Economics Bulletin

Volume 36, Issue 1

Risk Preference and Student Behavior on Multiple-Choice Exams

Zhou Yang
Robert Morris University

Martin Tackie
Ernst & Young LLP

Abstract

This paper modifies standard multiple-choice form of assessment by adding an intertemporal bonus point mechanism in the assessment. The modified mechanism enables us to uncover traditionally unobserved student behavior and provides insights into strategic behavior of students. Based on panel data on students enrolled in an economics course, our results suggest that risk preference plays an important role in explaining student strategic behavior on multiple-choice exams. In addition, the findings shed light on the possibilities to enhance student performance through improvements on their learning habits based on their risk preferences.

Citation: Zhou Yang and Martin Tackie, (2016) "Risk Preference and Student Behavior on Multiple-Choice Exams", *Economics Bulletin*, Volume 36, Issue 1, pages 58-67
Contact: Zhou Yang - yangz@rmu.edu, Martin Tackie - martin.tackie@ey.com.
Submitted: September 07, 2015. **Published:** February 04, 2016.

1. Introduction

Multiple-choice assessment format and incentive mechanisms are common elements of course syllabi for large classes. These features often have some undesirable characteristics that may encourage students to indulge in strategic behavior. Moreover, students with different characteristics may respond differently to incentives. There has been limited research regarding student responses to a combination of multiple-choice and bonus incentive contracts. This paper is the first to employ the combination of multiple-choice tests and intertemporal bonus point mechanism to explore student strategic behavior driven by risk preference.

Currently, multiple-choice exams are the most common assessment tools used in large-scale classes at many higher education institutions. Despite the widespread use of this exam format, there have been ongoing debates on its effectiveness in assessing students' knowledge. Existing studies argue that the relative benefits of multiple-choice tests are debatable (Walstad, 1998). On one hand, essay tests are claimed to potentially provide an opportunity for students to showcase their originality and creativity; however, these formats are more subjective. On the other hand, multiple-choice tests are efficient and much easier to grade (Walstad and Becker, 1994). Nevertheless, there is no clear evidence to support the superiority of one assessment tool over the other (Bennett et al., 1991; Wainer and Thissen, 1993). Despite this claim, one drawback of multiple-choice assessments is that these assessments create an opportunity for strategic behavior. Specifically, multiple-choice exams create a chance for students to engage in unobserved guessing on these exams. In this paper, we introduce an intertemporal bonus point mechanism to reveal traditionally unobserved student behavior. The bonus point mechanism allows us to uncover certain strategic behavior of students with different risk preferences.

Risk preferences play a vital role in shaping attitudes and choices of students. Walker and Thompson (2001) investigate whether students exhibit a preference for risk aversion when they are allowed to demonstrate their preferences on exams. They compare a standard multiple-choice format to a modified format to elicit students' risk preferences as well as student confidence.

Following Walker and Thompson (2001), we investigate behavior of students with different risk preferences using a modified multiple-choice exam format. Unlike their paper, we provide students with an opportunity to make intertemporal choices of allocating a given amount of bonus points across assessment periods. One would expect that students are more likely to gamble on a multiple-choice exam when they don't understand course materials due to lack of study. However, if there is an opportunity to hedge against the risk, one would expect risk preferences to play a role in driving different behaviors. Note that the tendency to engage in risk taking behavior (guessing) is more inherent and less evident in multiple-choice exam format than other forms of assessment. Therefore, the use of multiple-choice questions in the assessment of student performance provides a greater opportunity to investigate the role that risk preference plays in explaining student strategic behavior.

Using data from principles of economics course, we study how risk preference influences student behavior during multiple-choice exams with an intertemporal bonus point mechanism. The intertemporal bonus point mechanism helps unveil information about student strategic behavior. Our results show that risk preference plays a significant role in explaining student behavior during assessment periods. The findings suggest that risk-loving students are more likely to be unprepared for their earlier exams until they are warned by their low scores. This behavior is reflected in the way that they allocate available bonus points across all exams during a given semester. Risk-loving students appear to allocate a greater share of the bonus points to the first exam and the use of bonus points tends to be less effective compared to other students.

The rest of the paper is organized as follows. Section 2 introduces the intertemporal bonus point mechanism and discusses expected results. Section 3 presents the empirical specification, followed by data description in Section 4. Section 5 reports the results. Section 6 concludes with a discussion on possible extensions.

2. Intertemporal Bonus Mechanism

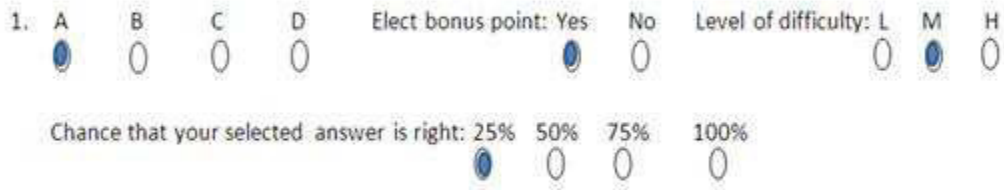
This paper blends an intertemporal bonus point mechanism into standard multiple-choice tests. The mixed design serves as a revelation mechanism through which traditionally unobserved behaviors in multiple-choice exams are now elicited. An anecdotal example of unobserved behavior is that a student gambles on a multiple-choice question and makes the correct guess. Without a properly functioning mechanism, this risk taking behavior is hard to identify. This paper considers an intertemporal bonus incentive mechanism to uncover certain information about participating students.

Standard multiple-choice exam format is altered by introducing an intertemporal bonus point mechanism which can potentially inflate students' scores. Thus, at the beginning of the principles of economics course, each student is offered an incentive contract, $S(e, B)$. The contract is part of the course syllabus and the instructor provides detailed instructions about the incentive contract during the first class meeting. The contract could be interpreted as a grade function, $S(\cdot)$, which depends on a student's unobservable effort (e) and the intertemporal use of bonus points (B). Note that the bonus point mechanism only provides a fixed amount of points (40 bonus points which could be applied to a total of 20 independent multiple-choice questions) that can be allocated to three exams over the duration of a semester.

Throughout the semester, there are three assessment periods (three exams) where a student could apply the bonus points. During each assessment period, students are tested based on 30 multiple-choice questions coupled with essay problems. Each multiple-choice question receives equal weight. It is important to note that the bonus point mechanism is only applicable to multiple-choice questions. Further, bonus points are only valid in inflating the score of a student if the student completes the following tasks: first provide his/her preferred answer to the multiple-choice question; second, elect to apply bonus points on a question; third, self-report difficulty level of the question; and finally indicate his/her own guess on the probability that the preferred answer is right.

The decision making process for each student on a multiple-choice question can be translated into choices in a lottery space. The lottery space comprises four different gamble options. Only one of the possible four lottery options is correct. The likelihood of a student selecting one gamble option over the other three options depends on how much effort the student invests in preparing for each exam. If the student is unsure of the right answer, the student can fully hedge against the risk of reducing his or her score by electing bonus points on a specific question. Since each question is worth two points, the remaining amount of bonus points available to the student is reduced by two. The less confident the student is about his or her answer, the more attractive the bonus point mechanism becomes. However, the more bonus points the student uses on a single exam, the less left for subsequent exams; therefore, the student faces the tradeoff between using bonus points now and using them in the future. Figure 1 below displays a sample response sheet of a hypothetical student for a given multiple-choice question. As shown in Figure 1, one can observe the preferred answer chosen by the hypothetical student, the use of bonus points on that question, the perceived difficulty level of the question, and the self-reported chance of being right. In this example, the student receives full credits for the question, but the available amount of bonus points is reduced by two.

Figure 1: Sample Exam Answer Sheet



At the end of each assessment period, the Holt and Laury (2002) risk elicitation gamble exercise is conducted. This is a typical exercise that asks respondents to choose between a sure return and a gamble, providing a means of eliciting risk preferences of the respondents. A typical risk-averse student would have at least six safe choices whereas a risk-loving student would have no more than four safe choices. In addition, a post-exam survey is conducted after the final exam to collect additional information about the students. The survey asks about time spent on preparing for each exam, clarity of the instructions on bonus points, understanding of exam questions, perceived difficulty of each exam, gender and other student characteristics.

The modified multiple-choice exams with the intertemporal bonus point mechanism allow us to observe how students allocate their available bonus points across exams, which implicitly reveals student strategic behavior. For instance, risk-loving students may be more engaged in guessing and thus are less prepared for early exams and rely more on bonus points to inflate their scores. In contrast, risk-averse students spend more time on studying and use the bonus points effectively, that is, only using bonus points when they know that their answers are probably wrong. The post-exam survey indeed demonstrates that risk-loving students spend less time on studying on average compared to risk-averse and risk-neutral students. Given the design of the assessments, we expect that risk-loving students to use the bulk of bonus points at early stages of course assessments, in particular, the first exam. However, their use of the bonus points is less effective compared to other students.

3. Empirical Specification

To test the hypothesis that students with different risk preferences behave differently in the way they allocate the bonus points across three exams, the following fixed effects model is used as the baseline empirical model:

$$B_{it} = B_0 + \beta' X_{it} + E_t + B_1 Risk_{it} + B_2(E_1 * Risk_{it}) + \mu_i + v_{it}, \quad (1)$$

where B_{it} is the outcome measure indicating the use of bonus points in exam *t* by student *i*. Two measures are considered and discussed in the data section. The explanatory variables include the following: X_{it}, a vector of student characteristics in a given exam; E_t , a vector of exam dummies; $Risk_{it}$, a dummy variable indicating whether the student is risk-loving based on responses to the risk elicitation exercise at the end of each assessment period[1], and it is

[1] See discussions about the Holt and Laury (2002) risk elicitation exercise in the previous section.

interacted with exam 1 dummy variable; μ_i, a vector of individual student fixed effects and v_{it}, the disturbance term.

One may be concerned about unobservable individual characteristics that might affect the allocation of bonus points across exams. The empirical model in this paper includes student fixed effects to control for unobservable individual heterogeneity that affects student behavior. For comparison purpose, we also run a pooled regression.

4. Data

The bonus point mechanism was employed for three sections of Principles of Economics during Spring and Fall terms in 2011. There were three exams for each section and the exams were the same across sections. Each exam included 30 multiple-choice questions and a few essay questions. The total number of questions to which students can apply bonus points over three exams was 20. In addition, a post-exam survey was conducted after the final exam to collect more information about the students. A total of 158 students from three class sections participated in the modified assessments.[2] About 57.8% of the students are male and 99.3% of the students indicated that the instructions on the bonus point mechanism were clear.

For the analysis on the allocation of bonus points, two outcome measures are considered. The first one measures the share of bonus points that a student allocates to a given exam, and the second one measures the share of bonus points that are used effectively in a given exam. As discussed before, the use of bonus points is effective if a student misses the question and elects to use bonus points.[3] These two measures enable us to better examine strategic behavior of students with different risk preferences.

The empirical model includes a set of explanatory variables. One important variable is the risk preference dummy indicating whether a student is risk loving based on the responses to the risk elicitation exercise at the end of each exam period, and it is interacted with exam 1 dummy variable. The coefficient on this interaction term indicates differences in strategic behavior driven by risk preference when students allocate available bonus points across exams. In addition, the share of multiple-choice questions that a student correctly answered in a given exam is included. Student performance on essay questions in each exam is used to account for the student's ability. This variable can also serve as a proxy for understanding of class materials by this student. In addition, time spent on studying and preparing for each exam is utilized to control for student learning efforts. A dummy variable indicating whether an exam is perceived to be very difficult and the share of questions that this student has trouble understanding are also included. Exam dummies are employed to account for unobservable common factors associated with each exam.

Table 1 reports summary statistics of our sample data as well as the definition of each variable. It is interesting to observe that some students almost use up all of their bonus points on a single exam. On one hand, this behavior could occur during the final exam because extremely risk-averse students who expect the final exam to be harder than the other two exams are likely to save more bonus points and use them all on the final exam. Evidence from our data indicates such behavior. On the other hand, an alternative explanation is that risk-loving students tend to use up the bonus points quickly on the first exam given that they are less likely to study hard unless they are forced to do so. Our sample data also shows evidence of this behavior: risk-loving students on average allocate significantly more bonus

[2] Six students withdrew from the course in the middle of a semester.

[3] As shown in the sample exam answer sheet (Figure 1), students have to indicate their preferred answers to a given question.

points to the first exam than risk-averse or risk-neutral students. In our sample, about 37% of the students are risk-loving. Students, on average, spend about seven hours studying and preparing for an exam. Further examination of the data indicates that students spend more time on the final exam due to the following reasons: first, the material tested on the final exam is cumulative and thus, students are by default forced to study harder; second, students may have to make more efforts to reach a target score in order to increase their overall grades. Moreover, detailed breakdowns of the post-exam survey data imply that risk-loving students spend less time studying and preparing for an exam on average.

5. Results

Table 2 reports the results based on regressions that utilize the first outcome measure, that is, the share of bonus points allocated to a given exam. Results from the pooled regression and the fixed effects model are displayed in the table, but our preferred fixed effects model can control for unobservable student characteristics that affect the allocation of bonus points, providing consistent and more accurate estimates. The regression results indicate that students with different risk preferences behave differently in their intertemporal allocation of bonus points. As expected, the coefficient on the interaction between risk-loving dummy and exam 1 dummy is positive, suggesting students who are risk-loving tend to allocate more bonus points to the first exam than those risk-averse and risk-neutral students. It is interesting to note that students, on average, use more bonus points on the third exam (final exam) and less bonus points on the first exam. One reason is that our student pool consists of a larger share of risk-averse and risk-neutral students than risk-loving ones. Risk-averse students are more likely to save the bonus points for the final exam; hence, less bonus points, on average, are used on the first exam and more on the final exam. Another reason for the positive coefficient on the third exam is that it is rational for all students to use up all bonus points left on the final exam, leaving nothing on the table. The negative coefficient on exam 1 dummy and positive coefficient on its interaction with risk-loving dummy reflects differences in strategic behavior driven by risk preference when students allocate the bonus points across exams.

Other control variables also help explain the use of bonus points. As expected, students with a better performance on multiple-choice questions in a given exam allocate fewer bonus points to that exam. In addition, students use more bonus points on the exam perceived to be very difficult, which is consistent with our expectation.

To further investigate student behavior driven by risk preference, we replace the dependent variable with the second outcome measure, the share of bonus points that are used effectively in a given exam. As defined early in the paper, an effective use of bonus points means that a student elects to use bonus points on the multiple-choice questions that he or she answers incorrectly. Results are presented in Table 3. As discussed before, the fixed effects model can control for unobservable student heterogeneity. Compared with the estimates from the pooled regression, the fixed effects estimates regarding our key variables of interest are larger in magnitude.

Interestingly, students, on average, have a higher share of bonus points that are used effectively on the first exam. However, risk-loving students appear to use bonus points less effectively on the first exam. The results are consistent with strategic behavior discussed earlier. As demonstrated in the survey responses, risk-averse students spend more time on studying; thus, only using bonus points when they know that their answers are probably wrong. Anticipating that the final exam is cumulative, these students are especially careful when allocating available bonus points to the first exam since they want to save more for the final exam. In contrast, risk-loving students are less likely to study hard and rely more on bonus points to inflate their scores before running out of available points. Further, they may not have a clear idea about how to use the bonus points effectively due to lack of study. As a result, the use of bonus points on the first exam by risk-loving students tends to be less effective. Given that the majority of students are risk-averse or risk-neutral in our student pool, the average share of effective use of bonus points is higher on the first exam compared to other exams. Results on all other control variables are consistent with our expectation.

To sum up, results from two alternative specifications (shown in Tables 2 and 3) suggest that risk-loving students appear to allocate more bonus points to the first exam but the use of bonus points is less effective compared to students who are not risk loving. The findings support our hypothesis that risk preference drives differences in student behavior. The results may imply that risk-loving students are prone to guessing on multiple-choice tests and rely more heavily on bonus points to inflate their scores compared to other students.

6. Conclusion

With the help of an intertemporal bonus point mechanism, we investigate student behavior in multiple-choice tests over three assessment periods. The results suggest that risk preference plays a significant role in influencing student behavior and helps us understand student learning attitudes. The paper shows that students with different risk preferences differ in their strategic behavior when allocating available bonus points across exams. As discussed above, risk-loving students are more likely to be unprepared for their earlier exams until they are warned by their low scores. This behavior is reflected in the way that they allocate the bonus points across three exams: risk-loving students allocate a significantly larger share of total bonus points to the first exam, but the use of bonus points is less effective compared to other students. The findings inspire us to think about the ways to effectively enhance student performance through improvements on their learning habits given their risk preferences. If an instructor happens to have more students who are risk-loving, the instructor might consider encouraging students to form good learning habits at the early stage of the course duration.

Instructors often provide various incentives to motivate students to study hard and learn more. Nevertheless, the effects of those incentives are unclear. An incentive mechanism that is not well designed could end up encouraging inappropriate behavior. How to design an effective incentive mechanism and help students perform better is a very interesting and important question; however, it is beyond the scope of this paper. One thing worth noting is that risk preference drives differences in student behavior; therefore, a sound incentive mechanism should take into account risk preferences of students and provide opportunities for dynamic feedback between the instructor and students. One possible extension of this paper is to design an incentive mechanism based on student risk preference, and conduct field experiments to test student responses to different incentives. This would allow us to examine the impacts of various incentives on students and provide possible suggestions on the design of a good incentive mechanism.

References

Bennett, R. E., D. A. Rock, and M. Wang (1991). Equivalence of free-response and multiple choice items. *Journal of Educational Measurement*, 28(1), 77-92.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.

Walker, D. M. and J. S. Thompson (2001). A note on multiple choice exams with respect to students' risk preference and confidence. *Assessment & Evaluation in Higher Education*, 26(3), 261-267.

Wainer, H. and D. Thissen (1993). Combining multiple-choice and constructed-response test scores: toward a Marxist theory of test construction. *Applied Measurement in Education*, 6(2), 103-118.

Walstad, W. B. and W. E. Becker (1994). Achievement differences on multiple-choice and essay tests in economics. *American Economic Review,* 84(2), 193-96.

Walstad, W. B. (1998). Multiple choice tests for the economics course. In Walstad, W. B. and P. Saunders (ed.), *Teaching Undergraduate Economics: A Handbook for Instructors*. New York: McGraw-Hill/Irwin, pp. 287-305.

Table 1: Summary Statistics

Variable	Obs	Definition	Mean	Min	Max
Share of bonus points used in an exam	459	Bonus points used in a given exam divided by 40	0.32	0.00	0.85
Share of bonus points used effectively in an exam	439	Amount of effectively used bonus points in a given exam divided by total bonus points applied in that exam	0.59	0.00	1.00
Risk loving	459	Dummy variable: 1, if a student chooses no more than four safe choices in the risk elicitation exercise at the end of each assessment period; 0, otherwise	0.37	0	1
Share of correct MC answers	459	Percent of multiple choice questions answered correctly in an exam	0.68	0.00	1.00
Performance on essay questions	455	Percent of essay questions scored in an exam	0.69	0.10	1.00
Exam preparation	455	Hours spent on studying and preparing for an exam	6.60	0	90.00
Share of troubling questions	459	Percent of questions that a student has trouble understanding in an exam	0.20	0	0.85
Difficult exam	459	Dummy variable: 1, if a student perceives the exam to be very hard; 0, otherwise	0.16	0	1

Table 2: Results on Allocation of Bonus Points across Exams

Variables	Pooled Regression	Fixed Effects Model
Exam1	**-0.2247*****	**-0.2279*****
	(0.0240)	(0.0312)
Exam3	0.1150***	0.1225***
	(0.0177)	(0.0238)
Risk loving	-0.0004	0.0064
	(0.0194)	(0.0325)
Exam1*Risk loving	**0.1262*****	**0.1514*****
	(0.0309)	(0.0440)
Share of correct MC answers	-0.1680***	-0.2990***
	(0.0566)	(0.0931)
Performance on essay questions	0.0461	-0.0432
	(0.0455)	(0.0786)
Difficult exam	0.0445**	0.1863***
	(0.0181)	(0.0488)
Exam preparation	0.0004	0.0012
	(0.0006)	(0.0012)
Share of troubling questions	0.0374	0.0286
	(0.0492)	(0.1014)
Constant	0.3942***	0.5090***
	(0.0435)	(0.0892)
Observations	455	455
Adjusted R-squared	0.4232	0.4778

Notes: Entries are coefficient estimates with standard errors in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table 3: Results on Effective Use of Bonus Points in a Given Exam

Variables	Pooled Regression	Fixed Effects Model
Exam1	**0.1011*****	**0.1343*****
	(0.0387)	(0.0453)
Exam3	-0.0600**	-0.0525
	(0.0273)	(0.0328)
Risk loving	0.0095	-0.0072
	(0.0297)	(0.0449)
Exam1*Risk loving	**-0.1402*****	**-0.1686*****
	(0.0490)	(0.0623)
Share of correct MC answers	-1.0938***	-1.0764***
	(0.0902)	(0.1361)
Performance on essay questions	0.1753**	0.1555
	(0.0704)	(0.1089)
Difficult exam	-0.0485*	-0.1264*
	(0.0282)	(0.0677)
Exam preparation	0.0000	0.0004
	(0.0009)	(0.0016)
Share of troubling questions	-0.0847	-0.1375
	(0.0760)	(0.1411)
Constant	1.2508***	1.2729***
	(0.0688)	(0.1265)
Observations	439	439
Adjusted R-squared	0.3121	0.2320

Notes: Entries are coefficient estimates with standard errors in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.