



UNSW
THE UNIVERSITY OF NEW SOUTH WALES

Faculty of
Business

**School of Economics
UNSW, Sydney 2052
Australia**

<http://www.economics.unsw.edu.au>

**Confidentially is Not Enough: Framing Effects in Student
Evaluation of Economics Teaching**

Kieron J. Meagher

School of Economics Discussion Paper: 2007/14

The views expressed in this paper are those of the authors and do not necessarily reflect those of the School of Economic at UNSW.

ISSN 1323-8949

ISBN 978 0 7334 2450 2

Confidentially is Not Enough: Framing Effects in Student Evaluation of Economics Teaching

Kieron J. Meagher¹
University of New South Wales

Abstract: Contrary to previous research we show lack of anonymity is associated with large positive shifts in student evaluation of teaching. The results are consistent with the simple observation that due to higher expected future earning economics and business students have more at stake in terms of potential retaliation by an instructor. Our analysis is based on both a comparison of distributions and ordered probit multi-variate regression. These methods overcome the statistical problems associated with previous studies which looked at differences in means for ordinal responses.

Keywords: evaluation, bias, survey design, teaching economics.

JEL Codes: A20, A22, C81

Introduction

Student evaluation of teaching (SET) is now common in universities. Although the goal of improving teaching quality is natural the process of evaluation is widely criticised, directly on the basis of reliability and interpretation, and more tangentially because of the way SET enter into hiring and tenure/promotion decisions.

The way in which questions are asked (framing effects), including confidentiality and anonymity, are generally well understood as a potential source of bias in data collection in the social sciences. Presumably, it is for this reason that many universities, including my own, collect SET data anonymously. This practice of anonymity is being eroded by two teaching interest groups. First many teaching professionals currently advocate the use of more in-depth data collection methods, such as face-to-face discussions/focus groups, which may promise confidentiality to participants but do not guarantee effective anonymity. See for example the discussions on evaluation methods in Harvey (1998).

Secondly concerns about the use of SET as a personnel management tool raise serious equity issues:

"...anonymous student evaluations of teachers may serve as vehicles for transmitting popular misconceptions, expectations and prejudices, to the disadvantage of, for example, women and visible minorities." CAUT (2006)

The assertion is that by signing evaluations students will feel more responsibility and will more truthfully evaluate the teaching rather than distorting their evaluations according to their prejudices.

Both of these arguments, and indeed any other proposal that does not guarantee students anonymity, assumes that the distortions from non-anonymous responses are negligible. This is not an unreasonable point of view since the seminal papers in the field, Stone, Spool and Rabinowitz (1977) and Feldman's (1979) survey, failed to find clear evidence of significant positive bias introduced from non-anonymity.

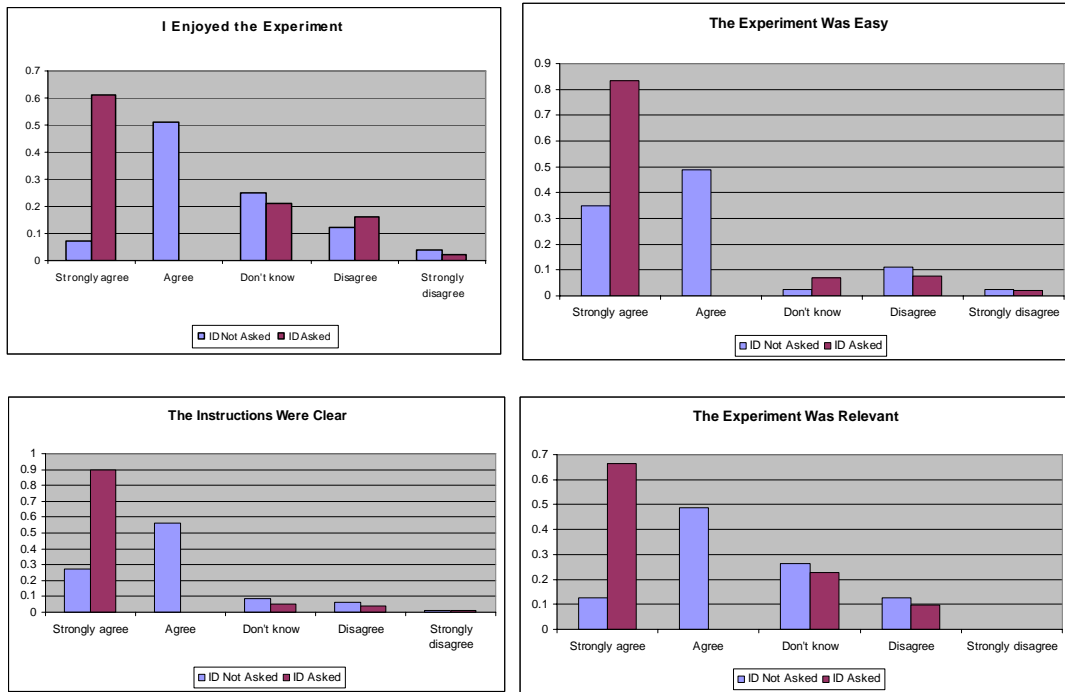
The main framing issue with anonymity and confidentiality is that students will be concerned about possible negative consequences for themselves from giving low teaching ratings and hence will bias their responses upwards. There is good reason to question the applicability of previous studies to the contemporary teaching of economics and business students. First, thanks in part to Becker and other labour economists education is now viewed more strongly as an investment by students than it was in the 1970's. Possibly causing students to be concerned not just with immediate negative retaliation by teachers but also the possible impact it might have on their life time earning.

¹ School of Economics, UNSW, Sydney, NSW 2052, Australia. *email:* k.meagher@unsw.edu.au. *Fax:* +61-2-9314 6337. *Tel:* +61-2-9385 1145.

If students are concerned about the potential impact of poor teaching evaluations on grades then students whose studies lead to the highest income occupations should be the most likely to exhibit a positive bias. However Stone et al (1977) investigated bias in an evening class of jurisprudence students who already had jobs. Fries and McNinch (2003) have shown that contemporary sociology students produce small but more biased responses than the Stone et al study, but there has been no study of economics students. We predict that the higher expected incomes of the group of economics and business students we investigate will lead them to exhibit even greater positive bias.

Methodologically the studies cited above are unsatisfactorily simplistic in their statistic analysis. All the relevant data is from Likert scales, which are ordinal but not cardinal. However the previous studies have all imposed a cardinal scale and analysed differences in the constructed means. We instead focus on differences in distributions caused by anonymity effects. Furthermore we also offer a multivariate analysis to see if other individual characteristics play a role in explaining the anonymity effect.

Figure 1: Distributions of Student Responses by Question and Framing Method



Empirical Results

In the first teaching session of 2006 we introduced an economic experiment on ultimatum bargaining in to our first year microeconomics class. We used Charlie Holt's excellent online Veconlab software to run the experiment and everything ran smoothly². It is the evaluation of this teaching innovation which provides the data for the following analysis.

Students were asked to complete an online survey. The ethics information provided before the students completed the questionnaire indicated that the information was being collected solely for the purpose of teaching research by Dr Meagher (I was not an instructor for the course), participation was voluntary, all responses were confidential and only aggregate results would be made available to the course instructor (although no date for the release of these results was given).

² See Meagher and Chan (2007) for a discussion of using the ultimatum bargaining experiment in an international classroom.

We collected basic personal data from the students (gender, foreign/domestic student, reason for taking the course). We also asked students about their experience of participating in the experiment: enjoyment, ease of participation, clarity of instructions and relevance to their careers. Responses were on a standard five-point Likert scale.

It was announced that the questionnaire would be available for two weeks to complete. During the first week we did not request student identification numbers (student ID's) in the survey. Without notification in the second week we did request student ID's in the online survey. Thus throughout the survey students responses were always confidential but it was technically possible to identify students during the second week of data collection. *The main result of this paper is to show that despite confidentiality student responses were significantly more positive without anonymity.* This fundamental result is readily apparent from the distributions of student responses in Figure 1.

Table 1: Chi-Squared Goodness of Fit Test for Constant Distributions Across Framing³

Question	Test Statistic	1% Critical Value
<i>I enjoyed the experiment</i>	435.41	$\chi^2(4 \text{ df}) = 13.28$
<i>The experiment was easy</i>	126.87	$\chi^2(4 \text{ df}) = 13.28$
<i>The instructions were clear</i>	202.66	$\chi^2(4 \text{ df}) = 13.28$
<i>The experiment was relevant to my studies⁴</i>	284.18	$\chi^2(3 \text{ df}) = 11.34$

H₀: With ID distribution is the same as the Without ID Distribution
H₁: With ID distribution is different to the Without ID Distribution
Right Tail Test at 1% Significance Level

80 students responded in the first week when ID was not requested while 103 students responded in the second week when was ID asked. 6 did not provide ID when requested. Since they fell in the ID asked treatment their responses are included in that category for the construction on the sample distributions. Common sense suggests that this group of 6 students felt the need for confidentiality because they had negative opinions they wanted to express. This issue is pursued in the ordered probit analysis below through the inclusion of a dummy variable *declined to give ID when asked*.

The sample distributions in figure 1 tell a consistent story: student responses were more positive, on average, to every question when they were also asked to provide their student ID. The variation in the neutral and negative responses is reasonably small, most of the change in the distributions is due to the increased weight on the *strongly agree* category compared to the *agree* category.

The chi-squared goodness of fit tests reported in Table 1 confirm that there is a very highly statistically significant difference between the distribution when ID was not asked compared to when ID was asked.

Although student responses were confidential and were not supplied to a course instructor there appears to be strong evidence that the framing effect caused by asking for student ID's had a significant effect on responses between the two treatments.

Students were assigned to treatments on the basis of time. This does not guarantee that the observable characteristics of the two groups are the same. To investigate if the apparent treatment effects are due to the composition of the two groups we estimated an ordered probit for each question. The sample is restricted to the 178 students who answered all the questions rather than the 183 who answered any of the questions. The results on enjoyment are reported in Table 2. The ordered probit results on the positive effect of anonymity for the other questions are similar or stronger.

³ Expected frequencies are low in some categories but do satisfy the weak criteria of Doane and Seward (2007, p668). Furthermore the test results on difference in distribution are driven the *strongly agree* category in which the expected frequencies are high.

⁴ For the relevance question only four categories were used in the test because there were no responses in the *strongly disagree* category.

For interpretation a more positive value of the latent variable is associated with a more positive response to the question, i.e. strongly disagree is the lowest coded category and strongly agree is the highest coded category. The probit results indicate that anonymity, whether provided through the treatment (*ID asked*) or self selected by students (*Declined to give ID when asked*) is associated with more negative responses. None of the other control variables are significant in the estimation equation. The marginal effects of a change in *ID asked* show a transfer of probability weight to the most positive category, *strongly agree*.

Table 2: Ordered Probit Results for Student Evaluation of Enjoyment of Experiment (Standard errors in parenthesis).

Coefficient Estimates			Marginal Effect of a One Unit Change in <i>ID asked</i>		
Variable	Coefficient	p-value	Category	Marginal Effect	p-value
<i>ID asked</i>	0.683 (0.170)	0.000	<i>Strongly agree</i>	0.251 (0.061)	0.000
<i>male</i>	-0.231 (0.169)	0.171	<i>Agree</i>	0.007 (0.014)	0.630
<i>local</i>	-0.282 (0.197)	0.152	<i>Don't know</i>	-0.086 (0.025)	0.000
<i>Course relevant to career</i>	0.148 (0.216)	0.493	<i>Disagree</i>	-0.134 (0.037)	0.000
<i>Course of interest</i>	0.017 (0.285)	0.952	<i>Strongly disagree</i>	-0.038 (0.018)	0.041
<i>Declined to give ID when asked</i>	-1.565 (0.562)	0.005			
<i>N</i>	178				
Pseudo R ²	0.05				

Marginal effects evaluated at the means of the independent variables.

Conclusion

A student evaluation of teaching applied to the used of an ultimatum bargaining experiment in a first year was conducted by a third party with strong confidentiality guarantees. Student responses were found to be systematically significantly more positive when their responses were confidential but not anonymous.

The evidence suggests that contemporary economics and business students, a group for whom the anonymity issue has not been previously investigated, are more concerned about possible negative outcomes of giving bad SETs than are the groups studied in the existing literature. This result is consistent with the prevailing view of education as an investment and the greater income streams at risk for business and economics students.

The results of this study suggest that anonymity concerns can, contrary to the prevailing wisdom, have a very large impact on student responses. The positive bias introduced by confidentiality without anonymity suggests that the additional information available through focus groups and other personal interactions may be significantly tainted. Furthermore, the techniques involving third party data collection used in this study are not sufficiently insulated to remove the bias in student responses.

References

- CAUT, 2006, "Policy on the Use of Anonymous Student Questionnaires(1) in the Evaluation of Teaching", *Canadian Association of University Teacher (CAUT)*, retrieved 22 January 2007 (<http://www.caut.ca/en/policies/questionnaires.asp>).
- Doane, DP, and LE Seward, 2007, "Applied Statistics in Business and Economics", New York: McGraw-Hill Irwin.

- Feldman, K.A., 1979, "The Significance of Circumstances for College Students' Ratings of the Teachers and Courses", *Research in Higher Education*, Vol. 10, No. 2, 149-172.
- Fries, C.J. and R.J. McNinch, 2003, "Signed versus Unsigned Student Evaluations of Teaching: A Comparison", *Teaching Sociology*, vol. 31(3), 333-344.
- Harvey, J, 1998, *Evaluation Cookbook*, J. Harvey (Editor), Edinburgh: Learning Technology Dissemination Initiative.
- Meagher, K.J. and A. Chan, 2007, "Online Economic Experiments: Diversity and Differing Student Experiences?", mimeo, University of New South Wales.
- Stone, E.F., M.D. Spool and S. Rabinowitz, 1977, "Effects of Anonymity and Retaliatory Potential on Student Evaluations of Faculty Performance", *Research in Higher Education*, Vol 6, 313-325.