# A STUDENT ASSIGNMENT TO COMPLEMENT IN-CLASS INSTRUCTION IN REGRESSION ANALYSIS

Penny Verhoeven, Kennesaw State University, MB 0403, 1000 Chastain Rd., Kennesaw, GA 30144, pverhoev@kennesaw.edu

Hope Baker, Kennesaw State University, MB 0403, 1000 Chastain Rd., Kennesaw, GA 30144, hbaker@kennesaw.edu

## ABSTRACT

This paper describes an assignment designed to both reinforce in-class instruction in regression analysis and demonstrate to students the broad applicability of regression theory to practice. The assignment involves students identifying an application of regression analysis in the published literature. This paper further reports on student (n = 82) performance on the assignment and student perceptions of the assignment's "value added."

## INTRODUCTION

In our college, all business majors take a junior-level quantitative methods course having as one of its major curricular components an introduction to simple and multiple linear regression. Approximately ten hours of class time are devoted to (ordinary least squares) regression. In-class instruction and associated homework exercises address: translating a verbal description of relationships between variables into a model specification; using Excel to perform a regression analysis on sample data; assessing the validity of model assumptions; testing for the significance of model coefficients; interpreting estimates of model parameters; and using a sample regression equation for estimation and prediction.

With the intention of both strengthening students' understanding of regression concepts and demonstrating to students the broad applicability of regression theory to practice, we designed an assignment whereby students are instructed to locate an application of regression analysis in the published literature. The assignment was offered as an optional bonus problem on a test to students in three sections of the aforementioned course; the students picked up a copy of the assignment upon turning in their tests. Approximately two-thirds (82) of the students elected to submit the assignment. In the sections that follow, we describe the assignment, summarize the students' performance on the assignment, and share the students' evaluations of the assignment.

## ASSIGNMENT FOR STUDENTS

The assignment instructed students to locate, within a scholarly journal, an article that: (a) posed a research question related to business; and (b) presented the results of a regression analysis that was performed to assist in addressing that research question. Instructions provided for students are as follows:

Submit a hard copy of:
(1) The first page of the article (so as to convey the name and author(s) of the article, the journal and year of publication, and an abstract of the article).

(2) The page(s) containing the specification of a regression model subsequently assessed. Note: The specification may be in the form

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \qquad (1)$$

or may be expressed or implied in words. For time-series data, a model uses the subscript t to denote the time period. If the article contains more than one regression model, select one to focus upon.

(3) The page(s) containing estimates of the model coefficients. Note: It is highly likely that those estimates will be listed in a table along with associated measures.

Mark up by hand your submission as follows:
(1) Draw a vertical line in the right margin to indicate where the regression model is specified.
(2) Draw a vertical line in the right margin to indicate where estimates of the model coefficients are provided.

The above instructions were accompanied by further assistance and guidance in the form of the following "For your information" notes, the first of which is shown in modified form so as to remove online access instructions specific to University System of Georgia institutions and the second of which refers to the study by Bebchuk and Grinstein (2005) highlighted in the Appendix.

For your information:
(1) One approach for locating an appropriate article is as follows: Access the ABI/INFORM Complete (at ProQuest) database. Enter a search term for a business-related topic or issue of interest to you.* Check the "Full text documents only" and "Scholarly journals, including peer-reviewed" boxes before requesting the search. Request the search. Peruse the listed articles for indications that a regression analysis was performed.
(2) A sample submission—related to the issue of executive pay—is attached and "marked up" as requested in the instructions.
(3) You may, of course, encounter in an article equations that are not regression models. In this regard, note that every regression model contains an error term $\varepsilon$.

*For example, you may be interested in one or more of the following: leadership styles; customer retention; quality assurance; health care costs; corporate bankruptcy; corporate ethics; leveraging technology; Sarbanes-Oxley; auditing; employee stock ownership; portfolio selection; energy conservation; developing economies; privatization.

## STUDENT PERFORMANCE ON ASSIGNMENT

We employed a grading rubric that awarded four points for each of: (1) submitting an abstract of an article related to business that employed a regression analysis, (2) correctly marking the location of a regression model specification, and (3) correctly marking the location of the estimates of the model coefficients. Partial credit (two points instead of four) was awarded for item (2) above in the case of either "extra" marking (i.e., marking extraneous discussion or an extraneous equation in addition to the actual specification of a regression model) or "incomplete" marking (e.g., when the author's regression model specification is expressed verbally as opposed to conveyed in equation form, marking reference to some but not all the variables used in the model). Partial credit (two points instead of four) was awarded for item (3) for marking an extraneous table (e.g., a table summarizing the mean and standard deviation of each variable used in the model) in addition to the table containing estimates of the model coefficients.

The distribution of student scores (out of a maximum of 12 points) on the assignment is provided in Table 1. Fifty percent of the students earned a perfect score, and all but one student submitted an application of regression analysis related to business (the one exception submitted a didactic article on regression that employed an application in education as an example). Errors related to identifying the specification of a regression model fell into four categories: no attempt at identification; incomplete identification; identification with "extra" marking; and total misidentification. Errors related to identifying the estimates of model coefficients fell into three categories: no attempt at identification; identification with "extra" marking; and total misidentification. It is noteworthy that no guidance (e.g., no in-class discussion of the assignment) was afforded students beyond the aforementioned instructions, "for your information" notes, and example of a "correctly marked up" submission.

| Score | # of Students | % of Students |
|-------|---------------|---------------|
| 12    | 41            | 50%           |
| 10    | 8             | 10%           |
| 8     | 23            | 28%           |
| 6     | 7             | 9%            |
| 4     | 3             | 4%            |

**Table 1. Student Performance on Assignment**

## STUDENT PERCEPTIONS OF ASSIGNMENT'S VALUE ADDED

To informally assess the value added by the assignment, we additionally requested students to, on the first page of their submission:

> Indicate your level of agreement (A = strongly disagree, B = disagree, C = slightly disagree, D = slightly agree, E = agree, F = strongly agree) with each of the following questionnaire items. Your responses will not affect your grade; honest responses are being requested.
>
> (1) This assignment reinforced the fact that regression analysis is an approach for assessing the relationship between a dependent variable and a set of independent variables.
> (2) This assignment reinforced what we discussed in class about model specification.
> (3) After completing this assignment, I better understand why the study of regression analysis is part of the business core curriculum.
> (4) After completing this assignment, I am more confident that I, too, could specify a regression model and gather sample data to test the model.

The distribution of student responses to each item is given in Table 2. We realize that lack of anonymity may have led some students to inflate their levels of agreement with the questionnaire items; note, however, that we explicitly requested honest responses and gave assurances that the responses would not affect the grade. The percentage of students strongly agreeing or agreeing with questionnaire items 1, 2, and 3 were 79%, 78%, and 72%, respectively. The considerably lower percentage (56%) strongly agreeing or agreeing with item 4 is to be expected and, as such, suggests that student responses across the four items involved reflection and discernment. We believe that the student responses suggest that, overall the students perceived that the assignment added value to the in-class instruction on regression analysis.

| | Questionnaire Item | | | |
|---|---|---|---|---|
| Response | 1 | 2 | 3 | 4 |
| Strongly Agree | 41% | 34% | 29% | 13% |
| Agree | 38% | 44% | 43% | 43% |
| Slightly Agree | 9% | 7% | 16% | 26% |
| Slightly Disagree | 4% | 2% | 4% | 10% |
| Disagree | 5% | 6% | 5% | 6% |
| Strongly Disagree | 4% | 6% | 4% | 2% |

**Table 2. Percentage distribution of responses to questionnaire items**

## CONCLUSION

In light of the results of our above described "trial run" of the assignment, we intend to require (rather than make optional) the assignment of all students in future sections of the course in question. Furthermore, to better prepare students for the assignment, we intend to (rather than merely provide one sample submission in writing) illustrate and discuss in class how to correctly identify a regression model specification and correctly identify the estimates of the model's coefficients with respect to three published studies; three of the innumerable studies that could be used are highlighted in the Appendix.

## REFERENCES CITED

[1] Bebchuk, Lucian, and Grinstein, Yaniv. 2005. The growth of executive pay. *Oxford Review of Economic Policy* 21(2):283-303.

[2] Chen, Ming-Chin, Cheng, Shu-Ju, and Hwang, Yuhchang. 2005. An empirical investigation of the relationship between intellectual capital and firms' market value and financial performance. *Journal of Intellectual Capital* 6(2):159-176.

[3] Sulek, Joanne M., and Hensley, Rhonda L. 2004. The relative importance of food, atmosphere, and fairness of wait: the case of a full-service restaurant. *Cornell Hotel and Restaurant Administration Quarterly* 45(3):235-247.

## APPENDIX

**Highlights of three applications of regression analysis in published research**

**Bebchuk & Grinstein (2005)**

Portion of abstract on page 283: "This paper examines both empirically and theoretically the growth of US executive pay during the period 1993-2003. During this period, pay has grown much beyond the increase that could be explained by changes in firm size, performance, and industry classification."

Regression model specification, equation (1) on page 286:

$$Log(compensation_{i,t}) = a_0 + a_1 Log(sales_{i,t-1}) + a_2 Log(1 + ROA_{i,t-1}) + a_3 Log(1 + \operatorname{Re}turn_{i,t-1}) +$$
$$a_4 Log(1 + \operatorname{Re}turn_{i,t-2}) + Yeardummies(1994 - 2003) + f_i + \varepsilon_{i,t} \qquad (A.1)$$

Estimates of model coefficients using each of two operationalizations of the dependent variable--Log(total CEO compensation) and Log(total top-5 compensation)--are in Table 2 on page 287.

**Sulek & Hensley (2004)**

Portion of abstract on page 235: "Food quality came out as the only one of nine factors being tested that had a significant effect on intent to return for 239 diners at an Irish-pub-style full-service restaurant in the southeastern United States....Examining customer satisfaction, food quality again was at the top of the list, but the restaurant's atmosphere and the fairness of the seating procedures also had significant effects."

Verbal reference on page 239 to a regression model specification: "The first full model regressed the design attributes against overall dining satisfaction..."

Estimates of model coefficients are in Exhibit 2 on page 242.

**Chen et al (2005)**

Portion of abstract on page 159: "...the authors construct regression models to examine the relationship between corporate value creation efficiency and firms' market-to-book value ratios, and explore the relation between intellectual capital and firms' current as well as future financial performance.

Regression model specification on page 164:

$$M / B_{it} = \alpha_0 + \alpha_1 VACA_{it} + \alpha_2 VAHU_{it} + \alpha_3 STVA_{it} + \varepsilon_{it}$$ (A.2)

Estimates of model coefficients are in Table VII on page 169.