Learning Political Science with Prediction Markets: An Experimental Study

Cali Mortenson Ellis, *University of Michigan, Ann Arbor*
Rahul Sami, *University of Michigan, Ann Arbor*

ABSTRACT   Prediction markets are designed to aggregate the information of many individuals to forecast future events. These markets provide participants with an incentive to seek information and a forum for interaction, making markets a promising tool to motivate student learning. We carried out a quasi-experiment in an introductory political science class to study the effect of prediction markets on student engagement with the course topics. Although we found no significant improvement in students' enthusiasm or extent of topical reading, we did find that those students who were already reading broadly at the course start were more likely to trade actively in the markets. These findings indicate that prediction markets may be most successful as an education tool in settings, like graduate education, where individuals are already knowledgeable about the topics of the market, instead of an introductory learning context.

Prediction markets (also known as information markets or decision markets) are designed to aggregate the information of many individuals to forecast future events. Such markets have been used in a variety of contexts. For example, the Iowa Electronic Market forecasts political and economic events including election outcomes, and the Defense Advanced Research Projects Agency's (DARPA) proposed Policy Analysis Market was intended to forecast world political, economic, and military actions. The motivation for using markets for forecasting is that traders reveal some privately held information through the trades they carry out, and the market price aggregates the information of uncoordinated traders through a "wisdom of crowds" effect. These markets have been shown to closely approximate the predictive capabilities of public polls and expert opinions (Wolfers and Zitzewitz 2004). They have been studied for their effect on public policy choices (Hahn and Tetlock 2003), predictive abilities (Servan-Schreiber et al. 2004), and accuracy in the face of severely limited information (Forsythe et al. 1992).

Prediction markets, however, have not been studied for their effect on traders themselves. Pedagogically, these markets are part of a larger trend in education toward including interactive and technological resources in the classroom (Buckley, Garvey,

and McGrath 2011). Among the key advantages of prediction markets, researchers have noted that they provide incentives to motivate traders to "ferret out accurate information" and "not amplify individual errors, but eliminate them" (Sunstein 2006). These strengths align well with our goals as instructors: we want to train our students to search for relevant information and critically analyze received information.

Prediction markets can be a useful in-class teaching tool for a variety of political science subfields, including American elections and campaigns (Abramson 2010) and public policy (Wolfers and Zitzewitz 2004). In the Iowa Electronic Markets, for example, the greatest volume of trading occurs prior to the presidential primary and general elections, and instructors may be interested in using prediction markets in the classroom as early as the fall term of 2011 (i.e., the time leading up to the first primaries of 2012 in Iowa and New Hampshire). Prediction markets also provide interesting information for professors of comparative politics about a variety of international elections (for example, InTrade hosted a market about the presidential elections of Brazil in 2010). Yet, until now, we are unaware of any controlled studies of the use of prediction markets as learning tools.

We carried out a quasiexperiment in an introductory political science class at a large midwestern university to study the effect of prediction markets on student engagement with the course material. We used a prediction market website, built around the open-source Zocalo software, and created markets relevant to the topic of the course (world politics, for the study reported here). We intentionally picked markets that were tangential rather than

Cali Mortenson Ellis *is a PhD candidate in public policy and political science at the University of Michigan, Ann Arbor. She can be reached at cmortens@umich.edu.*
Rahul Sami *is an associate professor in the School of Information, University of Michigan, Ann Arbor. He can be reached at rsami@umich.edu.*

central to the class syllabus because we did not want to give a subset of students an advantage in class performance. We conducted surveys of the entire class at the start and end of the course, with questions intended to measure their engagement with the course topic as well as their knowledge of specific topics. We selected a random group of students to be invited to trade in the market, tracked their individual participation, and provided small monetary rewards based on performance.

The results on the effects of prediction markets on student enthusiasm were disappointing: We detected no significant improvement in students' enthusiasm for the topics of the course or extent of extra reading among students who traded in the markets relative to the control group. However, a deeper analysis revealed that, among the students who were eligible to trade, those who chose to trade actively had a significantly higher prior level of reading in the area. In other words, the students who were already reading broadly at the start of the course were more likely to trade actively in the market. Active traders self-reported a high level of satisfaction with the prediction markets. Our findings are consistent with those of Whitton (2007), who cautions that not all students are receptive to interactive technology in the classroom, warning that, "the sole reason for using [computer-based games] should not be because they are perceived to be motivational." Our results suggest that instructors may benefit from using prediction markets to engage students with selected

students (traders) were allowed to trade in the market with virtual points. Surveys at the beginning and end of the semester evaluated student enthusiasm for the subject matter. Unfortunately, we only obtained a 7% compliance rate among the treated (6 traders), despite 92% awareness of the potential for cash prizes. We also found underreporting of treatment status by more than half (i.e., students assigned to be traders did not report that they knew they were in the treatment group; 41 reported treatment, 83 actually treated). Low compliance among the treatment group, a feature typical of experimental studies of prediction markets (Forsythe et al. 1992; Wolfers and Zitzewitz 2004), limited our ability to make inferences from the pilot study.

As a result, in fall 2009, we performed a nonrandomized study in a smaller ($N = 22$) elective upperlevel undergraduate course that focused on war in international relations. We changed the interface and markets, implemented a different incentive structure, and administered more detailed and indepth surveys to understand what students claim would be motivating to them. We collected unique identifiers to create a panel dataset including both survey responses and trading behavior. We also developed a new quiz about topics related exclusively to the markets to determine changes in knowledge levels at the subject level. Finally, we introduced students to the prediction market experiment using a live in-class demonstration, during which they were able to ask the researchers questions. Both phases of the pilot study informed

Our results suggest that instructors may benefit from using prediction markets to engage students with selected topics, but should do cautiously: this tool is enjoyable and useful for some students, but it does not appear to motivate the entire spectrum of students.

topics, but should do cautiously: this tool is enjoyable and useful for some students, but it does not appear to motivate the entire spectrum of students.

The remainder of this article is structured as follows: First, we outline observations from our earlier pilot experiments that led to conclusions about how to motivate participation in markets. Second, we describe our experimental methodology. Third, we present descriptive statistics of our student population. Next, we provide reports on our data analysis procedures and the results of this analysis. Finally, we discuss the insights yielded by this analysis, the limitations of this study, and interesting directions for future research.

TWO-PHASE PILOT STUDY

A pilot study was carried out in the spring semester of 2009 to gather baseline data about participation in and response to decision markets in a pedagogical setting. This pilot used customized online prediction markets in a large undergraduate course on world politics. Our original primary hypothesis was that participation in prediction markets would increase students' enthusiasm for the topic of world politics. We supposed that students would research the market topics to increase their information advantage and the likelihood of receiving cash rewards. In our pilot phase, 166 students were randomly divided into treatment (trader) and control (nontrader) groups. Presurveys and postsurveys were conducted with both groups at semester start and finish. Treated

the design of our final experiment, which is reported later in this article.

METHODOLOGY

For the data reported here, we used a nonequivalent comparison group design, a quasiexperimental design using both control groups and pretests as per Shadish, Cook, and Campbell (2002). Using a computer algorithm, students were randomized over the entire class into treatment (trader) and control (nontrader) groups. At the start of the semester, we administered an in-class pretest survey, pretest knowledge quiz, and statement of informed consent to each student. We explained the purpose of the research and provided a live demonstration of the market interface projected on a large screen using a simple example market of a major movie's likelihood of winning an Academy Award (Oscar). Based on the feedback from the students in the second phase of our pilot study, we expressed enthusiasm about the project and asked students directly for their help. With the permission of the instructor, we also developed an extra credit question for one of the midterm exams. To ensure fairness and address potential concerns about diffusion, both treatment and control groups were given access to view the markets via the experimental website to learn the information needed for the extra-credit question. However, only the treated group was given access to trade in the market for cash prizes. At the end of the semester, all students were provided with a posttest survey and posttest knowledge quiz. Surveys and

Table 1

Market Topics

- Iran will conduct a nuclear weapons test on or before April 15, 2010.
- Israel will announce a total freeze in settlement construction on or before April 15, 2010.
- Osama Bin Laden will be captured or neutralized on or before March 31, 2010.
- Cuba will be removed from US State Department list of State Sponsors of Terrorism on or before February 15, 2010.
- The Prime Minister of Japan will visit the Yasukuni Shrine by March 1, 2010.
- The trial for Khalid Sheikh Mohammed will begin in the US on or before April 15, 2010.
- Palestinian President Mahmoud Abbas will be announced the winner in the Palestinian elections by January 31, 2010.
- Google will announce that it has ceased all operations in China on or before April 15, 2010.
- Legislative elections in Haiti will be held on schedule on February 28, 2010.
- Germany will win the largest number of gold medals in the 2010 Winter Olympics.

quizzes were administered in class, on paper, to ensure a greater level of participation than often occurs with web-based surveys. As there was some attrition of students who did not attend the last class session, we report results only for the 129 students for whom we collected a complete panel of information.

Specification Checks

Specification checks, using an independent two-tailed group means *t*-test, with an assumption of unequal variances and the Satterthwaite approximation of degrees of freedom, indicated that the randomization process produced two groups that were generally equivalent on survey questions and quiz performance. Thus, the groups were likely to be equivalent on other unobservable characteristics.

Market Topics

By definition, prediction markets are about current events with future uncertainty. Therefore, we created 10 markets with topics that fit the general content of an introductory world politics course. (See table 1.) We attempted to select a range of relevant topics that would be of interest to a variety of students, and included a market with only a tangential connection to world politics—the 2010 Winter Olympics—to get students interested in trading on a topic of general interest.

Market Structure and Interface

We provided a market built using Zocalo, an open-source software package for running prediction markets. In prior trials, however, we found this trading interface to be slightly complex for new traders. Hence, we implemented a custom-user interface using the Drupal open-source application platform. In addition to providing a simple trading interface, we included a sidebar in which students could see a news feed consisting of recent headlines relevant to the market and a comment box.

Students in the treatment group received a budget of 1,000 virtual points at the start of the semester to buy and sell securities based on the market topics. For each of the market topics, there were two securities: a "yes" security that eventually paid off one virtual point if the event happened, and a "no" security that paid one point if the event did not happen by the closing date. Students could use their point balances to buy or sell shares of either "yes" or "no" for the event to drive the price to a point that they thought was correct. For simplicity, we gave students the option of buying and selling shares in blocks of 10 or 100 points, with an accompanying comment about the meaning of the transaction (see figure 1). Each market had a closing date that could be either the date of the supposed future event (for example, elections in Haiti that were scheduled to be held on February 28, 2010), a date in the middle of the term chosen by the researchers, or the end of the term. It was important to have some markets close in the middle of the term, as the students' correct predictions in such a market could generate additional points to be used for trading in another market. In other words, their balances could increase and they could trade more and more frequently if they wished, or bank their "winnings" for the cash prize payout.

Students did not directly trade securities with each other; instead, we used a format in which an (automated) market maker buys securities from students who want to sell, and sells securities to students who want to buy. The market maker quotes a price at each point in time, and adjusts the prices for each security bought or sold. We used the market scoring rule market maker introduced by Hanson (2003). Using a market maker is very helpful in thin market settings such as those we had: it eliminates the need for a buyer and seller to be simultaneously interested in trading, and it allows instantaneous trades to take place at any time. A market maker also reduces the strategic complexity for traders and has attractive incentive properties (Hanson 2003).

Incentive Structure

Students were told that their trading account with 1,000 points would translate into cash at the end of the semester at a rate of 1 cent per point (for an effective cash balance of $10) but the cash could only be collected if the students logged in and executed at least one trade. Of the 66 students in the treatment group for whom we were able to collect a full panel, 45 logged in to trade, and 21 did not. Only three students executed only one trade and, of those, only one student actually picked up the cash prize. The mean number of trades was 29.16 and the mean number of final points was 1052.02. Students were told at the beginning of the semester that their balance could go up or down depending on the success of their trading activity. In other words, trades based on more accurate predictions would be rewarded with a higher cash payout, and poor predictions could mean that students end up with less than their original $10 balance. Seventeen students ended up with point balances lower than their initial endowments.

Communication with Students

We sent a number of e-mails to students in the treatment group to facilitate communication. Each e-mail was written in an enthusiastic tone, reminded students to trade in the market, provided the direct link to the market, as well as the current trading price and volume for selected markets. The e-mails also included links to news stories about selected markets and a commentary on the trend of trading behavior. The e-mails were carefully written to

Figure 1
Screenshot of Market Interface

Note: This screenshot was taken long after the market trading period, and hence the headlines shown here are not identical to those read to the subjects.

take a tone that would encourage students to trade in the selected markets, but not influence trading one way or the other.

Knowledge Quizzes and Surveys

All students—in both the treatment and control groups—completed "Knowledge Quizzes" at the beginning and end of the semester to gauge if learning had occurred on topics relevant to the markets. The quiz at the start of the semester consisted of 24 questions with three options—true, false, or "I don't know." Students were encouraged to answer "I don't know" if they truly did not know the answer to discourage false positive responses. In addition to the Knowledge Quizzes, students filled out surveys that elicited information about their background, their enthusiasm for the course topic, and the sources they relied on for information. For the treatment group, the final survey had additional questions to gauge the students' own perceptions of the prediction markets and their reasons for trading or not trading. The descriptive statistics from the initial survey and knowledge quiz are detailed next.

DESCRIPTIVE STATISTICS FROM INITIAL KNOWLEDGE QUIZ AND SURVEY

Here, we report values for the overall population of the study, with specification checks indicating the results of two-tailed *t*-tests for mean differences between nontrader (control) and trader (treatment) groups in parentheses unless otherwise indicated. For example, the class was 44% female ($\bar{x}_{control} = 0.4677$, $\bar{x}_{treated} = 0.4179$, $|t| = 0.5654$, $p = 0.5728$). *P*-values exceeding our preferred $\alpha = 0.05$ indicate that we cannot reject the null hypothesis that the

difference of means between traders and nontraders is zero, while an α less than 0.05 indicates a statistically significant difference in the means for the two groups.

Knowledge Quiz

We present an overview of the salient features of the initial knowledge quiz. We noted that there was a high negative correlation between correct answers and the number of students answering "I don't know"—in other words, students did not generally attempt to answer questions for which they did not know the correct answer, consistent with our intentions in developing the quiz. All questions were based on prediction market topics. There was no significant difference in mean performance between the control and treatment groups for the initial quiz ($\bar{x}_{control} = 0.3031$, $\bar{x}_{treated} = 0.3377$, $|t| = 1.2861$, $p = 0.2008$). The knowledge quizzes at the end of the semester included the same questions as well as three additional questions on current events topics not part of the experiment (the revolution in Kyrgyzstan, earthquake in Chile, and coup of the president of Niger) to assess if traders gained a greater overall understanding of events of world politics.

MSLQ Survey Questions—Cognitive and Meta-Cognitive Strategies

To control for intrinsic motivation, we administered a modified version of the Motivated Strategies for Learning Questionnaire (MSLQ, from Pintrich et al. 1993), specifically the section, "Cognitive and Meta-Cognitive Strategies: Self-Regulation." As used in educational research, the MSLQ locates students who are more

Improvement in Key Dependent Variables over One Semester

QUESTION—CHANGES FROM START TO END SURVEYS	NONTRADERS START	NONTRADERS END	NONTRADERS DIFFERENCE	TRADERS START	TRADERS END	TRADERS DIFFERENCE	DIFFERENCE-IN-DIFFERENCE (NON-TRADERS MINUS TRADERS)
How interested are you in world politics? (1, completely unenthusiastic—6, completely enthusiastic)	5.6290	5.2903	−0.3387* ($\lvert t \rvert = 1.8002$ $p = 0.0768$)	5.6119	4.9701	−0.6418 ($\lvert t \rvert = 3.1011$ $p = 0.0028$)	0.3031 ($\lvert t \rvert = 1.0836$ $p = 0.2806$)
Do you read news about the politics of other countries at least once a week in a newspaper or on the internet?	0.6613	0.7313	0.0806* ($\lvert t \rvert = 1.2981$ $p = 0.1991$)	0.7419	0.8060	0.0746* ($\lvert t \rvert = 1.6896$ $p = 0.0958$)	0.0060 ($\lvert t \rvert = 0.0790$ $p = 0.9372$)
Quiz Scores (mean percent correct answers)	0.3031	0.3542	0.0511 ($\lvert t \rvert = 2.1891$ $p = 0.0511$)	0.3377	0.3936	0.0559 ($\lvert t \rvert = 4.6728$ $p = 0.0559$)	−0.0048 ($\lvert t \rvert = 0.2108$ $p = 0.8334$)

likely to "use more deep-processing strategies such as elaboration and organization and who attempt to control their cognition and behavior through the use of metacognitive planning, monitoring, and regulating strategies [and] are more likely do better in their course assignments, exams, and papers as well as overall course grade" (Duncan and McKeachie 2005). The mean MLSQ score for the class overall was 3.6843 out of 7.

Survey Questions on Dependent Variables of Interest: Enthusiasm, Outside Reading, and Knowledge

As we were primarily interested in how trading in a prediction market would improve students' knowledge of and interest in political science, our key dependent variables of interest pertained to enthusiasm, outside reading on world politics, and general knowledge of world politics. Using a 7-point Likert scale to measure student's enthusiasm for world politics, we found an initial average level of enthusiasm of 5.6202 ($\bar{x}_{control} = 5.6290$, $\bar{x}_{treated} = 5.6119$, $\lvert t \rvert = 0.0763$, $p = 0.9393$), corresponding to a point between "slightly enthusiastic" and "fairly enthusiastic." We also asked if students read news about the politics of other countries at least once a week in a newspaper or on the Internet. Almost 70% of students indicated that they did ($\bar{x}_{control} = 0.6613$, $\bar{x}_{treated} = 0.7313$, $\lvert t \rvert = 0.8591$, $p = 0.3919$), which may indicate a selection effect for students who elect this specific course. The knowledge quiz was the final component of interest, measuring actual student knowledge about the topic(s) and not just student enthusiasm.

RESULTS

In this section, we present results on the effect of student exposure, obtained by comparing the initial and final surveys and knowledge quizzes of control (nontrader) and treatment (trader) group.

Our original hypothesis was that participation in prediction markets would increase student enthusiasm for and knowledge of the subject matter of an undergraduate political science course. Therefore, we expected to see improvements among traders in our primary dependent variables of interest: enthusiasm, outside reading, and knowledge.

Unfortunately, all students reported lower levels of enthusiasm for the subject of the course, and traders in the treatment group reported even lower levels of enthusiasm than their control group counterparts. As the end of semester survey occurred shortly before the final exam, we were not surprised to see an overall drop in course enthusiasm. The first six columns of table 2 report the results of a within-group paired t-test indicating the significance of the changes within each group before and after the intervention (or lack thereof, for the control group.) Specifically, we test the null hypothesis that the difference between the measure from the beginning to the end of the semester is equivalent to zero. Each star therefore indicates that the change is significantly different from zero at our preferred $\alpha = 0.05$. For the control group, the drop in enthusiasm and increase in reading of news from other countries both showed significant within-group changes, and for the treatment group, only the increase in reading was significant. The last column of table 2 shows the results of an independent two-group t-test of the differences between the groups. This difference-in-difference between traders and nontraders was not significant for any of the key dependent variables. This means that although some of the changes from the beginning to end of the semester were significant for each group, the changes in levels between traders and nontraders was not. Nonetheless, the greater drop in enthusiasm among traders versus nontraders is surprising and led us to ask further questions about the trader group. We next decompose the treatment group into active participants to gain a more fine-grained perspective on the data driving these results.

Indicators of Active Trading

The following section reports results only for students in the treatment group (traders). Here, we differentiate between *active* traders and *inactive* traders to determine if there are differences between students in the treatment group who chose to participate in the market. Active traders are defined as those who logged in and executed at least one trade, as was required to receive the cash payment. Inactive traders are those students in the treatment group who never logged in. In the treatment group (those randomly assigned to be traders), 45 students were active traders and 22 were inactive traders. Of the active traders, only three students actually executed only one trade and, of those, only one picked up the cash incentive. Among the 42 students with

Table 3

Reading about the Politics of Other Countries at Least Once per Week

	SEMESTER START	SEMESTER END	DIFFERENCE (START-END) WITHIN-GROUP t-TEST
Nontraders	0.6612	0.7419	0.0806 ($|t| = 1.2981$ $p = 0.1991$)
Traders	0.7313	0.8059	0.0746 ($|t| = 1.6896$ $p = 0.0958$)
Difference (Non-Traders—Traders) independent 2-group t-test with unequal variance			0.0060 ($|t| = 0.0790$ $p = 0.9372$)
Active Traders	0.8222	0.8444	0.0222 ($|t| = 0.4432$ $p = 0.6598$)
Inactive Traders	0.5454	0.7272	0.1818* ($|t| = 2.1602$ $p = 0.0425$)
Difference (Inactive—Active) independent 2-group t-test with unequal variance			0.1596 ($|t| = 1.6291$ $p = 0.1119$)
Total	0.6976	0.7751	0.0775*
Difference (Start-End) within-group t-test			($|t| = 2.0670$ $p = 0.0408$)

multiple trades, the mean number of trades was 8.18.[1] In this section, we report values for the overall group of treated students, and indicate differences between active and inactive traders in parentheses unless otherwise indicated.

Active traders had higher MSLQ scores than did inactive traders ($\bar{x}_{active} = 3.7696$, $\bar{x}_{inactive} = 3.4008$, $|t| = 1.5051$, $p = 0.1446$), although their score was comparable to the mean for the control group (3.7228). It is possible that participation in the market is related to self-regulatory cognitive strategies, and that participation in the prediction markets was seen as an additional metacognitive strategy used to supplement course preparation.

We detected a possible gender bias among active traders that may be worth considering when using prediction markets in a classroom setting. First, recall that the class was 44% female, with no statistically significant difference between students randomized into control and treatment groups. Of the 22 inactive traders, 12 were female and 10 were male; of the 45 active traders, 29 were male and 16 were female. The proportion of female active traders was slightly lower than the trader group as a whole, but the difference is not statistically significant. In terms of number of trades, there is a statistically significant difference: The average number of trades for male traders was 15.5172, but the average for female traders was 6.125 ($|t| = 0.3943$, $p = 0.6941$).

Another notable finding among active traders is that they had a significantly higher prior level of reading in the area. 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. By comparison, only 54% of inactive traders reported such reading, compared to 70% of the class overall. In other words, the students who were already reading broadly at the start of the course were more likely to trade actively in the market. As seen in table 3, by the end of the semester, the class as a whole exhibited a statistically significant improvement on this measure, with inactive traders actually making the most gains. Inactive traders exhibited a statistically significant improvement of 18%. Active traders improved their reading by 7%, which is significant at the $\alpha = 0.1$ level, and remained well ahead of other groups, with 84% reporting this activity by the end of the term. There was no statistically significant difference in improvement between active and inactive traders or traders (treatment) and nontraders (control).

The modal answer for nonparticipation among inactive traders was "too busy," with 13 out of 22 inactive traders providing this response. The next most common answer, "not interested in the topics," was given by only four students. When asked directly in a separate question, inactive traders indicated being the busiest both at the start and end of the term. At the same time, inactive traders exhibited lower mean MSLQ scores than either active traders or the class as a whole. Recall that the MSLQ used here is a measure of student's self-regulation about study habits, and a higher score is positively correlated with higher grades. The mean MSLQ score for inactive traders was 3.4008, while the mean for active traders was 3.7697, and for the entire class was 3.6843. We note that self-perceived levels of "busyness" may affect student's willingness to participate in an educational experiment without a direct impact on their coursework or grade, even if the "busyness" is not necessarily directly connected with preparation for the course.

The entire class showed statistically significant improvement in their quiz score results (see table 4). It is not possible from our data to determine precisely which component of the improvement was a direct result of trading. We do note, however, that, despite starting with the lowest mean quiz scores, active traders had the greatest improvement on their scores in the knowledge quiz, increasing the mean percentage of correct answers from 32.96% to 38.68%, a statistically significant 5.72% improvement. Further, among treated students, the inactive traders' improvement was only borderline significant ($p = 0.09$). Although inactive traders started with the highest mean quiz score, and ended up with the highest scores, their amount of improvement was about the same as the class as a whole, and their improvement was not statistically significant.

These results taken together indicate that the prediction markets may have some value in improving content learning for students who are motivated by the instrument, but further experiments are needed to investigate the extent to which they enable learning beyond the specific topics of the markets they trade in.

Motivations and Behavior of Active Traders

The following section only reports findings for students in the group of active traders ($N = 45$), defined as those who logged in and executed one or more trade. Active traders were asked to report their reasons for participation. Percentages here add up to

Table 4
Improvement in Knowledge Quiz Scores by Group

	START QUIZ SCORE	END QUIZ SCORE	DIFFERENCE (START-END) WITHIN-GROUP t-TEST
Active Traders	0.3296	0.3868	0.0572* ($|t|$ = 2.6699 p = 0.0106)
Inactive Traders	0.3542	0.4074	0.0532 ($|t|$ = 1.7485 p = 0.0950)
Difference (Inactive—Active) independent 2-group t-test with unequal variance			−0.00397 ($|t|$ = 0.1069 p = 0.9154)

greater than 100 as students were given the option to choose multiple answers. Of the students who did trade in the experiment, 68% reported that their reason for participating was that they "wanted to win money," and 44% of students were "interested in learning about the topics." Some 20% were interested in learning about decision markets, but only 17% wanted to improve their understanding of world politics. Seventy-eight percent of traders reported that the experiment seemed fun, and 56% stated that it was challenging. Of these students, 80% found the experiment relevant to the topic of the course in which the experiment was conducted.

We also asked active traders if the experiment encouraged them to read about market topics outside of class. This gives us finer-grained data than our survey question for all students, which only asked whether students read some international news sources. Students (51%) reported reading more about the Google/China dispute, and 40% reported reading more about the Olympics. Only 11% reported reading about the military or Palestinian elections, and 9% reported reading about the elections in Haiti. Overall, there was some evidence that the markets prompted additional reading, but that it was not distributed evenly across all the topics. This is consistent with the original design of our experiment, in which we attempted to provide markets on a range of topics to accommodate different interests.

Trading Behavior

Of active traders 69% understood how to use the experimental interface, and 67% found the instructions given by the researchers to be clear. We asked students a categorical question on how they made their trading decisions. Some (58%) stated 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%) or based on the trades of others (i.e., the price reported on the graph—6%). Forty percent of students believed that their predictions were more correct than those of their classmates. This is in line with the claim by Sunstein (2006) that market trading encourages individual information processing and reporting, rather than groupthink or herd behavior.

IMPLICATIONS FOR POLITICAL SCIENCE PEDAGOGY AND IMPLICATIONS FOR FUTURE RESEARCH

As Whitton (2007) warns, not all students will be responsive to interactive technological projects such a prediction markets. One of our most striking findings is that active traders had a significantly higher prior level of reading than other students in the topics of the markets, outside of their course requirements, both at the beginning and end of the semesters. Another finding is students' higher level of MLSQ characteristics, indicating students who have more self-regulated study habits. The students who did actively trade reported that they enjoyed doing so, and that it prompted them to read more widely on some of the market topics. These findings together hint that prediction markets may be best used 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.

For reasons of fairness in our randomized study, we chose a controlled set of market topics that were outside of the specific topics studied in the class. One unanswered question from our research involves how students would interact with a decision market that is more closely connected to their interests. For example, it is possible to design the interface so that students can create markets based on their own interests in certain topics of a class. We asked students directly if they had the opportunity to create their own market on any topic (for example, who will win the NCAA championships), if they would do so, and 44% stated that they would. It would also be interesting to conduct a future study with markets linked to the course syllabus, for verification of whether more immediacy in the market topics would result in learning gains.

Although this experiment was conducted in a class on world politics, we believe our results may still be useful for an instructor considering using prediction markets as an interactive teaching tool for other courses, such as a US campaigns and elections course, especially in conjunction with the 2012 presidential elections. Furthermore, the use of one or two narrowly focused markets, such as those used in the Iowa Electronic Markets, may lead to "thicker" trading behavior, or more informative predictions than those observed in our study. With these caveats, we encourage instructors to consider using prediction markets as a hands-on technology supplement to traditional resources in political science courses.

NOTES

A version of this work was presented at the 9th International Conference on Computer Supported Collaborative Learning (CSCL): Connecting Computer Supported Collaborative Learning to Policy and Practice, July 4–8, 2011, at The University of Hong Kong.

1. With our market interface, a student could buy or sell either 10 units or 100 units at any time. It is therefore possible that a student would have to initiate

multiple trading actions in to move the price to precisely the right level, and so we need to be careful about how we count trades. We use the following methodology in counting trades: Any sequence of trading actions on the same market by the same student within 1 minute was considered a single "trade" for counting purposes. The results would be almost identical if we used a 10-minute limit instead. The mean number of raw trading actions per active trader was 30.47.

REFERENCES

Abramson, Paul R. 2010. "Using Intrade.com to Teach Campaign Strategies in the 2008 U.S. Presidential Election." *PS: Political Science & Politics* 43 (1): 139–44.

Buckley, Patrick, John Garvey, and Fergal McGrath. 2011. "A Case Study on Using Prediction Markets as a Rich Environment for Active Learning." *Computers & Education* 56: 418–28.

Duncan, T. G., and Wilbert J. McKeachie. 2005. "The Making of the Motivated Strategies for Learning Questionnaire." *Educational Psychologist* 40 (2): 117–28.

Forsythe, Robert, Forrest Nelson, George R. Neumann, and Jack Wright. 1992. "Anatomy of an Experimental Political Stock Market." *American Economic Review* 82 (5): 1142–61.

Hahn, Robert W., and Paul C. Tetlock. 2003. "Using Information Markets to Improve Public Decision Making." *Harvard Journal of Law and Public Policy* (Summer); AEI-Brookings Joint Center Working Paper No. 04-18.

Hanson, R. 2003. "Combinatorial Information Market Design." *Information Systems Frontiers* 5 (1): 107–19.

Pintrich, Paul R., David A. F. Smith, Teresa Garcia, and Wilbert J. McKeachie. 1993. "Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ)." *Educational and Psychological Measurement* 53 (3): 801–13.

Servan-Schreiber, Emile, Justin Wolfers, David M. Pennock, and Brian Galebach. 2004. "Prediction Markets: Does Money Matter?" *Electronic Markets* 14 (3): 243–51.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton-Mifflin.

Sunstein, Cass R. 2006. "Deliberating Groups versus Prediction Markets (or Hayek's Challenge to Habermas)." *Episteme: A Journal of Social Epistemology* 3 (3): 192–213.

Whitton, Nicola. 2007. "Motivation and Computer Game Based Learning." In *ICT: Providing Choices for Learners and Learning. Proceedings Ascilite Singapore 2007.* http://www.ascilite.org.au/conferences/singapore07/procs/whitton.pdf

Wolfers, Justin, and Eric Zitzewitz. 2004. "Prediction Markets." *Journal of Economic Perspectives* 18 (2): 107–26.