MERIT-BASED SCHOLARSHIPS AND STUDENT EFFORT

Rey Hernández-Julián
Department of Economics
Metropolitan State College
of Denver
P.O. Box 173362
Denver, CO 80217-3362
rherna42@mscd.edu

Abstract

Twenty-one states offer merit scholarships that require students to maintain a minimum grade point average (GPA). Using a comprehensive administrative database from Clemson University, this study estimates the relationship between the incentives created by a South Carolina merit scholarship (LIFE) and students' academic performance. I hypothesize that being at risk of gaining or losing this scholarship will lead to increased effort and, as a consequence, higher grades. The results suggest that the incentives created by the scholarship increase GPAs by as much as 0.101 on a four-point scale, controlling for student and course characteristics. Moreover, the results indicate that for men the relationship between the risk of gaining or losing the scholarship and grades is large and statistically significant; for women, however, there is little evidence that the scholarship is related to grades.

1. INTRODUCTION

Approximately 18 million students are currently enrolled in institutions of higher education in the United States at a current annual cost of over $200 billion (NCES 2006). In 2001, South Carolina spent $2.2 billion on higher education, most of which went to universities and colleges, but about 6 percent was given directly to the students through merit programs (SCCHE 2006a, 2006b). South Carolina's main merit-based scholarship, called LIFE (Legislative Incentives for Future Excellence), together with other smaller merit-based scholarships, gives qualifying students vouchers they can use at any in-state college or university, public or private.

Academics have paid a fair amount of attention to the effect of merit-based scholarships on student outcomes. For instance, Cornwell, Lee, and Mustard (2005) studied student responses to Georgia's merit-based HOPE scholarship and showed that it increased strategic withdrawals from courses. Their research also found evidence to suggest that HOPE led to students' enrolling in fewer, and often easier, courses.[1] Or, to take another example, Singell, Waddell, and Curs (2006) showed that Georgia's merit-based scholarship did not crowd out Pell Grant students.

The current study examines whether there is evidence to support the hypothesis that a merit-based scholarship, South Carolina's LIFE, encouraged potential recipients to perform better academically. In the empirical analysis below, I compare the effect of the LIFE scholarship on in-state students who were at risk of losing or obtaining the scholarship with its effect on out-of-state students who could not receive the scholarship but who had similar grade point averages (GPAs).

There is good evidence that grades at institutions of higher education have increased over time (Rojstaczer 2003). Rosovsky and Hartley (2002) offered a number of explanations. One possibility is that instructors have become increasingly lenient graders; other potential factors are the Vietnam War, changes in curriculum, increases in class size, and the adoption of student evaluations. University administrators may also prefer higher grades because they encourage students to remain in school, leading to higher retention rates

1. Other work in this area includes Heller and Marin (2002), who edited a book of essays on the effects of scholarships on underrepresented and low-income students, and Dynarski (2004), who surveyed the impacts of merit-based aid on enrollment in several states. A study of New Mexico's merit-based scholarship found that receipt was associated with a small increase in the grades of freshmen, but the focus was on the composition of the freshman class as opposed to its performance (Binder, Ganderton, and Hutchens 2002). Angrist and Lavy (2002) found that, under certain conditions, cash payments were positively associated with high school matriculation in Israel. In addition, several K–12 public school districts in the United States, most notably the New York City school district, have begun to reward their students' good performance with cash (Medina 2008).

and improved institutional rankings.[2] In 1996 President Clinton proposed that students who maintained a minimum GPA would receive a new government scholarship. Concern over grade inflation led Thomas Bickel, registrar of Dartmouth College, to criticize Clinton's proposal by arguing that the scholarship "would be the end of any effort to control grade inflation" (Bulkeley 1997, p. B1). This article can be thought of as examining a potential mechanism through which Bickel's concern might be realized.

The LIFE scholarship program began in 1998 granting $2,000 to in-state freshmen who satisfied the SAT and high school GPA requirements and to returning in-state students who maintained a GPA of 3.0 or higher. Since then, the LIFE scholarship has increased to $5,000. This amount is substantial: LIFE covers at least half of tuition and fees at all public institutions in South Carolina and over half of the tuition of students observed in this study. I hypothesize that the 3.0 GPA cutoff provided an incentive to students who were at risk of either obtaining or losing the LIFE scholarship to exert more academic effort. The empirical analysis below tests this hypothesis.

Section 2 describes the LIFE scholarship and the estimation procedure. The scholarship has two basic requirements. First, the student must have at least a 3.0 GPA. In the empirical analysis below, only courses taken by students with GPAs sufficiently close to this cutoff were considered. Courses were excluded from the sample if they were taken by students whose GPA was high enough that the loss of the scholarship was not a possibility. Similarly, courses taken by students whose GPA was low enough that obtaining the scholarship was not feasible were also excluded from the sample. The second requirement of the scholarship, that the student must be a resident of the state of South Carolina, separates the treatment group from the control group. In-state students, as potential recipients, are expected to have increased their effort in an attempt to increase their likelihood of receiving the scholarship. Because out-of-state students could not receive LIFE, their effort and grades were presumably unrelated to the scholarship. Thus my estimation strategy relies on comparing the academic performance of in-state students at risk of losing or obtaining the LIFE scholarship with that of out-of-state students with similar GPAs.

Section 3 provides more detail with regard to Clemson University students. The data set combines the students' application and enrollment information with their choice of major and their grade in each class. Due to selection concerns, classes taken by students who entered Clemson after the fall of 1998, the LIFE start date, were excluded from the analysis. As a result, no classes

2. Former Stanford president Donald Kennedy added another reason while defending his institution's grades. He argued that the admission standards increased to the point that very few students ever did any C work.

taken by freshmen are included in the post-LIFE observations. The earliest grades in the data are from spring 1990 and the last are from fall 2000, the final semester of attendance by students admitted before LIFE began.

Section 4 presents the estimation results. Difference-in-difference estimation shows a positive effect of the merit-based scholarship on the effort of students who were at risk of losing or obtaining the scholarship as measured by grade outcomes. Results from a censored regression that controls for student and course characteristics suggest that the incentives created by the LIFE scholarship led to a significant increase in GPAs of 0.101. Estimates by gender show that the incentive effect of the LIFE scholarship was stronger for males than females. In fact, there is little evidence that females responded to the incentives created by the LIFE scholarship. After additional sensitivity and falsification tests, section 5 concludes.

2. THE LIFE SCHOLARSHIP AND ESTIMATION PROCEDURE

The LIFE scholarship provides a natural experiment to study the relationship between merit-based scholarships and student grades. The legislation signed on June 19, 1998, by South Carolina governor David Beasley provided a set of eligibility requirements for receipt of the scholarship. The scholarship was made available beginning in fall 1998 to all entering and returning college students who met the scholarship criteria. To qualify for LIFE, a student had to satisfy South Carolina residency requirements and had to have graduated from a South Carolina high school. To collect the scholarship in the freshman year, a student also needed to fulfill two of the following three requirements: a 3.0 high school GPA, an SAT score of 1000 or better, or a ranking in the top 30 percent of their high school graduating class.[3] For sophomores and upperclassmen, only college performance mattered: to receive the scholarship they needed at least a 3.0 cumulative GPA in their college courses. Clearly, students who received LIFE payments had a strong incentive to maintain a 3.0 GPA and as a consequence could have opted to take a lower course load. In fact, Cornwell, Lee, and Mustard (2005) found that students in Georgia did exactly that in order to maintain their eligibility for the HOPE scholarship. However, unlike the HOPE scholarship, LIFE required students to average thirty credits per year to maintain eligibility.[4]

3. South Carolina began offering LIFE scholarships in fall 1998. Initially students could receive $2,000 per school year. In 2000, the amount of the scholarship was increased to $3,000 and was further increased to $5,000 in 2002. Relative high school rank was added as a criterion in 2002; before that students had to fulfill both SAT and GPA requirements. The SAT requirement increased to 1050 starting in 2000 and to 1100 starting in 2002.

4. The three main findings of Cornwell, Lee, and Mustard (2005) were a lower-course load, an increase in withdrawals, and an increase in summer enrollment. An attempt to replicate these findings using the data in this study finds that Clemson students increase their course load, have no significant

Although the LIFE scholarship required (and still requires) a minimum high school GPA of 3.0, in general there is little evidence that merit-based university scholarships lead to grade inflation in high schools. Bugler, Henry, and Rubenstein (1999) and Henry and Rubenstein (2002) found that Georgia's HOPE scholarship did not lead to grade inflation in Georgia high schools, although they did find some evidence that it increased student effort. High school grade inflation is a concern because it may have changed the composition of freshmen entering Clemson. In fact, the LIFE scholarship could have affected the composition of the freshman class in a variety of ways: for instance, students who in the absence of the scholarship would have attended an out-of-state college might remain in state for the relatively lower tuition. Or LIFE could have enabled students to go to college who otherwise would not have had the resources. Finally, Clemson may have received a larger number of in-state applicants as a result of the LIFE scholarship, allowing it to be more selective. In order to avoid any potential bias due to the possibility that LIFE changed the composition of the student body, the sample is restricted to students who entered Clemson before LIFE began. As a result, there are no freshmen included among the post-LIFE observations.[5]

In the empirical analysis below, the timing of the scholarship's introduction and its eligibility requirements are used to separate into three mutually exclusive categories courses earned by students who entered Clemson before fall 1998. The first is composed of courses taken by Clemson students whose GPAs did not put them at risk of reaching or falling below the 3.0 GPA cutoff. These students would not have been affected by the scholarship, so their courses are excluded from the analysis. The second is the treatment group: courses taken by in-state students whose grades put them at risk of reaching or falling below a cumulative 3.0 GPA. The final category, the control group, consists of courses earned by out-of-state students whose grades would have put them at risk of reaching or falling below a cumulative 3.0 GPA. Students in treatment and control groups are observed both before and after LIFE's initiation in fall 1998, the date the treatment begins.[6]

As noted, in-state students are considered to have been "at risk" if their academic performance during the semester in which they were observed could have potentially affected whether the LIFE scholarship was received in subsequent semesters. For instance, first-semester in-state sophomores whose GPAs were above 2.33 (and who entered Clemson before fall 1998) were

change in withdrawals, and have a small statistically significant (but of very low magnitude) decrease in summer enrollment.

5. Excluding pre-LIFE observations of grades earned by freshmen generates coefficient estimates similar to those reported.
6. I use the terms *pre-LIFE* and *post-LIFE* to designate the periods before and after the treatment.

included in the analysis. Because first-semester sophomores with GPAs below 2.33 and a normal course load could not have met the 3.0 cutoff even if they received As in all their classes, they were not considered at risk and were therefore excluded from the analysis.[7] Finally, to be a potential recipient, a student must also have completed the required thirty credits per year. The focus of this analysis is on the incentive effect of being at risk of losing or gaining the scholarship. It is not on the effects of actually having received the scholarship.[8]

A difference-in-difference analysis can give a rough sense of the magnitude of the LIFE incentive effect. As stated above, the first difference is that between in-state students within the at-risk range (the treatment group) and comparable out-of-state students (the control group). The second difference is between the pre- and post-LIFE periods. The post-LIFE period begins in fall 2008, when the scholarship was instituted. Several previous studies have used comparable ineligible students as a control group. Kremer, Miguel, and Thornton (2004) used male students as a control group when examining the effects of a scholarship that is offered only to females. Binder, Ganderton, and Hutchens (2002) separated students by their high school graduation date, the eligibility factor for New Mexico's merit-based scholarship.

Difference-in-differences estimation cannot control for student characteristics or for other margins on which students make course choices. For instance, grades may increase as a result of taking easier classes or choosing an easier major. Multivariate regression analysis separates the effects of student quality and course choice from the effect of increased effort. Specifically, the following equation is estimated using course-level data on grades from in-state students who were at risk of either reaching or falling below the 3.0 GPA cutoff and grades from out-of-state students with comparable GPAs. Only grades earned by students admitted before the LIFE scholarship began are included:

$$grade_{ict} = \boldsymbol{X_{ict}}\beta + \theta time_t + \alpha(in\text{-}state)_I + \delta(LIFE)_t + \kappa(in\text{-}state * LIFE)_{it} + \varepsilon_{ict}, \quad (1)$$

where $grade_{ict}$ is student i's grade in course c in period t. $\boldsymbol{X_{ict}}$ is a vector of controls that includes characteristics of the student and the course, such as

7. Students were included in the at-risk category if they are first-semester freshmen with any GPA, second-semester freshmen with GPAs between 2 and 4, third-semester students with GPAs between 2.33 and 3.66, fourth-semester students with GPAs between 2.5 and 3.5, fifth-semester students with GPAs between 2.66 and 3.33, and sixth-semester students with GPAs between 2.8 and 3.2. However, since only students admitted before LIFE began add observations to the regression, all freshmen observations after fall 1998 are excluded. Students who did not satisfy the load requirement are not eligible, and neither were seniors. Results are robust to the variation of this range.

8. Nine percent of students who began college receiving the scholarship eventually lost it. Almost 2 percent of students who began without the scholarship gained it. Although I do not compare those who kept the LIFE scholarship with those who lost it, Dee and Jackson (1999) studied retention in the context of Georgia's merit-based scholarship.

gender, race, SAT score, major, course subject, and class size, and captures the effect of the student's background, characteristics, and course choices. The variable denoted as *time* is a counter: for spring 1990 *time* equals one, for fall 1990 time equals two, and so on. This trend controls for omitted determinants of grades that change smoothly over time—for instance, grade inflation. Although grade inflation could be caused or aggravated by the scholarship, it is impossible to isolate grade inflation induced by the scholarship from the existing trend in grade inflation.

The focus of this analysis is on the variables *in-state*, *LIFE*, and their interaction. *In-state* equals one if the student was a South Carolina resident and as a result was in the treatment group. *LIFE* equals one in fall 1998 and after, when the LIFE scholarship became available. Thus the interaction of these two variables, *In-state* × *LIFE*, equals one for students in the treatment group after the treatment began. Since all the courses in the sample are contributed by students who were within the at-risk range, the adoption of LIFE is expected to be related to effort. If LIFE had the effect of increasing effort and grades, the estimated coefficient of the interaction of *LIFE* and *in-state*, kappa, will be positive.[9]

A possible complication could arise if variables correlated both with the scholarship and grades were omitted from the regression. Measures of student ability in $\boldsymbol{X_{ict}}$ control for the fact that in-state students tended to have lower SAT scores and high school ranks than out-of-state students. A student's semester course load is also included in $\boldsymbol{X_{ict}}$ to control for effort and drive. Other student and course characteristics may capture variation in grades not due to LIFE but potentially correlated with when the scholarship was instituted. These extra controls present a small cost in terms of degrees of freedom but offer in exchange additional certainty that the key estimates do not suffer from omitted variable bias.

Controls for choice of major and course load are necessary because merit-based scholarships may also affect these choices (Cornwell, Lee, and Mustard

9. An alternate specification would include all students (as opposed to only students in the at-risk category) and a dummy variable specifying students who were at risk of falling below or attaining the 3.0 GPA cutoff:

$$\begin{aligned}\text{grade}_{ict} = {} & X_{ict}\beta + \theta(\text{time})_t + \alpha(\text{in-state})_I + \delta(\text{LIFE})_t + \Upsilon(\text{at-risk})_{it} \\ & + \varphi(\text{in-state} * \text{at-risk})_{it} + \kappa(\text{in-state} * \text{LIFE})_{it} \\ & + \lambda(\text{in-state} * \text{at-risk} * \text{LIFE})_{it} + \psi(\text{at-risk} * \text{LIFE})_{it} + \varepsilon_{ict}\end{aligned}$$

This estimation strategy generates nearly identical results as using the equation above. An additional alternative specification would include a fixed effect for each semester and have these fixed effects be interacted with *in-state*. This would loosen the requirement imposed by the *time* variable that time-dependent variations are constant through all the semesters. One would then compare the coefficients on these interactions for the post-LIFE semesters with the pre-LIFE semesters. The results of such a specification show the same pattern of results as those presented in table 3.

2005). Certain majors, such as music and English, are associated with higher GPAs than others (Anglin and Meng 2000). Students may have chosen their major strategically to reduce their risk of losing the scholarship. In fact, Dee and Jackson (1999) showed that students in different majors had different probabilities of losing Georgia's HOPE scholarship. The estimation also controls for the subject and level of the class because Sabot and Wakeman-Linn (1991) found that course choices were affected by grade inflation. These articles suggest that students consider scholarship requirements when choosing majors, classes, and course loads. The variable *tuition* controls for the effects of price on performance: as the price of college increases, students should, in theory, work harder to decrease the likelihood of having to attend an additional semester (Garibaldi et al. 2007). Finally, the mean SAT score of the other students in each classroom is intended as a control for peer effects.

Ideally, equation 1 would control for characteristics of the student's high school, the instructor of the course, and family income. Students from different high schools likely enter with different levels of preparation, and this might explain a part of the variation in grades. Detailed information on family income would also be useful because the scholarship may have a stronger impact on families with lower incomes. However, to protect the privacy of students, the financial information was not made available for this project.[10]

Data on the instructor of a course would be particularly helpful because students may have improved their chances of earning a better grade by seeking out professors with reputations for being easy. I attempt to address this concern by estimating a separate regression that includes only observations from sections that had fifty or more students. Grades tend to be lower in larger sections (Kokkelenberg, Dillon, and Christy 2008), so the students who were more likely to strategically choose the easiest professors and sections would presumably be excluded from this sample. In addition, I estimate a regression that includes only one-section courses where students could not choose their professor.[11] To the extent that professors actually respond to the scholarship by becoming more sensitive to the in-state students' pleas for grades, these estimates of student effort will be biased upward.

10. I do not have access to this information, but including controls for the students' zip codes, which capture socioeconomic status and the student's likely high school, slightly increases the magnitude of the estimates, and they remain significant.
11. As noted by an anonymous referee, graduate record examination (GRE) scores would provide an excellent objective measure of human capital that students could not adjust by negotiating with professors. However, these are available only for the small proportion of students who enroll in graduate school at Clemson. Any estimate that focuses on these students would not be representative because this sample consists only of students who choose to enroll in graduate school at the same institution where they completed their undergraduate education.

Because grades are bounded by zero and four, ordinary least squares (OLS) may produce biased coefficient estimates (Greene 1981). This problem is corrected by estimating a Tobit regression (Tobin 1958) using zero as a lower bound and four as the upper bound. Standard errors are corrected for clustering at the student level.

3. DATA: CLEMSON STUDENTS

Clemson University made available extensive data on its students, including every grade received and all application and enrollment information.[12] Clemson uses a four-point grading system without pluses or minuses. All grades are reported as four, three, two, one, or zero representing A, B, C, D, and F.[13] A unit change in the dependent variable corresponds to a full letter grade in a course.[14] Remedial courses are offered by the several community colleges in the area. These were not included in the sample.

As noted, the sample is limited to grades earned after 1990 by students admitted before fall 1998. The regression includes in-state courses taken by students whose current GPA placed them at risk of gaining or keeping the LIFE scholarship, courses taken by out-of-state students with comparable GPAs, and courses taken by pre-LIFE students with comparable GPAs. Thus the regression uses information from 21,895 students with an average of 13.94 grade observations for a total sample of 305,292. Table 1 shows summary statistics for the primary variables used in the analysis. The first set of columns includes all students, the middle set the treatment group, and the final set corresponds to the control group, out-of-state and pre-LIFE students whose GPAs would have put them at risk had they been eligible for LIFE.

Clemson is a public school, and most students (68 percent) are South Carolina residents. About 52 percent of the students are male, and 26 percent have had a sibling, parent, or ancestor who attended Clemson in the past. The average student in the regression is a third- semester student at Clemson. The observations in this sample do not include all the grade observations from any semester because the majority of grades were earned by students not in the at-risk category. Grades earned by at-risk students range from 30 percent of all grades in fall 1990 to about 1 percent in spring 2000.[15] This percentage tended to be lower in the spring than in the fall. Out-of-state students earned

12. Names and personal identifiers have been removed. Randomly generated record labels have been assigned to match students to grades.
13. P (pass) and W (withdrawal) are also valid grades but were not included in the regression because the rules on pass/fail courses and withdrawals have changed several times within the sample period.
14. In the summary statistics and regressions, all grade observations are weighted by their credit value.
15. By then, only a few students admitted before the LIFE scholarship remained on campus. Using only grades earned after 1994 instead of 1990 does not significantly affect any of the findings of this study.

Table 1. Summary Statistics

	All			*Treatment Group*			*Control Group*		
	N	***Mean***	***Std. Dev.***	***N***	***Mean***	***Std. Dev.***	***N***	***Mean***	***Std. Dev.***
Grade	305,292	2.908	0.986	206,299	2.89	0.997	98,993	2.947	0.961
In-state (d), by student	21,895	0.681	0.466	14,902	1	–	6,993	0	–
Percent of observations post-LIFE	305,292	0.066	0.249	206,299	0.068	0.252	98,993	0.062	0.241
SAT math, by student	21,895	570.65	72.043	14,902	566.88	71.87	6,993	579	71.77
SAT verbal, by student	21,895	557.416	72.727	14,902	555.27	72.86	6,993	561.99	72.25
Relative high school rank, by student	21,895	0.207	0.169	14,902	0.187	0.16	6,993	0.251	0.185
Class size, by grade	305,292	51.23327	43.366	206,299	51.413	43.48	98,993	50.858	43.124
Tuition, by term	–	–	–	11	$3,357.94	$169.05	11	$9,014.47	$503.41
Credit load, by grade	305,292	15.28	2.16	206,299	15.208	2.202	98,993	15.423	2.059
Fall term (d), by grade	305,292	0.538	0.498	206,299	0.541	0.498	98,993	0.532	0.499
Male (d), by student	21,895	0.522	0.499	14,902	0.528	0.499	6,993	0.512	0.5
Legacy (d), by student	21,895	0.264	0.441	14,902	0.33	0.47	6,993	0.123	0.329

Notes: Grade is measured in points: A = 4, B = 3, C = 2, etc. SAT values before 1996 are adjusted by re-centering criteria. (d) designates {0,1} variable. "Relative high school rank" is the rank divided by size of graduating high school class. "Legacy" designates a student with prior family ties to Clemson. "Class size" is the number of people in the class.

Table 2. Difference-in-Differences

Mean Grades (Standard Error) [Observations]	**Eligible Students *(Treatment)***	**Out-of-State Students within Eligibility Range *(Control)***	**Difference**
Course taken after LIFE	3.120 (0.87) [14,118]	3.124 (0.85) [6,127]	−0.004 (0.01)
Course taken before LIFE	2.873 (1.01) [192,181]	2.935 (0.97) [92,866]	−0.062* (0.004)
Difference	0.246* (0.008)	0.189* (0.013)	0.058* (0.016)

Notes: All students in the sample enrolled before the adoption of the LIFE scholarship. Grade is measured in points: A = 4, B = 3, etc.
* significant at 1%

higher grades and entered Clemson with higher SAT scores than in-state students.[16] The grades and SAT scores for both groups increased over time, but they did so at rates that are not statistically different from each other at conventional levels.[17]

4. ESTIMATION RESULTS

Difference-in-Differences

Table 2 compares the change in grades of in-state students who were at risk of gaining or losing the LIFE scholarship with the change in grades of the control group (out-of-state students with comparable grades). On average, the grades of at-risk in-state students, the treatment group, increased by 0.246 grade points, from 2.873 to 3.120. Over the same period, the grades of the control group increased by 0.189 points, from 2.935 to 3.124. Before LIFE, the grades of treated students were significantly lower than those of the control group, but after LIFE this difference is no longer significant at conventional levels.

This pattern of results is consistent with the hypothesis that the LIFE scholarship encouraged students to work harder; however, the increase could

16. For in-state students in the regression sample, the mean SAT math score before the pre-LIFE observations is 566.95, increasing to 567.63 after the scholarship. For out-of-state students in the sample, SAT math scores fall from 578.72 to 569.19. The pattern for SAT verbal scores is similar. Out-of-state student quality falls by a significantly larger amount than the in-state student quality. If the retention decision is correlated with the scholarship, in-state students who previously would have left Clemson may decide to stay. To the extent that this is the case, part of my estimated effort effect may be a result of increased retention.
17. If in-state students are improving at a significantly faster rate than out-of-state students, my estimates would be biased. The difference-in-differences estimate would capture the grade improvement from a pre-existing trend. Regressions of SAT scores and student grades on a set of controls combined with an interaction of in-state and a semester counter show no significant difference in the rate at which out-of-state and in-state student GPAs are increasing.

also be driven by changes in student characteristics or course choices. Both groups of students saw an increase in their grades; both groups also saw an increase in their SAT scores. Some portion of these changes may be due to student quality, but some may also be due to grade inflation. The numbers reported here are consistent with the nationwide grade inflation average of 0.146 grade points per decade estimated by Rojstaczer (2003). To isolate the effect of the scholarships from these other potential factors, the next section presents the estimation results from a multivariate regression at the course level that includes additional controls such as a student's race, gender, SAT scores, relative high school rank, major, and size and subject of the class.

OLS and Tobit Regressions

The first column of table 3 shows the OLS estimation of equation 1. Tobit results are shown in the next three columns.[18] The estimated coefficients of the control variables are by and large consistent with those of previous studies (Maloney and McCormick 1993). They suggest that males on average received lower grades than their female counterparts. As expected, SAT scores are positively associated with grades, and so is a student's relative rank in his or her high school class. On average, students at the bottom of their high school classes received marks almost a full letter grade lower than those at the top. Grades in the students' first semesters were lower, possibly because some students had difficulty making the transition into college or did not find Clemson to be a good institutional fit. The estimate of the second-semester dummy is also negative but not as large. These estimates are based entirely on pre-LIFE variation because there are no post-LIFE observations on freshmen.

The mean SAT score of a student's classmates is negatively related to grades. The relationship between class size and grades is also negative, possibly because professors felt more confident giving lower grades to students when they had more students in their class. Larger classes are also more likely to be lecture oriented, and as a consequence the professor might not develop as strong a relationship with his or her students. A student's credit load was positively related to grades earned. This counterintuitive result could be due to omitted variable bias: students who were unobservably better may also have chosen heavier schedules.

Finally, the relationship between the semester counter and grades is positive and significant. Grades increased over time independently of the LIFE

18. Some of the student characteristics may affect freshmen differently from other students; SAT scores, in particular, are expected to be more predictive of a student's performance as a freshman than as an upperclassman. I interact these variables with a dummy for first and second semester, although these are not reported. The effect of these variables may also vary as the student progresses through his or her studies; interactions between these variables and the number of completed hours are included, though not reported.

Table 3. Regression with Grade as Dependent Variable, Weighted by Credit Hours

Variable	OLS	Tobit	Tobit, only graduating students	Tobit
In-state	−0.103	−0.125	−0.204*	−0.223**
	(1.11)	(1.36)	(2.08)	(2.98)
LIFE	−0.035*	−0.047**	−0.014	−0.064**
	(1.97)	(2.38)	(0.71)	(2.27)
In-state × LIFE[a]	0.073**	0.101**	0.066**	0.104**
	(3.55)	(4.59)	(2.93)	(3.28)
Tuition (in thousands)	−0.007	−0.006	−0.028	0.017
	(0.39)	(0.38)	(1.56)	(1.24)
Male	−0.072**	−0.104**	−0.072**	−0.118**
	(6.32)	(8.83)	(5.80)	(7.15)
Fall	0.009	0.012*	0.011	−0.019**
	(1.59)	(1.69)	(1.55)	(2.38)
Legacy	0.008	0.012*	−0.002	0.009
	(1.39)	(1.98)	(0.32)	(1.00)
SAT math	0.003**	0.005**	0.007**	0.006**
	(3.30)	(5.33)	(6.46)	(6.00)
SAT verbal	0.005**	0.008**	0.011**	0.008**
	(5.45)	(9.50)	(11.53)	(8.00)
Rank/HS class	−0.904**	−1.240**	−1.200**	−1.298**
	(22.98)	(33.01)	(29.59)	(21.63)
First semester	−2.048**	−2.745**	−2.193**	−2.283**
	(21.95)	(31.29)	(22.35)	(18.90)
Second semester	−0.427**	−0.264**	−0.050	−0.493**
	(5.72)	(3.65)	(0.63)	(4.04)
Mean SAT math of rest of class	−0.018**	−0.026**	−0.029**	−0.024**
	(15.77)	(17.96)	(19.04)	(14.37)
Mean SAT verbal of rest of class	−0.006**	−0.007**	−0.006**	−0.007**
	(5.17)	(4.39)	(3.57)	(3.50)
Class size	−0.002**	−0.002**	−0.002**	−0.003**
	(29.36)	(29.32)	(29.36)	(29.53)
Time[b]	0.003**	0.005**	0.006**	0.003**
	(3.57)	(5.29)	(5.68)	(2.36)
Credit load	0.024**	0.032**	0.024**	.
	(18.81)	(24.48)	(16.47)	.
Intercept	3.694**	4.042**	3.569**	3.158**
	(16.37)	(8.91)	(7.62)	(28.71)

Table 3. Continued

Variable	OLS	Tobit	Tobit, only graduating students	Tobit
Major fixed effects	Yes	Yes	Yes	No
Race fixed effects	Yes	Yes	Yes	Yes
Subject and level fixed effects	Yes	Yes	Yes	Yes
Observations	305,292	305,292	243,020	305,292

Notes: All students in the sample enrolled before the adoption of the LIFE scholarship. The LIFE dummy variable equals one after the scholarship begins, zero before it. The dependent variable (grade) is measured in points: A = 4, B = 3, etc. The regression includes but does not report measures of academic progress, measured as completed credits, along with interactions of academic progress with the variables listed above. These results are available upon request.
[a] "×" indicates that the two variables are interacted.
[b] Time is a regular semester counter starting in spring 1990.
* significant at 5%; ** significant at 1%

scholarship at a rate of 0.005 grade points per semester. Over a ten-year period, this translates to an increase of 0.1 grade points, a fraction of the total actual increase. The race, major, subject, and course-level effects are jointly significant at the 1 percent level.

The relationship of primary interest is that between the LIFE scholarship and grades (top panel of table 3). Consistent with the hypothesis that LIFE increased student effort, both the OLS and Tobit estimates show a positive and significant relationship between *In-state* × *Life* and grades. Specifically, the coefficient of the interaction term is 0.101, equivalent to a letter grade for every ten courses a student completed. The effect is not large: a student who before LIFE earned ten Bs now earns nine Bs and one A, about one-tenth of a standard deviation in grades. If the sample is limited to courses taken by students who eventually graduated from Clemson, the estimate falls to 0.066 but is still significant at the 1 percent level.

The final regression drops the student's credit load and the fixed effects for the student's major. Since students can respond to the scholarship by adjusting their credit loads or their majors, a large change in coefficient estimates when these estimates are excluded would suggest that students are acting strategically rather than increasing their effort. Removing these variables, however, leaves the coefficient estimates nearly unchanged.

Sensitivity Analysis

Previous studies have found that merit-based scholarship and tuition assistance have differential effects on men and women. For instance, Dynarski (2004) found a positive relationship between tuition assistance and college completion for women, but none for men. An experiment in Kenya that

offered cash and tuition assistance to the highest performing female students led to improvements in their educational attainment but also improved the performance of the ineligible male students (Kremer, Miguel, and Thornton 2004). Angrist, Lang, and Oreopoulos (2009) examined a Canadian experiment that randomly offered tuition assistance and tutoring to students. They found evidence that such assistance was positively related to performance of females, though the relationship diminished over time. There was no evidence that males responded to the aid.[19]

Table 4 presents estimates of equation 1 splitting the sample by gender. The first two columns correspond to the full sample; the third and fourth columns limit the sample to courses taken by students who eventually graduated from Clemson. The results suggest that males responded to the LIFE scholarship by increasing their effort. Specifically, the estimated coefficient of *In-state* × *LIFE* is positive and statistically significant at the 1 percent level. In contrast, the coefficient of *In-state* × *LIFE* for courses taken by females is small and not statistically distinguishable from zero.

This pattern of results runs counter to what some previous studies have found. However, it should be pointed out that these studies focused on test scores, the decision to enroll, and semester GPA, whereas the focus of this study is the grade received in a particular course, which controls for the potential that males and females respond to scholarships along different margins. For instance, Angrist, Lang, and Oreopoulos (2009) examined the determinants of semester GPA, an outcome that can be altered by taking easier courses, fewer credits, or an easier major. It could also be the case that men are more likely to negotiate with their instructors. The negotiation hypothesis finds support in the fact that men tend to negotiate on salary more often than women (Babcock and Laschever 2003).

This possibility that in-state students increased their level of negotiation after LIFE began highlights one remaining concern: professors could selectively raise the grades of in-state students after the adoption of the LIFE scholarship. Such an assignment of grades could potentially ruin the natural experiment. An increase in successful negotiation by in-state students after the introduction of LIFE would have raised the grades of the treatment group, making it impossible to separate the increase due to the additional negotiation from that due to effort. However, if instructors can be persuaded to change their grades, there are several reasons why it would affect all students, not just potential LIFE recipients. First, instructors who are more likelyto respond to student

19. Garibaldi et al. (2007) found evidence that higher tuition was more positively related to the completion rates of women than to those of men, and Angrist and Lavy (2002) found that women responded more strongly to monetary rewards for high school matriculation than did men.

Table 4. Sensitivity Tests

Estimates by gender of student	***Tobit***		***Tobit, only graduating students***	
Variable	***Male***	***Female***	***Male***	***Female***
In-state	−0.108	−0.216	−0.082**	−0.290
	(0.81)	(1.71)	(0.31)	(2.14)*
LIFE	−0.128**	0.038	−0.157	0.057
	(4.22)	(1.49)	(1.11)	(2.22)*
In-state × LIFE [a]	0.175**	0.026	0.136**	0.004
	(5.18)	(0.92)	(3.92)	(0.29)
Observations	147,994	157,298	113,850	129,170
Pseudo R^2	0.103	0.117	0.100	0.117

Classes with 50 or more students	***Tobit***		***Tobit, only graduating students***	
Variable	***Male***	***Female***	***Male***	***Female***
In-state	−0.168	−0.101	−0.344	−0.135
	(0.76)	(0.50)	(1.44)	(0.62)
LIFE	−0.071	0.097*	−0.046	0.105**
	−1.1	(2.02)	(0.7)	(2.14)
In-state × LIFE [a]	0.201**	0.045	0.193**	0.027
	(2.84)	(0.85)	(2.68)	(0.49)
Observations	45,932	49,066	34,344	39,316
Pseudo R^2	0.120	0.131	0.125	0.134

Classes with only one section	***Tobit***		***Tobit, only graduating students***	
Variable	***Male***	***Female***	***Male***	***Female***
In-state	−0.510	0.352	−0.609	0.186
	(1.37)	(1.09)	(1.52)	(0.54)
LIFE	−0.109	0.074	−0.108	0.062
	(1.58)	(1.36)	(1.52)	(1.09)
In-state × LIFE [a]	0.043	0.002	0.070	0.005
	(0.57)	(0.04)	(0.88)	(0.08)
Observations	23,865	28,107	19,067	23,753
Pseudo R^2	0.111	0.130	0.111	0.132

Table 4. Continued

False-LIFE in 1993	*Tobit*		*Tobit, only graduating students*	
Variable	***Male***	***Female***	***Male***	***Female***
In-state	−0.323	−0.587**	−0.353	−0.342
	(1.52)	(2.85)	(1.52)	(1.56)
False-LIFE	0.007	0.025	0.029	0.044
	(0.24)	(0.91)	(0.93)	(1.57)
In-state × False-LIFE [a]	−0.008	0.013	0.042	0.004
	(0.23)	(0.43)	(1.20)	(0.12)
Observations	96,095	98,384	74,984	81,304
Pseudo R^2	0.103	0.122	0.103	0.122

Notes: The dependent variable (grade) is measured in points: A = 4, B = 3, etc. The regression includes but does not report student characteristics, course characteristics, and measures of academic progress, measured as completed credits, along with interactions of academic progress with the variables listed above. These results are available upon request.

[a] "×" indicates that the two variables are interacted.

* significant at 5%; ** significant at 1%

entreaties could be approached by any student, not just those who needed the grade for the scholarship. Second, verifiable information about who was in-state or out-of-state was not readily available to instructors. Out-of-state students could have used LIFE as a pretense for a higher grade as easily as in-state students. Third, professors may have wanted the final grades to reflect the rank in the students' performance. If two students had identical grades and one asked the professor for a higher grade, there would have been pressure to change the grades of both. These assumptions are key to my identification. Without them, the entire increase in student grades could be due to in-state students increasing requests for grade increases from their professors.

To the extent that professors gave in to student demands and violated the conditions listed above, the estimate of student effort will be biased upward. In addition, in-state students may have been more likely to register in classes with professors who had a reputation for being easy. This registration effect should be dampened, however, because the order in which students registered was defined by random draw of student identification numbers, and both out-of-state and ineligible students also had an incentive to register for classes with easier professors.

Larger classes tend to give lower grades than smaller ones, ceteris paribus (Kokkelenberg, Dillon, and Christy 2008). They could also change the attitude of the students and make it more difficult for students to negotiate with their professors on grades because students have less access to the professor (Glass

et al. 1982).[20] Thus one strategy for assessing the extent to which negation is driving the results in table 3 is to restrict the sample to large classes. The second panel of table 4 limits observations to courses that had an enrollment of fifty or more students (31 percent of the observations). The results for this subsample suggest that LIFE worked through effort as opposed to negotiation. Specifically, the estimated coefficient of 0.201 for this subsample is statistically significant and even larger than the one reported in table 3.

The third panel in table 4 limits the observations to courses that had only one section per semester (17 percent of the observations). If only one section was being offered, then presumably students were more limited in their choice of time slot and professor and therefore could not increase their grades by choosing an easier arrangement. In these regressions, the estimated coefficient of the variable *In-state* × *LIFE* is positive but not statistically significant.[21] This lack of significance suggests that part of the estimate on the full sample could stem from students choosing sections with professors who are more likely to give them higher grades. However, the lack of significance could also be due to other differences in the characteristics of the course. For instance, these courses were also more likely to be upper level and taught by tenured and tenure-track professors instead of higher-grading adjuncts (Moore and Trahan 1998).

The final panel of table 4 presents the results of a falsification test. Prior to fall 1998, the LIFE scholarship had not yet been proposed or instituted. Thus choosing an arbitrary start date for LIFE prior to 1998 and estimating a regression similar to equation 1 should provide no evidence of an incentive effect. The falsification test restricts the sample to courses taken before and after 1993 by at-risk students admitted before 1993. This regression is otherwise identical to equation 1, but *LIFE* is replaced by *False-LIFE* and *In-state* × *LIFE* by *In-state* × *False-LIFE*, where *False-LIFE* equals one for all terms from fall 1993 on. All courses taken by students admitted after fall 1993 are excluded from the regression. If the increase is truly associated with the LIFE scholarship, then estimating the same regression in a year when the scholarship did not exist

20. The fact that large classes are related to lower grades may have led students in the treatment group to avoid them. This does not seem to be the case: a regression of class size on *In-state, LIFE, and In-state x LIFE* shows that the estimated coefficient of *In-state* × *LIFE* is not statistically significant. This suggests that large classes did not experience a change in in-state enrollment concurrent with the timing of the scholarship, providing additional evidence that the increase in grades is a result of increased learning or effort rather than selection into courses.

21. Another method of exploring this issue is to restrict the sample to classes in which the grading should be most objective. When the sample is restricted to first-year mathematics courses, the estimated coefficient of *In-state* × *LIFE* is 0.342 and significant at the 10 percent level. One lingering concern with this result is that, since post-LIFE observations include only students who were admitted before the scholarship, this regression oversamples students who may have been behind on their math requirements.

should have no impact on student effort, and the magnitude of the estimate should be small.

In fact, the estimation shows just that—the coefficient of interest is not significant and is of much smaller magnitude than in the base regression. Although the variables of interest show no significance, the control variables from this regression are significant and of similar magnitudes as those from the actual LIFE year in table 3. Replicating this falsification test using data from other years before the scholarship began gives results similar to those presented.[22]

In regressions not reported, section fixed effects are included in the regression to control for the fact that students may choose those sections taught by sympathetic professors. The results from this estimation still show a positive and significant increase in the grades of in-state students in the critical range, with decreases in the grades of out-of-state students. The estimated effect, however, falls to 0.02, presumably because the variation of grades within a section is smaller than the overall variation in grades.

5. CONCLUSION

South Carolina's LIFE scholarship is related to a significant increase in the grades of in-state students of 0.101 relative to a control group of out-of-state students with comparable GPAs, and the effect occurs primarily on male students. Merit-based scholarships are associated with higher grades, but this relationship explains only a fraction of the increase in grades over time. Merit-based scholarships significantly accelerate the historical increase in grades, but grade inflation exists even without them.

Differences in scholarship details between the LIFE scholarship and Georgia's HOPE, particularly the requirement of a minimum course load for receipt of LIFE, led students to respond on different margins. Cornwell, Lee, and Mustard (2005) showed that students withdraw from more courses, take lower course loads, and take easier courses. Here I show that, controlling for course choices and course loads, the grades of the students who are on the margin of gaining or losing the scholarship increase significantly. Policy makers must take these student responses into consideration when defining the requirements and estimating the costs of merit-based scholarships. Although the magnitude of these findings may not be generalizable to other states, this study corroborates the findings of previous studies, suggesting that students adjust their behavior as a response to merit-based scholarships.

22. The estimates of *In-state* × *False-LIFE* on other years range from −0.008 to 0.06 and are never significant at conventional levels.

Although I argue that these estimates are due to student effort, they may be biased upward if the students who are on the margin of gaining or losing the scholarship become particularly keen at identifying the professors who give higher grades or who are particularly responsive to a sob story. A final mechanism for the increase in grades could be an income effect—the students who receive the scholarship are now able to work fewer hours and have more time for their studies.

I would like to thank Brett Dalton, Michael Maloney, and Robert McCormick, who generously offered the data. I also thank Daniel I. Rees, Doru Cojoc, Angela K. Dills, Sean Mulholland, Christina Peters, and Robert Tollison for their constructive criticism. Finally, I appreciate the research support provided by the Earhart Foundation. All data used in this article are proprietary to Clemson University and cannot be made publicly available. All remaining errors are my own.

REFERENCES

Anglin, Paul M., and Ronald Meng. 2000. Evidence on grades and grade inflation at Ontario's universities. *Canadian Public Policy* 26 (3): 361–68.

Angrist, Jousha, Daniel Lang, and Philip Oreopoulos. 2009. Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics* 1 (1): 136–63.

Angrist, Joshua, and Victor Lavy. 2002. The effect of high school matriculation awards: Evidence from randomized trials. NBER Working Paper No. 9389.

Babcock, Linda, and Sara Laschever. 2003. *Women don't ask: Negotiation and the gender divide*. Princeton, NJ: Princeton University Press.

Binder, Melissa, Philip T. Ganderton, and Kristin Hutchens. 2002. Incentive effects of New Mexico's merit-based state scholarship program: Who responds and how? In *Who should we help? The negative social consequences of merit-aid scholarships*, edited by Donald E. Heller and Patricia Marin, pp. 41–56. Cambridge, MA: Civil Rights Project, Harvard University.

Bugler, Daniel T., Gary T. Henry, and Ross Rubenstein. 1999. An evaluation of Georgia's HOPE scholarship program: Effects of HOPE on grade inflation, academic performance and college enrollment. Atlanta: Council for School Performance, Georgia State University.

Bulkeley, William M. 1997. Would tax plan further inflate college grades? *Wall Street Journal*, 22 April.

Cornwell, Christopher M., Kyung Hee Lee, and David B. Mustard. 2005. The effects of merit-based financial aid on academic choices in college. *Journal of Human Resources* 40 (4): 895–917.

Dee, Thomas S., and Linda A. Jackson. 1999. Who loses HOPE? Attrition from Georgia's college scholarship program. *Southern Economic Journal* 66 (2): 379–90.

Dynarski, Susan. 2004. The new merit aid. In *College choices: The economics of which college, when college, and how to pay for it*, edited by Caroline Hoxby, pp. 63–100. Chicago: University of Chicago Press.

Garibaldi, Pietro, Francesco Giavazzi, Andrea Ichino, and Enrico Retorre. 2007. College cost and time to obtain a degree: Evidence from tuition discontinuities. NBER Working Paper No. 12863.

Glass, G., L. Cahen, M. L. Smith, and N. Filby. 1982. *School class size: Research and policy*. Beverly Hills, CA: Sage.

Greene, William. 1981. On the asymptotic bias of the ordinary least squares estimator of the Tobit model. *Econometrica* 49 (2): 505–13.

Heller, Donald E., and Patricia Marin, eds. 2002. *Who should we help? The negative social consequences of merit-aid scholarships*. Cambridge, MA: Civil Rights Project, Harvard University.

Henry, Gary T., and Ross Rubenstein. 2002. Paying for grades: Impact of merit-based financial aid on educational quality. *Journal of Policy Analysis and Management* 21 (1): 93–109.

Kokkelenberg, Edward C., Michael Dillon, and Sean M. Christy. 2008. The effects of class size on student grades at a public university. *Economics of Education Review* 27 (2): 221–33.

Kremer, Michael, Edward Miguel, and Rebecca L. Thornton. 2004. Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics* 1 (1): 112–35.

Maloney, Michael T., and Robert E. McCormick. 1993. An examination of the role that intercollegiate athletic participation plays in academic achievement: Athletes' feats in the classroom. *Journal of Human Resources* 28 (3): 555–70.

Medina, Jennifer. 2008. Next question: Can students be paid to excel? *New York Times*, 5 March. Available www.nytimes.com/2008/03/05/nyregion/05incentive.html. Accessed 16 June 2009.

Moore, Melanie, and Richard Trahan. 1998. Tenure status and grading practices. *Sociological Perspectives* 41 (4): 775–81.

National Center for Education Statistics (NCES). 2006. Digest of education statistics (tables 173, 347, and 349). Available nces.ed.gov/programs/digest/d06/tables_1.asp. Accessed 3 March 2008.

Rojstaczer, Stuart. 2003. *Grade inflation at American colleges and universities*. Available www.gradeinflation.com. Accessed 8 September 2008.

Rosovsky, Henry, and Matthew Hartley. 2002. Evaluation and the academy: Are we doing the right thing? Grade inflation and letters of recommendation. Cambridge, MA: American Academy of Arts and Sciences Occasional Paper.

Sabot, Richard, and John Wakeman-Linn. 1991. Grade inflation and course choice. *Journal of Economic Perspectives* 5 (1): 159–70.

Singell, Larry D., Glen R. Waddell, and Bradley R. Curs. 2006. HOPE for the Pell? Institutional effects in the intersection of merit-based and need-based aid. *Southern Economic Journal* 73 (1): 79–99.

South Carolina Commission on Higher Education (SCCHE). 2006a. LIFE scholarship program. Available www.che.sc.gov/New_Web/GoingToCollege/LIFE_Hm.htm. Accessed 4 December 2006.

South Carolina Commission on Higher Education (SCCHE). 2006b. LIFE scholarship statistics: Fall 2006. Available www.che.sc.gov/StudentServices/LIFE/Fall2006.htm. Accessed 17 June 2009.

Tobin, James. 1958. Estimation for relationships with limited dependent variables. *Econometrica* 26 (1): 24–36.