Student performance in business and economics statistics: Does exam structure matter?
Krieg, Randall G;Uyar, Bulent
Journal of Economics and Finance; Summer 2001; 25, 2; ProQuest Central
pg. 229

Student Performance in Business and Economics Statistics: Does Exam Structure Matter?

Randall G. Krieg and Bulent Uyar *

Abstract

We investigate whether exam structure really matters in an economics and business statistics course by (1) determining how well performance on multiple choice questions alone explains overall performance on exams composed of both multiple choice questions and problems, and (2) discovering whether various student characteristics have significantly different impacts on student performance on multiple choice questions versus problems. Our findings suggest that student scores on multiple-choice portions of exams do not adequately determine overall student performance, and that some students are predisposed to do better on multiple choice or problems depending on the characteristics they possess. (*JEL* A22, A29)

Introduction

Every instructor administering examinations must determine their structure. Alternatives include fixed-response (e.g., multiple-choice), constructed response (e.g., problems and/or essays), or oral questions. Time constraints generally render oral exams impractical at the college level. Thus, instructors are left weighing the pros and cons of fixed-response versus constructed-response tests or some combination. Our study focuses on whether exam type really matters in an economics and business statistics course. More specifically our objective is to seek answers to the following questions: (1) Does exam structure matter to the professor, i.e., would fixed-response questions explain student performance as well as constructed-response questions? (2) Does exam structure matter to students, i.e., does one form of question enable some students to better demonstrate their understanding of the subject matter than the other form? Instructors administering multiple-choice exams have often heard students comment, "I prefer essay style (multiple choice) exams," or "I do not perform well on multiple choice (essay style) exams." We investigate which students may be predisposed to do well on both constructed response and multiple-choice exams. Moreover, we proceed to test whether students with certain individual

* Randall G. Krieg, University of Mary, 7500 University Drive, Bismarck, ND 58504, rkrieg@umary.edu. Bulent Uyar, University of Northern Iowa, Economics Department, Cedar Falls, IA 50614-0129, uyar@uni.edu. The authors would like to thank two anonymous referees and Professor J. Zietz, the editor, for very helpful comments and suggestions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

characteristics tend to perform significantly better on exams of one structure as compared to the other.

We investigate these questions empirically for an economics and business statistics course by (1) determining how well performance on multiple choice questions alone explains overall performance on exams composed of both multiple choice and constructed response questions, and (2) discovering whether various student characteristics have significantly different impacts on student performance on multiple choice versus constructed response questions. Since we are studying performance in a business and economics statistics course, our constructed-response questions are in the form of problems, and our fixed response questions take the form of multiple-choice questions. Our findings suggest that student scores on multiple-choice portions of exams do not adequately determine overall student performance, and that some students are predisposed to do better on multiple choice or problems depending on the characteristics that they possess.

Exam Performance and Exam Structure

Multiple-Choice Format as a Sufficient Determinant of Student Performance

Instructors have often mulled over the advantages and disadvantages of constructed response versus multiple-choice questions. Multiple-choice questions may encourage memorization of information, rather than the understanding of concepts, are more likely to be misconstrued by students, and leave more to chance in the form of guessing. Moreover, it has been argued that constructed-response questions better allow students to show the depth of their understanding (Saunders and Walstad 1990). On the other hand, multiple-choice questions are less time consuming to grade and are graded more objectively than problems or essays. Walstad and Becker (1994) argue that if constructed response questions measure unique aspects of subject comprehension, then the extra cost of giving this type of exam may be justified. There is a lack of empirical evidence as to which is a better measure of student performance. The limited empirical evidence that does exist indicates there may well be very little difference in what the two types of exams measure (Wainer and Thissen 1993; Bennett et al. 1991; Taub 1993; Bridgeman 1991).

Walstad and Becker (1994) research the need for the use of constructed response questions in economics exams. Using data from the Educational Testing Service for the College Board AP exams, they employ ordinary least squares to measure how well performance on the multiple-choice questions explains overall performance on economics exams. They include a constant term along with the multiple-choice score as regressors to explain the total score, which is the sum of the scores on problems (33 percent) and multiple-choice questions (67 percent). The R^2 obtained from the estimations reflects the extent to which multiple-choice scores explain overall exam performance. Walstad and Becker run one regression for Principles of Microeconomics and one for Principles of Macroeconomics obtaining R^2s of 0.94 and 0.90, respectively. They conclude that "these results suggest that the one-third contribution of the essay to the composite score adds minimal information about the student for determining the AP grade beyond what is already contained in the multiple-choice score (p. 195)." To our knowledge, this issue has not been explored for Business and Economics Statistics.

In the first stage of our study, we investigate whether the fixed-response score is sufficient to measure student performance, or whether constructed-response questions can add to our assessment of student performance. We attempt to discover how well multiple-choice scores can proxy overall student performance. For this purpose, we first apply the Walstad-Becker (1994) model to our data. Then, we estimate an alternative model to account for the fact that the weights assigned to multiple-choice and problem scores are different in the two studies. In the second stage

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of our study, we further expand the literature by investigating whether some students have a propensity to perform differently on different types of exams.

Propensities for Performance Differentials

Is it true that some students are predisposed to do better on either multiple-choice or constructed-response questions? We seek to determine whether students with particular individual characteristics have a propensity to perform significantly better on problems as compared to multiple-choice questions. If such differences do exist, it would buttress the argument for including both types of questions in examinations. For example, if only multiple-choice questions were included on an exam, this might be perceived as unfair to students with characteristics that predispose them to lower scores on such exams. Of course, we are concerned with how measurable characteristics affect performance by exam type, rather than student perceptions.

Data and Model

Data

Our sample consists of 223 students who enrolled in a business statistics course taught by one of the authors in the fall 1992, spring 1993, and fall 1993 semesters at a Midwestern university. The course is a required core course in the College of Business. We use three mutually exclusive data sources. The first and most direct source is the instructor's records. These records include student scores for both problem and multiple-choice questions on each of the four examinations, attendance record, homework record, and gender.

Our second source is questionnaires that were administered in class to obtain student-specific information on human capital, academic and socioeconomic background, financial constraints, environment, place of residence, etc.

Our third source of data is the official student records. We use these records to supplement the first two sources. Official student records provide data for such variables as ACT scores, GPA, whether a student is repeating a course, etc. Using official student records for these variables also eliminates possible overstatement of achievements by some students when self-reporting such data (Maxwell and Lopus 1994). For a more detailed discussion of these data, see Krieg and Uyar (1997).

Models

Performance as a Function of Multiple-Choice Score

In the first stage of our study, we follow the methodology of Walstad and Becker (1994) in studying exam performance as a function of performance on multiple-choice questions. Exams were not comprehensive, and students were not allowed to keep the exams. Using ordinary least squares we model the total score on each of four exams as a function of the multiple-choice score on that exam and a constant term. The value of the R^2 gives us an estimate as to how well performance on multiple-choice questions alone is likely to explain overall performance. It also indicates the extent to which constructed response questions provide significant additional information in determining student performance.

There is a potential problem, however. In these equations, the total score is the sum of multiple-choice and problem scores. Multiple-choice scores make up 67 percent of the total score in Walstad-Becker regressions. In our models, they constitute 56 percent of the total score for the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

first three exams and 57 percent of the fourth. So it is difficult to properly ascertain whether the differences between the Walstad-Becker coefficients of determination and our coefficients of determination reflect the differences in weights assigned to multiple-choice scores as opposed to differences in actual explanatory power. Therefore, as an alternative model, we estimate the percentage problem score on each exam as a function of the percentage multiple-choice score on that exam and a constant term. ("Percentage multiple-choice score" for an exam is defined as the ratio of the multiple-choice points earned by a student to the total number of multiple-choice points on that exam. The "percentage problem score" is defined similarly.)

Estimation of Problem and Multiple Choice Performance

Our next model examines the determinants of student performance on multiple-choice questions and problems based on the concept of an education production function as first put forth by Bach and Saunders (1965). We model student performance as a function of student characteristics including ability, background, attitude, peer-related factors, exam-specific factors, and environment. A total of eight equations are estimated using ordinary least squares. More specifically, we estimate an equation for the percentage multiple-choice score for each of four exams and an equation for the percentage problem score on each of the four exams.

The variables used in this study are listed and defined in Table 1. As the last column of Table 1 indicates, we expect the following variables to have a positive influence on performance on both problem and multiple-choice questions: *Homework, ColGPA, Repeat, Funding*, and *ACTmath*. Homework is intended to prepare students for upcoming exams. Student GPA (*ColGPA*) is included to capture expectations of students based on past performance, while math ACT scores (*ACTmath*) are included because of the math intensive nature of Business and Economics Statistics. Both *ACTmath* and *ColGPA* also measure student ability. Students may retake the course (*Repeat*) in an attempt to raise their grade. Such students have had prior exposure to the material and are expected to be highly motivated to succeed. Students who receive significant funding from their parents (*Funding*) may be held to high standards and receive more encouragement and support from their parents during their studies, or may be subject to stronger parental pressure to perform well. *Funding* may also be a proxy for socioeconomic status during a student's upbringing.

We expect *Absent, Transfer*, and *Workhours* to negatively influence performance on both problem and multiple-choice questions. Students who transfer from two-year colleges (*Transfer*) may have lower academic potential and/or may have received lower quality instruction (Laband and Piette 1995). Control variables in the model with no *a priori* expectations include *Gender*, *Upperclass*, *Dorms*, *Workhrs*Semload*, and *Semload*.

Estimation of Performance Differentials

In the third part of our study, we use results from the estimations of problem and multiple-choice exams to test whether students possessing particular characteristics are prone to perform better on one exam type or the other. We conduct two types of tests for this purpose. The first is a t-test for differences in individual estimated coefficients for each explanatory variable across exam types. For example, the difference between the estimated coefficients for *ACTmath* in the multiple-choice versus problems equations is tested for significance. In this case, a significantly higher estimated coefficient in the multiple-choice exam would mean that students with higher *ACTmath* scores are predisposed to perform relatively better on multiple-choice exams. The second test we use is a Chow test for the overall structural equivalency of the problem and multiple-choice regressions for each exam.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 1. VARIABLE NAMES AND DEFINITIONS

Explanatory Variables	Definitions	Means	Expected Signs
Gender	= 1 if male = 0 if female	0.52	+/-
Homework	percentage of satisfactory homework assignments turned in for the i-th unit in the course (i = 1...4)	64.29 57.30 59.40 30.23	+
Upperclass	= 1 if student is not a freshman, = 0 otherwise	0.78	+/-
Absent	percentage of Friday classes missed for the i-th unit in the course (i = 1...4). (This is considered to be a commuter school, consequently most absenteeism is on Fridays.)	5.93 10.75 12.36 12.80	-
ColGPA	College GPA at the time students were enrolled in the course	2.92	+
Repeat	1 if retaking the course in order to substitute for the grade received for it in the past, 0 otherwise	0.09	+
Dorms	1 if student is residing in a dorm, 0 otherwise	0.60	+/-
Transfer	1 if student transferred from a two-year college, 0 otherwise	0.10	-
Funding	1 if parents are the primary source of funding for education, 0 otherwise	0.43	+
Workhrs	Number of hours worked per week at a job	11.01	-
*Workhrs*Semload*	Workhrs times credit hours enrolled in current semester	161.14	+/-
ACTmath	Mathematics score on ACT	24.10	+
Semload	Credits student is taking this semester	15.09	+/-

Empirical Results

Performance as a Function of Multiple Choice Score

Following Walstad and Becker (1994), we estimated composite score as a function of multiple-choice score and an intercept term for each of the four exams administered during the semester. This reveals to what extent the multiple-choice score determines the composite score for each exam. T-statistics are given in brackets beneath the coefficients.

EXAM1 = 29.37 + 1.21 (Multiple Choice Score on Exam 1)
[21.70] [35.19]
Adjusted $R^2 = 0.84$ n = 223

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

EXAM2 = 21.55 + 1.31 (Multiple Choice Score on Exam 2)
[11.28] [28.01]
Adjusted $R^2 = 0.76$ n = 223

EXAM3 = 20.22 + 1.39 (Multiple Choice Score on Exam 3)
[15.29] [36.90]
Adjusted $R^2 = 0.85$ n = 223

EXAM4 = 19.82 + 1.35 (Multiple Choice Score on Exam 4)
[14.62] [37.67]
Adjusted $R^2 = 0.85$ n = 223

Thus, with the exception of the second exam, multiple-choice score explains approximately 85 percent of student performance on exams. For the second exam the effect is limited to 76 percent. In order to compare these results with Walstad-Becker's results and determine whether it is essential to use problems as well as multiple-choice questions on exams, we employ Fisher's z-transformation test. Test results show that our multiple-choice scores have significantly less explanatory power than that obtained by Walstad and Becker, leading us to conclude that it is valuable to employ essay/problem questions as well as multiple-choice questions in Business and Economics Statistics.

Walstad and Becker's (1994, p. 195) R^2s are 0.94 and 0.90 for their micro and macro regressions, respectively. Based on these, they contend that the constructed response scores provide "minimal information about the student for determining the (composite grade) beyond what is already contained in the multiple choice score." Using Fisher's z-transformation, we tested the null hypothesis that $\rho_1 = \rho_2$ against the alternative that $\rho_1 < \rho_2$. Fisher's z-transformation is

$$z = \frac{(X_1 - X_2)}{\sqrt{\frac{1}{(n_1 - 3)} + \frac{1}{(n_2 - 3)}}}.$$

This z is approximately normal for relatively large sample sizes and

$$X_i = \frac{1}{2}\ln\left(\frac{1+R_i}{1-R_i}\right)$$

where R_i is the i-th sample correlation coefficient. For our regressions, we have $R_1 = 0.92195$ (from our highest $R^2 = 0.85$) and $n_1 = 223$. For Walstad-Becker (1994) runs, we have $R_2 = 0.94868$ (from their lowest $R^2 = 0.90$) and $n_2 = 4{,}876$. We find that the results of Fisher's z test are significant at $p < 0.001$ which leads us to reject the null hypothesis that the population correlation coefficients for our test scores (ρ_1) and Walstad and Becker's (ρ_2) are the same. The results indicate that our samples come from a population with a significantly lower correlation coefficient. For details on Fisher's z test, see Kenkel (1981, Ch. 15.5).

Because of concerns noted above, we also estimate the percentage problem score on each exam as a function of percentage multiple-choice score. This allows us to control for the differing weights given to multiple choice and problem questions in our study. The results reveal to what extent the multiple choice score determines the problem score on each exam. T-statistics are given in brackets beneath the coefficients.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Percentage Problem Score on Exam 1 = 0.667 + 0.265 (Percentage Multiple Choice Score on Exam 1)
[21.70] [6.07]
Adjusted $R^2 = 0.13$ n = 223

Percentage Problem Score on Exam 2 = 0.490 + 0.397 (Percentage Multiple Choice Score on Exam 2)
[11.27] [6.67]
Adjusted $R^2 = 0.15$ n = 223

Percentage Problem Score on Exam 3 = 0.459 + 0.499 (Percentage Multiple Choice Score on Exam 3)
[15.28] [10.40]
Adjusted $R^2 = 0.31$ n = 223

Percentage Problem Score on Exam 4 = 0.461 + 0.465 (Percentage Multiple Choice Score on Exam 4)
[14.62] [9.78]
Adjusted $R^2 = 0.28$ n = 223

We see that the coefficients of determination are consistently small. For the first two exams, the adjusted coefficients of determination are as small as 13 and 15 percent, respectively. Explanatory power increases for exams three and four but only to 31 and 21 percent, respectively. These results show that it is important to use both multiple choice and essay/problem questions in Business and Economics Statistics.

As suggested by an anonymous referee, we also estimated the percentage multiple-choice score as a function of the percentage problem score and a constant term for each exam. The results are shown below. (Adjusted R^2s and sample sizes are, of course, the same as for the corresponding equations given in the text.)

Percentage Multiple Choice Score on Exam 1 = 0.259 + 0.498 (Percentage Problem Score on Exam 1)
[3.69] [6.07]
Percentage Multiple Choice Score on Exam 2 = 0.410 + 0.390 (Percentage Problem Score on Exam 2)
[8.88] [6.67]
Percentage Multiple Choice Score on Exam 3 = 0.133 + 0.618 (Percentage Problem Score on Exam 3)
[2.88] [10.40]
Percentage Multiple Choice Score on Exam 4 = 0.179 + 0.609 (Percentage Problem Score on Exam 4)
[3.72] [9.78]

Future studies concerned with the importance of including problem sets in Business and Economics Statistics exams might find it fruitful to investigate the relative predictive powers of multiple choice and problem scores. If, for example, multiple-choice scores predict future performance better than problem scores, "we might judge them to be a better measure of learning," as noted by a referee. This, however, is beyond the scope of the present paper.

Determinants of Problem and Multiple Choice Performance

Estimation results from the percentage score on multiple-choice questions and the percentage score on problem questions are given in Tables 2 and 3, respectively. Variables with a significantly positive influence on multiple-choice scores include *Gender* (exam 1), *ColGPA* (all exams), *Homework* (exams 1 and 4), *Upperclass* (exam 2), *Repeat* (exams 1 and 3), *Funding* (exam 4), and *ACTmath* (all exams). Variables that have a significantly negative effect on multiple-choice performance are *Absent* (exams 1, 3, and 4), *Transfer* (exams 1 and 3), *Workhrs* (exam 2), and *Workhrs***Semload* (exam 2).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.

TABLE 2. ESTIMATED COEFFICIENTS MULTIPLE-CHOICE (T-STATISTICS)

	Exam 1	Exam 2	Exam 3	Exam 4
Gender	0.0390*	0.0203	-0.0186	-0.0199
	(2.06)	(1.11)	(0.90)	(1.06)
Absent	-0.0013**	-0.0003	-0.0014**	-0.0010**
	(2.52)	(0.73)	(2.47)	(2.48)
ColGPA	0.1294**	0.1337**	0.1332**	0.1129**
	(6.19)	(6.54)	(5.88)	(5.31)
Homework	0.0005*	0.0002	4.4E-05	0.0009**
	(1.64)	(0.74)	(0.14)	(3.07)
Upperclass	0.0326	0.0384*	0.0191	-0.0080
	(1.34)	(1.64)	(0.73)	(0.33)
Repeat	0.0705*	0.0192	0.1034**	0.0357
	(1.93)	(0.55)	(2.67)	(1.00)
Dorms	-0.0176	-0.0200	-0.0082	0.0031
	(0.74)	(0.87)	(0.32)	(0.13)
Transfer	-0.0952**	-0.0203	-0.0854*	-0.0223
	(2.47)	(0.55)	(2.06)	(0.59)
Funding	0.0146	0.0018	0.0047	0.0257*
	(0.76)	(0.10)	(0.23)	(1.36)
Workhrs	-0.0027	-0.0157*	-0.0079	-0.0006
	(0.37)	(2.24)	(1.00)	(0.09)
*Workhrs**	0.0002	0.0010*	0.0005	1.4E-06
Semload	(0.36)	(2.08)	(0.88)	(0.003)
Semload	0.0023	-0.0108	-0.0010	0.0133
	(0.25)	(1.22)	(0.11)	(1.47)
ACTmath	0.0136**	0.0101**	0.0093**	0.0078**
	(4.63)	(3.58)	(2.90)	(2.71)
Constant	-0.1181	0.2145	0.0251	-0.0837
	(0.78)	(1.48)	(0.15)	(0.56)
Adjusted R^2	0.333	0.267	0.248	0.330
d. f.	209	209	209	209

Notes: ** indicates significance at the 1 percent level. * indicates significance at the 5 percent or 10 percent level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 3. ESTIMATED COEFFICIENTS FOR PROBLEMS (T-STATISTICS)

	Exam 1	Exam 2	Exam 3	Exam 4
Gender	-0.0026	0.0230	-0.0224	-0.0236
	(0.17)	(1.22)	(1.19)	(1.37)
Absent	-0.0005*	-0.0018**	-0.0012**	-0.0009**
	(1.36)	(3.88)	(2.46)	(2.43)
ColGPA	0.0562**	0.1154**	0.0748**	0.0930**
	(3.28)	(5.49)	(3.62)	(4.78)
Homework	0.0006**	0.0003	0.0008**	0.0008**
	(2.34)	(0.87)	(2.71)	(2.89)
Upperclass	-0.0033	-0.0017	-0.0167	-0.0253
	(0.16)	(0.07)	(0.69)	(1.14)
Repeat	0.0341	0.0895**	0.0281	0.0172
	(1.14)	(2.49)	(0.79)	(0.53)
Dorms	-0.0095	-0.0174	0.0322	0.0174
	(0.49)	(0.74)	(1.39)	(0.81)
Transfer	-0.0268	-0.0439	0.0121	-0.0074
	(0.85)	(1.15)	(0.32)	(0.21)
Funding	0.0334*	0.0205	0.0008	0.0339*
	(2.13)	(1.09)	(0.04)	(1.96)
Workhrs	-0.0144**	-0.0070	-0.0179**	-0.0021
	(2.40)	(0.97)	(2.48)	(0.31)
*Workhrs**	0.0010*	0.0003	0.0012**	0.0002
Semload	(2.44)	(0.68)	(2.48)	(0.35)
Semload	-0.0076	-0.0104	-0.0146	0.0035
	(1.00)	(1.14)	(1.61)	(0.42)
ACTmath	0.0094**	0.0124**	0.0096**	0.0023
	(3.88)	(4.28)	(3.28)	(0.89)
Constant	0.5307	0.3065	0.4968	0.3698
	(4.27)	(2.06)	(3.35)	(2.69)
Adjusted R^2	0.213	0.290	0.238	0.256
d. f.	209	209	209	209

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As reported in Table 3, performance on problem sets increases significantly with *ColGPA* (all exams), *Homework* (exams 1, 3, and 4), *Repeat* (exam 2), *Funding* (exams 1 and 4), *Workhrs*Semload* (exams 1 and 3), and *ACTmath* (exams 1, 2, and 3). Variables with a significantly negative influence on problem set score include *Absent* (all exams), and *Workhours* (exams 1 and 3).

Differing Performance between Problem and Multiple Choice Questions

Next we examine whether the estimated coefficients of the explanatory variables differ significantly across the multiple-choice and the problem components for each of the four exams. For this purpose, we conduct both t-tests and Chow tests.

The t-tests are for differences in individual coefficients for each variable across the two exam types. More formally, let:

$$y_{1i} = \beta_{0i} + \Sigma_j \beta_{ji} x_{ji} + v_{1i}$$
$$y_{2i} = \alpha_{0i} + \Sigma_j \alpha_{ji} x_{ji} + v_{2i} .$$

y_{1i} and y_{2i} are the percentage multiple-choice and percentage problem scores, respectively, on exam i (= 1, ..,4); x_j is personal characteristic j (= 1, ..., J); β_{ji} and α_{ji} are the parameters for the j-th variable in the multiple choice and problems equations, respectively, for exam i; and v_{1i} and v_{2i} are the residuals. The t-test is for differences in individual coefficients for each variable across the two exam types. Therefore, the hypotheses are:

$$H_0 : \beta_{ji} = \alpha_{ji} \text{ and } H_1 : \beta_{ji} \neq \alpha_{ji} \text{ for each individual variable j, given exam i.}$$

The test statistic is:

$$t = (b_{ij} - a_{ij}) /([var(b_{ij}) + var(a_{ij})] /n)^{1/2} .$$

Var (b_{ij}) and var(a_{ij}) are the variances of the coefficients estimated from the two equations. We have a large sample, so we actually use the standard normal table for the critical values.

The results of these tests are given in Table 4. They indicate that the estimated coefficients for all variables, without exception, differ significantly for multiple-choice and problem questions by exam for at least three of the four exams. In cases where a given explanatory variable contributes to superior performance on multiple-choice questions for a particular exam, *MC* is inserted into Table 4. For example, since high college GPAs predispose students to do significantly better on multiple-choice questions (as opposed to problems) on exams, *MC* was inserted in column 1, row 3. Similarly, entries of PR in Table 4 indicate a statistically significant predisposition to do relatively better on problems than multiple-choice questions.

It is interesting to note that statistically significant differences exist for the great majority of comparisons (48 of 52). This illustrates that personal characteristics do have a substantial bearing on student performance by exam structure. One's propensity to perform relatively better on multiple-choice questions is enhanced consistently by *ColGPA* and *Upperclass*. Since multiple-choice questions (computer graded) are not as prevalent in primary and secondary schools, *Upperclass* may reflect one's experience in taking such exams; *ColGPA* may reflect past success or proven ability in taking multiple-choice exams, which others (with low GPAs) are not yet comfortable with.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 4. PERFORMANCE DIFFERENCES BETWEEN PROBLEMS AND MULTIPLE CHOICE QUESTIONS

	Exam 1	Exam 2	Exam 3	Exam 4
Gender	*MC***	-----	*MC**	*MC**
Absent	*PR***	*MC***	*PR***	*PR**
ColGPA	*MC***	*MC***	*MC***	*MC***
Homework	*PR***	*PR***	*PR***	*MC***
Upperclass	*MC***	*MC***	*MC***	*MC***
Repeat	*MC***	*PR***	*MC***	*MC***
Dorms	*PR***	----	*PR***	*PR***
Transfer	*PR***	*MC***	*PR***	*PR***
Funding	*PR***	*PR***	*MC**	*PR***
Workhrs	*MC***	*PR***	*MC***	*MC**
Workhrs Semload*	*PR***	*MC***	*PR***	*PR***
Semload	*MC***	----	*MC***	*MC***
ACTmath	*MC***	*PR***	----	*MC***

Notes: "*MC*" indicates that students possessing the characteristic are predisposed to perform better on multiple-choice exams. "*PR*" indicates that students possessing the characteristic are predisposed to perform better on problems.

Other explanatory variables that lend themselves to relatively superior multiple-choice performance include *Repeat*, *Workhrs*, *Semload*, *Gender*, and *ACTmath*. Students repeating the course have added experience with similar multiple-choice questions. Students who perform well on the math ACT have demonstrated their ability to successfully complete multiple-choice questions.

Attributes that contribute to relatively superior performance on problems include *Absent*, *Homework*, *Dorms*, *Transfer*, *Funding*, and *Workhrs*Semload*. Virtually all homework assignments consist of problems, explaining the significance of *Homework*. Living on campus gives students easier access to classmates and professors, possibly facilitating their ability to work through problems. It is not surprising that attendance is important since problem solving is stressed in lectures.

We also conduct Chow tests for the overall structural equivalency of the equations for percentage multiple choice score and percentage problem score for each exam. The F-values for the four exams are 14.78, 2.53, 10.81, and 7.03, respectively. These F-values are highly significant

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.

($p \leq 0.005$), lending further support to our contention that explanatory variables differ significantly across multiple choice and problem type exams, and that some students are predisposed to perform differently on different types of exams depending on the characteristics they possess.

Conclusion

A major contribution of our paper is that, to our knowledge, it is the first study that extends the literature of student performance and exam structure into the area of Business and Economics Statistics. First, we investigated how well multiple-choice question performance by students proxies overall performance. This is important since it is much more expedient for instructors to use such objective exam structures. If multiple choice scores adequately proxy composite performance on both types of questions, there would be no need to include problem sets. In this study, we estimated the Walstad-Becker (1994) model along with an alternative model that controls for the weighting of question types. Results confirm that both multiple-choice and essay/problem type questions are important in evaluating student performance in Business and Economics Statistics.

Second, we tested whether students with particular personal characteristics may have an advantage when taking either problem or multiple-choice questions. That is, are some students predisposed to perform better on multiple-choice versus constructed response questions? We found that personal characteristics have a substantial bearing on student performance by exam structure. In fact, statistically significant differences exist for all explanatory variables in at least three of the four exams. Our most consistent results are that high college GPAs, gender, freshman status, and semester load contribute strongly to a relatively better performance on multiple choice questions and living on campus enhances performance on problems.

These results buttress our contention that, in Business and Economics Statistics, it is important to incorporate both types of exam questions. First, since multiple-choice questions do not adequately explain overall exam performance, we would miss crucial information in assessing student performance if we omitted problem sets. Second, since some students are predisposed to perform better on one type of question than another, it may be seen as unfair to some students if problem sets are omitted.

This is a campus-specific study based on a "local" sample, which may limit the extent to which its results can be generalized to all introductory Business and Economic Statistics courses. Yet the local nature of the data also allows us to control for some very important variables (e.g., instructor, textbooks, curricular expectations, etc.) and introduces consistency in design and methodology (see Krieg and Uyar 1997; Romer 1993; Myatt and Waddell 1990). Because there is a lack of such studies in Business and Economics Statistics courses, future studies of Business and Economics Statistics are needed to provide a more complete picture of exam structure and student performance.

References

Bach, G. L., and P. Saunders. 1965. "Economic Education: Aspirations and Achievements." *American Economic Review* 55: 329-356.

Bennett, Randy E., Donald A. Rock, and Minhwei Wang. 1991. "Equivalence of Free-Response and Multiple-Choice Items." *Journal of Educational Measurement* 28: 77-92.

Bridgeman, Brent. 1991. "Essays and Multiple Choice Tests as Predictors of College Freshman GPA." *Research in Higher Education* 32: 319-332.

المنارة للاستشارات
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Kenkel, James L. 1981. *Introductory Statistics for Management and Statistics.* Boston, Mass: PWS/Duxbury.

Krieg, Randall G., and Bulent Uyar. 1997. "Correlates of Student Performance in Business and Economic Statistics." *Journal of Economics and Finance* 21: 65-74.

Laband, D. N., and M. J. Piette. 1995. "Does Who Teaches Principles of Economics Matter?" *AEA Papers and Proceedings* 85: 335-338.

Maxwell, Nan L., and Jane S. Lopus. 1994. "The Lake Wobegon Effect in Student Self-Reported Data." *AEA Papers and Proceedings* 84: 201-205.

Myatt, A., and C. Waddell. 1990. "An Approach to Testing the Effectiveness of the Teaching and Learning of Economics in High School." *Journal of Economic Education* 21: 355-363.

Romer, D. 1993. "Do Students Go to Class? Should They?" *Journal of Economic Perspectives* 7: 167-74.

Saunders, Phillip, and William Walstad. 1990. *The Principles of Economics Course: A Handbook for Instructors.* New York: McGraw-Hill.

Taub, Ross E. 1993. "On the Equivalence of the Traits Assessed by Multiple Choice and Constructed-Response Tests." In *Construction versus Choice in Cognitive Measurement,* edited by R. E. Bennett and W. C. Ward. Hillsdale, NJ: Lawrence Elbaum.

Wainer, Howard, and David Thissen. 1993. "Combining Mutiple-Choice and Constructed-Response Test Scores: Toward a Marxist Theory of Test Construction." *Applied Measurement in Education* 6: 103-118.

Walstad, William B., and William E. Becker. 1994. "Achievement Differences on Multiple-Choice and Essay Tests in Economics." *AEA Papers and Proceedings* 84: 193-196.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission. www.