Loss Aversion, Distributional Effects, and
Asymmetric Gender Responses in Economics Education*

Maria Apostolova-Mihaylova**

William Cooper§

Gail Hoyt§

Emily Marshall+

April 30, 2014

Abstract

Do students behave differently when faced with alternative grading systems? This paper examines heterogeneous gender effects in response to a loss aversion-grading scheme in the economics classroom. Over the course of two semesters, we conducted an experiment with undergraduate students at the University of Kentucky that frames their final grade and all of its components as a loss rather than a gain of points. We find that, on average, students in the treatment class with the loss aversion grading scheme score approximately 1.20 percentage points higher on the final course grade compared to students in the control group. In addition, we conclude that males in the treatment group perform about 1.86 percentage points better than males in the control group. Using an ordered probit model, we evaluate the effect of this grading scheme on the probability distribution of final course grades. We expand on the finding of Apostolova-Mihaylova, Cooper, Hoyt, and Marshall (2015) of an asymmetric gender response of the loss framing of the grade by observing an economically significant favorable effect on the grade distribution for male students. Framing the grade as a loss increases the probability of receiving a B by 8 to 11% and decreases the probability of receiving a D by 4 to 9% for male students in the treatment classes compared to male students in the control classes. There is no evidence that the loss framing of the grade affects the grade distribution for female students.

PRELIMINARY AND INCOMPLETE – DO NOT CITE WITHOUT AUTHOR'S PERMISSION

* We would like to thank Carlos Asarta, Chris Bollinger, William Hoyt, Carlos Lamarche, Georg Schaur, and participants at the 2014 Conference on Teaching and Research in Economics Education for helpful comments.
** Department of Economics, University of Mary Washington, Fredericksburg, VA 22401. Most of the work was performed while the author was affiliated with the Department of Economics at the University of Kentucky.
§ Department of Economics, Gatton College of Business and Economics, University of Kentucky, Lexington, KY 40506.
+ Primary contact at emily.marshall@uky.edu. Department of Economics, Gatton College of Business and Economics, University of Kentucky, Lexington, KY 40506 and Department of Economics, Bates College, Pettengill Hall, Lewiston, ME 04240 (emarshal@bates.edu).

We use a field experiment to investigate the effect of a small change in the grading scheme, framing grades as a loss of points, on student performance. We find a positive pure treatment effect, implying that the loss aversion-grading scheme improves the final course grades of students. In addition, we examine heterogeneous gender effects in response to a loss aversion-grading scheme. On average, males in the treatment class score about 1.86 percentage points higher than males in the control class. We also use an ordered probit to discover where in the distribution of grades this effect is occurring. The loss aversion framing increases the probability that male students receive a B by 8 to 11 percentage points and decreases their probability of receiving a D by 4 to 9 percentage points. We find no statistically significant effect of the innovative grading scheme on females.

REVIEW OF THE LITERATURE

Traditional Determinants of Student Outcomes

The literature on student outcomes has a rich and long tradition of exploring different determinants of student success. Ability, effort, demographic and educational background, and various socio-economic characteristics are some of the well-established factors affecting student performance in college-level courses. Many studies have clearly demonstrated the power of college grade point average (GPA), as a proxy for ability, to predict student outcomes in college economics classes (Rochelle and Dotterweich 2007; Borde et al., 1998; Raehsler and Yang 2005). Other ability measures often tested in the literature are previous performance in high school classes or college math classes (Anderson et al. 1994; Rochelle and Dotterweich 2007) and performance on standardized tests (Elzinga and Melaugh 2009; Raehsler and Yang 2005), both of which are found to be positively correlated with outcomes. A limited number of studies focus on achievement in prerequisite courses as a determinant of student grade. For example, Borde et al. (1998) analyze success factors for introductory corporate finance courses and find a positive relationship between grades in prerequisite accounting courses and final grades, while Green et al. (2009) confirm the importance of prerequisite math courses for student outcomes.

The role of gender and race for achievement in college economics courses is examined by many investigators, including Borde et al. (1998), Borg and Stranahan (2002), and Elzinga and Melaugh (2009). These studies find that gender matters for student success in the college economics course, such that male students tend to do better than female students. A few studies have attempted to explain the observed gender differences in student success. Graddy and Yang (2010) argue that the observed gender-specific achievement differences can be explained by differing concentrations of particular brain-types by gender, noting that type E, or empathizing, brain-types are much more common among females and type S, or systemizing, brain-types are much more common among males. This fact, along with the observation that type S brain-types find economics more appealing, can explain the well-documented differences in gender-based outcomes. A similar idea is explored in Borg and Shapiro (1996), Ziegert (2000), and Borg and Stranahan (2002), which identify personality types to be of particular importance for student achievement in upper-level economics courses. Borg and Stranaham (2002) test for an independent effect of personality type but also interact it with gender. They find that students with a sensing/judging (SJ) temperament consistently outperform students with a sensing/perceiving (SP) temperament, but that this effect is largely due to SJ males. The authors also find evidence that introverted students perform better than extroverted students, but this effect is mainly driven by female introverts.

Among the more established factors affecting achievement is parental educational level, having been confirmed not only for early childhood development, but also for student outcomes later in life (Fan and Chen 2001; Sirin 2005). Time allocated to other activities is also a critical determinant of student success. Borde et al. (1998) analyze the impact of membership in student organizations and self-reported hours worked. They conclude that while membership does not seem to affect performance of the average student, an increase in the number of hours worked during the semester negatively affects a student's final course grade.

Alternative Determinants of Student Outcomes

In addition to traditional factors, alternative influences on performance have emerged from the fields of psychology and behavioral economics, which offer novel insights into individual behavior. More specifically, the behavioral bias of loss aversion and its effectiveness in improving performance are the topic of several recent studies. Some of these studies emphasize the role of loss aversion in the workplace (Hossain and List 2012) while others focus on motivating student performance in school (Fryer Jr. et al. 2012; Levitt et al. 2012). Overall, most of the existing literature suggests that loss aversion bias can be used as an effective method to improve performance. Hossain and List (2012) test the existence of the loss aversion bias in a production setting, where they frame a performance-related bonus as a gain or as a loss. They find evidence that workers' productivity increases more when the productivity reward is framed as a loss rather than a gain. Fryer et al. (2012) conduct an experiment in a school setting to evaluate the power of loss aversion on teachers' productivity. In their experiment design, the student outcome-related bonus received by teachers is either given to them at the beginning of the school year and then taken away if their students do not meet specific year-end improvement targets (treatment group), or it is given to them at the end of the school year based on year-end student performance (control group). In this case, the fear of losing the already received bonus motivates teachers in the treatment group to improve their teaching effectiveness, which results in improvement in their students' math scores by 0.4 standard deviations. Rather than incentivizing teachers to perform better, as in Fryer et al. (2012), Levitt et al, (2012) offer financial and nonfinancial incentives framed as losses directly to school-age students. In this setting students do not respond significantly to either the loss or the gain framing, but this could be due to the small size of the rewards.

Loss Aversion in Economics Courses

A natural extension of this literature is the application of behavioral economics principles to improving student outcomes in the college classroom. It is therefore surprising that there is only one study evaluating the possible effect of loss aversion in a college setting. In this study, Apostolova et al. (2015) conduct a field experiment where they structure the grading scheme as a loss or a gain of points. More specifically, in the treatment classes, the grading scheme is such that students receive the maximum number of points at the beginning of the semester and lose points as they complete the different grade components throughout the semester. On the other hand, in the classes that serve as the control group, the grading structure is based on accumulation of points. Overall, there is no evidence that the loss-grading scheme has an effect on the final class grade; however, the results point to a differential effect on males and females. The course grade for male students in the treatment classes is, on average, higher by 3.17 to 4.05 percentage points than the course grade for male students in the control classes. In contrast, the course grade for female students in the treatment classes is 3.61 to 4.36 percentage points lower than the course grade for female students in the control classes. The stark contrast between male and female responses to the loss-grading scheme warrants a more detailed examination of heterogeneous gender effects; this is the main goal of the present study.

FIELD EXPERIMENT DESIGN

The experiment design here is isomorphic to the setup of Apostolova-Mihaylova et al. (2015). In the semester following the data used in Apostolova-Mihaylova et al. (2015), two instructors again taught two different courses (Principles of Macroeconomics and Economics and Business Statistics).[1] Each instructor taught two sections of his or her respective course—a treatment and a control section.[2,3] On all basic demographic components, including class size, the two sections of each course were similar (see Tables 2 and 3 for details and difference in means tests).[4] The experiment was designed to randomly distribute the student population across sections.[5]

The purpose of this experiment was to provide an explanation for the heterogeneous gender effect in response to a loss aversion-grading scheme discovered in Apostolova-Mihaylova et al. (2015) by employing the behavioral economics concepts of loss aversion and risk aversion.

---

[1] Henceforth, we will refer to the data collected in Apostolova-Mihaylova et al. (2015) as Semester 1 data, and the data gathered in the subsequent semester as Semester 2 data. Institutional Review Board (IRB) provisions do not allow us to provide the reader with the precise year or term the data were collected.

[2] The two sections of Principles of Macroeconomics were taught on Monday, Wednesday, and Friday from 10:00 a.m. to 10:50 a.m. and 11:00 p.m. to 11:50 p.m., and the two sections of Economics and Business Statistics were taught on Monday, Wednesday, and Friday from 12:00 p.m. to 12:50 p.m. and 1:00 p.m. to 1:50 p.m. Due to scheduling constraints, we could not have identical course times.

[3] We will use the same notation for the two instructors as in Apostolova-Mihaylova et al. (2015), where instructor X taught Principles of Macroeconomics and instructor Y taught Economics and Business Statistics. For the data collected in Apostolova-Mihaylova et al. (2014), the treatment section for instructor X was the second section and for instructor Y was the first section. For the data used in this analysis, the treatment group for instructor X was the first section and for instructor Y was the second section.

[4] The final enrollment of Principles of Macroeconomics was 49 students in the treatment section and 45 in the control section. At the conclusion of the semester, both sections of Economics and Business Statistics had 45 students enrolled.

[5] Our randomization technique was successful. Test of mean differences reveal that the treatment and control groups from the data collected during the most recent semester do not differ in composition based on pre-treatment characteristics. Combining the data used in Aposolova et al. (2015) and the data gathered in the subsequent semester, the treatment and control classes are not statistically different (except for the proportion of Asian students and age).

We implement a straightforward change in the grading scheme in which students are given 500 points at the beginning of the semester and points are subtracted as students complete assignments (with less than perfect scores).[6] We refer to “treatment” sections as those that utilize the loss aversion grading system and “control” sections as those that calculate student grades using the traditional system. Grades were contemporaneously updated using Blackboard, so that students could periodically see their final course grade changing.

The two sections for each course were indistinguishable (with the exception of how grades were entered); they received the same assignments, quizzes, lectures, project components, and exams. The only distinction between the treatment and control sections was that the treatment sections began the course with 500 out of 500 possible points (an “A”), and the control sections started the semester with 0 of out 500 possible points (this additive approach is traditionally how college students are graded). The final grade was comprised of five equally weighted parts (20% each or 100 points): assignments and quizzes, group project, exam 1, exam 2, and exam 3. For more details regarding the experiment design, please see Apostolova-Mihaylova et al. (2015).

The assignments and quizzes category was made up of twenty-four quizzes worth five points each, meaning the maximum score was 100.[7] In addition, students completed a group project over the course of the semester, also worth a total of 100 points. Students were assessed on two dimensions regarding the group project: their contribution to the group (evaluated by their peers) and the quality of their work (graded by the instructor). The two scores were

[6] The original study, protocol number 12-0953-P4S, was approved on December 10, 2012 by the University of Kentucky IRB. Modifications and extensions for continuation of the study were approved on August 19, 2013. On the third day of class, a proctor administered the informed consent form and pre-course survey. Upon signing the informed consent form, students were notified that data on major, minor, cumulative GPA, SAT and ACT score would be collected from the registrar’s office.
[7] Students received full credit for completing the assignment, not necessarily answering the question correctly.

averaged in order to calculate a student's final project grade.[8] The individual contribution grade was assessed based on peer surveys; due to its highly subjective nature, we have removed it from our analysis.

All exams were given on the same day in both courses and in both sections. Major project deadlines were also uniform across courses and sections. Final course grades were the weighted average of the group project grade, three exams, and the quizzes/assignments score with the lowest four grades dropped. All categories were equally weighted (20% each). We refer to this variable as the final course grade, and it is the main dependent variable used in our analysis.[9]

DESCRIPTIVE STATISTICS

Data used in the subsequent analysis were gathered in Semester 1 and Semester 2 from the pre-course survey, post-course survey, instructor records, and the registrar's office. The pre-course survey was administered on the third day of class in each respective term. The post-course survey was completed at the time of the final exam during each semester. Standardized test scores, GPA, major, and class data were provided by the registrar's office. Students were informed that if they signed the consent form this information would be retrieved.

Instructor grade sheets were used to collect student course performance records. Table 1 provides definitions of the variables used in the subsequent statistical analysis.

(Table 1: Variable Descriptions)

[8] The provisions of the group project grade are slightly more complex than outlined above; however, for the purposes of the analysis the specifics are irrelevant. Please see Appendix A for complete details.

[9] Also, the final course grade used in the analysis does not include any curves designated by the instructor or bonus points awarded as a requirement of the IRB documentation.

The full sample (data from Semester 1 and Semester 2) consists of 327 observations, of which 304 are used in the final regression specification.[10] The demographic composition of the full sample is comparable to the Gatton College of Business and Economics population at the University of Kentucky.[11] Of the full sample, about 36.39% are female and approximately 80.12% are white. The average age of students in the sample is 21.18. The majority of the sample, approximately 54.13%, is comprised of students with a junior class status.[12] The average cumulative GPA is 3.08 and composite ACT score is 24.91. Approximately 9.17% of the full sample consists of economics majors. The pre-course survey asked students to report the highest level of education completed by their mother and father, categorized as less than high school, high school, some college, completed college, or graduate school. Responses were coded 2 through 6, with 2 being the lowest education category (less than high school) and 6 being the highest (graduate school). Based on these parameters, the average educational attainment for mothers is 4.61 and for fathers is 4.60, implying that the mean level of parental education is between some college and completed college.

Table 2 displays complete descriptive statistics for the variables used in the empirical estimation. Column 1 shows mean values for the entire sample. Columns 2 and 3 provide variable averages and number of observations for the four treatment sections and four control sections, respectively, over the course of two semesters. Columns 4 and 5 display the same information for the Principles of Macroeconomics and Economics and Business Statistics courses.

[10] We have one missing observation for father's education and two for mother's education. In addition, we could not recover a composite ACT score for 21 students.

[11] About 80.77% of the Gatton population are white (non-Hispanic only), 35.92% are female, and the average cumulative GPA is 3.14.

[12] There are very few freshmen and professional degree seekers in our sample. About 0.61% of the observations are freshmen and 1.53% are professional degree seekers.

(Table 2: Descriptive Statistics – Semester 1 and 2 Data)

Shaded cells in Table 2 indicate statistically different means between treatment and control groups, as well as between the two courses.

Following the two semesters, we had student consent to use information for 160 individuals in the treatment sections and 167 in the control sections. The only statistically significant differences between the treatment and control sections are the percent of Asians and age; however, the difference in age is not intuitively meaningful (20.93 in the control group and 21.45 in the treatment group). In the treatment class, 6.88% of the sample self-identified as Asian, compared to 14.37% in the control. We control for race (among other demographic characteristics and measures of educational background) in the statistical analysis, thus accounting for the statistically different proportion of Asian students in the treatment and control groups.

Not surprisingly, there are also statistically significant differences in composition between the Principles of Macroeconomics and Economics and Business Statistics courses. The Economics and Business Statistics course is an upper-level class meant for business majors and minors; naturally, older students will filter into this course. The mean age in business statistics is 21.43 while the average age in macroeconomics is 20.97. In addition, the class rank of students in the two courses is statistically different. About 62.99% of the students in the business statistics course are juniors and 27.27% are seniors. In contrast, approximately 46.24% of the students in macroeconomics are juniors and only 10.98% are seniors. There is also a larger proportion of females in the macroeconomics course compared to business statistics—41.62% as opposed to 30.52%. This is not surprising as students from a wider range of major types take principles courses. A greater percentage of the class is economics majors in the higher-level course. In business statistics, 16.23% of the class is economics majors while 2.89% of the macroeconomics course identifies their primary major as economics. We include a course indicator variable in our empirical estimation to control for these differences.

EMPIRICAL METHODOLOGY

Ordinary Least Squares Estimation

In this section, we repeat the analysis of Apostolova-Mihaylova et al. (2015) using the augmented dataset to test the validity of the results obtained using only one semester of data. Only slight modifications to the list of explanatory variables are made and we expand the previous analysis exploiting the ordinal nature of the outcome variable. The dependent variable in regression specifications (1) and (2) is the final course grade, using the same calculation as Apostolova-Mihaylova et al. (2015). Prior to the statistical analysis, the final course grade is converted to a percentage.

Following Apostolova-Mihaylova et al. (2015), we use the following ordinary least squares (OLS) regression specification to analyze the effect of the loss aversion treatment on student performance.

$$Final\ grade_{ij} = \beta_0 + \beta_1 treat_{ij} + \beta_2 macro_j + \beta_3 \underline{X_i} + u_{ij} \qquad (1)$$

The coefficient of interest is β_1, which represents the marginal effect of being in the treatment class on final course grade. Subscripts indicate the observation student *i* in course *j*. Instructor fixed effects are captured by the $macro_j$ dummy variable. The vector $\underline{X_i}$ is a set of individual specific explanatory variables intended to control for all factors that may affect final course grade other than the loss aversion treatment. See the regression output in Table 3 for a complete list of control variables.[13]

We explore the heterogeneous gender effect found in Apostolova-Mihaylova et al. (2015) with the larger sample using the following specification:

$$Final\ grade_{ij} = \beta_0 + \beta_1 treat_{ij} + \beta_2 female_i + \beta_3 treatXfemale_{ij} + \beta_4 macro_j + \beta_5 \underline{X_i} + u_{ij} \quad (2)$$

Regression specification (2) includes $treatXfemale$, an interaction term between the treatment dummy variable and gender.

Ordered Probit Estimation

The actual grades reported on student transcripts are letter grades, not numerical grades. At the University of Kentucky, students receive an A, B, C, D, or E if they complete the course; plusses and minuses are indistinguishable. In this portion of the analysis, we exploit the discrete nature of grades using ordered probit estimation. Ordered probit models are commonly used in the educational outcomes literature, as letter grades are often viewed as a measure of course material comprehension.[14] As a result, we have a polychotomous dependent variable with 5 levels that have a natural ranking order. It is important to use an ordered probit model in order to take advantage of all the information provided by the ranking of letter grades. An ordered probit allows us to differentiate between the effect of the loss aversion treatment on receiving an A versus a B, C, D, or E. In other words, the marginal impact of the treatment on various letter grades may be different. Simply estimating via OLS would produce only one set of marginal

[13] We note a few deviations from Apostolova-Mihaylova et al. (2014). Instead of the count variable indicating number of math courses taken in high school and college, we use composite ACT score as a proxy for mathematical, as well as verbal ability. We lose 21 observations as a result. In the previous analysis, we did not have a sufficient number of observations to justify the loss; however, we now have a significantly greater sample size. Also, we use mother's education and father's education as additional control variables, neither of which are individually statistically significant but they do increase the overall explanatory power of our regression.

[14] See Borg and Stranaham (2002), Elzinga and Melaugh (2009), Graddy and Yang (2010), Green et al. (2009), Raehsler and Yang (2005), for example.

effects whereas an ordered probit model reports marginal effects for each of the possible outcome alternatives. A multinomial probit model would not be appropriate because the extra information implied by the ordinal nature of the dependent variable would be ignored.

For each student, we observe the final letter grade ($Letter\ grade_i$) and assign numerical values to each as follows: E=1, D=2, C=3, B=4, A=5. By using ordered probit and designating five discrete outcomes (A, B, C, D and E), we are implying that a score of 80 is equivalent to a score of 89.99 and that all grades below 60 are equivalent. Henceforth, let j refer to the number of alternatives in $Letter\ grade_i$. The values of $Letter\ grade_i$ are random and inconsequential, provided they satisfy the order conditions if $Letter\ grade_i^* < Letter\ grade_j^*$ then $Letter\ grade_i < Letter\ grade_j$.

The model to be estimated is:

$$Letter\ grade_i^* = \beta_0 + \beta_1 treat_{ij} + \beta_2 female_i + \beta_3 treatXfemale_{ij} + \beta_4 macro_j + \beta_5 \underline{X}_i + u_{ij} \quad (3)$$

where $Letter\ grade_i^*$ is a latent variable. We can only observe when $Letter\ grade_i^*$ crosses the threshold according to the following:

$$Letter\ grade_i = \begin{cases} E\ if\ Letter\ grade_i^* \leq \gamma_1 \\ D\ if\ \gamma_1 < Letter\ grade_i^* \leq \gamma_2 \\ C\ if\ \gamma_2 < Letter\ grade_i^* \leq \gamma_3 \\ B\ if\ \gamma_3 < Letter\ grade_i^* \leq \gamma_4 \\ A\ if\ Letter\ grade_i^* \geq \gamma_4 \end{cases}$$

The threshold parameters are each γ_1 to γ_M where M is equal to $j - 1$. Given that $y_i - x_i'\beta = u_i$, the probabilities of observing each $Letter\ grade$ value in terms of the parameters to be estimated are given by:

$$\Pr(Letter\ grade_i = E|x_i, \beta, \gamma) = F(\gamma_1 - x_i'\beta)$$

$$\Pr(Letter\ grade_i = D|x_i, \beta, \gamma) = F(\gamma_2 - x_i'\beta) - F(\gamma_1 - x_i'\beta)$$

$$\Pr(Letter\ grade_i = C|x_i, \beta, \gamma) = F(\gamma_3 - x_i'\beta) - F(\gamma_2 - x_i'\beta)$$

$$\Pr(Letter\ grade_i = B|x_i, \beta, \gamma) = F(\gamma_4 - x_i{}'\beta) - F(\gamma_3 - x_i{}'\beta)$$

$$\Pr(Letter\ grade_i = A|x_i, \beta, \gamma) = 1 - F(\gamma_4 - x_i{}'\beta)$$

where F is the cumulative distribution function of u_i. In the subsequent analysis, we assume u_i is normally distributed, with mean zero and variance one.[15] The coefficients and the threshold parameters must be estimated by maximizing the log likelihood function.

RESULTS

Ordinary Least Squares Results

Table 3 shows the output from the OLS regression specification (1) for the full sample including Semester 1 and Semester 2 data. When we control for all factors that have been shown to affect student performance, a statistically significant pure treatment affect arises. The implication of this result is that a relatively costless change to the grading scheme (a framing of points as a loss opposed to a gain), improves average student performance by approximately 1.20 percentage points on the final course grade (see column 1.3). Over an entire percentage point increase in the final course grade is an economically meaningful result and merits serious consideration as a more common teaching practice. Apostolova-Mihaylova et al. (2015) find no statistically significant pure treatment effect as a result of the loss aversion grading scheme; however, the magnitude of the treatment coefficient is similar to what we find here. The pure treatment effect is 1.12 percentage points after controlling for a similar set of explanatory variables in Apostolova-Mihaylova et al. (2015).

(Table 3: Regression Specification (1))

[15] The assumed distribution of u_i distinguishes an ordered probit from an ordered logit. It is common in the educational outcomes literature to use an ordered probit.

The coefficients on the demographic variables have the anticipated signs. On average, females score about 2.74 percentage points lower than males, holding all other factors constant. The effect of gender on final grade is statistically significant at the 0.01 level. The dummy variables for race are not statistically significant. Students in principles of Macroeconomics, the lower-level course, had a final grade that was, on average, 3.71 percentage points higher than students in business statistics.

Indicators of past academic performance also have the expected impact on the final course grade. A one-point increase in college GPA increases the final course grade by 3.14 percentage points. A higher ACT score also results in a higher final course grade. Both of these effects are statistically significant at the 0.01 level. There is no statistically significant effect of attending a private high school, mother's education, or father's education on final course grade. These results are comparable to Apostolova-Mihaylova et al. (2015).

Now, we explore the heterogeneous gender effect of the loss aversion treatment by including an interaction term between the $treat$ and $female$. Table 4 shows the results for regression model (2).

(Table 4: Regression Specification (2))

After controlling for demographic factors and various measures of academic ability, males in the treatment group, on average, score about 1.86 percentage points higher on the final course grade than males in the control class. There is no statistically significant impact of the treatment on females. We further investigate the differential impact of the loss aversion grading scheme on males and females in the next section.

Ordered Probit Results

Table 5 presents the coefficients from the ordered probit regression.

(Table 5: Coefficients from Ordered Probit Regression)

The statistically significant coefficient on the Treatment variable implies that the loss framing of the grade has a favorable effect on the grade distribution for male students in the treatment classes relative to the grade distribution of male students in the control classes.

To assess the actual impact of being in the treatment class on the probabilities to receive a given letter grade, we next examine the likelihood of receiving each of the 5 letter grades for the treatment and control sections of the principles of macroeconomics course and the economics and business statistics course for the modal student (Table 6).

(**Table 6: Probabilities of Receiving a Given Grade by Course and Group)**

We define the modal student as a white 21-year-old male in his junior year who has chosen a major different than economics. His cumulative GPA is 3.0 and his ACT composite score is 25. He has not attended a private high school, works 9 hours a week, and both parents have college degrees.[16]

For the principles of macroeconomics class, controlling for other factors affecting performance, it is more likely for a student in the treatment section to receive a higher grade (A or B). Consequently, if the student is in the treatment class, the probability of receiving a C or a D grade is reduced by 7.4% and 4.4%, respectively. It appears that the same is true for the

[16] For ease of interpretation we rounded the sample means to the nearest whole number

economics and business statistics course even though the differences between treatment and control probabilities are less pronounced there for high grades and more pronounced for low grades.

Figure 1 depicts the grade probability distributions for both courses and both groups. For both treatment grade distributions it is clear that they are shifted to the left, which implies an increased probability to receive a higher grade and a decreased probability to receive a lower grade.

(Figure 1: Grade Probability Distributions)

To examine the statistical significance of these differences, in Table 7 we present the marginal effects of the treatment with their corresponding standard errors.

(Table 7: Marginal Effects of the Treatment for the Modal Student)

For both the principles of macroeconomics and the economics and business statistics courses the probability of receiving the lowest and the highest grade was actually not affected by the different framing of the grading scheme because, presumably, students who are at the high and low ends of the grade distribution are already highly motivated/demotivated. However, we observe that the loss framing changes the probability of receiving grades in the middle of the grade distribution. In the principles of macroeconomics course, the higher probability for a male student to receive a B (+11%) is statistically significant at the 0.10 level, and the probability of receiving a C is lower (-7.4%) and also statistically significant at the 0.10 level. The probability of receiving a D also decreases but by a smaller magnitude. Interestingly, in the economics and business statistics courses, only the probabilities of receiving a B and a D are affected in a statistically significant way while the chances of receiving any other grade are virtually unchanged by the type of grading scheme.

Overall, these results point to a favorable effect of the loss framing of the grade on the grade distribution. Our conclusions are consistent with Apostolova-Mihaylova et al. (2015) who observe a heterogeneous gender effect from the loss framing on the course grade so that male students in the treatment class earned higher grades than their male counterparts in the control class. The contribution of this paper is to identify the parts of the grade distribution where these effects are most pronounced.

CONCLUSION

In this paper, we extend the work of Apostolova-Mihaylova et al. (2015) who evaluate the effect of a loss aversion grading scheme on the final grade in principles of macroeconomics and economics and business statistics. We augment the sample with an additional semester of data and find a pure treatment effect of 1.20 percentage points. Apostolova-Mihaylova et al. (2015) investigate an asymmetric response to the treatment based on gender and discover that the loss aversion grading scheme has a favorable effect on male grades and a negative impact on female grades. With the extended sample, our findings only show the existence of a male effect. We use a larger sample and an ordered probit model to evaluate the effect of the type of grading scheme on the probability distribution of the grades. Our results show that the loss framing of the grade does not affect the probability to receive the lowest and the highest grade but it does increase the probability to receive a B (11%), which is offset by a lower probability to receive a

C or a D (-7.4% and -4.4% respectively) for the principles of macroeconomics courses, and a lower probability to receive a D in the economics and business statistics courses (-9%).

Several insights can be drawn from these results. First, while there is a minimal cost to implement such a grading scheme, the benefits are economically significant and unlike the previous work of Apostolova-Mihaylova et al. (2015), negative effects on the female grade are not observed. Therefore, we can conclude that the overall impact on class performance is positive and there is no reason to use a traditional (gain) grading scheme when this loss aversion scheme can produce overall positive results. Second, our findings suggest that students in the principles of macroeconomics benefit more from the loss framing of the grade, which implies that it is more beneficial to use this grading methodology in low-level courses.

This paper does not evaluate the effect of the loss aversion grading scheme on long-term knowledge of economics, nor does it provide an explanation of the observed results. These questions represent important avenues for future research that can help us understand how to use novel methods and insights from behavioral economics to better motivate student learning in the economics classroom.

References

Anderson, Gordon, Dwayne Benjamin, ad Melvyn A. Fuss. “The Determinants of Success in University Introductory Economics Courses.” *The Journal of Economic Education* 25, no. 2 (1994): 99-119.

Apostolova-Mihaylova, Maria, William Cooper, Gail Hoyt, and Emily Marshall. 2015. “Heterogeneous Gender Effects under Loss Aversion in the Economics Classroom: A Field Experiment.” Forthcoming: *Southern Economic Journal* (2015).

Borde, Stephen F., Anthony K. Byrd, and Naval K. Modani. “Determinants of Student Performance in Introductory Corporate Finance Courses.” *Journal of Financial Education* 24 (1998): 23-30.

Borg, Mary O. and S. L. Shapiro. “Personality Type and Student Performance in Principles of Economic Education. *Journal of Economic Education* 27, no. 1 (1996): 3-25.

Borg, Mary O. and Harriet A. Stranahan. “Personality Type and Student Performance in Upper-Level Economics Courses: The Importance of Race and Gender.” *The Journal of Economic Education* 33, no. 1 (2002): 3-14.

Elzinga, Kenneth G. and Daniel O. Melaugh. “35,000 Principles of Economics Students: Some Lessons Learned.” *Southern Economic Journal* 76, no. 1 (2009): 32-46.

Fan, Xitao and Michael Chen. “Parental Involvement and Students’ Academic Achievement: A Meta-Analysis.” *Educational Psychology Review* 13, no. 1 (2001): 1-22.

Fryer Jr, Roland G., Steven D. Levitt, John List, and Sally Sadoff. 2012. “Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment.” National Bureau of Economic Research Paper No. w18237.

Graddy, Duane B. and Fang Yang. “Brain-Type, Gender, and Student Success in the Principles of Economics.” *Journal of Economics and Finance Education* 9, no. 1 (2010): 20-30.

Green, Jeffrey J., Courtenay C. Stone, Abera Zegeye, and Thomas A. Charles. “How Much Math Do Students Need to Succeed in Business and Economics Statistics? An Ordered Probit Analysis.” *Journal of Statistics Education* 17, no. 3 (2009).

Hossain, Tanjim and John A. List. "The behavioralist visits the factory: Increasing productivity using simple framing manipulations." *Management Science* 58, no. 12 (2012): 2151-2167.

Levitt, Steven D., John A. List, Susanne Neckermann, and Sally Sadoff. 2012. “The behavioralist goes to school: Leveraging behavioral economics to improve educational performance.” National Bureau of Economic Research Paper No. w18165.

Raehsler, Rod D. and Chin W. Yang. "An Economic Analysis on Intermediate Microeconomics: An Ordered Probit Model." *Journal for Economic Educators* 5, no. 3 (2005): 1-11.

Rochelle, Carolyn F. and Douglas Dotterweich. "Student Success in Business Statistics." *Journal of Economics and Finance Education* 6, no. 1 (2007): 19-24.

Sirin, Selcuk R. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational* Research 75, no. 3 (2005): 417-453.

Ziegert, Andrea L. "The Role of Personality Temperament and Student Learning in Principles of Economics: Further Evidence." *The Journal of Economic Education* 31, no. 4 (2000): 307-322.

Tables and Figures

Table 1: Variable Descriptions

Variable	Description	Source
Final grade	Final course grade (excludes project peer evaluation, bonus points, and exam curves)	Blackboard
Macro	Course (1 for Principles of Macroeconomics)	
Female	Gender (1 for female)	Beginning of semester survey
White	Race (1 for White)	Beginning of semester survey
Asian	Race (1 for Asian)	Beginning of semester survey
Black	Race (1 for Black)	Beginning of semester survey
Hispanic	Race (1 for Hispanic)	Beginning of semester survey
Age	Age (calculated from birth year and month)	Beginning of semester survey
Assignments zero	Number of assignments not turned in	Blackboard
Hours work	Hours per week of work	Beginning of semester survey
Private HS	Type of high school (1 for private)	Beginning of semester survey
College GPA	College cumulative GPA including the previous semester, but excluding the semester of the study	Registrar's office
ACT composite	Composite ACT score	Registrar's office
Class	Year in college (Freshman=1, Sophomore=2, Junior=3, Senior=4, Professional, second year=5)	Registrar's office
Educ father	Father's highest level of educational attainment of father (less than high school=2, high school=3, some college=4, completed college=5, graduate degree=6)	Beginning of semester survey
Educ mother	Mother's highest level of educational attainment of father	Beginning of semester survey
Econ major	Primary major economics (1 if the student's primary major is economics)	Registrar's office

Table 2: Descriptive Statistics – Semester 1 and 2 Data Combined

Variable	Full Sample (1)		All Treat (2)		All Control (3)		All Macro (4)		All Statistics (5)	
	n	mean	n	mean	n	mean	n	mean	n	mean
Final grade	327	75.802	160	76.441	167	75.191	**173**	**76.681****	**154**	**74.816****
Female	327	0.364	160	0.338	167	0.389	**173**	**0.416****	**154**	**0.305****
White	327	0.801	160	0.844	167	0.761	173	0.832	154	0.766
Asian	327	0.107	**160**	**0.069****	**167**	**0.144****	173	0.075	154	0.143
Hispanic	327	0.018	160	0.019	167	0.018	173	0.017	154	0.020
Black	327	0.077	160	0.075	167	0.078	173	0.081	154	0.071
Age	327	21.183	**160**	**21.447****	**167**	**20.929****	**173**	**20.965****	**154**	**21.427****
Assignments zero	327	3.914	160	4.100	167	3.737	173	4.081	154	3.727
Hours work	327	8.959	160	9.053	167	8.868	173	8.962	154	8.955
Private HS	327	0.257	160	0.250	167	0.264	173	0.225	154	0.292
College GPA	327	3.078	160	3.107	167	3.051	173	3.054	154	3.105
ACT composite	306	24.909	150	25.033	156	24.789	165	24.958	141	24.851
Class	327	2.954	160	3.031	167	2.880	**173**	**2.740*****	**154**	**3.195*****
Educ father	326	4.598	160	4.694	166	4.506	173	4.549	153	4.654
Educ mother	325	4.609	159	4.610	166	4.608	172	4.669	153	4.543
Econ major	327	0.092	160	0.100	167	0.084	**173**	**0.029*****	**154**	**0.162*****

Note: Shaded cells and bold text indicate statistically different means between the indicated samples. ** *and* *** *indicate significance at the 5% and 1% levels respectively.*

Table 3: Regression Specification (1)

	(1.1)	(1.2)	(1.3)
Treatment	1.453	1.179*	1.199*
	(0.919)	(0.667)	(0.675)
Female	-1.234	-2.501***	-2.740***
	(0.947)	(0.682)	(0.674)
Macro	2.100**	3.331***	3.708***
	(0.924)	(0.670)	(0.765)
White	4.792***	1.620	1.724
	(1.831)	(1.405)	(1.476)
Asian	7.255***	2.674	2.825
	(2.380)	(1.849)	(1.933)
Hispanic	3.762	3.185	2.893
	(2.626)	(2.296)	(2.351)
Age	-0.210	0.428*	0.264
	(0.319)	(0.225)	(0.297)
College GPA		3.083***	3.141***
		(0.757)	(0.791)
ACT composite		0.884***	0.876***
		(0.098)	(0.104)
Assignments zero		-1.185***	-1.205***
		(0.112)	(0.121)
Class			0.173
			(0.631)
Private HS			0.416
			(0.726)
Educ father			-0.221
			(0.363)
Educ mother			-0.257
			(0.373)
Econ major			2.004
			(1.285)
Hours work			0.007
			(0.030)
Observations	327	306	304
Adjusted R-squared	0.035	0.532	0.532

*Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The omitted race category is black.*

Table 4: Regression Specification (2)

	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)
Treatment	1.583	1.678	1.973*	1.758**	1.862**
	(1.181)	(1.168)	(1.160)	(0.859)	(0.872)
Female	-0.159	-0.279	-0.549	-1.747*	-1.853*
	(1.289)	(1.279)	(1.286)	(0.994)	(1.017)
TreatXFemale	-1.012	-1.266	-1.458	-1.605	-1.876
	(1.905)	(1.905)	(1.913)	(1.342)	(1.394)
Macro		2.010**	2.142**	3.373***	3.772***
		(0.925)	(0.923)	(0.671)	(0.768)
White			4.874***	1.733	1.865
			(1.844)	(1.397)	(1.461)
Asian			7.283***	2.720	2.884
			(2.388)	(1.840)	(1.921)
Hispanic			3.970	3.406	3.134
			(2.657)	(2.213)	(2.264)
Age			-0.205	0.415*	0.243
			(0.327)	(0.230)	(0.302)
College GPA				3.066***	3.124***
				(0.758)	(0.791)
ACT composite				0.877***	0.863***
				(0.099)	(0.105)
Assignments zero				-1.186***	-1.207***
				(0.113)	(0.121)
Class					0.187
					(0.625)
Private HS					0.348
					(0.725)
Educ father					-0.170
					(0.371)
Educ mother					-0.291
					(0.379)
Econ major					2.045
					(1.260)
Hours work					0.006
					(0.030)
Observations	327	327	327	306	304
Adjusted R-squared	-0.001	0.010	0.033	0.533	0.533

*Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The omitted race category is black.*

Table 5: Coefficients from Ordered Probit Regression

	Coefficient	Robust St. Error	P-value
Treatment	0.309	0.171	0.070
Female	-0.461**	0.200	0.021
TreatXFemale	-0.118	0.277	0.672
Macro	0.695***	0.148	0.000
Asian	0.523	0.350	0.135
White	0.302	0.245	0.218
Hispanic	0.578	0.446	0.195
Age	0.088	0.060	0.144
Class	0.056	0.123	0.651
College GPA	0.719***	0.161	0.000
Private HS	0.106	0.151	0.482
Education father	-0.041	0.075	0.590
Education mother	-0.083	0.074	0.264
Econ major	0.417**	0.211	0.048
Assignments zero	-0.212***	0.025	0.000
Hours work	0.001	0.006	0.818
ACT composite	0.149***	0.022	0.000
γ_1	4.527	1.341	
γ_2	6.187	1.364	
γ_3	7.918	1.375	
γ_4	9.982	1.400	
Observations		304	
Pseudo R-squared		0.268	

Table 6: Probabilities of Receiving a Given Grade by Course and Group

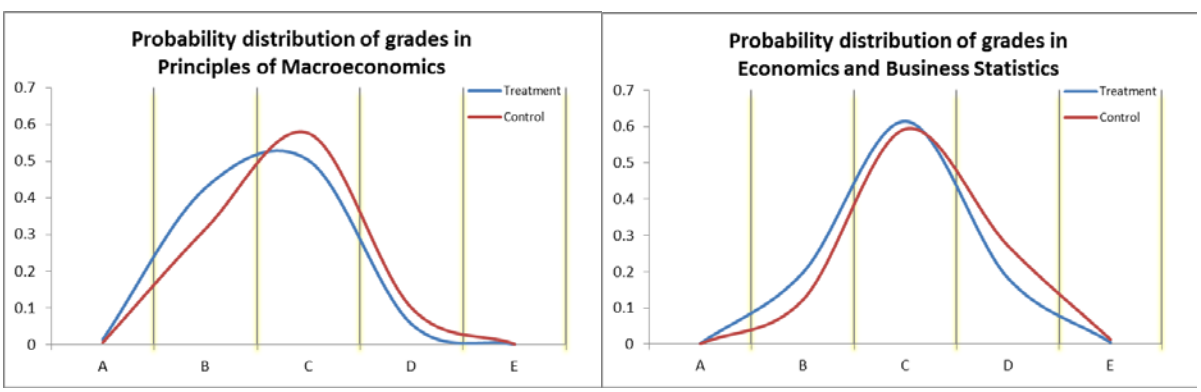
Course	Group	A	B	C	D	E
Principles of Macroeconomics	Treatment	0.014	0.428	0.502	0.056	0.001
	Control	0.006	0.316	0.576	0.100	0.002
Difference (marginal effect)		*0.008*	*0.111*	*-0.074*	*-0.044*	*-0.001*
Economics and Business Statistics	Treatment	0.002	0.197	0.614	0.182	0.005
	Control	0.001	0.122	0.593	0.271	0.013
Difference (marginal effect)		*0.001*	*0.074*	*0.019*	*-0.087*	*-0.007*

Table 7: Marginal Effects of the Treatment for the Modal Student

Course	A	B	C	D	E
Principles of Macroeconomics	0.008 (0.006)	0.111* (0.060)	-0.074* (0.042)	-0.044* (0.025)	-0.001 (0.001)
Economics and Business Statistics	0.001 (0.001)	0.074* (0.042)	0.019 (0.024)	-0.087* (0.048)	-0.007 (0.007)

*Notes: Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

Figure 1: Grade Probability Distributions



Appendix A

Syllabus Excerpt

Project Grades

Below I discuss two separate components of the overall project grade: the group project grade and the individual project grade. Your overall project grade, used to calculate your final course grade, will depend on your relative group project grade and individual project grade. See details in the following paragraphs.

Your group project grade will depend on how well your group performs the five parts of the project detailed below. Your individual-contribution grade depends on how much you contribute to your group project. To determine your individual contribution, I will observe you throughout the semester, and at the end of the semester I will ask each group member to evaluate each other group member's contribution. A checklist is provided at the end of the syllabus on what is expected of each group member.

If your individual contribution grade is above the group project grade, your overall grade will be your group project grade. For example, if your group's grade is 94% and your individual grade is 100%, then your overall project grade will be 94%.

If your individual contribution grade is below the group project grade but above a B- (82%), your overall project grade will be the simple average of two grades: your group's grade and your individual-contribution grade. For example, if your group's grade is 94% and your individual grade is 90%, then your overall project grade will be 92% [(94+90)/2=92%].

If you receive a B- (82%) or below on your individual-contribution grade, then your overall project grade will only be your individual-contribution grade. For example, if your group's grade is 94% and your individual-contribution grade is 76%, then your overall project grade will be 76%, not [(94 + 76)/2 = 85%].