

GRADE EXPECTATIONS AND OVERCONFIDENCE: IS ECONOMICS DIFFERENT?

Dirk S. Yandell¹

Abstract

Do students in economics courses predict their course grade more accurately than students in other business school courses? Data over several semesters and from a variety of disciplines suggest that business school students in all disciplines are overconfident. After controlling for other factors, however, Economics and Accounting students predict their grades more accurately than other students.

Key Words: actual grade, expected grade, overconfidence

JEL Classification: A22

Introduction

The causes and consequences of grade inflation, and the effects of grade inflation on both the grades that students expect and their evaluations of professors have been the topics of hundreds of published studies. Although this paper is not intended to summarize this literature, some background is informative.

College and university grade inflation in the U.S. has been a topic of interest since at least the Vietnam era in the mid-1960's. Many time-series studies provide similar estimates of the magnitude of the change in average grades. An early study by Juola (1980) examined data from 180 colleges and reported a GPA increase of 0.432 grade points between 1960 and 1974. Kuh and Hu (1999) find an average increase of 0.27 grade points between 1984-87 and 1995-97. Grove and Wasserman (2004), for the four-year period from 1998 - 2002, report an increase of 0.11 of a grade point. The average annual increase is between 0.024 and 0.028 grade points in each of these studies. The website www.gradeinflation.com, with data maintained and summarized by Stuart Rojstaczer, suggests that the mean grade point average at U.S. universities and colleges has risen slightly slower, by about 0.15 points per decade over a longer time span, from an average of about 2.7 in 1983 to 3.15 in 2013. Over that period, A's have become the most common grade, and the prevalence of C and D grades are at historic lows.

Lowe, Borstorff and Landry (2008), using data from the Baccalaureate and Beyond Longitudinal Study conducted by the National Center for Educational Statistics, find that for the nine year period from 1993 to 2000/01 the average student cumulative GPA increased 0.23 grade points, and major GPA rose 0.21 grade points. They found a higher level of within-major grade inflation for business graduates than for students with majors in engineering, life sciences, mathematics, and physical sciences, but lower than for graduates in education.

Many papers have examined the relationship between expected grades or actual grades on student evaluation of teacher performance. For example, Krautmann and Sander (1999) find that

¹ Professor of Economics, School of Business, University of San Diego, 5998 Alcalá Park, San Diego, CA 92110.

grades affect an instructor's evaluations. Andrew Ewing (2012), using multiple estimation models, also finds that instructors have an incentive to inflate grades to achieve higher evaluation scores. Buying higher course evaluations by giving higher grades can contribute to grade inflation and diminish the effectiveness of student evaluations as a means of evaluating the quality of teaching.

The focus of this paper is not specifically on grade inflation, but on grade expectations and whether students are accurate or overconfident when predicting their final grade. Hossain and Tsigaris (2015) surveyed 169 students in six sections of Business and Economic Statistics over three semesters to examine students' expectations about their final grade. They find that overconfidence is typical, but the difference between expected and actual grades diminishes over the semester as students receive new information about their performance. David Burns (2007) evaluated predicted grades from the start of the semester and from the time of the final exam and found that the accuracy of predictions improved over time and were related to students' class attendance rates, but not related to self-reported study time.

Overconfidence can have both positive and negative effects on student performance and effort. The accuracy of grades expected by students may have implications on the selection of a major and the amount of time dedicated to a course. Grimes (2002) notes that the accuracy of grade predictions influences how students select study material and exert effort to succeed in a course. Jensen and Owen (2000) report that higher expected grades are associated with an increase in students' confidence to succeed in economics. More recently, Main and Ost (2014) find no evidence that course grades affect students' course behavior or major decisions.

Nowell and Alston (2007) examine overconfidence in economics courses and find that male students and low-GPA students exhibit greater grade overconfidence than females and high-GPA students. More to the point of this paper, they also find that students in lower-division courses are more overconfident than upper-division students, a result consistent with widespread overconfidence among principles of economics students reported by Grimes (2002) and Grimes, Millea, and Woodruff (2004).

An interesting question is whether economics courses differ from other disciplines. A study of grading patterns at Wellesley College showed that economics grades were lower than in any other department in the social sciences and humanities, and 5th lowest overall from the 28 departments studied (Butcher, McEwan, and Weerapana, 2014). After examining grades at the University of Michigan across 25 different departments within the College of Literature, Science, and the Arts, Achen and Courant (2009) report that lower-division Economics grades were lower for 2005-2007 than for all but Mathematics and Statistics, and also 3rd lowest overall for upper-division grades (only Biology and Statistics were lower). Schulz and Thöni (2016), examining data from two European universities, find that students of politics, law, business and economics are more overconfident than students in other disciplines.

This paper extends the discussion of grade overconfidence and examines the accuracy of students' expected grades by discipline within a business school. The goal is to see whether the results for economics classes differ from other business school disciplines, after controlling for class-level variation.

Data

Data were collected from over 1600 individual undergraduate courses taught over eight semesters within the Business School at the University of San Diego. The university is a selective private school with mostly resident students of traditional age. Business School students from

2013-2015 had an average SAT score of 1208 and high school GPA of 3.84, compared to 1212 and 3.87 for the entire university.

All courses from each semester were included unless the course enrollment was below 10 students. Eliminating low enrollment classes excludes a few special topics courses, practicums, and internship courses where grading standards might differ from the norm. The median class size in this study is 35 students, with a range from 10 to 55. The unit of observation for this study is the individual course, so data values such as expected grades and actual grades represent an aggregated value from the individual responses of the enrolled students in each class. The ending sample size is 1541.

Expected grades are captured from student evaluation forms, which are completed by students in every class during the last few days of each semester, but before the final exam. The Business School uses the University of Washington Instructional Assessment System evaluation forms, which include the question, "What grade do you expect in this course?" Students select one of nine grade options ranging from A to F (including plus and minus distinctions). Actual grade distributions for each course were obtained from the Registrar's office after the end of each semester. Courses were categorized into six disciplines: Business/Management, Decision Science/Information Technology, Accounting, Finance/Real Estate, Marketing, and Economics. Eight semesters are represented, from Fall 2011 to Spring 2015. Of the 1541 courses, 302 (19.6%) were from Economics. The final sample of 1541 courses represents 87 unique course numbers (11 from Accounting, 18 from Business/Management, 6 from Decision Science, 22 from Economics, 16 from Finance/Real Estate, and 14 from Marketing). These courses were taught over the eight semesters by 154 different instructors.

It is important to note that this is class-level data. For each course, the summary statistics from the course evaluation were obtained, as well as the actual distribution of assigned grades. It is the policy in the Business School that evaluations are completed and submitted confidentially. The instructor is not present when evaluations are completed and results are not made available until after final grades are submitted. For confidentiality and the protection of human subjects, individual evaluations are not available, only the summary results for each course. Although it would be interesting to explore how gender, GPA, and other specific demographic variables influence overconfidence, no student-specific characteristics are available. Some student details can be inferred, however. For example, freshman and sophomore students represent the majority of enrollment in lower-division courses (courses numbered in the 100's and 200's), and upper-division courses are restricted to junior and senior students by school policy. A dummy variable for lower- versus upper-division status is included in the regression model to determine whether there are differences by class level.

The weighted mean response rate on student evaluations is 82.2%. This mean is obtained by dividing the total number of completed evaluations by the sum of all class enrollments. The standard deviation of the response rate over the 1541 courses is 10.8%, and the means and variation across departments are similar. Students who did not complete evaluations were either not present on the day the evaluations were administered or had officially withdrawn from the class within the allowable withdraw period. Withdrawing with a W grade is uncommon. Only 970 W grades were reported (1.99%) out of the 48,891 grades assigned in these 1541 courses. Failing grades are also rare (332 reported, or 0.68%). Although there is slight variation by department as shown in Table 1, the withdraw and failure proportions are too small to have a material impact in this study.

Table 1: W and F grades by department

	# Courses	# Grades	W	F	% W	% F
ACCT	255	8544	284	85	0.0332	0.0099
BUS/MGT	364	11008	123	30	0.0112	0.0027
DSCI	202	6211	149	63	0.0240	0.0101
ECON	302	9952	262	106	0.0263	0.0107
FINA	201	6793	85	31	0.0125	0.0046
MKTG	217	6273	67	17	0.0107	0.0027
Overall	1541	48781	970	332	0.0199	0.0068

Students who major in the business school take seven required lower division preparation courses. Calculus is taught by the Mathematics department outside the business school, but the remaining six are represented here: two economics principles courses, two accounting principles courses, information systems, and business statistics. Principles of Microeconomics is the only course offered by the Business School that students can count in the core curriculum, but the great majority of Micro students are business majors or minors, so any influence from students entirely outside the business school is minimal. Major requirements differ slightly but each BBA major includes at least six required upper division courses and between four and nine additional upper-division courses.

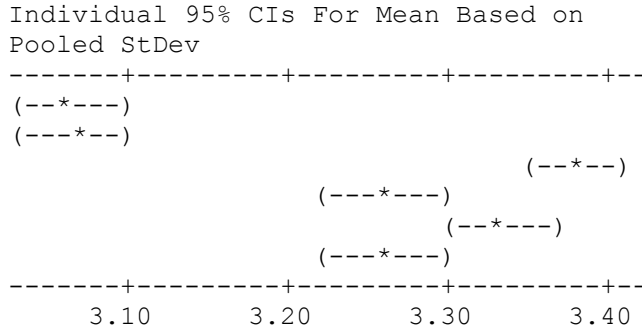
Preliminary Results

The average economics GPA was 3.06, the lowest of the six departments. The school-wide average GPA was 3.23. In every discipline students expected higher grades than they achieved, but the gap was larger in Economics than in any other area. Across all disciplines, students expected an average GPA of 3.40; the average overconfidence was 0.1765 grade points. The correlation between expected grades and actual grades was strong at 0.81, but expected grades were consistently overestimated by students. Figure 1 provides details by discipline.

Figure 1: Actual grades, expected grades, and overconfidence by discipline

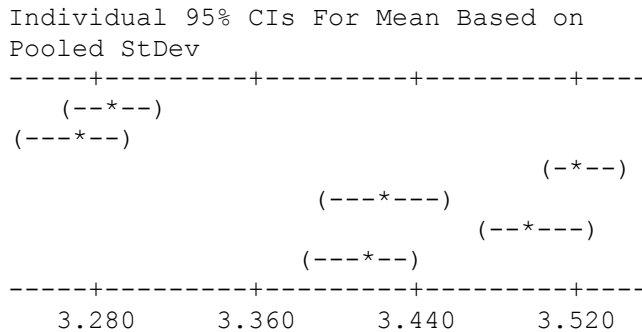
Actual Grade by Discipline

Level	N	Mean	StDev
ECON	302	3.0645	0.2812
ACCT	255	3.0672	0.2424
BUS/MGT	364	3.3792	0.2945 @
DSCI/IT	202	3.2563	0.3507 @
FINA/RE	201	3.3348	0.2813 @
MKTG	217	3.2579	0.2419 @



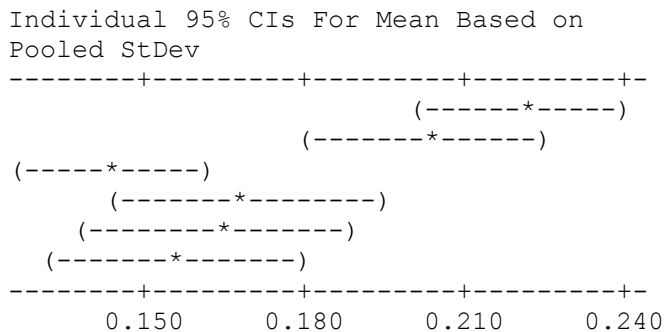
Expected Grade by Discipline

Level	N	Mean	StDev
ECON	302	3.2853	0.2251
ACCT	255	3.2698	0.1914
BUS/MGT	364	3.5237	0.1980 @
DSCI/IT	202	3.4252	0.2905 @
FINA/RE	201	3.4983	0.2167 @
MKTG	217	3.4150	0.1870 @



Overconfidence (Expected Grade - Actual Grade) by Discipline

Level	N	Mean	StDev
ECON	302	0.2208	0.2056
ACCT	255	0.2026	0.1538
BUS/MGT	364	0.1445	0.1793 @
DSCI/IT	202	0.1689	0.2023 @
FINA/RE	201	0.1635	0.1961 @
MKTG	217	0.1571	0.1473 @

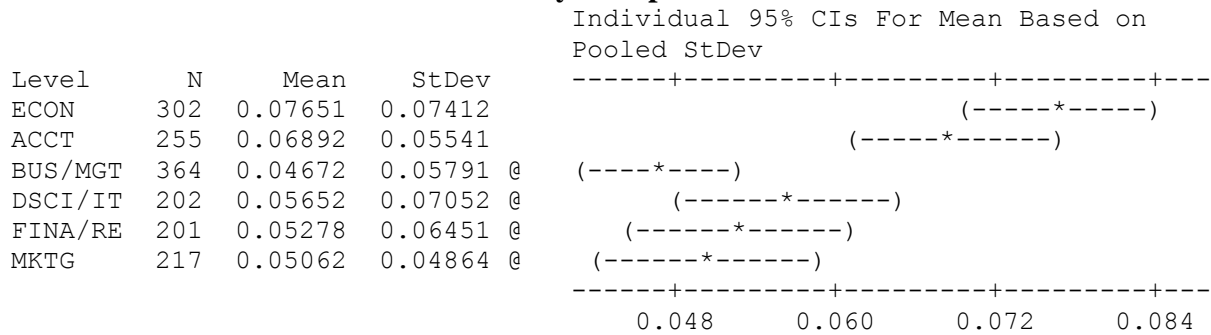


@ denotes that the mean is significantly different from the Economics mean at the 5% level of significance, using a Tukey-Kramer multiple comparison test.

The actual grades in economics and accounting were significantly lower than grades assigned in the other four disciplines, but expected grades were also significantly lower. When overconfidence is examined instead as a percentage of the actual grade rather than as a simple difference between expected and actual grades, the results do not change, as shown in Figure 2.

Figure 2: Overconfidence as a percentage of the actual grade by discipline

Overconfidence as % of Actual Grade by Discipline



@ denotes that the mean is significantly different from the Economics mean at the 5% level of significance, using a Tukey-Kramer multiple comparison test.

Compared to other departments, Economics and Accounting have a higher proportion of lower-division courses in the Business School. Prior studies have found that students in lower-division courses are more overconfident than upper-division students. Does this explain the large difference in expected vs. actual grades for Economics and Accounting? A regression model is estimated to explore this question.

Model

Overconfidence, defined as the difference between the average expected grade and the average actual grade, is calculated for each course and is used as the dependent variable. A positive value for an individual class indicates that students are overly optimistic when predicting their performance, a negative value occurs when actual grades are higher than expected in that class. Out of 1541 classes over the eight semesters examined, the gap is zero or negative in 252 classes and positive in 1289 cases.

On average, students in all disciplines expected higher grades than they actually received, but the gap was largest for Economics. Can this gap be explained by other factors, such as the mix of upper- vs. lower-division classes, the course enrollment, instructor characteristics, or the perceived rigor of the course? The following model is estimated:

$$GAP_{ijk} = \beta_0 + \beta_1 UD_i + \beta_2 CEI_i + \beta_3 Enroll_i + \beta_4 Actual_i + \beta_5 Eval_i + \beta_6 ResponseRate_i + \beta_7 PartTime_i + \beta_8 FT-non-TT_i + \gamma_i Dept_j + \delta_k Semester_k$$

where GAP_{ijk} represents overconfidence, defined as the GPA difference between the expected grade and the actual grade for course i in department j during semester k . The independent variables are:

UD	a dummy variable equal to one if the course was upper division (300 or 400 numbered courses), and zero if the course was lower division (100 and 200 numbered)
CEI	the Achallenge and engagement index score reported from the course evaluations, measured from student responses, used as a proxy for perceived course rigor
Enroll	the number of students enrolled in the course
Actual	the GPA based on grades actually assigned by the professor in that class
Eval	the average course rating from student input on evaluations (0-5 scale)
ResponseRate	the number of completed evaluations for the course divided by the reported enrollment
PartTime	a dummy variable equal to one if the instructor was a part-time adjunct professor
FT-non-TT	a dummy variable equal to one if the instructor was full-time but not tenured or tenure track
Dept	dummy variables for the five departments other than Economics
Semester	dummy variables for the semesters from Spring 2012 to Spring 2015 (Fall 2011 is the default semester)

Summary statistics for all regression variables are shown in Table 2.

Table 2: Summary statistics for regression variables

<i>Variable</i>	<i>Mean</i>	<i>St.Dev.</i>	<i>Minimum</i>	<i>Median</i>	<i>Maximum</i>
GAP	0.1765	0.1842	-0.4662	0.1686	0.8154
UD	0.7236	0.4474	0.0000	1.0000	1.0000
CEI	4.8824	0.5322	3.4000	4.8600	6.6667
Enroll	32.034	8.2310	10.000	35.000	55.000
Actual	3.2269	0.3111	2.0978	3.2316	3.9500
Eval	3.9896	0.5717	1.4810	4.0909	5.0000
ResponseRate	0.8269	0.1078	0.5556	0.8333	1.3571
PartTime	0.2291	0.4204	0.0000	0.0000	1.0000
FT-non-TT	0.2323	0.4225	0.0000	0.0000	1.0000
ACCT	0.1655	0.3717	0.0000	0.0000	1.0000
BUS MGT	0.2362	0.4249	0.0000	0.0000	1.0000
DSCI IT	0.1311	0.3376	0.0000	0.0000	1.0000
FINA RE	0.1304	0.3369	0.0000	0.0000	1.0000
MKTG	0.1408	0.3480	0.0000	0.0000	1.0000

Results

There is no apparent trend or semester effect, none of the dummy variables for semester are significant at even the 10% level. The model is quite robust to model specification, so the semester dummies were removed for simplicity. The OLS estimates of the reduced model are in Table 3.

Table 3: OLS results, dependent variable: GAP (expected GPA - actual GPA)

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
constant	1.74515	0.05950	29.331	0.0000	***
UD	0.02382	0.01073	2.219	0.0266	**
CEI	-0.07955	0.00708	-11.228	0.0000	***
Enroll	-0.00013	0.00042	-0.320	0.7488	
Actual	-0.51951	0.01287	-40.362	0.0000	***
Eval	0.12810	0.00682	18.774	0.0000	***
ResponseRate	-0.07390	0.03105	-2.380	0.0174	**
PartTime	0.00072	0.00838	0.086	0.9313	
FT non TT	0.03271	0.00831	3.938	0.0001	***
ACCT	0.00286	0.01155	0.248	0.8046	
BUS MGT	0.02623	0.01318	1.991	0.0467	**
DSCI IT	0.07990	0.01302	6.137	0.0000	***
FINA RE	0.04707	0.01427	3.298	0.0010	***
MKTG	0.01610	0.01395	1.154	0.2486	

Mean dependent var 0.1765 S.D. dependent var 0.1842

R-squared 0.5494 Adjusted R-squared 0.5456

F(13, 1527) 143.226 P-value(F) 0.0000

Significance level: *** 0.01 ** 0.05 * 0.10

After accounting for other factors, the gap between expected grades and actual grades is actually smaller for Economics than for any of the other departments, as shown by the positive coefficients for the other discipline dummy variables. Overconfidence in Economics and Accounting is essentially the same, but the expectations gap is larger for all of the other disciplines, and significantly larger in Business/Management, Decision Science, and Finance/Real Estate, compared to Economics. The average grade assigned in upper-division courses is higher than the average grade in lower-division courses. This compression of the upper-division grade scale should make it easier for such students to more accurately estimate their grade. The opposite is found here, the impact is small but significant. Students in upper-division courses, all else equal, are more optimistic (less accurate) when predicting their ending grade. This may be true if most upper division courses are in the students' intended major such that they have self-selected into these courses and expect to do well in their area of interest. But, this result is not consistent with Nowell and Alston's (2007) findings that lower-division students were more overconfident than upper-division students.

It is not surprising that as the perceived rigor of the course increases (a higher CEI value), overconfidence decreases. This confirms the expectation that students expect lower grades in harder courses. As expected, higher average actual grades assigned in a course reduce the expectations gap. The results show that each 0.10 grade point increase in the actual average grade, all else constant, leads to a 0.052 grade point decrease in overconfidence. The size of the class did not have a statistically significant effect on the expected grade gap, but the overall course evaluation had a positive and significant effect on the gap. This is consistent with the notion that student evaluations tend to be higher for courses in which students expect to get good grades. Since expected grades and actual grades are highly correlated, this supports the claim that professors can buy better evaluations by grading easier.

There is no difference in the expectations gap between students in courses led by part-time (adjunct) faculty and full-time tenured and tenure track faculty. Students with full-time non-tenure-track faculty (clinical and visiting professors) were the most overconfident.

Alternative model specifications do not change the results. Changing the dependent variable by defining overconfidence as a percentage of the actual grade instead of in GPA units does not change the significance of individual variables. A semi-log model in which the non-binary independent variables are transformed to log values also does not change the results. No multicollinearity is evident. Variance inflationary factors for the thirteen variables are between 1.1 and 3.1, and none of the pairwise correlation coefficients is above 0.39.

To further examine the effect of expected grades on course evaluations, a regression model is estimated using EVAL as the dependent variable. Results are shown in Table 4.

Table 4: OLS results, dependent variable: EVAL

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
constant	-2.08921	0.24508	-8.729	0.0000	***
UD	-0.13689	0.03610	-3.792	0.0002	***
CEI	0.31978	0.02355	13.580	0.0000	***
Enroll	-0.00279	0.00141	-1.973	0.0487	**
Expected	1.35518	0.05213	25.994	0.0000	***
ResponseRate	0.20039	0.10380	1.931	0.0537	*
PartTime	-0.05019	0.02824	-1.777	0.0757	*
FT-non-TT	-0.05468	0.02817	-1.941	0.0524	*
ACCT	0.04170	0.03906	1.068	0.2859	
BUS_MGT	0.10109	0.04452	2.271	0.0233	**
DSCI IT	-0.54545	0.04233	-12.886	0.0000	***
FINA_RE	-0.02309	0.04840	-0.477	0.6334	
MKTG	-0.02986	0.04721	-0.633	0.5271	
Mean dependent var	3.9896	S.D. dependent var	0.5717		
R-squared	0.4640	Adjusted R-squared	0.4598		
F(12, 1528)	110.223	P-value(F)	0.0000		
Significance level:	*** 0.01	** 0.05	* 0.10		

Course evaluations are positively and significantly related to expected grades. The direction of causality remains an interesting question. Do students assign higher course evaluations because they expect higher grades, or are grade expectations higher because they enjoyed and valued the course or instructor more highly? Prior literature predominately assumes the former. It is interesting to note that CEI is positive and significant. This supports the alternate view that students offer higher ratings to courses that provide more value, substance, and engagement, not just higher grades.

Summary

Although at first glance the level of grade overconfidence appeared to be high in economics courses, this analysis shows that the difference between actual grades and expected grades is smaller in Economics than in other disciplines, and significantly less than in three of the five other areas, after accounting for class level, rigor, class enrollment, and other factors.

Economics faculty assigned the lowest actual grades, and significantly lower than in all areas except Accounting. This is consistent with reported results from Wellesley College, where Economics grades were also low compared to other disciplines. Lower actual grades, all else equal, should increase the expected grade gap, but students in Economics classes were actually less overconfident after accounting for other factors. These results add additional evidence to other studies in which economists and economics students have been shown to be different. A summary article by Frank, Gilovich, and Regan (1993), for example, reports that economics students are different in risk attitudes, behave less cooperatively in trust games, and defect more often in prisoners' dilemma games than other students, and that Economics professors give less money to charity. Results presented here suggest that economics faculty and students also are different in the way they assign grades and estimate expected grades.

Although all students are overconfident, it would be nice to believe that our economic lessons about rationality and efficiency make our students less irrationally exuberant than students in other disciplines. Instead, since the teaching methods in Economics classes have been reported to be more traditional than in other disciplines (Becker, 1997), it may simply be that it is easier for economics students to judge their classroom performance. Disciplines other than Economics in the Business School tend to use more teamwork, and rely on projects and cases more than exams, for example, so students may have a harder time estimating their grade when completing these less traditional activities.

References

- Achen, Alexandra C., and Paul N. Courant, 2009. What Are Grades Made Of? *Journal of Economic Perspectives*, 23(3): 77-92.
- Barth, Michael M., Jun Liu and William H. Wells, 2009. A Comparative Analysis of Grading Practices by Discipline within a College of Business, *Academy of Educational Leadership Journal*, 13(4): 93-107.
- Becker, William E., 1997. Teaching Economics to Undergraduates, *Journal of Economic Literature*, 35(3): 1347-1373.
- Burcher, Kristin F., Patrick J. McEwan and Akila Weerapana, 2014. The Effects of an Anti-Grade-Inflation Policy at Wellesley College, *The Journal of Economic Perspectives*, 28(3): 189-204.
- Burns, David J., 2007. An Examination of the Accuracy of Students' Expected Grades, *Academy of Educational Leadership Journal*, 11(3): 45-58.
- Ewing, Andrew M., 2012. Estimating the Impact of Relative Expected Grade on Student Evaluations of Teachers, *Economics of Education Review*, 31(1): 141-154.
- Frank, Robert H., Thomas Gilovich and Dennis T. Regan, 1993. Does Studying Economics Inhibit Cooperation?, *The Journal of Economic Perspectives*, 7(2): 159-171.
- Grimes, Paul W., 2002. The Overconfident Principles of Economics Student: An Examination of a Metacognitive Skill, *Journal of Economic Education*, 33(1): 15B30.
- Grimes, Paul W., Meghan J. Millea and Thomas W. Woodruff, 2004. Grades:Who's to Blame? Student Evaluation of Teaching and Locus of Control. *Journal of Economic Education*, 35(2): 129B47.
- Grove, Wayne, and Tim Wasserman, 2004. The Life-Cycle Pattern of Collegiate GPA: Longitudinal Cohort Analysis and Grade Inflation, *Journal of Economic Education*, 35(2): 162-174.
- Hossain, Belayet, and Panagiotis Tsigaris, 2015. Are Grade Expectations Rational? A Classroom

- Experiment, *Education Economics*, 23(2): 199-212.
- Juola, A. E., 1980. Grade inflation in higher education-1979. Is it over? Retrieved from ERIC database (ERIC Number: ED189129).
- Krautmann, Anthony C., and William Sander, 1999. Grades and Student Evaluations of Teachers, *Economics of Education Review*, 18: 59B63.
- Kuh, G. D., and S. Hu, 1999. Unraveling the complexity of the increase in college grades from the mid-1980s to the mid-1990s, *Educational Evaluation and Policy Analysis*, 21: 297B320.
- Lowe, S. Keith, Patricia C. Borstorff and Robert J. Landry, III, 2008. An Empirical Examination of the Phenomenon of Grade Inflation in Higher Education: A Focus of Grade Divergence between Business and Other Fields of Study, *Academy of Educational Leadership Journal*, 12(1): 15-33.
- Main, Joyce B., and Ben Ost, 2014. The Impact of Letter Grades on Student Effort, Course Selection, and Major Choice: A Regression-Discontinuity Analysis, *The Journal of Economic Education*, 45(1): 1-10.
- Millea, Meghan., and Paul W. Grimes, 2002. Grade Expectations and Student Evaluation of Teaching, *College Student Journal*, 36(4): 582-591.
- Nowell, Clifford and Richard M. Alston, 2007. I Thought I Got an A! Overconfidence Across the Economics Curriculum, *The Journal of Economic Education*, 2007, 38(2): 131-142.
- Schulz Jonathan F., and Christian Thöni, 2016. Overconfidence and Career Choice. *PLoS ONE*, 11(1): e0145126. doi:10.1371/journal.pone.0145126