, we are able to perform a hypothesis test regarding the course as a whole.

3   METHODS AND RESULTS

To study forum participation, we constructed a series of networks where students and instructors are nodes, and responses become directed ties. Forum participation data was segmented by week,

and the number of responses a user received in the preceding weeks was modeled as a node covariate using an ERGM to test our hypothesis in a given week.
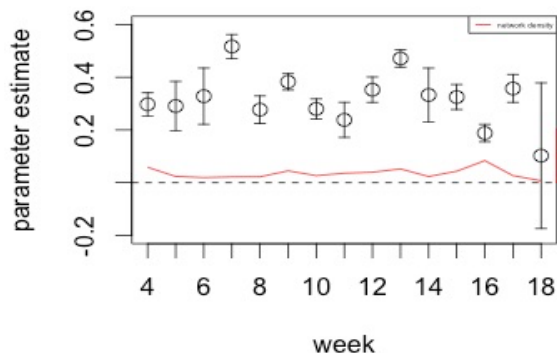
Figure 1: "Pay-it-forward" parameter estimates in DM dataset. Estimates begin in week 4 with a look-back window of three weeks.

Table 1: Two-stage least squares meta-analysis of "pay-it-forward" parameter estimates

	Meta-analytic effect size μ	Standard error
DM_course	0.336***	0.017
DM_wk6	0.367***	0.018
BDE_course	0.155***	0.020

*** indicates p < 0.0001

Our base model controls for density, degree distribution, asymmetry of ties, and instructor status. Previous replies are aggregated over a three-week window to account for the fading significance of past replies. In addition to reporting meta-analysis results for the duration of the course, we also perform a sensitivity analysis on the size of the look-back window. If the "pay-it-forward" hypothesis holds, we expect the a positive parameter estimate. Figure 1 shows a plot of the estimates by week. The estimated mean meta-analytic effect is shown in Table 1. Also shown are results for a particular week (DM_wk6) by varying the look-back window from 1-5 weeks, and results for the BDE course.

In sum, we show a statistical method for testing hypotheses about network behavior in online Q&A forums and a meta-analysis to temporal effects by means of a. It works for courses where student populations are stable as well as online courses with significant attrition effects.

REFERENCES((

1. Haythornthwaite, C. (2001). Exploring multiplexity: Social network structures in a computer-supported distance learning class. *The information society*, *17*(3), 211-226.
2. Rabbany, R., ElAtia, S., Takaffoli, M., & Zaïane, O. R. (2014). Collaborative learning of students in online discussion forums: A social network analysis perspective. In *Educational data mining* (pp. 441-466). Springer International Publishing.
3. Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & De Kereki, I. F. (2016, April). Translating network position into performance: importance of centrality in different network configurations. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 314-323). ACM.
4. Zhu, M., Bergner, Y., Zhang, Y., Baker, R., Wang, Y., & Paquette, L. (2016, April). Longitudinal engagement, performance, and social connectivity: a MOOC case study using exponential random graph models. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 223-230). ACM.