

Prediction Markets for Education: An Experimental Study

CALI MORTENSON ELLIS

Departments of Public Policy and Political Science, University of Michigan

and

RAHUL SAMI

School of Information, University of Michigan

In this letter, we report the results of a quasi-experimental study of prediction markets as a pedagogical tool in an undergraduate setting.

Categories and Subject Descriptors: K.3.1 [**Computing Milieux**]: Computer Uses in Education—*Collaborative Learning*

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: prediction markets

In this letter, we summarize our recent work on using a popular forecasting tool, prediction markets, to supplement classroom learning. A full description of the experiment and results is forthcoming [Ellis and Sami 2012]; a short preliminary version appeared in the proceedings of the Computer Supported Collaborative Learning (CSCL) 2011 conference [Ellis and Sami 2011]. In contrast to most prior empirical research on prediction markets, which has focused on market outcomes and accuracy, we concentrate on the effect of the market on the traders themselves, as well as on characterizing the self-selected group of traders within the larger group of potential traders. Thus, these results may also be of interest to practitioners outside of the educational domain.

Prediction markets are widely used as forecasting tools, in a variety of commercial and non-commercial settings. In arguing for the use of prediction markets, proponents emphasize that they provide incentives that motivate traders to “ferret out accurate information” and “not amplify individual errors, but eliminate them” [Sunstein 2006]. These strengths match our goals as instructors: we want to train our students to search for relevant information, and critically analyze received information. Prediction markets also fit within the larger trend of integrating more interactive and technological resources into classroom learning. In order to test the performance of prediction markets as a learning tool, we carried out a controlled semester-long experiment in an introductory undergraduate political science class.

We used a nonequivalent comparison group quasi-experimental design using both control groups and pretests as per Shadish, Cook and Campbell [Shadish et al. 2002]. Half the class (traders) was randomly selected and granted permission to trade in the markets, while the other half served as the control. At the start and

Author’s addresses: rsami@umich.edu, cmortens@umich.edu

the end of the course, we administered in-class surveys and quizzes. The surveys included questions to gauge student demographics, interest in the class topic, and reading behavior. In addition, we included questions from the Motivated Strategies for Learning Questionnaire (MSLQ) [Pintrich et al. 1993]; these questions have been developed in the education literature to assess students' skill at self-regulating their own learning. The quizzes contained general-knowledge questions related to the market topics. Our final complete dataset included data on 129 students.

Our market deployment built on pilot studies from two previous semesters to increase student comprehension, participation, and trading activity. We deployed a market that used the Zocalo open-source market software (*zocalo.sourceforge.net*), with a custom user interface developed using the Drupal application platform (*www.drupal.org*). Due to the thin market setting, we used the market scoring rule market maker introduced by Hanson [2003]. We created 12 markets relevant to the topic of the course, world politics. In order to avoid impacting grades (even indirectly) with our randomized experiment, we chose market topics to be tangential, although still relevant, to the syllabus of the class. At the end of the semester, students could cash in their trading budgets based on their performance for a modest amount of money (about \$10).

Our first set of results are based on comparisons between the trading and control groups. Unfortunately, we did not find evidence that the trading group had significantly greater improvements in enthusiasm for the subject, independent reading, or quiz knowledge. In fact, all students reported lower levels of enthusiasm for the subject of the course at the end of the semester; one factor may be that the end of semester survey occurred shortly before the final exam. The entire class showed statistically significant improvement in their quiz score results. In both cases, however, the differences between traders and control was not significant.

Our second set of results makes a finer distinction between those who were randomized into the trader group but did not choose to trade (inactive traders) and those who chose to trade more actively in the markets (active traders). There were 22 inactive traders and 45 active traders. One striking finding was that active traders had a higher level of broad information gathering: At the start of the semester, 82% of those students who would eventually become active traders reported reading about the politics of other countries at least once per week; this was significantly higher than inactive traders (54%) and the overall class (70%). Active traders also had higher average MSLQ scores, indicating that they were more self-motivated learners. We noted that active traders' quiz scores improved the most of any group, although the differences-in-differences with inactive traders were not statistically significant.

We detected a possible gender bias among active traders which may be worth considering when deploying prediction markets in a classroom setting. The class was 44 percent female, with no statistically significant difference between students randomized into control and treatment groups. Of the 45 active traders, 29 were male and 16 were female. The proportion of female active traders was slightly (but not significantly) lower than the trader group as a whole. In terms of number of trades, however, there is a statistically significant difference: The average number of trades for male traders was 15.5, but the average for female traders was 6.1.

Active traders indicated a high level of enthusiasm for the use of prediction markets in class. Of the active traders, 68 percent reported that their reason for participating was that they “wanted to win money,” and 44 percent were “interested in learning about the topics.”

We also gathered information on what they relied on for their trading decisions. 58% reported that trading decisions were made “based on personal beliefs,” followed by 51% based on news reports. The smallest number of students reported making trades based on the outcome they wanted (4 percent) or based on the trades of others (i.e. the price reported on the graph - 6 percent). The results of our studies yield several new insights about the use of prediction markets as learning tools.

While we found no significant improvement in students’ enthusiasm or extent of topical reading, we did find that those already reading broadly at the course start were more likely to trade actively in the markets, and those who did trade actively reported that they enjoyed the addition of markets to the class. These results taken together indicate that the prediction markets may be best deployed in a classroom of students who are highly motivated and already engaged in the subject matter. An elective upper-level undergraduate course or a graduate course may be more appropriate settings for using prediction markets as an educational tool. Further, instructors should be aware of the possibility of gender-biased participation. One limitation of our study is that we intentionally picked market questions that were tangential to the course content. An important direction for future research is to study (in a non-randomized design) the use of prediction markets that are closer to the core of the syllabus as a tool for student engagement.

Acknowledgements:

This research was partially supported by the National Science Foundation under grant CCF-0728768. We are very grateful to Allan Stam for supporting these experiments in his courses; Skip Lupia, Stephanie Teasley, Kara Makara, and Rick Wilson for numerous helpful suggestions and feedback; Stan Dimitrov for assistance with the market interface; and Mohammed Hadhrawi, Yung-Ju Chang, and Yi-Chin Wu for helpful discussions on an earlier project.

REFERENCES

- ELLIS, C. AND SAMI, R. 2011. Learning with prediction markets: An experimental study. In *Proceedings of the 9th International Conference on Computer Supported Cooperative Learning (CSCL '11)*.
- ELLIS, C. AND SAMI, R. 2012. Learning political science with prediction markets: An experimental study. *PS: Political Science and Politics (forthcoming)*.
- HANSON, R. 2003. Combinatorial information market design. *Information Systems Frontiers* 5, 1, 107-119.
- PINTRICH, P., SMITH, D., GARCIA, T., AND MCKEACHIE, W. 1993. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement* 53, 3, 801-813.
- SHADISH, W. R., COOK, T. D., AND CAMPBELL, D. T. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton-Mifflin, Boston.
- SUNSTEIN, C. R. 2006. Deliberating Groups versus Prediction Markets (or Hayek’s Challenge to Habermas). *Episteme: A Journal of Social Epistemology* 3.3, 192-213.