Does Cooperative Learning Improve Student Learning Outcomes?

Steven Yamarik

Abstract: What is the effect of small-group learning on student learning outcomes in economic instruction? In spring 2002 and fall 2004, the author applied cooperative learning to one section of intermediate macroeconomics and taught another section using a traditional lecture format. He identified and then tracked measures of student learning outcomes. Using multivariate regression analysis, he found that students taught by cooperative learning achieved greater academic performance in the form of higher exam scores.

Keywords: academic achievement, cooperative learning, economic instruction, group learning
JEL codes: A20, A22, I21

The use of active and group learning techniques by economic instructors lags behind other disciplines. In a 1994 survey of academic economists, 86 percent cited lecture as their predominant teaching method, whereas only 7 percent reported using small-group activities regularly (Benzing and Christ 1997). Similarly, in a 2000 survey of economic instruction, the median time devoted to student discussion in the economics classroom was just over 5 percent (Becker and Watts 2001a; 2001b). However, Sax et al. (1999) and Lindholm et al. (2001) reported that active and group learning was used extensively in most other disciplines.

Steven Yamarik is an associate professor of economics at California State University at Long Beach (e-mail: syamarik@csulb.edu). This study was conducted at Tufts University. The author thanks his students for participating in the study, the Registrar’s Office for use of student records, and Peter Kennedy, David Garman, Hirsch Kasper, Jay Shimshack, Seiji Steimetz, and three anonymous referees for their insightful comments. Copyright © 2007 Heldref Publications
Of particular relevance, Brawner et al. (2002) reported that 60 percent of engineering faculty surveyed used assigned group learning at some point in their classes.

My objective in this article is to understand better the effect of small-group learning on student learning results in economic instruction. I applied cooperative learning to a course in intermediate macroeconomics in spring 2002 and again in fall 2004. During the same semester, I taught another section of the course, using a traditional lecture approach. I tracked five types of student learning outcomes: interest, preparation, participation, attendance, and academic performance. Using multivariate regression analysis, I then tested the effect of cooperative learning on learning outcomes.

Researchers in undergraduate science, mathematics, engineering, and technology (SMET) instruction have found that group learning can improve learning performances. For example, Felder (1995) and Felder, Felder, and Dietz (1998) taught five consecutive chemical engineering classes to a cohort of students using cooperative and other active learning techniques. Felder reported that the experimental cohort achieved higher academic performance (retention, grades) and interest levels compared to an instructor-taught cohort. Springer, Stanne, and Donavan (1999) conducted a meta-analysis of 39 studies and found that “various forms of small-group learning are effective in promoting greater academic achievement, more favorable attitudes toward learning, and increased persistence through SMET courses and programs” (p. 21).

In economic instruction, the evidence on the effectiveness of group learning has been promising but far from conclusive. Johnston (1997) found that introductory microeconomic students tutored in a group setting performed better on examinations. Similarly, Moore (1998) showed that students who participated in cooperative learning labs outside of the classroom reported the labs worthwhile and enjoyable. In more recent articles, Johnston et al. (2000) and Brooks and Khandker (2002) incorporated a cooperative-learning approach in weekly recitations (labs). Johnston et al. found that students in cooperative-learning recitations spent more time preparing for the tutorials and were more interested but did not perform any better on the examinations. However, Brooks and Khandker, found that students in small cooperative-learning labs scored higher on the final exam. In a survey of 34 liberal arts colleges, Jensen and Owen (2001) reported that less lecture and more group learning in the classroom encouraged students to take more economic courses and become economics majors.

My application of group learning differed from earlier studies of economics instruction in that I paid more careful attention to group formation and group dynamics. First, I established cooperative base groups of three to four students. Base groups stayed together during the entire course and thus could provide each student the support, encouragement, and assistance needed to progress academically (Johnson, Johnson, and Smith 1991). Second, I had the students work in their base groups both inside and outside the classroom. Inside the classroom, students worked together on problem-solving exercises, where as outside the classroom, students solved problem sets. Third, I established the groups, facilitated the cooperative-learning exercises, and evaluated the outcomes. In most of the
previous economic studies, the group-learning component was incorporated in the recitation periods by the teaching assistants. I believe that these three modifications increased the credibility and effectiveness of cooperative learning and thus had the potential to yield better learning outcomes.¹

To test the effect of cooperative learning, I estimated an empirical model where each learning outcome depended upon teaching pedagogy, demographic factors, economic knowledge, and other academic factors. I found that the use of cooperative learning increased academic performance. The point estimates, which were statistically significant, implied that group learning raised combined (midterm plus final) exam scores by five to six points, *ceteris paribus*. As a percentage of the mean score on the exams, the point estimates translated into a 3 to 4 percent improvement in exam performance. As a consequence, the results of my study suggested that small-group learning could raise academic achievement in economics instruction similar to that found in other disciplines.

**COOPERATIVE LEARNING**

*Cooperative learning* is a teaching method where students work in small groups to help one another learn academic material. In the groups, students are expected to help each other find answers to questions, rather than seeking answers from the instructor. Cooperative work rarely replaces teacher instruction but rather replaces individual lecture and drill. If implemented properly, students in cooperative groups work with each other to make sure that everyone in the group understands the concepts being taught. Ultimately, the success of the group depends on its ability to make certain that everyone grasps the key ideas (Slavin 1995).

Johnson, Johnson, and Smith (1998) listed five elements essential for successful cooperative learning groups. First, there must be *positive interdependence* in that members of the group understand that they should learn together to accomplish their goal. Second, there must be *promotive interaction* in that students interact face-to-face in the group. Third, there must be *individual and group accountability* in that members are held responsible for their own contribution to the group’s success. Fourth, there must be *group processing* in that members reflect on their collaborative efforts and decide on ways to improve effectiveness. Fifth, there must be the *development of small-group interpersonal skills* such as giving constructive feedback, involving each member, and reaching a consensus.

Some educational researchers such as Bruffee (1995) have insisted there is a clear and important distinction between cooperative learning and collaborative learning. To these educators, *cooperative learning* supports the traditional role of teacher as a subject matter expert and classroom authority whereas *collaborative learning* has the instructor work directly with students to discover and create knowledge. In cooperative learning, the instructor sets the task and has the students work in groups to find the “correct” answer. In collaborative learning, however, knowledge is not set by the instructor but rather is acquired through consensus among students and the teacher (Barkley, Cross, and Major 2005).
Cooperative Learning Application

In my experiment, I applied cooperative learning to a course in intermediate macroeconomics. I used the group-learning method of Johnson and Johnson (1987), which requires the instructor to: (1) make a number of preinstructional decisions, (2) explain the task and the positive interdependence, (3) monitor students’ learning and intervene to provide task assistance, and (4) assess students’ learning (Johnson and Johnson 1999). I set aside nine class sessions for group problem-solving exercises. The task and reward system for the groups was established at the beginning of the course. The students knew that they were to produce a group answer to each question and that each answer was then presented to the rest of the class. I monitored the task by circulating throughout the classroom and answering questions of clarification.

I established base groups of three to four students in classes that ranged from 22 to 35 students. On the first day of class, I administered a 10-question test to gauge mathematic skills. In the third class, I gave the preexperiment questionnaire and asked for demographic and scholastic information. I asked students to e-mail me their individual preferences for group members. Using the math test results, demographic information, and e-mail responses, I established groups in the fifth class. I attempted to form groups that were heterogeneous in aptitude and demographics but also fulfilled some of the students’ preferences.

The students met in their groups both inside and outside the classroom. In class, the groups applied economic theory to a new situation. The new economic situation was presented in the form of additional readings and handouts made available beforehand. Each handout contained a series of questions and was made available beforehand on the class homepage. The groups went over the questions in class. Initially, some students were somewhat hesitant to interact with fellow group members. In certain groups, students worked individually on the answers. However, as the semester progressed, students became more comfortable in the group setting and began to truly collaborate in their groups. At the end of the class period, a representative from each group presented the group answer to the rest of the class. An eight-sided dice was used to determine which group presented which answer.

Outside the classroom, the groups worked on the problem sets. For each problem set, one group member was designated as the leader. The leader was responsible for coordinating the group and making sure everyone understood the answers before they were handed in. The role of leader rotated through the group members. There is, of course, a huge potential for free riding on the problem sets. I had the students report on their fellow group members at the end of the semester but was aware of the limitations of this type of self-reporting.

Experimental Design

I conducted the experiment in spring 2002 and fall 2004. In each semester, students could register for one of four sections of intermediate macroeconomics: Monday and Wednesday (MW), early afternoon or late afternoon, Tuesday and
Thursday (TTh), morning or early afternoon. I taught one MW section (control) as a traditional lecture style and the other MW section (experimental) using cooperative learning. Because the early afternoon sections were more popular, the time of the control and experimental sections in 2004 were switched to keep the sample balanced. The two TTh sections were taught by other instructors who chose not to participate in the experiment. The students were only aware of the time and not the control/experiment grouping when registering.

I used the same course organization and content in the control and experimental sections. The course was divided into five main parts: introduction (with national income and product accounts), long-run static model, long-run growth model, short-run IS-LM model (with business cycle theory), and macroeconomic policy. Mankiw’s (2003) *Macroeconomics* was the required text in both sections. In addition, I assigned identical additional readings, handouts, and problem sets, and proctored nearly identical exams. The final grade for the course was calculated as 10 percent for attendance, 20 percent for problem sets, 30 percent for the midterm exam, and 40 percent for the final exam. This breakdown followed the recommendation of Cross and Steadman (1996) that grades from peer group work make up a small part of the overall course grade.

The purpose of the additional readings and handouts was to provide students with a relevant economic situation to apply and extend the theoretical models developed in class. There were nine problem-solving exercises in the course: three for the long-run static model, two for the long-run growth model, two for the short-run IS-LM model, and two for macroeconomic policy. In the problem-solving exercises, the students first answered objective-type questions and then moved on to application-type questions.

The teaching method practiced in each section differed. In both sections, I developed the theoretical model on the blackboard. The development of the model was then followed by a problem-solving exercise in the form of an additional reading and a handout of questions. In the traditional lecture section, I led the problem-solving exercise. I spent the first 15 minutes introducing the topic and then asked for volunteers from the entire class to answer the questions on the handout. In the experimental section, the students discussed the reading and handout in their groups. I circulated throughout the classroom to encourage interaction and answer questions of clarification, but the students themselves answered the questions in the handout. During the final 15 minutes of the class, a representative from each group presented the answer to the rest of the class.

**Limitations of the Study**

A few weaknesses in the study design limit the generality of the results. First, there is a real possibility that unmeasured biases can occur when one instructor teaches the same course. For example, differences in instructor excitement across the two groups could have contributed to student outcomes. Moreover, the uniqueness and possible awkwardness of being in an experiment could have affected student outcomes. In their meta-analysis of SMET instruction, Springer, Stanne, and Donavan (1999) found that studies with the investigator as the instructor
reported significantly greater effects for small-group learning. Second, the application of cooperative learning to one course limits its generality. The students in the sample were typically sophomores and juniors with three prior economic classes, on average. One wonders if the impact of cooperative learning would be the same on a principles class with mostly freshman or an upper-level elective with mostly seniors. Third, cooperative learning may not work for all instructors. For any teaching method to prove effective, the instructor must be convinced of its merits. These limitations can be addressed if the experiment is repeated by a number of instructors; the learning results could then be compared across different teaching environments and instructors.

DATA

Student Learning Outcomes

I tracked five different types of student learning results: interest, participation, preparation, attendance, and performance. The first three results were collected from a preexperiment and postexperiment questionnaire. I measured preparation as the willingness to talk to other students, the willingness to talk to the instructor outside the classroom, and the amount of time spent reading and studying the material outside the classroom. Participation was measured by the preference for working together, the regularity in the students discussion of their work with other students, and the number of times they spoke in the last three classes. Interest was seen if the student enjoyed the intellectual challenge of economics, if the student was interested in economics, if the student enjoyed economic theory, and if the student applied economics to real-life situations.

I collected the outcomes for attendance and academic performance from daily attendance records and graded course material. I measured class attendance as the percentage of total classes attended. The midterm exam covered the first three sections of the course (introduction, long-run static model, and long-run growth model), and the final exam covered the last two sections of the course (short-run IS-LM model and macroeconomic policy). I measured academic performance as the sum of the raw scores on the midterm exam and the final exam to minimize the potential measurement error of each exam. Because of the potential for free riding, I excluded the problem set grade as a learning outcome.

Sample

A total of 116 students were enrolled in the four intermediate macroeconomic courses. In 2002, 35 students were in the early afternoon control section and 22 in the late afternoon experimental section. In 2004, 35 students were in the early afternoon experimental section and 24 in the late afternoon control section. In each semester, one student from the control section dropped the course after taking the midterm exam. In addition, one student from the experimental section did not take the midterm exam in 2004. For student interest, participation, and preparation measures, 104 students completed the preexperiment questionnaire, 103 completed the postexperiment questionnaire, and 93 completed both.
RESULTS

Student Interest, Preparation, and Participation

The sample means of the student interest, preparation, and participation outcomes are presented in Table 1, with sets of columns for preexperiment, postexperiment, and the difference between the post- and preexperiment. Each set of columns contains the values for the control group, the treatment group, and the difference between the treatment and control groups.

To test for a significant difference in means across the control and treatment groups, I estimated the following regression equation:

$$outcome_i = \beta_0 + \beta_1 \text{experimental section}_i + \varepsilon_i,$$

where experimental section is a dummy variable recording cooperative learning, $\beta$'s are the coefficients to be estimated, and $\varepsilon$ is the error term. The point estimate (and $t$ statistic) of $\beta_1$ measures the difference in mean values, which is reported in each Diff column. The last column reports the “difference in differences” result where the population average difference over time in the control group is subtracted from the population average difference over time in the treatment group (Meyer 1995).

There are some significant differences between the treatment and the control groups in the pre- and postexperimental values but only one significant difference in differences between the treatment and control groups (Table 1). In the preexperiment set of columns, the treatment group was more interested in economics (items 1 and 4) and more inclined to study in groups (items 11 and 12). Similarly, in the postexperiment set of columns, the treatment group was more interested in economics (items 1–3) and better prepared (item 10). However, in the difference set of columns, there was only one significant difference in differences (item 14). As a result, the use of cooperative learning in my experiment appears to have had little to no impact on student interest, participation, and preparation outcomes.

I also estimated a proportional-odds ordered logit model (McKelvey and Zavoina 1975), using the difference between the pre- and post experimental values as the dependent variable. The dependent variable ranged from $-4$ (strongly agree to strongly disagree) to 4 (strongly disagree to strongly agree). Along with the dummy for the experimental section, I included data for class time, demographic factors, current academic factors, and economic knowledge as independent variables. For the 14 outcomes, I found that the experimental dummy variable was significant only once (item 2). Consequently, the ordered logit results confirmed the difference in differences results of Table 1 that cooperative learning had a negligible effect on student interest, participation, and preparation outcomes.

Attendance and Academic Performance

Sample statistics of the students in the experiment are given in Table 2. Recall that students were only aware of the time and not the control/experiment grouping when registering. There are three sets of columns for the control group, the
| No. | Question                                                                 | Preexperiment |         |         | Postexperiment |         |         | Difference |         |         |         |
|-----|--------------------------------------------------------------------------|---------------|---------|---------|---------------|---------|---------|------------|---------|---------|---------|
|     |                                                                          | Treat (n = 51) | Control (n = 51) | Diff. (t stat) | Treat (n = 51) | Control (n = 52) | Diff. (t stat) | Treat (n = 48) | Control (n = 45) | Diff. (t stat) |
| 1.  | I enjoy the intellectual challenge of studying economics.                | 4.25          | 3.98    | 0.27**  | (2.25)         | 4.35          | 4.00    | 0.35**  | (2.56)         | 0.083          | 0.022          | 0.061   |
| 2.  | I am interested in economics.                                           | 4.53          | 4.37    | 0.16    | (1.54)         | 4.55          | 4.27    | 0.28**  | (2.13)         | 0.021          | −0.133         | 0.154   |
| 3.  | I enjoy economic theory.                                                | 4.23          | 3.98    | 0.25*   | (1.88)         | 4.33          | 4.08    | 0.25**  | (2.04)         | 0.104          | 0.068          | 0.036   |
| 4.  | I apply economic theory to real-life situations.                        | 4.15          | 3.76    | 0.39*** | (2.89)         | 4.06          | 3.88    | 0.18    | (1.06)         | −0.063         | 0.114          | −0.177  |
|     |                                                                          |               |         |         |               |         |         |            |         |         |         |
|     |                                                                          |               |         |         |               |         |         |            |         |         |         |
|     |                                                                          |               |         |         |               |         |         |            |         |         |         |
|     |                                                                          |               |         |         |               |         |         |            |         |         |         |

Table 1. Sample Means of Student Interest, Participation, and Preparation
| No. | Question                                                                 | Preexperiment | Postexperiment | Difference |
|-----|---------------------------------------------------------------------------|---------------|----------------|------------|
|     |                                                                           | Treat (n = 51)| Control (n = 51)| Diff. (t stat) | Treat (n = 51)| Control (n = 52)| Diff. (t stat) | Treat (n = 48)| Control (n = 45)| Diff. (t stat) |
| 9.  | I usually do the required reading before each class.                      | 3.83          | 3.67           | 0.16 (0.28)  | 3.64          | 3.35           | 0.29 (1.28)    | -0.298        | -0.356           | 0.058 (0.28)    |
| 10. | I usually try to understand the material before each class.              | 3.58          | 3.41           | 0.17 (0.99)  | 3.53          | 3.12           | 0.41** (2.22)  | -0.146        | -0.267           | 0.121 (0.59)    |
| 11. | I regularly discuss my studies with other students in the class.         | 3.34          | 2.61           | 0.73*** (4.12)| 3.57          | 3.23           | 0.34 (1.59)    | 0.208         | 0.511            | -0.302 (−1.30)  |
| 12. | I prefer to work alone.                                                  | 2.92          | 3.35           | -0.43** (−2.33)| 2.90          | 3.06           | -0.16 (−0.71)  | -0.042        | -0.267           | 0.225 (0.99)    |
| 13. | I seek help from other students when I need it.                          | 3.96          | 3.86           | 0.10 (0.59)  | 3.78          | 3.75           | 0.03 (0.18)    | -0.167        | -0.155           | -0.012 (−0.06)  |
| 14. | I seek help from the instructor when I need it.                          | 3.64          | 3.92           | -0.28** (−1.62)| 3.84          | 3.73           | 0.11 (0.61)    | 0.208         | -0.178           | 0.386* (1.84)   |

Notes: Each entry is the sample mean. The values for the preexperiment and postexperiment are measured on a 1 to 5 scale where 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree, except for item 7. The values for item 7 are the self-reported number of times each student spoke in the last three classes. The values for the difference columns are the change in values from the preexperiment to the postexperiment. The t statistics are in parentheses and have been corrected for heteroscedasticity. *p = .10, **p = .05, ***p = .01.
### Table 2. Descriptive Statistics

| Characteristic          | Treatment $n = 57$ | Control $n = 59$ | Difference | Mean | SD | Mean | SD | $t$ stat |
|-------------------------|--------------------|------------------|------------|------|----|------|----|----------|
| Attendance (%)          |                    |                  |            |      |    |      |    |          |
| Attendance before midterm | 90.9 14.5          | 90.6 14.0        | 0.25       | (0.10) |
| Attendance after midterm | 87.9 16.8          | 86.7 16.9        | 1.16       | (−0.37) |
| Attendance during the semester | 88.9 13.5 | 88.2 14.4 | 0.70 | (0.27) |
| Academic performance    |                    |                  |            |      |    |      |    |          |
| Grade on midterm exam   | 80.8 13.0          | 79.7 14.3        | 1.12       | (0.44) |
| Grade on final exam     | 78.7 14.4          | 78.6 13.6        | 0.05       | (0.02) |
| Combined grade on two exams | 159.5 25.1      | 159.2 24.7       | 0.20       | (0.04) |
| Demographic factors     |                    |                  |            |      |    |      |    |          |
| Female                  | 0.351 0.481        | 0.407 0.495      | −0.056     | (−0.62) |
| African American        | 0.035 0.186        | 0.068 0.254      | −0.033     | (−0.79) |
| Asian                   | 0.228 0.423        | 0.135 0.345      | 0.093      | (1.29) |
| Latino                  | 0.053 0.225        | 0.102 0.305      | −0.049     | (−0.99) |
| White                   | 0.684 0.469        | 0.661 0.477      | 0.023      | (0.26) |
| Upper classman          | 0.596 0.495        | 0.339 0.477      | 0.157**    | (2.85) |
| International students  | 0.088 0.285        | 0.169 0.378      | −0.081     | (−1.32) |
| Transfer students       | 0.070 0.258        | 0.051 0.222      | 0.019      | (0.43) |
| Current academic factors|                    |                  |            |      |    |      |    |          |
| Concurrent classes      | 3.43 0.608         | 3.42 0.622       | 0.01       | (0.05) |
| Prior economic knowledge|                    |                  |            |      |    |      |    |          |
| Cumulative GPA          | 3.19 0.468         | 3.25 0.386       | −0.06      | (−1.02) |
| Number of past economic classes | 3.19 1.420      | 2.80 1.471     | 0.39       | (1.48) |
| Economic grades         | 3.11 0.667         | 3.24 0.498       | −0.13      | (−1.13) |
| Principles of macroeconomics grade | 3.14 0.751      | 3.29 0.619     | −0.15      | (−1.09) |

Notes: Each entry is the sample mean and standard deviation (or $t$ statistic) of the control, treatment, and difference between the two. The data are drawn from classroom and academic records. *$p = .10$. **$p = .05$. ***$p = .01$. 

The results for the difference column are from estimating equation (1) for each variable. The attendance and academic performance results for the two groups are quite close in mean values. In fact, the control and treatment groups are similar in most characteristics, but there are a few exceptions. First, there were more upper classman in the treatment group. Second, there were more African-American, Latino, and international students in the control group. Third, and potentially important, the students in the control group had higher values for the prior (or pretest) economic knowledge measures: cumulative GPA, economics grades, and principles of macroeconomics grade.
The lack of statistically significant difference in mean attendance and test scores between students in the control and treatment groups does not necessarily imply that cooperative learning had no effect on attendance and academic performance. Consider the difference in the prior economic knowledge measures. The lower value in prior economic knowledge in the treatment group is likely to be correlated with reduced attendance and test scores. If cooperative learning had a positive effect on the treatment group, then that positive effect could have been offset by the negative effect of the prior economic knowledge, resulting in an undistinguishable difference in the mean outcomes of the control and treatment groups.

To control for this possibility, I estimated the following multivariate regression equation:

$$outcome_i = \beta_0 + \beta_1 \text{ experimental section}_i + \sum_{j=2}^{k} \beta_j X_j + \epsilon_i.$$  \hspace{1cm} (2)

The independent variables $X$ measure classroom factors, demographic information, current academic factors, economic knowledge, and prior levels. I used two dummy variables for students in 2004 and in an early afternoon, smaller size class to capture classroom factors.\(^6\) For demographic information, I included dummy variables for students that were African-American, Asian, Latino, international, upper classman, and transfer. I included the number of concurrent classes and total attendance for current academic factors. For prior economic knowledge, I used cumulative grade point average (GPA), the number of past economic courses, grades of past economic courses, and the grade for principles of macroeconomics. I also included the preexperiment values for all preparation questions for class attendance and the cumulative GPA for academic performance to measure the prior or pretest values of attendance and knowledge.\(^7\)

For the OLS estimates for attendance and academic performance (Table 3), I considered three different specifications where prior economic knowledge was measured as: (1) cumulative GPA, (2) number and grades of past economic courses, and (3) grade for principles of macroeconomics. The $R$-squared indicate that my model explained around 50 percent of differences in class attendance and 70 percent of the variation in exam scores. For classroom factors, students in 2004 attended class more often. For demographic factors, Latino students did better on the exams, and African-American and transfer students did worse. For current academic factors, the number of concurrent classes was insignificant whereas attendance was positively linked to exam performance. All measures of prior economic knowledge were positively linked to attendance and exam performance. For attendance, three of the preexperiment questionnaire coefficients were positive and two were negative—suggesting that work outside of class served more as a complement rather than a substitute to attending class.

Turning to the variable of interest, I found that cooperative learning was positively related to exam scores. For attendance, the coefficient for the experimental section was insignificant under all three specifications. For academic performance, however, the coefficient for the experimental section was positive and significant in all three instances. The point estimates implied that the use of cooperative learning
### TABLE 3. Determination of Student Attendance and Academic Performance

| Variable                      | Attendance (1) | Attendance (2) | Attendance (3) | Exam scores (1) | Exam scores (2) | Exam scores (3) |
|-------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Cooperative learning          |                |                |                |                |                |
| Experimental section (1 = yes, 0 = no) | −0.92          | −1.79          | −0.31          | 4.37*          | 4.98*          | 5.47*          |
| Classroom factors             |                |                |                |                |                |
| Year (1 = 2004, 0 = 2002)     | 7.89**         | 8.87**         | 10.21**        | −5.02          | −2.99          | −0.74          |
| Class time/Class size (1 = early/large, 0 = late/small) | −2.50          | −3.35          | −1.93          | −0.62          | 1.18           | −1.20          |
| Demographic factors           |                |                |                |                |                |
| Gender (1 = female, 0 = male) | 1.19           | 2.62           | 0.92           | 4.50           | 3.26           | 0.91           |
| African American (1 = yes, 0 = no) | 0.03           | 2.14           | −1.95          | −11.58**       | −10.73**       | −11.73**       |
| Latino (1 = yes, 0 = no)      | −0.92          | −1.36          | −0.40          | −2.03          | 0.02           | −0.21          |
| International student (1 = yes, 0 = no) | 4.26           | 3.61           | 0.87           | 9.14*          | 11.70**        | 6.44           |
| Class (1 = upper, 0 = lower)  | −2.51          | −0.86          | −4.16          | −0.43          | −1.60          | −1.26          |

*Table continues*
|                          |        |        |        |        |        |
|--------------------------|--------|--------|--------|--------|--------|
| **Transfer student**     | −0.59  | 1.29   | 4.13   | −13.52*** | −14.13*** | −7.59*  |
| (1 = yes, 0 = no)        | (−0.17)| (0.42) | (0.88) | (−3.29) | (−3.81) | (−1.72) |
| **Current academic factors** |        |        |        |        |        |
| Concurrent classes       | 0.53   | 0.89   | 1.69   | 0.34   | 2.54   | 2.43    |
| (1−5)                    | (0.37) | (0.57) | (1.02) | (0.15) | (1.10) | (0.91)  |
| Attendance               | 0.52***| 0.46***| 0.47***|        |        |        |
| (0−100 percent)          | (4.01) | (3.85) | (3.52) |        |        |        |
| **Prior economic knowledge** |        |        |        |        |        |
| Cumulative GPA           | 13.88***| 34.97***| 12.90**| 25.71***|        |        |
| (0−4 scale)              | (3.79) | (6.76) | (2.20) | (3.72) |        |        |
| Past economic courses    | −1.36  |        | 2.62** |        |        |        |
| (0−7)                    | (−1.37)|        | (2.39) |        |        |        |
| Economic grades          | 8.49***|        | 20.38***|        |        |        |
| (0−4 scale)              | (3.67) |        | (4.46) |        |        |        |
| Principles of macro grade|        |        |        | 5.35***|        | 8.52***|
| (0−4 scale)              |        |        |        | (3.03) |        | (2.83) |
| **Prior attendance**     |        |        |        |        |        |
| I attend class regularly | 6.04***| 6.64***| 7.64***|        |        |        |
| (1−5 scale)              | (2.71) | (3.34) | (3.63) |        |        |        |
| I read before class      | 4.17***| 4.05***| 3.41*  |        |        |        |
| (1−5 scale)              | (3.19) | (2.74) | (1.88) |        |        |        |
| I study before class     | −5.63***| −5.59***| −3.81**|        |        |        |
| (1−5 scale)              | (−3.73)| (−3.66)| (−2.09)|        |        |        |
| I discuss with other students | 3.87***| 4.18***| 3.44** |        |        |        |
| (1−5 scale)              | (3.09) | (3.02) | (2.43) |        |        |        |

(Table continues)
| Variable                                    | Attendance (1) | Attendance (2) | Attendance (3) | Exam scores (1) | Exam scores (2) | Exam scores (3) |
|--------------------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| I prefer to work alone (1–5 scale)         | 0.58           | 0.33           | 0.69           |                |                |                |
| I seek help from other students (1–5 scale)| −1.92          | −3.21**        | −2.73*         |                |                |                |
| I seek help from instructor (1–5 scale)    | −2.54          | −1.91          | −2.40          |                |                |                |
| Regression diagnostics                      |                |                |                |                |                |                |
| $R^2$ squared                               | 0.549          | 0.551          | 0.522          | 0.719          | 0.782          | 0.748          |
| Adjusted $R^2$ squared                      | 0.441          | 0.434          | 0.392          | 0.680          | 0.747          | 0.705          |
| Sample size                                | 100            | 98             | 90             | 109            | 107            | 99             |

Notes: Equation (2) is estimated using OLS where the $t$ statistics are in parentheses. The standard errors used to compute the $t$ statistics have been corrected for heteroscedasticity. $^* p = .10$, $^{**} p = .05$, $^{***} p = .01$. 
was associated with an increase in combined exam scores of 4.4 to 5.5 points. As a percentage of the mean score on the exams, the point estimates translated into a 5.5 to 7.0 percent improvement in exam performance.

Next, I checked the sensitivity of the results to potentially influential outliers. I ran the second specification and excluded either (1) African American, (2) Asian, (3) Latino, (4) international, (5) freshman, or (6) transfer students. The results for the experimental dummy are provided in Table 4. In each instance but one, the coefficient for cooperative learning was positive and significant at the .10 Type I error level. The point estimates ranged from a low of 4.4 (without Latino students) to a high of 6.7 (without Asian students). Although the results suggest that the impact of cooperative learning may differ across demographic and ethnic lines, the positive coefficients argue that cooperative learning is associated with an increase in exam scores.

There is potential endogeneity of attendance in the academic performance regressions. Romer (1993) recognized that student attendance is not exogenous because students “choose whether to attend class” (p. 170). To control for the endogeneity of attendance, Schmidt (1983) used a latent variable approach where time and ability are treated as unobservable variables in the academic performance equation. Devadoss and Poltz (1996), on the other hand, posited a recursive system and then used a seemingly unrelated regression estimator.

I tested for the effects of cooperative learning on exam scores with attendance treated as endogenous in Table 5. I used a two-stage least squares (2SLS) estimator where attendance was instrumented with the answers to the preexperiment questions #8–14 and the other explanatory variables (included in the second stage). I considered the three earlier specifications in columns (1) to (3) and a preferred specification in column (4). I obtained the preferred specification by using the automated Hendry/LSE general-to-specific model selection criteria in PcGets (see Table 4. Sensitivity Analysis for Exam Scores

| Sample                  | Coefficient for experimental | t statistic for experimental | p value for experimental | Adjusted R squared | Observations |
|-------------------------|------------------------------|------------------------------|--------------------------|--------------------|--------------|
| Excluding African       | 5.61                         | 2.03**                       | 0.045                    | 0.727              | 101          |
| Americans               |                              |                              |                          |                    |              |
| Excluding Asians        | 6.73                         | 2.41**                       | 0.018                    | 0.764              | 87           |
| Excluding Latino        | 4.41                         | 1.59                         | 0.115                    | 0.752              | 98           |
| Excluding international | 6.73                         | 2.43**                       | 0.017                    | 0.742              | 95           |
| Excluding freshmen      | 5.26                         | 1.87*                        | 0.065                    | 0.742              | 102          |
| Excluding transfers     | 5.05                         | 1.82*                        | 0.073                    | 0.747              | 101          |

Notes: Equation (2) is estimated using OLS using cumulative GPA, past economics classes, and economics grade as measures of prior economic knowledge. Each row reports the results of the experimental variables for each sample. The standard errors used to compute the t statistics have been corrected for heteroscedasticity. *p = .10. **p = .05. ***p = .01.
### TABLE 5. Determination of Exam Scores with Endogenous Attendance

| Variable                          | (1)     | (2)     | (3)     | (4)     |
|-----------------------------------|---------|---------|---------|---------|
| **Cooperative learning**          |         |         |         |         |
| Experimental section             | 4.87**  | 5.67**  | 6.16*** | 5.63**  |
| (1 = yes, 0 = no)                 | (1.89)  | (2.32)  | (2.42)  | (2.45)  |
| **Classroom factors**             |         |         |         |         |
| Year                              | −3.59   | −0.94   | 1.78    |         |
| (1 = 2004, 0 = 2002)              | (−1.09) | (−0.30) | (0.56)  |         |
| Class time/Class size             | 0.70    | 2.16    | 0.36    |         |
| (1 = early/large, 0 = late/small) | (0.27)  | (0.91)  | (0.15)  |         |
| **Demographic factors**           |         |         |         |         |
| Gender                            | 2.31    | 1.24    | −1.47   |         |
| (1 = female, 0 = male)            | (0.82)  | (0.48)  | (−0.54) |         |
| African-American                  | −16.09*** | −16.43*** | −17.33*** | −15.95*** |
| (1 = yes, 0 = no)                 | (−3.05) | (−4.17) | (−4.14) | (−4.01) |
| Asian                             | −2.37   | −1.42   | 0.30    |         |
| (1 = yes, 0 = no)                 | (−0.68) | (−0.38) | (0.09)  |         |
| Latino                            | 8.67**  | 11.37*** | 4.70    | 11.95*** |
| (1 = yes, 0 = no)                 | (2.02)  | (2.63)  | (1.01)  | (2.87)  |
| International student             | −2.99   | −3.04   | 5.99    |         |
| (1 = yes, 0 = no)                 | (−0.62) | (−0.65) | (1.55)  |         |
| Class                             | −4.86   | −6.38*  | −6.93** | −6.31** |
| (1 = upper, 0 = lower)            | (−1.36) | (−1.79) | (−2.03) | (−2.21) |
| Transfer student                  | −12.52*** | −13.31*** | −4.62 | −13.63*** |
| (1 = yes, 0 = no)                 | (−3.00) | (−3.84) | (−0.86) | (−3.52) |
| **Current academic factors**      |         |         |         |         |
| Concurrent classes                | 1.30    | 3.33*   | 3.71*   | 3.50*   |
| (1–5)                             | (0.64)  | (1.61)  | (1.68)  | (1.72)  |
| Attendance                        | 0.61*** | 0.43*** | 0.58*** | 0.40*** |
| (0–100 percent)                   | (3.42)  | (2.74)  | (3.43)  | (2.73)  |
| Prior economic knowledge          |         |         |         |         |
| Cumulative GPA                    | 31.03*** | 13.65** | 17.38*** | 14.69*** |
| (0–4 scale)                       | (5.89)  | (2.08)  | (2.75)  | (2.17)  |
| Past economic courses             | 2.77*** | (2.78)** | (2.32)  | (2.30)  |
| (0–7)                             |         |         |         |         |
| Economic grades                   | 18.44*** | 18.34*** | (5.00) | (4.71)  |
| (0–4 scale)                       |         |         |         |         |
| Principles of Macro grade         |         |         | 9.96*** |         |
| (0–4 scale)                       |         |         | (3.42)  |         |
| **Regression diagnostics**        |         |         |         |         |
| R squared                         | 0.705   | 0.772   | 0.763   | 0.766   |
| Partial R squared                 | 0.351   | 0.361   | 0.327   | 0.401   |
| Hansen-Sargan test                | 0.672   | 0.580   | 0.153   | 0.713   |
| **Sample size**                   | 98      | 96      | 88      | 96      |

Notes: Equation (2) is estimated using 2SLS where the results of the second-stage are presented above. The answers to the preexperiment questions #8–14 and all other explanatory variables are used as instruments in the first-stage regression for attendance. The t statistics are in parentheses and have been corrected for heteroscedasticity. The R squared is the goodness of fit for the second-stage regression. The partial R squared is the Shea (1997) partial R squared of excluded instruments in the first stage. The Hansen-Sargan test is the p value of a test for valid instruments. *p = .10, **p = .05, ***p = .01.
Hendry and Krolzig 2001). The relatively high partial R squared of excluded instruments indicated that the instruments were relevant (i.e., highly correlated with attendance). Moreover, the results for the Hansen-Sargan test of overidentifying restrictions showed that the instruments were valid (i.e., uncorrelated with the error term and correctly specified).8

Results shown in Table 5 confirm the results found earlier. Latino and lower-level students scored higher on the exams whereas Africa American and transfer students scored lower. Attendance and prior economic knowledge was still positively correlated with exam scores. Most important, though, the coefficient for the experimental section was positive and significant. The point estimates again suggested that, all else equal, students in the experimental group scored five to six points higher on the combined exams than those in the control group.

Discussion

Why did students in the cooperative learning group perform better on the exams than students in the lecture group? I discuss three possible reasons for this result. First, cooperative learning raised student-instructor interaction. Students seemed less inhibited about asking questions in the small groups. I observed that students in the cooperative learning class (when working on assignments in groups) asked questions more often than students in the lecture class, even though I frequently asked for questions during lecture. Furthermore, students in the cooperative learning group often came to my office as a group whereas students in the experimental group came individually. As a result, more students from the cooperative learning sections sought help from me outside of class.

Second, cooperative learning increased group studying for the exams. Students in the cooperative learning sections were more likely to develop study groups. These study groups were parts of or the entire base groups used in the classroom. Furthermore, the rapport inside the study groups for the experimental group appeared better than that for control group. As one student stated, “I liked the use of groups in class because it gave me students to study with for the exams.”

Third, the novelty of working in small groups sparked greater interest in the material. In end-of-course evaluations, several students expressed positive attitudes toward cooperative learning. One student said that “[cooperative learning] was a great idea because it allowed me to learn the material from both the instructor and the other students.” Another student found it “very worthwhile and helpful . . . I learned the material much better by discussing it with my fellow group members.” However, the three reasons are mostly speculative and, as a result, further study is needed to decipher why group work raised student achievement.

CONCLUSION

My objective was to investigate the impact of cooperative learning in economic instruction. I used a two-group experimental design where one section of intermediate macroeconomics was taught using cooperative learning and the other section was taught using a traditional lecture format. I tracked five types of student learning
outcomes: interest, preparation, participation, attendance, and performance. Using multivariate regression analysis, I found that the experimental section scored four to six points higher on the combined exams when I controlled for classroom, demographic, and academic factors.

NOTES
1. The use of base groups inside and outside the classroom by the instructor is quite common in other disciplines. Examples include Kogut (1997) for chemistry, Davidson (1990) for mathematics, and Felder (1995) and Felder, Felder, and Dietz (1998) for engineering.
2. The exams in each class had identical short-answer questions and nearly identical problem questions, except that one class would have an increase in a specific exogenous variable, where as the other class would have a decrease in the same exogenous variable.
3. There are two time issues in the experiment: use of the problem-solving exercises and use of cooperative learning. By using the problem-solving exercises, I lost 2 classes (out of a total of 25) of material. For cooperative learning, the traditional lecture section typically finished the handout 5–10 minutes before the end of the class period, whereas the experimental section took the entire class period. I spent the last 5–10 minutes of the traditional lecture section elaborating on the issues raised and answering additional questions. I would like to thank one of the referees for alerting me to this distinction.
4. Student learning outcomes are measurable goals and results that are expected to occur from an academic experience. These outcomes can be classified as knowledge (cognitive), skills (behavioral), or attitudes (affective) gained from an individual class or academic program (Gronlund 2002).
5. These results are available upon request from the author.
6. Each early afternoon section had more students than the corresponding late afternoon section. As a result, I was unable to disentangle the effects of the class time and class size.
7. The independent variables included followed the previous work of Romer (1993), Borg and Shapiro (1996), Devadoss and Foltz (1996), Johnston et al. (2000), Ziegert (2000), and Borg and Stranahan (2002). I also included dummy variables for Asian students and economics majors and found them never to be significant.
8. The Shea (1997) partial R squared of excluded instruments records the increase in the goodness of fit from using those variables not included in the second stage. The partial R squared measures the correlation between the excluded instruments and the endogenous regressor. The Hansen-Sargan test of overidentifying restrictions has a joint null hypothesis that the instruments are uncorrelated with the error term and that the excluded instruments are correctly excluded from the estimated equation.

REFERENCES
Barkley, E. F., K. P. Cross, and C. H. Major. 2005. Collaborative learning techniques: A handbook for college faculty. San Francisco, CA: Jossey-Bass.
Becker, W. E., and M. Watts. 2001a. Teaching methods in U.S. undergraduate economics courses. Journal of Economic Education 32 (Summer): 269–80.
———. 2001b. Teaching economics at the start of the 21st century: Still chalk-and-talk. American Economic Review Papers and Proceedings 91 (2): 446–51.
Benzing, C., and P. Christ. 1997. A survey of teaching methods among economics faculty. Journal of Economic Education 28 (Spring): 182–88.
Borg, M. O., and S. L. Shapiro. 1996. Personality type and student performance in principles of economics. Journal of Economic Education 27 (Winter): 3–25.
Borg, M. O., and H. A. Stranahan. 2002. Personality type and student performance in upper-level economics courses: The importance of race and gender. Journal of Economic Education 33 (Winter): 3–14.
Brawner, C. E., R. M. Felder, R. Allen, and R. Brent. 2002. A survey of faculty teaching practices and involvement in faculty development activities. Journal of Engineering Education 91 (4): 393–96.
Brooks, T. B., and A. W. Khandker. 2002. A cooperative learning lab: Does the form matter? Contemporary Economic Policy 20 (3): 330–38.
Bruffee, K. A. 1995. Sharing my toys: Cooperative versus collaborative learning. Change 27 (1): 12–18.
Cross, K. P., and M. H. Steadman. 1996. Classroom research: Implementing the scholarship of teaching. San Francisco, CA: Jossey-Bass.
Davidson, N. 1990. *Cooperative learning in mathematics: A handbook for teachers*. Menlo Park, CA: Addison-Wesley.

DeVadoss, S., and J. Foltz. 1996. Evaluation of factors influencing student class attendance and performance. *American Journal of Agricultural Economics* 78 (3): 499–507.

Felder, R. M. 1995. A longitudinal study of engineering student performance and retention. IV. Instructional methods and student responses to them. *Journal of Engineering Education* 84 (4): 361–67.

Felder, R. M., G. N. Felder, and E. J. Dietz. 1998. A longitudinal study of engineering student performance and retention. V. Comparisons with traditionally-taught students. *Journal of Engineering Education* 87 (4): 469–80.

Gronlund, N. E. 2002. *Assessment of student achievement*. Boston, MA: Allyn & Bacon.

Hendry, D. F., and H. M. Krolzig. 2001. *Automatic econometric model selection using PcGets*. London: Timberlake Consultants.

Jensen, E. J., and A. L. Owen. 2001. Pedagogy, gender, and interest in economics. *Journal of Economic Education* 34 (Fall): 323–43.

Johnson, D. W., and R. T. Johnson. 1987. *Learning together and alone*. Englewood Cliffs, NJ: Prentice Hall.

——— 1999. Making cooperative learning work. *Theory into Practice* 38 (2): 67–73.

Johnson, D. W., R. T. Johnson, and K. A. Smith. 1991. *Cooperative learning: Increasing college faculty instructional productivity*. ASHE-ERIC Higher Education Report No. 4. Washington, DC: George Washington University.

——— 1998. *Active learning: Cooperation in the college classroom*. Edina, MN: Interaction Book Company.

Johnston, C. G. 1997. Cooperative learning of economics in a peer tutoring context at the University of Melbourne. *Australian Economic Papers* 0 (special edition): 56–68.

Johnston, C. G., R. H. James, J. N. Lye, and I. M. McDonald. 2000. An evaluation of collaborative problem solving for learning economics. *Journal of Economic Education* 31 (Winter): 13–29.

Kogut, L. S. 1997. Using cooperative learning to enhance performance in general chemistry. *Journal of Chemistry Education* 74 (6): 720–24.

Lindholm, J. A., A. W. Astin, L. J. Sax, and W. S. Korn. 2001. *The American college teacher: national norms for the 2001–2002 HERI faculty survey*. Los Angeles: Higher Education Research Institute, UCLA.

Mankiw, N. G. 2003. *Macroeconomics*. 5th ed. New York: Worth.

McKelvey, R. D., and W. Zavoina. 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* 4 (2): 103–20.

Meyer, B. 1995. Natural and quasi-natural experiments in economics. *Journal of Business and Economics Statistics* 13 (2): 151–61.

Moore, R. L. 1998. Teaching introductory economics with a cooperative learning lab. *Journal of Economic Education* 29 (Fall): 321–29.

Romer, D. 1993. Do students go to class? Should they? *Journal of Economic Perspectives* 7 (Fall): 167–74.

Sax, L. J., A. W. Astin, W. S. Korn, and S. K. Gilmartin. 1999. *The American college teacher: National norms for the 1998–1999 HERI faculty survey*. Los Angeles: Higher Education Research Institute, UCLA.

Schmidt, R. M. 1983. Who maximizes what? A study in student time allocation. *American Economic Review Papers and Proceedings* 73 (2): 23–28.

Shea, J. 1997. Instrument relevance in multivariate linear models: A simple measure. *Review of Economics and Statistics* 79 (2): 348–52.

Slavin, R. E. 1995. *Cooperative learning: Theory, research, and practice*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall.

Springer, L., M. E. Stanne, and S. S. Donavan. 1999. Undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research* 69 (1): 21–51.

Ziegert, A. L. 2000. The role of personality temperament and student learning in principles of economics: Further evidence. *Journal of Economic Education* 31 (Fall): 307–22.