DISCUSSION, DEBATE, AND DISAGREEMENT: TEACHING MULTIPLE REGRESSION BY CASE DISCUSSION

Peter G. Bryant
University of Colorado at Denver, Campus Box 165,
P. O. Box 173364, Denver, Colorado 80217-3364 USA
Peter.Bryant@cudenver.edu

1. Introduction

This paper describes my experience in teaching multiple regression in the required elementary business statistics courses at the University of Colorado at Denver by emphasizing case discussion. The approach is similar to standard discussion methods used in business curricula, but it is seldom used for statistics classes. I discuss:

- The course and the students;
- The case discussion approach and its application to multiple regression classes in particular; and
- The results from applying it.

2. Background: the course and the students

We require a single-semester course called "Data Analysis for Managers" of all graduate students in the Graduate School of Business Administration. Some students are enrolled in technically - oriented courses of study, such as accounting, finance or information systems. Others emphasize such less technical areas of business and administration as management or marketing, or are enrolled in a general business degree such as the M. B. A. Their backgrounds vary widely, as do their comfort levels with and enthusiasm for the material. Many students are returning to school after as long as thirty years. Their mathematical skills and confidence in their abilities to master the material often need bolstering.

The course includes basic descriptive statistics, some classical statistics such as confidence intervals and significance tests, and about 5 weeks (12.5 classroom hours) of multiple regression models, including dummy variables, autoregression, and approaches to variable selection. The remainder of the course includes such topics as statistical quality control, and index numbers. Case discussion is particularly helpful in dealing with the relatively technical material of multiple regression with such a heterogeneous group of students.

3. The case discussion approach

The case method of instruction has a long history in graduate professional education in the United States and elsewhere. In its original forms (in law and medicine) a *case* was a situation whose purpose to allow the student to practice analysis or diagnosis leading (if done correctly) inevitably to a single correct answer. Business educators, on the other hand, usually use cases differently. They use a case to provide a description of a situation, but the goal is to provide something for the class to discuss, not necessarily to lead to single correct answers. The cases and their discussion serve to raise awareness, elicit conflicting perspectives, assumptions, and points of view, and stimulate individuals in the group to think hard about their own approaches and answers and possible alternatives. These ideas are well illustrated and articulated in Christensen (1987), for example.

Case discussion as a basis for statistics courses (and technical courses more generally) is relatively recent, and many (perhaps most) instructors are accustomed to basing courses largely on lecture and reading. A recent article by Parr and Smith (1998) provides a thorough introduction to designing statistics courses around the theme of cases and gives appropriate references. Here we focus on one particular aspect of such a course: the *discussion* of the case, and in the context of multiple regression in particular.

Schedule. We typically assign one case per week, devoting between 30 to 75 minutes to the discussion. Cases are assigned at least a week in advance.

The case. The case usually consists of a situation, described in perhaps one or two pages, and a data set. Some typical cases are given in Bryant and Smith (1999) or in the references listed by Parr and Smith (1998).

The assignment. The general assignment is that students should determine what the data say about the situation and communicate their view in a manner appropriate to the setting. Typically we ask that they provide a one-paged summary of their results (and a lot of education happens in the attempt to condense their views into one page!), and bring it to class. This helps ensure they are prepared for the discussion, and offers practice in writing and communication skills.

We generally prefer to let students analyze the case using methods of their own devising – it gives them more ownership in the results, and leads to lively discussions. Occasionally, if the topic is particularly technical or the class not very advanced, we provide assignment questions to guide them. A mature class might work very well with a simple instruction like "How should the marketing manager proceed?" A less mature class might need such guidance as "What factors seem to influence sales? How can the manager take advantage of them? Do the same factors appear to apply in all regions?" A beginning class might need more explicit instructions: "Run the regression of sales on the demographic factors given. Interpret the coefficients, and decide which variables belong in the regression model. Compute the residuals from your final model, and interpret them."

The discussion. Many approaches to leading the discussion are possible and fruitful. The field of discussion methods in general is well developed [see Christensen (1987), for example] – most of its principles apply more or less unchanged here, though most of us in statistics have only recently begun to use them. The instructor's edition of Bryant and Smith (1999), for example, provides detailed discussion management suggestions for each case. Two possible approaches are:

- In regression-based cases involving variable selection, ask for competing models to be offered. Each can be written on the blackboard, and those proposing the different models can be asked to explain the basis for their choices. The class as a whole can then debate what the appropriate criteria should be, leading eventually to some sort of vote or consensus. It's usually important in this approach to make sure the students provide both technical and business-related justifications for their choices, not just equations.
- If regression is to be used to forecast future values of a time series, we like to ask the students to provide a forecast for the next one or two periods following the data they are given. We offer a prize (such as $20 contributed to a charity of the student's choice) for the forecast that comes closest to the actual value. The class can decide what criterion should be used in making this judgment.

4. Results

The method can be very successful, though it poses challenges for students and faculty alike. Some of the benefits include:

- Improved retention of material, presumably because the material has been used and interpreted in a context in which the student had some commitment to it, as opposed to simply memorizing it. Further, the case context reinforces the understanding of both the underlying theory and its interpretation. It avoids the common "I know many formulas but I don't know which one to apply here" complaint.

- Increased understanding of fundamental concepts, and earlier detection and correction of misunderstandings. If the climate for discussion is appropriate, students will ask the questions lurking in the background, whether simple ("I'm not sure about the difference between R^2 and adjusted R^2") or penetrating ("but does lack of statistical significance by itself mean a variable *isn't* important?") I find, for example, that students often try to interpret R^2 as the "percentage of the data points that are close to the line" or in other incorrect ways, and the discussion serves to correct these misunderstandings.
- Increased student interest in the material: the material has been *applied*, not just written down as equations. The students have discovered more of the results for themselves, and thus take more ownership in them. Finally, students from all backgrounds can contribute to the discussion. We notice, for example, that technically qualified students have a disturbing tendency to find the mathematically appropriate solution to the wrong problem, or to have difficulty interpreting the results in a general business setting. Students whose strengths are in other areas of business contribute by questioning or interpreting such analyses or presentations. They thus feel more included than in strictly lecture-oriented courses.
- More fun for the professor. The material in any elementary course becomes familiar, and faculty burnout is discouragingly real for this course, but no two classes or discussions are ever the same. They don't all go well, but they're seldom boring.
- Anecdotal evidence from instructors of subsequent courses also suggests that students are better prepared for later data analyses in those courses by virtue of being taught by discussion, too. They are used to the idea that their main value added in data analysis is the ability to look at problems and decide what kind of analysis might potentially be useful. They've had practice doing just that in their data analysis course.

There are, of course, some challenges and costs, too:

- Students and faculty must be educated to the expectations and culture of learning by discussion. Students must learn early in the course that data analysis is not going to be a spectator sport. They must come prepared to participate, to offer their views, and to listen actively to others' views. Many students appear at the first class expecting some sort of lecture, so it is important to communicate the new set of expectations quickly, before habits and attitudes harden.
- An appropriate climate must be established quickly, one in which participation is encouraged and where risk taking (in the form of offering tentative opinions) is rewarded, not punished. This usually happens (or doesn't) within the first two or three class periods, and instructors need to be sensitive to it.
- Instructors must let the students have more control over the material covered and how it's discussed. Students will make mistakes. They must discover them and be gently re-directed, either by their fellow students or by the instructor. The climate must become one in which the fundamental values are that we learn together, rather than that we come to class to be "given" the right answer.
- The discussion takes time from class, time which might be spent in lecture on other topics. This is not usually a big problem: the increased ownership in the material resulting from the discussions compensates for the lost lecture time. I felt the loss of control at first, but in retrospect, it was control I didn't really have in the first place: I could control what I *said*, but not what the students *learned*. Some may feel they "have to cover the material in the syllabus." Such policies are an unpleasant reality in many departments. On the other hand, just because I *talked* about it doesn't mean the students *learned* it. A little more motivation often means they are more likely to learn it when they need it later.

Acknowledgements

The experiences and approaches I report have been shared with and developed jointly with my colleague, Marlene A. Smith. I retain responsibility for the ideas reported here, but I cheerfully acknowledge her contributions to these ideas and to my development as a teacher.

REFERENCES

Bryant, P. G. and M. A. Smith (1999), *Practical Data Analysis: Case Studies in Business Statistics*, 2nd edition, Volumes I, II, and III. Burr Ridge, Illinois: Irwin/McGraw-Hill.

Christensen, C. R. (1987), *Teaching and the Case Method: Text and Cases*. Boston: Harvard Business School Publishing Division.

Parr, W. C. and M. A. Smith, (1998) "Developing Case-Based Business Statistics Courses," *The American Statistician 52,4*, 330 – 337.

SUMMARY

This paper describes recent experience in teaching multiple regression in the required elementary business statistics courses at the University of Colorado at Denver by emphasizing case discussion. Discussion methods are often used in business curricula, but seldom for statistics classes. The methods work, though they challenge students and faculty alike. They appear to lead to improved retention of material, increased understanding of fundamental concepts, earlier detection and correction of misunderstandings and increased student interest. Anecdotal evidence from instructors of subsequent courses suggests that students taught by discussion are better prepared for data analyses, too, but students and faculty must be educated to the expectations and culture of learning by discussion.

RÉSUMÉ

Ce travail décrit l'expérience récente d'enseignement de la regression multiple dans les cours introductoires de statistique pour gestion à l'Université de Colorado à Denver, en s'appuyant davantage sur la discussion de cas. Quoique l'approche soit de plusieures façons semblable aux méthodes habituelles de discussion utilisées dans beaucoup des curriculae de gestion, elle s'avère très rarement utilisée dans le contexte des cours de statistique. Notre expérience suggère que la méthode peut assurer un grand succés, bien qu'elle constitue un défit pour les étudiants, ainsi que pour la faculté. Elle semble conduire à une meilleure rétention du matériel, à une plus grande compréhension des concepts fondamentaux, à une plus prompte détection et correction de malentendus et à un interêt accru de la part de l'étudiant. L'expérience des enseignants des cours qui se suivent indique de même que ces étudiants sont mieux préparés pour l'analyse des données, du fait d'avoir été enseignés par la la methode de cas, mais il est important de reconnaitre que les etudiants et les professeurs doivent etre bien entrainés a utiliser les methodes de discussion et de cas.