INSURANCE OR SIGNALS? STUDENT PERCEPTION OF PRE-FINAL TESTS AN EXPERIMENTAL ANALYSIS

Abdulwahab Sraiheen, Kutztown University, Kutztown, PA 19530;(610) 683-4597.
James Ogden, Kutztown University, Kutztown, PA 19530;(610) 683-4596.
Ram Mudambi, University of Buckingham, Buckingham MK18 1EG, UK; 44-280814080

ABSTRACT

Do students perceive pre-final tests (quizzes, midterms etc.) as providing insurance against a poor final exam, or as means of signalling their "quality" to the instructor?
Theoretical work on insurance markets indicates that the individuals most at risk would seek the most insurance coverage (adverse selection) and that the acquisition of the insurance causes agents to alter their behavior, increasing their risk of loss (moral hazard). The insurance hypothesis predicts that poorer students would take pre-final tests (tests as insurance); and further that a greater number of pre-final tests would *ceteris paribus*, cause final performance to decline.
In the literature on market signalling, it is observed that the high-quality agents have an incentive to signal their type, as it costs them less to create a favorable signal. The signalling hypothesis predicts that better students would take more pre-final tests (tests as signals), and hence a greater number of pre-final tests would be associated with a better final performance.
Using data set of 528 students taking a single course over a six year period, we are able to strongly reject the insurance hypothesis. It would appear that in the aggregate students view pre-final tests as a signalling mechanism.

INTRODUCTION

A number of studies have investigated the impact of individual characteristics on student performance (Park and Kerr, 1990; Borg et al, 1989; Leppel, 1984 and Spector and Mazzeo, 1980). Further, a number of studies have examined the effect of differential test presentation on student performance (Chidomere, 1989; Gohmann and Spector, 1989; Schmitt and Scheirer, 1977 and Marso, 1970). In this paper the analysis considers students' behavior with respect to pre-final tests (i.e., quizzes, midterms etc.) and the effect of this behavior on their subsequent performance in the final examination and the course. Given that a course is a risky endeavor with a variable pay-off (a grade), pre-final tests which contribute to the course grade may be viewed by students in two different ways. They may either be viewed as *providing insurance* against a poor performance on the final examination or as a means of *signalling student abilities* to the instructor. These two perspectives lead to testable different theoretical predictions.
It is possible that extraneous psychological issues of self-perception may affect students' behavior toward tests. If students except a failure will indicate incompetence, they may intentionally reduce effort so that failure can be attributed to low effort, rather than low ability (Eshel and Kurman, 1991). There is evidence suggesting that such effort-reducing strategies are *not* pursued by students (Jagacinski and Nicholls, 1990). Thus, it is assumed that tests results are relatively unaffected by considerations of self-perception. Further, if students perceptions of their own performance change over the semester, this would affect their decisions with regard to pre-final tests, i.e., they may begin the semester with the expectation of doing well, but change this as they gain experience in the course (or vice versa). However, there is evidence indicating that students expectations of their own performance shows great stability through the semester (Ortinau and Bush, 1987). Thus, student decisions with regard to pre-final tests are considered to be chronologically unbiased.
If students view pre-final tests as providing insurance, then it follows that the students facing the highest risk of poor performance (loss) will tend to be disproportionately represented in the set of those seeking insurance (adverse selection). A further implication is that insured students will tend to modify their behavior so that their risk of poor performance or loss increases (moral hazard). On the other hand, if students view pre-final tests as a means of signalling their abilities, higher quality students (those facing the lowest risk of poor performance) have the strongest incentive to signal, as they can create a favorable signal with the least effort. The two perspectives lead to conflicting hypotheses. The insurance (IN) hypothesis leads to the predication that the poorer students would take more pre-final

tests, and further that a greater number of pre-final tests would, *ceteris paribus*, cause final exam performance to decline. The signalling (S) hypothesis leads to the prediction that the better students would take more pre-final tests and that a greater number of pre-final tests would be associated with a better final exam performance. The IN hypothesis of adverse selection and moral hazard is strongly rejected. This leads us to the conclusion that pre-final tests are generally *not* perceived as risk alleviating instruments by students. Rather, they appear to be seen as instruments of signalling superior ability.

These results indicate that increasing the number of grade-related activities in a course is unlikely to increase students' perceived welfare. This is because the increased work is *not* seen as an increase in availability of insurance and because a small number of pre-final activities is sufficient to create a favorable signal.

The paper is organized as follows. Section 2 describes the experimental procedure and the hypotheses under test. Section 3 contains a discussion of the results. Section 4 concludes the paper.

PROCEDURE

The sample consisted of students taking a statistical methods course from one of the authors over the period Fall 1985 to Spring 1991. A usable sample of 528 was generated.

There were two **Types** semesters. In type 1 semesters, subjects were allowed to choose the number of quizzes that they would take, but all the exams were mandatory. They were able to vary the weight of the final in their course grade from a minimum of 35% to maximum of 45%. In Type 2 semesters, subjects were allowed to choose both the number of quizzes and the number of exams that they would take. They were able to vary the weight of the final in their course grade from a minimum of 35% to a maximum of 85%. The variable CHOICE is defined to be the weight of the final in the overall grade. As described above, this variable has a minimum value of 0.35. It has a maximum value of 0.45 for Type 1 semesters and 0.85 for Type 2 semesters. An decreasing value for CHOICE implies the purchase of more insurance.

The first hypothesis test concerns the relationship between subject-specific risk exposure and choice with regard to insurance. In the experiment, subjects' risk exposure (or ability) was proxied by cumulative GPA on entry into the course; the higher the GPA, the lower the subject's risk of obtaining a poor grade. Both the IN and S hypotheses offer predictions with regard to the nature of the relationship between CHOICE and GPA. Thus in the relationship.

$$\text{CHOICE} = a_0 + a_1\,\text{GPA} + a_2\,\text{ATT} + e \qquad (1)$$

the IN hypothesis predicts that $a_1 > 0$. A secondary prediction would be that $a_0 > 0.35$, i.e., subjects with extremely high risk exposure should buy as much insurance as possible. If the number of pre-final tests is considered by the low-risk (high GPA) subjects, to be a signal indicating their risk-status to the instructor, then the S hypothesis predicts that $a_1 < 0$. The second hypothesis test concerns the relationship between the amount of insurance taken (number of signals emitted) and the outcome for the subject. Two measures of outcome for the subjects were used. The first measure used was the score obtained by the subjects on the (mandatory) final examination; this variable is called FINAL. The second measure was the overall semester score obtained by the subjects; this variable is called SCORE. Again, the IN and S hypotheses offer opposing predictions with regard to the relationship between dependent variables FINAL or SCORE and the independent variables GPA and CHOICE, normalizing for commitment using ATT. In the relationships

$$\text{FINAL} = b_0 + b_1\,\text{CHOICE} + b_2\,\text{GPA} + b_3\,\text{ATT} + e \qquad (2)$$

$$\text{SCORE} = c_0 + c_1\,\text{CHOICE} + c_2\,\text{GPA} + c_3\,\text{ATT} + e \qquad (3)$$

the IN hypothesis predicts that $b_1 > 0$ and $c_1 > 0$. And since the S hypothesis predicts that the low-risk (high ability) subjects would emit more signals, it implies that $b_1 < 0$ and $c_1 > 0$). And since the S hypothesis predicts that the low-risk (high ability) subjects would emit more signals, it implies that $b_1 < 0$ and $c_1 < 0$. In order to validate that GPA is indeed a measure of risk exposure (or subject ability), it is expected that $b_2 > 0$ and $c_2 > 0$, i.e., higher GPA is associated with better performance in the course. In order to take subjects' perception of 'good' and 'bad' outcomes into account, GPA is used to proxy a subject's expected grade; denoting the course grade obtaining by the subject by GRADE, the relationship

$$\text{NETGRADE} = d_0 + d_1\,\text{CHOICE} + d_3\,\text{ATT} + e \qquad (4)$$

is set up. The IN hypothesis then predicts that $d_1 > 0$, while the S hypothesis predicts the reverse. Finally, while variance is generally accepted to be a measure of risk, its significance as a measure of dispersion must be taken in account. Students' choosing to accept higher dispersion may not always coincide with their acceptance of higher risk exposure. In particular, students who expect to score at the high end of a particular grade may prefer a distribution of outcomes with greater dispersion to increase the probability of moving into a higher grade. For example, a student with an expected grade of 79 may prefer a distribution with a larger dispersion, since this may greatly increase the probability of obtaining a 'B' (rather than a 'C'), while the probability of obtaining a 'D' may increase by little, if at all. This effect causes a demand for increased dispersion (and fewer pre-final tests) by students who expect to score at the upper margin of any grade, and is independent of IN and S hypotheses. Therefore the sample is censored to remove these "upper marginal" students.

RESULTS

The first hypothesis test is based on an estimation of equation (1).

$$\text{CHOICE} = .486 - .041\text{GPA} + .0002\text{ATT} \qquad (1.1)$$
$$t \qquad (49.88) \quad (14.25) \quad (1.84)$$
$$\text{Adj. } R^2 = .518 \quad \text{F-Ratio} = 117.65$$

$$\text{CHOICE} = .8999 - .1417\text{GPA} + .00043\text{ATT} \qquad (1.2)$$
$$t \qquad (28.83) \quad (13.34) \quad (1.06)$$
$$\text{Adj. } R^2 = .599 \quad \text{F-Ratio} = 122.64$$

For all tabulations of the data, the IN hypothesis is soundly rejected. The GPA coefficient is always negative with a 't' statistic always at an extremely high level. This indicates that contrary to the IN hypothesis, increasing amounts of insurance are taken by subjects with successively lower exposure to risk. Further, while the IN hypothesis predicts that the constant term should approach 0.35, this prediction is strongly rejected, in favor of a larger value. *Thus, the experimental results overwhelmingly favor the S hypothesis.*

The second hypothesis test is based on the impact of CHOICE on FINAL and SCORE, after normalizing for GPA and ATT. Using SCORE as the dependent variable would appear to be preferable, since the grade is assumed to be the subjects' payoff. However, it suffers from the drawback that it is already partly determined on the basis of performance in pre-final tests. Thus, while FINAL is not the ultimate payoff, it has no pre-determined component. The results obtained from using either of these two variables as dependent are remarkably similar.

Regressing either FINAL or SCORE directly on CHOICE, GPA and ATT was problematic as the first two regressors are highly correlated. The correlation coefficient between CHOICE and GPA was -0.718 for Type 1 semesters and -0.775 for Type 2 semesters. Thus, using these regressors directly would pick up only the direct effects, and not reflect the effects of GPA on CHOICE. This would cause the effects of GPA on CHOICE. This would cause the effects of GPA to be understated. The researchers follow Park and Kerr (1990) by filtering CHOICE through GPA, i.e., using the residuals from the regression of CHOICE on GPA. These residuals are designated to be the filtered (net) values of CHOICE, and denoted by CHOICE'.

Thus, in order to run the second hypothesis test, equations (2) and (3) are estimated with CHOICE' in the place of CHOICE. On the basis of both versions of the test, the IN hypothesis is strongly rejected. Using either SCORE or FINAL as the dependent variable, the CHOICE' coefficient is always negative with a corresponding 't' statistic at an extremely high level. This indicates that contrary to the IN hypothesis, subjects taking on a high level of insurance are less likely to suffer the loss (i.e., a poor grade) than subjects taking on a low level. Further, GPA is indeed validated as a measure of risk as its coefficient is positive with a very high 't' statistic. *Again, the experimental results strongly favor the S hypothesis.*

$$\text{FINAL} = 0.262 + 17.441\text{GPA} - 7.346\ \text{CHOICE}'$$
$$t \qquad (0.09) \quad (21.28) \quad (13.79)$$
$$+.063\ \text{ATT} \qquad (2.1)$$
$$(1.67)$$
$$\text{Adj. } R^2 = .789 \quad \text{F-Ratio} = 407.81$$

$$\text{FINAL} = 1.166 + 18.409\text{GPA} - 4.878\ \text{CHOICE}'$$
$$t \qquad (0.53) \quad (24.43) \quad (9.85)$$
$$+.051\ \text{ATT} \qquad (2.2)$$
$$(1.78)$$
$$\text{Adj. } R^2 = .847 \quad \text{F-Ratio} = 451.96$$

$$\text{SCORE} = 10.953 + 13.186\ \text{GPA} - 2.64\ \text{CHOICE}' +$$
$$t \qquad (5.06) \quad (20.79) \quad (6.41)$$
$$.194\ \text{ATT} \qquad (3.1)$$

Adj. R^2 = .791 F-Ratio = 274.86

SCORE = 5.088 + 13.999 GPA - 3.455 CHOICE'+
t (5.06) (20.79) (6.41) (3.2)
.2225 ATT
Adj. R^2 = .791 F-Ratio = 274.86

The guard against the possibility that the above tests have ignored subject-specific perceptions of outcome quality, equation (4) is estimated.

NETGRADE= .439 - 4.857CHOICE + .013 ATT(4)
t (.95) (5.21) (5.63)
Adj. R^2 = .267 F-Ratio = 40.61

The results from this estimation merely reinforce the results described above; the IN hypothesis is again rejected in favor of the S hypothesis.

CONCLUSIONS

In this paper an experimental study to investigate students' perceptions of pre-final tests is described. Such tests are widely used in assigning overall course grades in US universities. It is pointed out that such pre-final exam testing serves two purposes: (a) It distributes the course grade over a number of tests and provides insurance against a poor performance in the final exam. (b) It signals the student's abilities to the instructor. The objective is to test which of these attributes is considered important by students in their decision-making. The hypothesis that pre-final tests are perceived as grade insurance implies the properties of adverse selection and moral hazard. The competing hypothesis that the tests are perceived as a signalling mechanism implies that the best students provide the most signals, since they can create favorable signals with the least effort. The experimental results strongly reject the insurance hypothesis, and consequently support the signalling hypothesis. These results suggest that the provision of optional grade-related work reinforces existing inequalities in student abilities. Since grading is typically done on a relative scale, such work will probably increase, rather than reduce the dispersion in the course grades. In addition, we have the advantage of CHOICE' and GPA being orthogonal to each other.

REFERENCES

Borg, M.O., Mason, P.M., and S.L. Shapiro (1989), "The Case of Effort Variables in Student Performance". The Journal of Economic Education, 19, 308-313.

Chidomere, R.C. (1989), "Test Item Arrangement and Student Performance in Principles of Marketing Examination: A Replication Study", Journal of Marketing Education, Fall 11, 37.

Eshel, Y. and J. Kurman (1991), "Academic Self-Concept, Accuracy of Perceived Ability and Academic Attainment", British Journal of Education Psychology, 61 (2), 187-196.

Gohmann, S.F. and L.C. Spector (1989), "Test Scrambling and Student Performance", Journal of Economic Education, 19, 235-238.

Jagacinski, C.M. and J.G. Nicholls (1990), "Reducing Effort to Protect Perceived Ability: They'd Do It but I Wouldn't", Journal of Educational Psychology, 81 (2), 15-17.

Leppel, K. (1984), "The Academic Performance of Returning and Continuing College Students: An Economic Analysis", Journal of Economic Education, 15, 46-54.

Marso, R.N. (1970), "Test Item Arrangement, Testing Time and Performance", Journal of Educational Measurement, 7, 113-118.

Ortinau, D.J. and R.P. Bush (1987), "The Propensity of College Students to Modify Course Expectations and its Impact on Course Performance Information", Journal of Marketing Education, Spring, 9, 45.

Park, K.H. and P.M. Kerr (1990), "Determinants of Academic Performance: A Multinomial Logit Approach", Journal of Economic Education, 20, 101-111.

Schmitt, J.C. and C.J. Scheirer (1977), "The Effect of Item Order on Objective Tests", Teaching of Psychology, 4, 144-145.

Spector, L.C. and M. Mazzeo (1980), "Probit Analysis and Economic Education", Journal of Economic Education, 11, 37-44.

Spence, A.M. (1973a), "Job Market Signalling", Quarterly Journal of Economics, 87, 355-374.