Research in Economic Education

In this section, the *Journal of Economic Education* publishes original theoretical and empirical studies of economic education dealing with the analysis and evaluation of teaching methods, learning, attitudes and interests, materials, or processes.

PETER KENNEDY, Section Editor

Can We Control Cheating in the Classroom?

Joe Kerkvliet and Charles L. Sigmund

A great deal of research reflects concern for the large and growing problem of cheating in academe (Maramark and Maline 1993; Collison 1990). The evidence indicates that many students cheat regularly and few students never cheat.

Most of this research is of little use to the university teacher because it does not inform a teacher about the extent of cheating in the class; it considers students' cheating over their entire university career (Kerkvliet 1994). Teachers use different classroom procedures and make varying demands on their students. Reacting to these differences, students may cheat more with some teachers than with others.

Furthermore, the extant research rarely examines the effectiveness of the measures that many teachers take to discourage cheating. Some authors exhort faculty to increase their deterrent efforts (Stevens and Stevens 1987; Singhal 1982; Singhal and Johnson 1983; Collison 1990). Others say that students' behavior mirrors a general decline in public morality (Fass 1990). Against this societal malaise, measures to discourage cheating are said to be ineffective, or worse, they may actually encourage cheating (Moffatt 1990; Michaels and Miethe 1989; Eble 1988).

Lacking empirical guidance, teachers almost haphazardly use different deterrent measures to counter the plethora of cheating methods students have invented (Stevens and Stevens 1989; Garred et al. 1991). These measures include sanctions, threats, classroom surveillance, multiple exam versions, and avoiding the

Joe Kerkvliet *is an associate professor of economics at Oregon State University (e-mail: kerkvlij @ucs.orst.edu).* ***Charles L. Sigmund*** *is with the Office for Oregon Health Plan Policy and Research.*

The authors thank the survey respondents. Clifford Nowell, Cassandra Manuelito-Kerkvliet, Art O'Sullivan, Lani Roberts, Carol Horton Tremblay, and Vic Tremblay provided helpful comments.

use of multiple-choice questions. Research would be more useful if it were directed toward knowing the efficacy of deterrent measures and guiding teachers in deciding which measures should receive scarce time and resources.

In this article, we explore the determinants of class-specific cheating behavior, including students' characteristics and behavior and deterrent measures. We present evidence on the level of cheating on examinations in 12 separate principles of economics classes and show that the level of cheating varies markedly from class to class. We test the claim that deterrent measures are ineffective and teachers have little choice but to accept passively classroom cheating. Finally, we present evidence on the relative effectiveness of different deterrent measures.

A MODEL OF ACADEMIC CHEATING BEHAVIOR

Many see academic dishonesty in the same light as other criminal activities (Tittle and Rowe 1974; Michaels and Miethe 1989; Bunn, Caudill, and Gropper 1992; Kerkvliet 1994; Mixon 1996). Becker (1968) proposed the first widely accepted economic model of crime, theorizing that the criminal is a rational decisionmaker. She weighs the expected benefits and expected costs of committing a crime and makes utility maximizing choices.

Bunn et al. (1992) and Kerkvliet (1994) adapted Becker's model to academic dishonesty. Straightforward extensions of the latter allow inclusion of measures used by university teachers to influence cheating behavior, with the resulting three predictions. First, an increase in the benefits associated with undetected cheating will encourage students to cheat; or, conversely, an increase in the benefits of not cheating will decrease the probability of cheating. Second, increases in the severity of penalties imposed on students who are detected cheating will decrease cheating. Finally, measures that increase the probability of detection will discourage cheating.

These predictions imply that the determinants of cheating include student characteristics (e.g., gender) or choices (e.g., study time). These determinants have been extensively examined in previous research. More important, another type of determinant reflects the choices of teachers and administrators. These choices affect the benefits and costs of cheating and detection probabilities and have been largely ignored in prior research.

GATHERING DATA ON ACADEMIC CHEATING

Cheating of any kind is sensitive behavior. Because people consider questions about sensitive behavior as threatening and/or as having socially undesirable answers, they may not respond honestly to salient questions about cheating (Sudman and Bradburn 1974). Because of the locale and time frame involved, queries about class-specific academic cheating are more threatening than questions about past cheating in any class; that is, the class in which the cheating occurred is explicit, so students are less able to distance themselves from their responses. Despite assurances of confidentiality, students may fear reprisals from teachers, administrators, or both. Moreover, respondents consider questions about current behavior more threatening than questions about past behavior (Sudman and Bradburn 1982).

Four methods have been applied to the gathering of information on academic cheating: (1) direct but surreptitious observation, (2) the error overlap method, (3) direct questions, and (4) and randomized response questions.

The observation method presents students with an opportunity to cheat while the researcher uses a surreptitious means of directly detecting cheating. For example, Tittle and Rowe (1974) used weekly quizzes that instructors collected, graded but did not mark, and then returned to students for self-grading. Students failing to mark incorrect answers as wrong were counted as cheaters. Using this method, Tittle and Rowe reported a .33 probability of a student cheating at each opportunity. Gardner et al. (1988) and Nowell and Laufer (1997) reported probabilities of .51 and .23, respectively.

The surreptitious method has the advantage of not relying on students' candidness. It is also well suited to focusing on class-specific cheating. Its disadvantage lies in placing students in a contrived situation that poorly mimics the typical examination environment. In addition, it can only be used on objectively scored tests. Further, it may have the appearance of entrapment.

The error overlap method tries to detect cheating by statistically comparing the incorrect answers of adjacent students with the incorrect answers of randomly chosen, nonadjacent students (Houston 1976a,1976b, 1983). This method is well suited to detecting class-specific cheating. It can be adapted to typical testing environments and does not rely on students' candidness. However, it only applies to objective tests and to one form of cheating—copying a neighbors' answers. Using this method, Houston (1976) found evidence of significant copying.

The most widely used method of data gathering is the direct question survey (DQS) method (see Michaels and Miethe 1989; Kerkvliet 1994). DQS does not account for the problems inherent in asking threatening questions, and evidence suggests that it leads to biased estimates of cheating (Kerkvliet 1994; Scheers and Dayton 1987). We are not aware of any DQS studies concerned with the question of class-specific cheating.

The randomized response survey (RRS) uses questions designed to reduce the problems of nonresponse or biased response to sensitive questions. The RRS method provides respondents with more anonymity by allowing them to answer a sensitive question without revealing with certainty their true status regarding the sensitive behavior. Anonymity is expected to encourage truthful responses (Chaudhuri and Mukerjee 1988). Three studies of academic misconduct compare DQS and RRS results. Scheers and Dayton (1987) and Kerkvliet (1994) found that DQS methods underestimate cheating. Nelson and Schaefer (1986) found opposing evidence.

Because our study focused on class-specific cheating, the questions asked were especially threatening. In addition, we sought evidence on the effectiveness of deterrent measures used in university test settings. The estimation method should also control for the effects of the students' characteristics. These goals led us to rely heavily on the RRS method to obtain information on cheating. For comparison, we placed one quarter of our respondents into a DQS control group.

THE EMPIRICAL MODEL AND THE DATA

The RRS and DQS questions about cheating that we used here were the same form as that used in Kerkvliet (1994), except that the cheating question was about in-class cheating. Surveys were given to 597 students at two public universities, in 12 separate classes, taught by seven different instructors in the 1993–1994 academic year. Of the 586 surveys returned (98.2 percent), 551 (92.3 percent) contained a usable response to the question of in-class cheating. For the 433 RR surveys, 398 (91.9 percent) contained a useable response, and all of the DQ surveys did so. However, only 3 (1.9 percent) of the 153 DQ respondents admitted to cheating. This small number made empirical estimation using DQS data impossible. Our estimates were based on the 393 (91 percent) RR surveys that contained responses to the cheating question and complete information on the other determinants of cheating.

The randomized-response cheating question invited a binary response from the *i*th student, denoted here as $y_i \,\varepsilon(0, 1)$. Students computed a random number equal to the sum of the last four digits of their Social Security numbers. On the basis of this random number, respondents were instructed to give an arbitrary 0 response with probability ρ_1 or an arbitrary 1 response with probability ρ_2. With probability $(1 - \rho_1 - \rho_2)$, respondents were instructed to give a 1 response if they had *ever* cheated on an examination in the class in which they were surveyed or an 0 response if they had *never* cheated in the class. The empirical ρ_i used for estimation was calculated from the students' actual Social Security numbers obtained from class rosters. The averages across all classes for the empirical ρ_j used were .142 and .551 for j = 1, 2.

The logit functional form for the probability of cheating at least once in a class, yielded the following likelihood of observing the data:

$$L = \prod_{y_i=1}[\rho_1 + (1-\rho_1-\rho_2)\frac{e^{\beta' X_{ik}}}{1+e^{\beta' X_{ik}}}]\prod_{y_i=0}[\rho_2 + (1-\rho_1-\rho_2)\frac{1}{1+e^{\beta' X_{ik}}}], \qquad (1)$$

where the vector $\boldsymbol{X_{ik}}$ consists of determinants of cheating for the *i*th student in the *k*th class, and the vector β contains unknown parameters to be estimated. For estimation, we specified

$$\begin{aligned}\beta' X_{ik} = {} & \beta_0 + \beta_1 NTESTS_{ik} + \beta_2 GRADE_{ik} + \beta_3 MULTICH_{ik} + \beta_5 CLA_{ik} \\ & + \beta_6 WARNING_{ik} + \beta_7 STUDY_{ik} + \beta_8 GTA_{ik} + \beta_{10} SQFOOT_{ik} \\ & + \beta_{11} VERSIONS_{ik} + \beta_{12} GENDER_{ik} + \beta_{14} GPA_{ik} + \beta_{15} YEAR_{ik}.\end{aligned} \qquad (2)$$

The first variable, *NTESTS*, controlled for the number of tests given in a class prior to the survey.[1] The effect of *NTESTS* on the probability of cheating at least once was ambiguous. On the one hand, increasing the number of exams increases the number of opportunities for cheating and may increase the probability of cheating at least once. On the other hand, increasing the number of exams decreases the importance of each exam which reduces the incentive to cheat on any given exam.

The variables *GRADE*, *CREDITHR*, and *CLA* measure the benefits of undetected cheating. *GRADE* was the student's expected grade in the class (A = 4, B = 3, ..., F = 0). Students with lower grades have more to gain from successful cheating, especially if they face the possiblity of receiving or not receiving credit for the class. The expected effect for *GRADE* was negative. *CREDITHR* is the number of credit hours the student is taking during the quarter. Relative opportunity cost argues for including this variable as a measure of the benefits of undetected cheating. Students carrying more credit hours face a higher opportunity cost for studying for an economics exam. This implies that students with more credit hours will be less prepared for exams and see higher benefits to cheating, *ceteris paribus*. *CLA* equaled 1 for students in the colleges of liberal arts and 0 otherwise. For both universities, the college of liberal arts had the lowest grade requirements of all the colleges. For example, in one of the universities, over half of the respondents were from the college of business, which requires students to receive a C+ or better in principles of economics to remain a business major. In contrast, liberal arts students need only attain a D or better to receive full credit for the course and fulfill college requirements. We expected that CLA students would cheat less than students in other colleges. Moffatt (1986) and Garred et al. (1991) found substantive differences in cheating behavior among students in different university majors and colleges.

To measure the cost of detected cheating, we used *GTA* (graduate teaching assistant) and *WARNING*. *GTA* equaled 1 or 0 depending on whether the instructor was a graduate teaching assistant or tenure-track faculty member. There are three reasons for expecting students to cheat more when *GTA* = 1. First, GTAs may be less skilled at increasing the expected costs of cheating than more experienced faculty. Second, GTAs are usually only a few years older than undergraduates and have only recently received the baccalaureate. As such, GTAs are more likely to share undergraduates' casual attitudes toward, and rationalizations of, academic dishonesty (Haines et al. 1986). Third, GTAs will be more reluctant to initiate a university's due process procedures by confronting a suspected cheater. The procedures are complicated, time-consuming, and often adversarial. They are daunting even to experienced faculty (Fass 1990). GTAs are less likely to understand the procedures and may see less of a chance of successfully penalizing cheaters. Exploring a related issue, Nowell and Laufer (1997) found more cheating with adjunct than with tenure-track faculty.

WARNING indicates whether teachers verbally warned students before exams that cheating was proscribed and actively discouraged. All teachers were required to include the universities' official statement proscribing cheating on their syllabuses. When WARNING = 1, students were given additional verbal reminders about cheating before each exam. When WARNING = 0, no additional mention was made. Tittle and Rowe (1974) found that verbal threats reduced cheating. Houston (1983) found that threats only inhibited cheating by the better students. We expected WARNING = 1 to discourage cheating.

The benefits of not cheating were measured by *STUDY*, the average hours per week spent studying for the class. More hours of *STUDY* provide better command of the class material at test time and increase the benefits of not cheating.

We expected cheating to decrease with *STUDY*.

PROCTOR, *SQFOOT*, *VERSIONS*, and *MULTICH* each influence the probability of detection. *PROCTOR* was the number of test proctors per student. A higher *PROCTOR* increases the probability of detection and inversely influences cheating. *SQFOOT* was square feet of classroom space per student. It measured the instructor's ability to spatially separate students during exams. More separation increases the probability that the copying student will be detected and negatively influences the probability of cheating. Houston (1976) found that answer copying decreased as spacing increased. *VERSIONS* was the number of test versions used by the instructor. More *VERSIONS* eases the detection of answer copying, makes copying more difficult, and should be negatively relate to cheating. Houston (1976), however, reported that having more versions was an ineffective deterrent method. *MULTICH* equaled 1 if the class's exams included any multiple-choice questions, and zero otherwise. Because multiple choice answers are relatively easy to copy from another student, *MULTICH* was expected to have a negative impact on the probability of detection and a consequent positive influence the amount of cheating (Maramark and Maline 1993).

GENDER, *GPA*, *ALCOHOL*, and *YEAR* represented students' characteristics. *GENDER* was 1 for males and 0 for females. We did not hypothesize a sign for this variable. Baird (1980) and Stern and Havlicek (1986) found a positive gender effect, but Kerkvliet (1994) found the opposite. Haines et al. (1986) and Michaels and Miethe (1989) found no effect. *GPA* and *ALCOHOL* measured the student's level of commitment to education. More committed students are expected to cheat less, *ceteris paribus* (Haines et al. 1986). *GPA* was the student's self-reported cumulative grade point average. Baird (1980), Haines et al. (1986), Gardner et al. (1988), Moffatt (1990), and Bunn et al. (1992) all found a negative correlation between cheating and *GPA*. Conversely, Singhal (1982) and Kerkvliet (1994) reported no significant *GPA* effect. *ALCOHOL* was the average weekly consumption of alcoholic beverages in ounces. It related directly to the extent of "party" behavior and inversely to the student's commitment to education. As in Kerkvliet (1994), we expected *ALCOHOL* to have a positive effect on cheating. *YEAR* was the student's year in college (1 = freshman, . . . , 4 = senior, 5 = graduate). We did not hypothesize sign for this variable. Stern and Havlicek (1986) and Haines et al. (1986) found little relationship between *YEAR* and cheating, but Baird (1980), Moffatt (1990), and Michaels and Miethe (1989) reported that upperclassmen were more likely to cheat. Kerkvliet (1994) found the direction of influence differred, depending on whether DQS or RRS data were used.

Descriptive statistics for all of the variables are given in Table 1. Teachers supplied data on *MULTICH*, *WARNING*, *PROCTOR*, and *VERSIONS*. We used building blueprints and class rosters to construct *SQFOOT*. Respondents provided the other variables.

The surveyed classes were representative of the modes of teaching economics principles classes at public universities. Six of the classes were taught by GTAs, three by a single full professor, and three by two different associate professors.

TABLE 1
Variable Definitions and Descriptive Statistics

Variable	Definition	Mean	*Sd.*	Min.	Max.
NTESTS	Number of midterm exams	2.56	.791	1.00	4.00
GRADE	Expected class grade (A = 4, . . . , F = 0)	3.03	.828	0.00	4.00
MULTICH	1 if any multiple choice questions; 0 otherwise	.791	.407	0.00	1.00
CREDITHR	Number of credit hours taken by student	14.96	2.62	3.00	28.00
CLA	1 if student in liberal arts; 0 otherwise	.129	.334	0.00	1.00
WARNING	1 if additional warnings given; 0 otherwise	.377	.488	0.00	1.00
STUDY	Hours of weekly study for class	2.89	2.28	0.00	21.00
GTA	1 if graduate teaching assistant; 0 otherwise	.583	.494	0.00	1.00
PROCTOR	Test proctors per student	.0033	.0151	0.014	0.077
SQFOOT	Square feet of classroom space per student	23.64	10.73	9.19	47.03
VERSIONS	Number of exam versions	1.75	.874	1.00	3.00
GENDER	1 if male; 0 if female	.59	.49	0.00	1.00
ALCOHOL	Student's weekly consumption of alcoholic beverages in ounces	40.51	80.78	0.00	576.00
GPA	Student's grade point average	2.95	.469	1.57	4.00
YEAR	Year in school (1 = freshman, . . . , 4 = senior, 5 = graduate)	2.41	1.01	1.00	5.00

Class sizes ranged from 16 to 97 students and averaged 48.8.

A wide variety of testing techniques were used. At one extreme, three instructors relied exclusively on multiple-choice questions. At the other extreme, only short-answer and essay questions were used in four classes taught by a tenure-track faculty. Elsewhere, the instructor used multiple-choice, short-answer, essay questions or some combination of the three testing techniques.

Only the instructor proctored exams in half of the classes, whereas either one or two additional proctors were used in the other six. In seven classes, one test version was used; two or three versions were used in the other classes. Instructors using only multiple-choice questions invariably prepared two or three test versions. Those using only short-answer and essay questions invariably wrote a single version. Teachers in three classes issued verbal warnings The remainder only placed a statement of university policy in their syllabuses.

ESTIMATION AND RESULTS

We now report estimates of the effects of the independent variables on the probability of cheating. All estimated coefficients and their *t* ratios are in given Table 2.

TABLE 2
Parameter Estimates

Variable	Model 1 Estimate	Model 1 *t*	Model 2 Estimate	Model 2 *t*
CONSTANT	0.778	.26	–0.800	.41
NTESTS	1.039	1.41	1.471	2.59
Benefits of cheating				
GRADE	–0.037	.11	—	—
CREDITHR	0.176	1.56	0.166	1.52
CLA	–0.314	.34	—	—
Benefits of not cheating				
STUDY	0.021	.18	—	—
Cost of detecting cheating				
WARNING	–1.463	1.40	–1.370	2.69
GTA	3.455	2.94	3.662	3.47
Probability of detection				
PROCTOR	–70.130	2.44	–52.285	3.61
SQFOOT	0.011	.25	—	—
MULTICH	–1.090	.86	—	—
VERSIONS	–1.849	1.89	–2.888	2.91
Student's characteristics				
GENDER	0.089	.20	—	—
ALCOHOL	0.002	1.08	0.003	1.22
GPA	–1.547	2.78	–1.551	3.02
YEAR	0.310	1.28	0.333	1.49
χ^2	36.86		35.50	

The Determinants of Cheating

Model 1 gives the parameter estimates for equation (1).[2] The results indicate that a variety of variables shed light on class-specific cheating. We discuss the variables that do not appear to have an influence on cheating and then those that seem to have an effect, including some variables teachers can influence.[3]

As to the benefits of cheating, the estimated coefficients for *GRADE* and *CLA* were negative, as expected, but not significant at conventional levels. The benefits of not cheating, as measured by *STUDY*, did not affect cheating behavior.

Two variables that instructors commonly use to reduce cheating were not effective. *SQFOOT*'s coefficient was small and insignificant. However, this could have been a result of a failure of some teachers to use the space available to them for student separation.[4] Contrary to expectations, the coefficient for *MULTICH* was negative, but its *t* statistic was only .86. Finally, of the four student's characteristics, *GENDER* did not influence cheating.

In Model 2, we eliminated the six variables with coefficients insignificantly different from zero and lacking a strong rationale for inclusion (*GRADE*, *MULTICH*, *CLA*, *STUDY*, *SQFOOT*, *GENDER*). Model 2 was supported over Model 1 by a likelihood ratio test that gave an χ^2 statistic of 1.36; the critical value was 12.59. Model 2 revealed that all of the conclusions from Model 1 were robust and that the *t* ratios increased for all but one of the remaining coefficients.

Turning first to the *NTESTS* control variable, the coefficient was relatively large and highly significant, suggesting that students with more chances to cheat were more likely to cheat at least once. Regarding benefits, students with larger credit loads (*CREDITHR*) had a greater probability of cheating.[5] The *p* value for the *CREDITHR* coefficient was .06 (one-tail test). A decrease in the cost of detected cheating, as measured by *GTA*, had a large positive and very significant influence on the probability of cheating. Similarly, *WARNING* had the expected negative, and a highly significant, effect on the probability of cheating.

Instructors'use of both multiple test *VERSIONS* and more *PROCTOR*s per student seemed effective in reducing cheating by increasing the probability of detection. The coefficients for these two variables were both negative and statistically significant ($\alpha < .01$, one-tail tests).

Of students' characteristics, more *ALCOHOL* consumption had a weak positive association with more cheating. The coefficient's *p* value was .11 (one-tail test). Confirming many previous studies, we found that students with higher GPAs had a lower probability of cheating. The *GPA* coefficient was negative and highly significant. *YEAR*'s positive coefficient suggested that the further along a student is in his or her academic career, the more likely he or she is to cheat. This coefficient had a *p* value of .07 (one-tail test).

Class Differences

The second purpose of our research was to provide evidence on the class-to-class variation in cheating. Using the parameter estimates from Model 2, we obtained predicted probabilities of cheating for each respondent. The predictions suggested that care must be taken not to tar all university classes with the same brush when it comes to cheating. For our sample of 12 principles of economics classes, the extent of the problem differed by a factor of 160. Overall, the probability that a student cheated at least once averaged .128 (*s.d.* = .127), but this masked large differences among classes. In the worst case, the probability that a student cheated was .32. This class contained many heavy-drinking upperclassmen with low GPAs and was taught by a GTA who used two test versions but employed no other measures against cheating. The second highest probability (.19) was in a class taught by an associate professor, who took no precautions against cheating.

Two small classes (13 and 22 students) taught by a full professor had the lowest probabilities of cheating (.002 and .025). The third lowest probability (.028) was for a class taught by a GTA who issued warnings, gave few exams, used an additional proctor, and wrote three versions of each exam.

TABLE 3
Incremental Effects on Cheating Probabilities

Variable	Change in probability
WARNING	–.125
GTA	+.318
PROCTOR	–.111
VERSIONS	–.253

Are Control Measures Effective?

Our third purpose was to test the claim that teachers can do nothing to reduce cheating. To test the null hypothesis that none of the control measures influence cheating behavior, we restricted the coefficients for *WARNING*, *GTA*, *PROCTOR*, and *VERSIONS* to zero and re-estimated. With Model 2 as a reference, the likelihood ratio test that all controls are ineffective yielded a test statistic $\chi^2 = 21.93$. Because this exceeded the critical value, $\chi^2_{df=6} = 9.49$ ($\alpha = .05$), we rejected the null.

The rejection of the null indicates that cheating was partially under the control of the faculty. This still left open the choice of control measures. Our fourth purpose was to provide evidence on the relative effectiveness of deterrent measures.

Relative Effectiveness of Controls

Although widely used, we did not find that avoiding multiple-choice questions and increased spacing of students were effective deterrent measures. Multiple-choice questions conserve on teachers' time; they are easier to administer and grade than other questions. Yet, it is argued (e.g., Moffatt 1986; Eble 1988; Maramark and Maline 1993) that using multiple-choice examinations encourages cheating. The results from Model 1 did not support this argument. Also, our findings did not support the argument that cheating is reduced by increasing the physical separation of students during exams. The insignificant coefficient for *SQFOOT* could have arisen from the failure of teachers to use the space available to separate students. It is also likely that separation is ineffective against the myriad cheating devices used by students besides copying from their peers.

Turning to the control measures that seemed to be effective, we first compared the incremental effects on the probability of cheating. The discrete variables, *WARNING*, *GTA*, and *VERSIONS* were computed as the difference in the average probability when the variable increased by one unit. For comparability, we computed the incremental effect of *PROCTOR* by computing the estimated change in probability assuming that one additional proctor was used. For the average class of 48.8 students, this meant that *PROCTOR* increased by .0205. Estimated incremental effects are presented in Table 3.

The most striking result was the incremental effect of using GTAs as instructors. Students taught by a GTA were 32 percent more likely to cheat than students taught by faculty. This is strong evidence that the use of GTAs is costly in terms of academic honesty. In many cases, however, it would be financially costly to avoid using GTAs as instructors. It appears, though, that similar decreased in the probability of cheating can be had by using combinations of relatively inexpensive deterrent measures. For example, writing an additional test version decreases the probability of cheating by an estimated 25 percent. Simply issuing a verbal warning at test time decreased cheating by 13 percent, and the use of an additional proctor reduced cheating by 11 percent.

SUMMARY

Our first purpose was to find the determinants of class-specific academic cheating on examinations. Using randomized response survey data, we found that a student's grade point average and year in school were important determinants of cheating. Important deterrent measures included the use of faculty instead of graduate teaching assistants as instructors, multiple test versions, more proctors during exams, and verbal warnings.

The second purpose was to estimate class-to-class differences in the severity of the cheating problem. For 12 different principles of economics classes, we found an overall probability of .13 that a student cheated at least once on a exam in a given class. Class-by-class probability estimates ranged from .002 to .32,and differences generally corresponded to the deterrent measures used.

Third, we tested and rejected the hypothesis that there are no effective means of controlling cheating. Fourth, we compared the efficacy of various control measures. Avoiding multiple-choice questions did not appear to have discouraged cheating, nor did increased physical separation of students during exams.

An estimated 12 percent reduction in cheating occurs with the simple use of verbal announcements that honesty is an enforced university policy or the use of an additional test proctor. Adding an additional test version reduced the probability that a student cheated by 25 percent. Finally, the control measure with the largest deterrent effect was having classes taught by tenure-track faculty instead of graduate teaching assistants. Students taught by faculty, rather than GTAs, were 32 percent less likely to cheat.

NOTES

1. We surveyed all classes during the last week of the term. However, the number of midterms,and therefore the number of opportunities to cheat, varied from class to class. To control for this, we included *NTESTS* as an explanatory variable.
2. A potential concern with the survey and estimation methods we used is that students prone to cheating may self-select into classes where cheating is more tolerated. If self-selection is important, our results may be biased. We argue here that respondents did not self-select to an important degree for five reasons. First, for 2 of the 12 classes, students had no choice but to enroll in the full professor's class,as the only two sections available were taught by him. Second, in two-thirds

of the cases, class schedules did not list the class's instructor. Students signed up for a class knowing only that it would be taught by "staff"; the instructor may turn out to be a faculty member or a GTA. After the first class meeting, enrolled students could self-select by transferring to another section. However, this occurred less than one-half of 1 percent of the time.

Third, at the time of our surveys, the individual GTAs had not taught sufficient classes to develop a reputation that students could use to self-select. Two of the GTAs had never taught before, and the remainder had taught only one class previously. Of course, it is possible that the inexperienced GTA was what the cheating-prone student was looking for, but information about experience was generally not available when students register.

Fourth, students have many other variables to juggle when registering for classes, including getting the other classes required for their academic major, hours for outside employment and family obligations, and an aversion to classes at certain times of the day or days of the week.

Fifth, we take the uncynical view that very few students enter a class with an intention to cheat. It is only after they have gotten into a class and reached certain circumstances (e.g., failed to study for an exam or found the class more difficult than anticipated) that they make the decision to cheat.

3. We also estimated models after eliminating possible outliers, students reporting *ALCOHOL* > 360 and *CREDITHR* > 20. This sample (N = 383) provided qualitatively identical results to those reported here.
4. For three of the seven teachers, we know that students were separated as much as possible, still leaving many students sitting quite close to one another. In these cases, additional separation would have been used if it had been available, and *SQFOOT* was an appropriate measure. However, in the other cases, we did not collect and were not able to obtain information about whether additional space would have been used for separation.

   We experimented with two other measures of physical separation. First, we created a variable that interacted *SQFOOT* with a dummy variable indicating that the instructor was known to use all of the space available for separating students. Second, we tried creating a variable that measured the percentage of students that could be located at least three feet from their nearest neighbor. In neither case was the associated coefficient statistically significant at conventional levels.
5. Many other demands on students' time other than credit hours could affect cheating. Our data set contained a variable measuring the number of hours per week spent working for wages. When we included this variable in the regression, its coefficient was small and statistically insignificant.

REFERENCES

Baird, J. 1980. Current trends in college cheating. *Psychology in the Schools* 17 (4): 515–22.

Becker, G. 1968. Crime and punishment: An economic approach. *Journal of Political Economy* 76 (2): 168–217.

Bunn, D., S. Caudill, and D. Gropper. 1992. Crime in the classroom: An economic analysis of undergraduate student cheating behavior. *Journal of Economic Education* 23(Summer): 197–207.

Chaudhuri, A., and R. Mukerjee. 1988. *Randomized response: Theory and techniques.* New York: Marcel Dekker.

Collison, M. 1990. Apparent rise in student cheating has college officials worried. *Chronicle of Higher Education* 36 (January 17): A33–34.

Eble, E. 1988. *The craft of teaching: A guide to mastering the professor's craft*. San Francisco: Jossey-Bass.

Fass, R. 1990. Cheating and plagiarism. In *Ethics in higher education.* W. May, ed. New York: Macmillan.

Gardner, W., J. Roper, C. Gonzales, and R. Simpson. 1988. Analysis of cheating on academic assignments. *Psychological Record* 38 (4): 544–55.

Garred, E., D. Bacon, N. Matsumoto, and S. Skonik. 1991. Schools for scandal. *People Weekly* 36 (May 13): 103–4.

Haines, V., G. Diekhoff, E. LaBeff, and R. Clark. 1986. College cheating: Immaturity, lack of commitment, and neutralizing attitude. *Research in Higher Education* 25 (4): 342–54.

Houston, J. 1976a. The assessment and prevention of answer copying on undergraduate multiple choice examinations. *Research in Higher Education* 53 (4): 301–11.

———. 1976b. Amount and loci of classroom copying, spaced seating, and alternative test forms. *Journal of Educational Psychology* 68 (December): 729–35.

———. 1983. College classroom cheating, threat, sex, and prior performance. *College Student Journal* 17 (Fall): 229–35.

Kerkvliet, J. 1994. Cheating by economics students: A comparison of survey results. *Journal of Eco -*

nomic Education 25 (Spring):121–33.
Maramark, S., and M. Maline. 1993. *Academic dishonesty among college students*. Washington, D.C.: U.S. Department of Education.
Michaels, J., and T. Miethe. 1989. Applying theories of deviance to academic cheating. *Social Science Quarterly* 70 (December): 870–85.
Mixon, F. 1996. Crime in the classroom. *Journal of Economic Education 27 (Summer*): 195–200.
Moffatt,M. 1990. Undergraduate cheating. Mimeograph. New Brunswick, N. J.:Rutgers University.
Nelson, T.,and N. Schaefer. 1986. Cheating among college students estimated with the randomized-response technique. *College Student Journal* 20 (Fall): 321–25.
Nowell, D., and D. Laufer. 1997. Undergraduate cheating in the fields of business and economics. *Journal of Economic Education* 28 (Winter): 3–12.
Roberts, D., and W. Rabinowitz. 1992. An investigation of student perceptions of cheating in academic situations. *Review of Higher Education* 15 (2): 179–90.
Scheers, N.,and M. Dayton. 1987. Improved estimation of academic cheating behavior using the randomized response technique. *Research in Higher Education* 26 (1):61–69.
Singhal,A. 1982. Factors in student dishonesty. *Psychological Reports* 51 (December):775–80.
Singhal, A., and P. Johnson. 1983. How to halt student dishonesty. *College Student Journal* 17 (Spring): 13–19.
Stern,E.,and L. Havlicek. 1986. Academic misconduct:Results of faculty and undergraduate student surveys. *Journal of Allied Health* 5 (2): 129–42.
Stevens, G., and F. Stevens. 1987. Ethical inclinations of tomorrow's managers revisited: How and why students cheat. *Journal of Education for Business* 63 (1): 24–29.
———. 1982 Asking questions London: Jossey-Bass.
Sudman, S., and N. Bradburn. 1974. *Response effects in surveys: A review and synthesis*. Chicago, Ill.:Aldine.
Tittle, C.,and A. Rowe. 1974. Fear and the student cheater. *Change* 6 (3): 47–48.