ACADEMIC ECONOMISTS BEHAVING BADLY? A SURVEY ON THREE AREAS OF UNETHICAL BEHAVIOR

JOHN A. LIST, CHARLES D. BAILEY, PATRICIA J. EUZENT and THOMAS L. MARTIN

This article measures the degree to which academic economists have engaged in unethical behavior and the degree to which academic economists believe the profession as a whole engages in unethical behavior. Three main types of unethical behavior are examined: (1) falsification of research; (2) expropriation of graduate student research or including an undeserving co-author on a research paper; and (3) exchange of grades for gifts, money, or sex. Using a unique data set gathered at the 1998 American Economic Association (AEA) meetings, we find that there is a significant amount of misconduct, particularly in the second category. (JEL A11, A13)

I. INTRODUCTION

Both the popular and academic press have noted the existence of falsified research in the hard sciences and in medicine. U.S. nuclear weapons tests conducted in the 1950s came under attack in the 1960s for allegedly using false data and defrauding federal courts concerning the health impact on downwind humans and livestock. The discovery of the deception of research subjects and the reporting of fabricated experiments led to a new emphasis in the 1960s and 1970s on the ethics of scientific research (Shrader-Frechette [1994]). Federally funded research, in particular, came under closer scrutiny of the White House and various executive branch agencies (LaFollette [1994]).

Recently, front-page stories in major newspapers have reported that Robert P. Liburdy, a cell biologist at Lawrence Berkeley Laboratories, resigned after investigators with the Office of Research Integrity of the Department of Health and Human Services concluded that he falsified key data tying cancer to electrical power lines (Schevitz [1999]; Broad [1999]).

In economics research, with no lives in danger, it would seem that such behavior would be easier to rationalize. An inflated t-ratio here, an undeserving co-author there—who really is hurt by such behavior? Academic economists, after all, do not have patients. Nevertheless, most scholars would agree that research is intended to further the pursuit of truth. Fraudulent or unethical research will poison the well of that scientific truth. Ethical conduct in research is not

1. This does not imply that economists' work fails to have important consequences. For example, see the recent debate on valuing environmental damages (e.g., List [2000]) or determining the correct auction mechanism to promote efficient allocations and maximize revenues (e.g., List and Licking-Reilly [2000]).

ABBREVIATIONS

AEA: American Economic Association
DR: Direct Response
RR: Randomized Response
LPM: Linear Probability Model
a matter of etiquette but a question of the moral principles by which any researcher is guided.

In the past few decades there have been efforts to examine the ethics of research in the areas of hard science, medicine, engineering, and national security (see, e.g., Hook [1977]; Bulmer [1982]; Sieber [1992]). A few decades ago our own discipline experienced an unfathomable event—the Quarterly Journal of Economics was forced to withdraw a published article after learning that its “author” had simply copied the work of another economist (QJE [1984]). A similar incident recently occurred in the economics journal Kyklos, forcing the editors to issue a public apology that included a strong statement: “Plagiarism jeopardizes scientific research” (Kyklos [1999, 311]). These pieces of evidence suggest that unethical research practices occur within economics, but very little nonanecdotal evidence exists to measure the extent of such behavior. Our study is an attempt to measure the degree to which individual academic economists have engaged in unethical behavior and the degree to which academic economists believe the profession as a whole engages in unethical behavior. Three main types of behavior are examined: (1) falsification of research; (2) expropriation of graduate student research or including an undeserving co-author; and (3) exchange of grades for gifts, money, or sexual favors. We examine the self-reported unethical behavior of the profession through anonymous responses. Because of the extremely sensitive nature of the questions, we split the sample by using direct-response questions for one group and the added protection of a “randomized response” technique for the other.

II. THE SURVEY METHOD

The ethics surveys were hand-delivered to approximately 1,000 academic economists (about 20% of the participants) attending the January 1998 meetings of the American Economic Association (AEA) in Chicago, Illinois. The survey was distributed to conference participants while they waited in line to register for the conference. As potential subjects were handed the survey they were told that “the survey was designed to obtain information about the ethics of professional economists.” We suggested that survey participants answer the questions in the privacy of their hotel room and place the survey in a drop box located near the registration desk or mail it back using the business reply envelope. No surveys were given orally. The survey instrument contained eight demographic questions and six questions on the respondent’s actual behaviors and the perceived behaviors of fellow economists.

The demographic information requested included age, rank, gender, and whether the participant’s university is primarily a research or teaching institution. In the section on ethics, the subjects were first asked “have you ever falsified research data?” They were then asked what percentage of research in the top 30 journals they believed had been the result of using falsified data. Next, participants were asked whether they had committed any of a variety of unethical acts, such as submitting a manuscript to two or more journals concurrently, expropriation of graduate student research, or giving a colleague unjustified co-authorship. As a follow-up question, they were asked what percentage of the profession they believed had committed such violations. Finally, subjects were asked if they had ever accepted gifts, money, or sexual favors in exchange for grades and what percentage of the profession they believed had committed such violations.

These would, of course, be quite difficult questions to answer about oneself if the response was in the affirmative. Because many types of research misconduct are difficult to detect by secondary methods of research, some means of self-reporting is a valuable approach to investigating the problem. To overcome subjects’ fears that the instruments can be traced to them (leading to underreporting), we use both a direct-response (DR) survey and a randomized response (RR) survey. In some situations, the

---

2. Other studies also exist (see, e.g., Bailey et al. [1998]; Mason et al. [1990]; Meier [1986]; Fisher [1986]). Related studies assess the ethics of university students (e.g., Nowell and Laufer [1997]; Kerkvliet [1994]; Kerkvliet and Sigmund [1997]).

3. Appendix 1 contains a copy of the randomized response survey. The direct-response survey was similar, except we excluded the randomization design (flipping the coin three times). Copies of this instrument are available on request.
RR technique has been more effective in generating honest responses than the DR questionnaire generates.\(^4\)

Under the RR method in our study, the subject uses coin tosses to determine whether to respond to the question or simply mark “yes.” A coin is flipped three times before each question is answered. If the coin turns up heads all three times, the respondent is instructed to answer the question with a “yes” regardless of the accuracy of that answer.

Under this particular RR design, we cannot know whether an individual respondent is answering “yes” to the sensitive question, but we can use established statistical techniques to estimate population parameters. This added security comes at a cost. Because randomness of the participant’s response must be taken into account, ceteris paribus, the variance of the parameter estimate is larger than variances of direct response estimates.

II. HYPOTHESES AND RESULTS

To provide a reference point for our analysis, we begin by examining whether responses to the RR technique are similar to those from the DR technique. If RR surveys induce more truthful responses from ethics offenders, we should find that a larger proportion of respondents answered “yes” to the sensitive questions. Accordingly, we formulate the following null hypothesis:

\[(H1) \quad \mu_{\text{YRR}|i} = \mu_{\text{YDR}|i}\]

where \(\mu_{\text{YRR}|i}\) represents the computed mean response to sensitive question \(i\) under the random (direct) response format. The null hypothesis in (H1) states that the proportion of “yes” responses to the sensitive questions is equivalent across the two survey methods. Rejection of (H1) and a finding that \(\mu_{\text{YRR}|i} > \mu_{\text{YDR}|i}\) could imply that a number of participants felt protected by the RR questionnaire design and provided more accurate answers.

Next, we move to our core hypotheses—estimating the amount of cheating behavior among academic economists:

\[(H2) \quad \mu_{Yij} = 0\]

where \(\mu_{Yij}\) is the mean self-reported cheating behavior for crime type \(i\), survey type \(j\), and \(\mu_{\text{perceptionsij}}\) is the mean perceived amount of cheating behavior in crime type \(i\), survey type \(j\). Null (H2) states that self-reported cheating behavior is not significantly different from zero. Rejection of (H2) implies that there is some degree of unethical behavior among academic economists. (H3) tests whether the trait of being ethical is less descriptive of self than of others. Rejection of (H3) and a finding that \(\mu_{\text{perceptionsij}} > \mu_{Yij}\) provides general evidence consistent with the so-called Muhammad Ali effect. The Muhammad Ali effect, which is discussed frequently in the psychology literature (see, e.g., Van Lange and Sedikides [1998]), suggests that the trait of honesty is more descriptive of oneself rather than one’s peers.

Next, we explore the determinants of unethical behavior. Becker’s (1968) crime model suggests that numerous factors may influence an individual’s decision to act unethically. We examine the dichotomous decision in a logistic framework and make the probability of unethical behavior a function of (i) demographic characteristics and (ii) the respondent’s perception of the ethics of the economics profession. If a significant relationship exists between individual unethical behavior and perceptions, then it may imply a snowball effect: one cheats, others perceive cheating occurs, so they rationalize cheating themselves. Given these goals, our alternative hypothesis is

\[(H4) \quad Y_i = f(X_i),\]

where \(Y_i\) takes on a value of 1 if person \(i\) has behaved unethically and 0 of not, and \(X_i\) is a vector of regressors presumed to affect the probability of behaving unethically.\(^5\) The null

\(^4\) Warner (1965) originally proposed the RR technique.

\(^5\) Although many variables may be influential in the decision to misbehave, when possible, we maintain consistency with past regression models of student cheating in the classroom to form vector \(X\) (see, e.g., Kerkvliet [1994] and the citations therein). Accordingly, vector \(X\) includes perceptions of other economists’ behavior, gender, academic rank, and ranking of employing institution, where institutional rankings are from Scott and Mitias (1996). Scott and Mitias (1996) partition economic departments into four categories: top 19 schools; schools ranked 20–49; schools ranked 50–100; and the rest of the institutions ranked 101–240. We gathered other variables, but they were excluded due to data
hypothesis of H4 premises that self-reported cheating behavior is not a function of perceptions, gender, academic rank, and rank of employing institution (e.g., \( \beta_i = 0 \forall i \)). If (H4) is rejected, important relationships can be conjectured about unethical behavior and the regressors. One can determine, for example, if unethical behavior could increase in the future based on the demographics of current offenders.

**Results**

Of the 1,000 surveys distributed, 600 were RR and 400 were DR. A larger number of RR surveys were distributed to overcome the disadvantage of the RR approach–inflated variance. We received 140 RR responses and 94 DR instruments (response rates of 23.3\% and 23.5\%). The response rates were almost identical, a surprising result in light of the additional protection that the RR method offers. Two explanations come to mind. The DR respondents may have felt fully protected because of the method of distribution and return, or the greater effort required by the RR technique might have offset its response-inducing effects. Furthermore, because the number of “guilty” subjects presumably is small, any resulting difference in response rates also should be small.

Table 1 presents some characteristics of the subjects who completed the surveys. Respondents across the two random samples appear similar as to gender, age, academic rank, and research rank of their institution of primary employment; a one-way ANOVA test suggests the sampled participants do not differ significantly across survey type at the \( p < .05 \) level. This result provides some assurance that the differences across treatments are not a function of the characteristics of the sample.

**Responses are Similar across Survey Formats**

Table 2 contains descriptive statistics of the six questions.\(^6\) Responses suggest that a nontrivial number of academic economists have committed certain “research crimes.” Although point estimates in Table 2 suggest that responses are somewhat different across the two survey formats, we find no statistical differences between answers obtained in the DR and RR survey techniques using standard large-sample two-tailed \( t \)-tests. This result is similar to Kerkvliet (1994) and suggests that in our study, the respondents were not acting heterogeneously across survey instruments. This is not surprising given that the DR surveys were completed in a fashion that guaranteed anonymity.

**Self-Reported Cheating Behavior Is Significantly Greater than Zero for Both Research Crime Types**

Estimates presented in column 1 of Table 2 indicate that there is a fair amount of “felony” crime types committed by academic economists—in both elicitation schemes, the point estimate is statistically different from zero at conventional levels, suggesting that a nonzero amount of research is plagued by falsification of data. In this case, the point estimates suggest that 4.26\% of direct respondents and 4.49\% of randomized survey respondents have committed this major offense at least once. Are these percentages economically significant? That is a very difficult question to answer. Whether the fraudulent 4\% of respondents are active publishers is not known, so the extent to which academic journals are contaminated is unclear. Indeed, research data may have been falsified for a dissertation or a paper that was never published or read. But in an academic environment in which business ethics is considered essential material for students, even 4\% of the profession admitting to these unethical behaviors must be considered important.

The reported amount of admitted research behavior in the list of “misdemeanors” (question 10) in column 3 of Table 2 is much greater than felony estimates, with point esti-

\[ n \approx \frac{\sigma^2 \cdot \pi(1-\pi)}{(1-P)\pi}/P \]

\[ \sigma^2_{\text{RR}} = Z(1-Z)/(\pi(1-\pi))^2 \]

\[ Z \text{ is the observed proportion of yes responses, } P \text{ is the probability of answering the sensitive question, } \pi \text{ is the proportion of yes responses to the nonsensitive question (in our case a series of coin flips, hence } \pi = 1), n \text{ is the sample size.} \]
TABLE 1
Demographic Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Randomized Response Mean (SD)</th>
<th>Direct Response Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (% Male)</td>
<td>77.9 (41.7)</td>
<td>84.0 (36.8)</td>
</tr>
<tr>
<td>Age</td>
<td>41.7 (11.1)</td>
<td>42.7 (10.9)</td>
</tr>
<tr>
<td>Respondent’s rank:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>full professor</td>
<td>32.1 (46.9)</td>
<td>37.2 (48.6)</td>
</tr>
<tr>
<td>Associate professor</td>
<td>20.7 (40.7)</td>
<td>18.1 (38.7)</td>
</tr>
<tr>
<td>Assistant professor</td>
<td>27.9 (45.0)</td>
<td>27.7 (45.0)</td>
</tr>
<tr>
<td>Instructor</td>
<td>4.2 (20.3)</td>
<td>4.3 (20.3)</td>
</tr>
<tr>
<td>Graduate</td>
<td>15.0 (35.8)</td>
<td>12.8 (33.6)</td>
</tr>
<tr>
<td>University’s Research Rank:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–19 (%)</td>
<td>22.1 (41.6)</td>
<td>24.5 (43.2)</td>
</tr>
<tr>
<td>20–49 (%)</td>
<td>24.2 (43.0)</td>
<td>19.2 (39.6)</td>
</tr>
<tr>
<td>50–101 (%)</td>
<td>22.1 (41.7)</td>
<td>18.1 (38.7)</td>
</tr>
<tr>
<td>101+ (%)</td>
<td>31.4 (46.6)</td>
<td>38.3 (48.9)</td>
</tr>
</tbody>
</table>

Note: Research rank is from Scott and Mitias (1996).

TABLE 2
Summary Statistics of Responses

<table>
<thead>
<tr>
<th>Research “Felony” (Falsification)</th>
<th>Research “Misdemeanor”</th>
<th>Selling Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self (Q 9)</td>
<td>Other (Q 9a)</td>
<td>Self (Q 10)</td>
</tr>
<tr>
<td>Have you ever falsified research data?</td>
<td>What percentage of research in the top 30 journals do you believe [is falsified]?</td>
<td>Have you ever committed any of four “minor” infractions?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomized response ($n = 140$):</td>
<td>4.49(0.30)</td>
<td>10.17(0.34)</td>
</tr>
<tr>
<td></td>
<td>7.04(0.85)</td>
<td>16.98(1.52)</td>
</tr>
<tr>
<td>Direct response ($n = 94$):</td>
<td>4.26(0.22)</td>
<td>5.13(0.73)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12.95(1.50)</td>
</tr>
</tbody>
</table>

Notes: Cell contents are means (standard errors) and represent percentages. For randomized response questions, we compute means and variances based on $\mu_{RR} = (Z - (1 - P)\pi)/P$; $\sigma_{RR} = Z(1 - Z)/(n - 1)P^2$, where $Z$ is the observed yes responses, $P$ is the probability of answering the sensitive question, $\pi$ is the proportion of yes responses to the nonsensitive question (in our case a series of coin flips, hence $\pi = 1$), $n$ is the sample size.
mates of 7.45% (DR) and 10.17% (RR). Although these practices are not as frightening as falsifying research data, it does not speak well for academia that an estimated 7–10% of academic economists have (i) submitted a manuscript to two or more journals in violation of journal policy, (ii) not given graduate students justified co-authorship, or (iii) given unjustified co-authorship to a colleague. Given that the punishment is presumably much less severe for this type of behavior, it is rational that these percentages are higher than those reported in the felony group.

Concerning the exchange of grades for money or sex, our mean estimates of 0% and 0.40% using the DR and RR technique (see column 5 of Table 2) suggest that a small amount of this activity takes place in academia.

A...
One possible explanation is that economists at lower-ranked schools have fewer resources with which to conduct valid, honest research and turn to the next best alternative. Second, there is some tendency for graduate students and full professors to behave differently, but this effect is only significant at the $p < .20$ level. Given that the questions were worded to inquire if respondents had ever engaged in unethical behavior, this is a curious result as one would believe that, ceteris paribus, graduate students would have less of a chance to behave unethically and therefore would appear “more honest” than professors. This finding may be suggestive of the underlying quality of graduate students’ research papers turned in as term projects. Nevertheless, this finding does not bode well for the believability of the future body of scientific knowledge but is consistent with findings in Bailey et al. (1998) who find that younger accounting faculty are more pessimistic about fraud and academic wrongdoings. Third, our data suggest there are no rank or gender effects. Although this could be a repercussion of the small number of respondents in certain gender and rank categories, these results suggest that misbehavior does not considerably vary over rank and gender.

### IV. CONCLUDING REMARKS

Policing research obviously places a burden on editors and reviewers. Some observers have recommended that editors should in the very least become cautious about data that falls “too well into place.” But, ultimately the problem may optimally be treated at the grassroots level of individual universities policing their own. In particular, as was observed in the 1981 congressional hearings, the “control of appointments, tenure and promotion is the real power of the university for the preservation of scientific ethics” (Steneck [1994], 213). In addition, universities could develop more effective administrative procedures for conducting investigations or establish ethics committees and ethical training programs.

Even without life-or-death consequences, falsified economics research can lead to misinformed or inappropriate policy recommendations that, if enacted, could degrade living standards. In this age in which economic analysis is a necessary part of the debate concerning legislation and executive branch action, the economics profession cannot afford to lose credibility with themselves or, more important, with the general public. That would be a sure prescription for irrelevancy. Are the percentages high enough to begin the journey down this road to irrelevancy? No one can say for certain.
This study, however, provides some initial evidence on the extent to which unethical research behavior exists among academic economists. Future research needs to track whether or not the ethical standards of the profession are rising or falling. To ensure a trend in the right direction, perhaps it is time for graduate programs to include moral or ethical training.

APPENDIX TABLE A1
Randomized Response Survey

The design of this questionnaire ensures your complete anonymity. Please circle the appropriate answer.

<table>
<thead>
<tr>
<th>Step 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Your gender: Female Male</td>
</tr>
<tr>
<td>2. Your rank: Pull Prof. Associate Prof. Assistant Prof. Instructor Graduate Student</td>
</tr>
<tr>
<td>3. Are you tenured? Yes No</td>
</tr>
<tr>
<td>4. Do you perceive your institution as primarily a research institution? Yes No</td>
</tr>
<tr>
<td>5. Do you perceive your institution as primarily a teaching institution? Yes No</td>
</tr>
<tr>
<td>6. Your age: ____________________________</td>
</tr>
<tr>
<td>7. Your major field of study: ____________________________</td>
</tr>
<tr>
<td>8. Based on the rankings provided on the attached sheet, identify the general ranking of your economics department: 1–19 20–49 50–100 101+</td>
</tr>
</tbody>
</table>

Step 2

After reading the first question, please secretly flip a coin three times. Note whether the coin comes up heads all three times.

If you do toss three heads, please answer “yes” to number 9 below and then answer 9a.

If you do not toss three heads, then please answer the following question:

9. Have you ever falsified research data? Yes (or coin came up heads three times) No

9a. What percentage of research in the top 30 journals do you believe is the result of falsified data? _______%

Step 3

Please secretly flip a coin again three times. Note whether the coin comes up heads all three times.

If you do toss three heads, please answer “yes” to number 10 below and then answer 10a.

If you do not toss heads, then answer the following questions:

10. Have you ever
   • submitted a manuscript to two or more journals in violation of journal policy, or
   • not given graduate students co-authorship on publications where co-authorship was justified, or
   • inappropriately given a colleague co-authorship status, or
   • presented the same research to more than one regional or annual meeting (against meeting policy)?
   Yes (or coin came up heads three times) No

10a. What percentage of research in the top 30 journals do you believe is affected by the behavior listed in number 10? _______%

Step 4

Please secretly flip a coin again three times. Note whether the coin comes up heads all three times.

If you do toss three heads, please answer “yes” to number 10 below and then answer 11a. If you do not toss heads, then answer the following questions:

11. Have you ever accepted sex, money, or gifts in exchange for grades?
    Yes (or coin came up heads three times) No

11a. What percentage of economics faculty members do you believe have accepted sex, money, or gifts in exchange for grades? _______%

Please place your response in the envelope provided and place it in the mail or in the UCF drop box at the conference registration desk.

Thank you very much for your time and your participation in this survey!
REFERENCES


