

RESEARCH ARTICLE

# Grade Inflation or Grade Increase

Anna Yeritsyan, James W. Mjelde\*  and Kerry K. Litzenberg

Department of Agricultural Economics, Texas A&M University, College Station, TX, USA

\*Corresponding author. Email: [j-mjelde@tamu.edu](mailto:j-mjelde@tamu.edu)

## Abstract

Academics worry grade inflation lowers standards; whereas, employers are concerned grades may not distinguish students' abilities. These concerns are dismissed if higher grades reflect improved academic achievement and not inflation. Potential grade inflation is examined using data from 17,696 classes between 1985 and 2019 in the College of Agriculture and Life Sciences at Texas A&M University. Evidence of systemic grade increase is found after controlling for institution, instructor, and student characteristics. Grade increases are partially explained by recruiting better and more female students. Grade inflation, however, may be associated with increased hiring of graduate students and other instructors.

**Keywords:** Grade increase; grade inflation

**JEL classifications:** I23

## Introduction

Concerns about the upward trend in university grades have been voiced by academia and industry. The questions raised include “Are students better prepared or are higher grades being granted easier?” History of concerns about grade increases dates back to at least 1894, when the “Report of the Committee on Raising the Standard, Harvard University, 1894” (as cited in Kohn, 2002, p. 1) stated students are sometimes receiving A and B grades too readily. This upward trend is often noted as grade inflation; however, not all grade increases should be labeled inflation. Of the multiple definitions of grade inflation (Schutz et al., 2015), the definition adopted here is the one used by Kostal, Kuncel, and Sackett (2016, p. 13) which states “. . . increases in grades across time that do not reflect changes in the quality of students' work (cf. Bejar and Blew, 1981).” The key component of this definition is not reflecting “changes in quality of student work.”

Grade inflation raises concerns about the credibility of academia's standards of excellence and accountability through the lowering of academic standards (Kolevzon, 1981). Lowering of standards through grade inflation raises the additional concern that grades are losing their ability to differentiate students' work (Kamber and Biggs, 2004; Pattison, Grodsky, and Muller, 2013; Peace, 2017). This inability to differentiate students' work arises because grades have a ceiling, and one cannot give students a grade higher than the university's highest grade (Kamber and Biggs, 2004). Studies indicate grade inflation is not only a concern in the U.S. (Marini et al., 2018; Baird, Carter, and Roos, 2019), but there is interest in the phenomenon internationally (Bachan, 2018).

Multiple forces have come together to produce grade inflation. Schutz et al. (2015) list 27 reasons for grade inflation that have been proposed in the literature. Within these forces, one finds evidence that instructor, university, and student characteristics may lead to grade inflation (Peace, 2017). Universities competing for reputation, competition for students, and increasing tuition and

costs may lead universities to permit grade inflation to positively influence their reputation and justify tuition increases (Jewell, McPherson, and Tieslau, 2013). Instructors may inflate grades for reasons such as favorably biasing the student evaluations, pressure from their system to maintain student satisfaction, and job security (Eiszler, 2002; Kezim, Pariseau, and Quinn, 2005; Jaschik, 2008; Peace, 2017). Students place pressure on the system for high GPAs because of financial aid considerations and job placement. The notion that with increased tuition comes the right for high grades may also lead to grade inflation (Peace, 2017). Studies such as Boretz (2004) contradict these consumerism conclusions, suggesting students would rather earn higher grades. She notes the widespread acceptance of grade inflation is potentially damaging to universities. Although grades maybe at an all-time high, Boretz (2004) argues faculty development programs and increases in study support services have led to the increases. Along these lines, Jephcote, Medland, and Lygo-Baker (2021) argue unexplained increases in grade should not automatically be viewed as grade inflation. Blindly labeling increases as grade inflation is damaging to higher education. Rather one should see opportunities to investigate the institutional context the grades are occurring.

Grade inflation is an empirical issue; factors, however, can give raise to higher grades that are not attributed to grade inflation as defined above. A positive shift in grade distribution that is caused by increases in achievement or learning is grade improvement, not grade inflation (Mostrom and Blumberg, 2012). Changing demographics and ability of students, increased student support services, improved instruction, and faculty development programs are some factors that can lead to grade improvement (Kuh and Hu, 1999; Boretz, 2004; Mostrom and Blumberg, 2012). Further, Boretz (2004) and Watjatrakul (2014) suggest that the move toward consumerism in higher education increases the quality of education through supporting student learning thus leading to higher grades. The idea of grade improvement was further developed by Flynn (1987), who came up with the phenomena known as “The Flynn Effect.” In his work, the results of IQ tests in schools are increasing on average approximately three points per decade. This finding is supported by Trahan et al. (2014); meta-analysis that finds grades for standardized tests is increasing, which may be attributed to an increase in the population’s intelligence quotient.

The current study uses a unique data set of university, instructor, and student-specific characteristics to examine grades in the College of Agriculture and Life Sciences (COALS) at Texas A&M University (TAMU) from 1985 to 2019. Specifically, three interrelated objectives are to examine: 1) has grade inflation occurred, 2) has the distribution of grades as measured by the standard deviation changed; and 3) how various factors influence class mean GPA and grade distributions.

This study’s contributions include the use of a detailed data set to identify the effect of specific variables. Further, this study explicitly examines a college of agriculture at the class level instead of department- or university-level GPAs. No previous study was found with this explicit, detailed focus at a class level. COALS is one of the largest colleges of agriculture, awarding the most degrees in agriculture in the U.S. (DATAUSA, 2020) in 2019, approximately 3.5% of the degrees awarded nationally. Further, being a land grant university, COALS missions are similar to other land grant universities.

## Literature Review

Given the importance of student success to both the student and university, it is not surprising the number of theories and factors proposed to explain student success. Kuh et al. (2006, p. 9) note “. . . no single view is comprehensive enough to account for the complicated set of factors that interact to influence student and institutional performance . . .” Student performance and grades are generally explained by the school climate including instructor characteristics and student preparedness and motivations (Kuh et al., 2006; Wang and Degol, 2016). Two good reviews are Wang

and Degol (2016) at the pre-university level which provides a comprehensive conceptualization including many factors also relevant at the university level and Kuh et al. (2006) which reviews university level studies with emphasis on student preparedness and behaviors.

Based on previous research, factors influencing grades are categorized into three groups: institutional, instructor specific, and student characteristics. Institutional factors under university, college, or departmental control include educational policies and curricula, resources devoted to teaching, and structural features ranging from quality of the classroom to class features. Instructor's factors include the experience and ability of instructors to convey information to students along with incentives associated with their position. Kuh et al. (2006) note theories of student success in college emphasize the importance of academic preparedness and experiences during college. Variables for these student characteristics, however, are difficult to measure or obtain especially over time. Proxies are commonly used. Further, to help account for institutional and instructor-specific characteristics, three-stage hierarchical models with two entity-specific levels are estimated (one model is developed for class mean GPA and the other for class GPA standard deviation). The two levels are department and instructor. Previous studies have also used fixed effects models to control for unobservable effects (Kokkelenberg, Dillon, and Christy, 2008; Beenstock and Feldman, 2016; Hernández-Julián and Looney, 2016). No study was found that explicitly considered only the college of agriculture with the level of detail in the current study.

### **Grade Inflation and Distributional Changes**

Juola (1979) shows grade increases of approximately 0.4 grade points using data from 1965–1973 with the largest increases occurring in between 1968 and 1970. Studies using data from the 1970s and early 80s show either a decrease or only a small increase in GPA (Juola, 1979; Bejar and Blew, 1981; Adelman, 2004). Later studies suggest grade inflation renewed in the late 1980s (Kuh and Hu, 1999; Rojstaczer and Healy, 2010; Kostal, Kuncel, and Sackett, 2016). Grade inflation leads to changes in grade distributions.

Suslow (1976) notes there has been a shift in grade distributions with a decrease in the proportion of C's, D's, and F's and an increase the proportion of A's and B's. Rojstaczer and Healy (2012) suggest that since the early 1980s, the largest change in grade distribution is an increase in proportion of As with corresponding decrease in Bs and Cs. The proportion of Cs shows the largest decrease. Proportion of Ds and Fs also show a slight decrease. The decrease in Ds and Fs may be a function of increasing withdrawals from classes. Adelman (2004), for example, mentions that an important issue in change in the distribution of grades is a doubling of the percent of withdrawal or noncredit between 1972 and 1992 cohorts. Jewell, McPherson, and Tieslau (2013) suggest grade inflation maybe nonlinear, they find trend, and trend squared variables are significant in explaining grades. In contrast, Kezim, Pariseau, and Quinn (2005) suggest a linear trend in grades over a 20-year period for a private college business school.

Not all studies, however, agree with these findings of grade inflation, suggesting grade inflation does not exist or is a non-issue (Adelman, 2008; Brighthouse, 2008; Pattison, Grodsky, and Muller, 2013). McAllister, Jiang, and Aghazadeh (2008), for example, show increasing grades in an engineering department match increasing achievement potential as measured by standardized test scores. Along these lines, Kostal, Kuncel, and Sackett (2016) note it is hard to indicate the extent of grade inflation, because of the many other factors that can contribute to raising grades; one must remove the influence of these factors to determine the severity of grade inflation. Kuh and Hu (1999) explore grade increases after controlling for student characteristics, such as family educational background, academic achievements in high school, and socioeconomic status. They show grade increases net of these specific controls is slightly larger than without controls. Kostal, Kuncel, and Sackett (2016) results, however, show an opposite effect of including control factors. These studies suggest the need to control for institutional, instructor, and student factors when examining potential grade inflation.

### ***Institutional Factors***

Studies such as Diette and Raghav (2015) and Hernández-Julián and Looney (2016) use fixed effects models to show there are differences associated with students' grades and their major. Danilowicz-Gösele et al. (2017) suggest one reason for differences in departments may be student self-selection. Further, Diette and Raghav (2015) and Hernández-Julián and Looney (2016) discuss different departments have different grading norms that may be used to manage demand for courses and majors. Bond and Mumford (2019) examine increase in grades across various colleges at Purdue University. Their results suggest differences exist between colleges. They show much of the increases in grades in the college of agriculture are driven by increasing student ability. Department/colleges characteristics and differences, however, are difficult to quantify. As such, to account for department-specific characteristics, departments are added as a level in the models.

Empirical studies and theoretical reviews find quantifiable variables under institutional control are important in student success (McElroy and Mosteller, 2006; Cuseo, 2007; Kokkelenberg, Dillon, and Christy, 2008; Henebry, 2010; Wang and Degol, 2016). Cuseo (2007, p. 12) in his review on class size concludes

The research reviewed in this article indicates that large class size is a contextual variable that has generally adverse effects on student learning, mediated primarily by lowering students' level of engagement (active involvement) with the course instructor, with classmates, and with the subject matter.

He goes on to state the only argument for large class size is fiscal. Class meeting times may also influence learning. McElroy and Mosteller (2006) believe "optimal learning times" depends on the students' circadian rhythm. Skinner (1985) finds class mean GPAs are lower for morning classes compared to afternoon and evening classes. Henebry (2010) finds students are more apt to pass a class that meets more than once a week. Skinner (1985), however, finds an opposite effect on the number of class meeting times. He suggests better attendance in classes meeting only once a week may improve grades. Student performance differs for length of class and meeting times be it because of student circadian rhythms or instructor ability.

### ***Instructor-Specific Factors***

The importance of the instructor and instructional method in students' learning is well documented (Kezim, Pariseau, and Quinn, 2005; Kuh et al., 2006; Joyce et al., 2014). Hoffmann and Oreopoulos (2009 p. 491) note "Hard-to-measure instructor qualities may matter more in predicting achievement, even for instructors that exhibit the same age, salary, rank, and gender." Given the importance and difficulty to measure these qualities especially for data going back to 1985, instructors are modeled as a level in the mixed effect models.

In addition to individual instructor effects, other quantifiable characteristics are shown in the literature to influence grades. Studies (Moore and Trahan, 1998; Ronco and Cahill, 2004; Kezim, Pariseau, and Quinn, 2005; Sonner, 2000) suggest grade inflation may be related to the increase in the use of non-tenure track faculty (term used loosely here as different studies used different terminology and length of contract). Kezim, Pariseau, and Quinn (2005) and Sonner (2000) indicate the use of non-tenured faculty is a cause of grade inflation, whereas Figlio, Schapiro, and Soter (2015) find students learn more from non-tenured faculty, which would not be grade inflation. On the other hand, Chen, Hansen, and Lowe (2021) state instructors hired on a temporary, part-time basis assign higher grades than full-time instructors without a discernable increase in learning. Hoffmann and Oreopoulos (2009) and Solanki and Xu (2018) find instructor gender plays a small but statistically significant role in student achievement. Both Hoffmann and Oreopoulos (2009) and Solanki and Xu (2018) indicate a same-sex instructor increases grade performance. Kapitanoff

and Pandey (2017) suggest at least for the first exam, females underperform at a higher rate than males if the instructor is also female. Stratton, Myers, and King (1994) find as instructors' experience measured as years teaching increases, grades also increase. In summary, both quantifiable and unquantifiable instructor characteristics have been shown to influence student's performance.

### **Student Specific Factors**

Kuh et al. (2006) review theoretical perspectives on student learning including sociological, organizational, psychological, cultural, and economic perspectives. They conclude the students "... wide path with twists, turns, detours, roundabouts, and occasional dead ends ..." is more realistic than a direct route to student achievement (Kuh et al., 2006, p. 7). In their framework, besides institutional factors, pre-college experiences and student behaviors or engagement also matter in student success. Pre-college experiences include enrollment choices, academic preparations, aptitude and college readiness, educational aspirations, family and peer support, motivation to learn, gender, race, ethnicity, and demographics. Of these experiences, Kuh et al. (2006, p. 31) conclude, "In fact, the best predictor of college grades is the combination of an individual student's academic preparation, high school grades, aspirations, and motivation." Kobrin et al. (2008) and Westrick et al. (2015) indicate the importance of pre-college preparation in explaining the first year success of college students. Rothstein (2004), however, finds that much of the SAT predictive power on grades is derived from SATs correlation with high school demographics. Studies also suggest changing demographics may lead to higher grades (Kuh and Hu, 1999; Kostal, Kuncel, and Sackett, 2016). Kuh and Hu (1999) suggest increased female enrollments in colleges may lead to higher grades because females tend to receive higher grades, while Bergtold, Yeager, and Griffin (2016) did not find any significant differences between performances of male and female students.

Although brief, this review shows the importance of controlling for factors that may influence grades, therefore, creating a better picture whether grade inflation is occurring. Further, it is not just the institution, instructor, or student factors that determines student success, but rather the interaction of all of these factors.

## **Data and Empirical Model**

### **Data**

Data for 14 departments (13 departments and one degree which is considered a separate department, but classes are taught by various departments) in COALS from Spring 1985 through Fall 2019 (the last pre-COVID-19 semester) are utilized giving 35 years of data, far more than many studies. The data starts in fall 1985 because this is the first semester grade data were available. The primary reason behind selecting COALS is access to the data. Another important factor is the wide range of diverse disciplines within COALS which ranges from basic science and technology to economics and agricultural communication. This wide spectrum of disciplines helps make the findings relevant to other universities. In addition, methods utilized are applicable to other universities. Data from the University are supplemented with information from undergraduate catalogs, faculty and department websites, e-mails to instructors, as well as discussions with staff and faculty in the various departments. Departments changing names and combining departments resulted in no data for Forestry and Rangeland Science Departments. A total of 17,696 usable observations are obtained. All data is at the class level. Grading scheme used by TAMU is a four-point scale based on receiving four points for an A, three for a B, two for a C, one for a D, and zero for an F (TAMU does not use grade modifiers of plus or minus). Class mean GPA is based on the number of students receiving each grade. Grade distribution is based on the standard deviation of the grades received in the class. Individual projects, study abroad,

and internship classes are not included. These classes are eliminated because students tend to receive A's or credit for completion in these classes. Distant learning or in-absentia classes are eliminated because they do not have a specific meeting time and number of classes per week, variables used in the model. Classes with variable credits are also eliminated. Further for privacy reasons, information for any classes of less than five students, are not available. Summer classes are also eliminated from the analysis, because most summer courses, especially early in the data set, fall into one of the categories of classes that are dropped. Further it is felt summer courses being shorter in length either five or 10 weeks instead of 14 or 15 weeks (changed over the data set) may not be directly comparable. Data from Texas A&M were given by sections. In some cases, classes in COALS might have different sections taught at the same time by the same instructor. Sections that were held at the same time and by the same professor were merged to reflect the fact that the sections meet as one class. The different sections are generally for accounting and reporting purposes and not a difference in the class delivery.

Student grade information is highly protected by students' privacy rights laws. Because of this privacy, this research uses class level data rather than individual student data. Over the data period, changes occurred in the data reporting formats and requirements from year to year, specifically in student admissions information and information reported in university catalogs. As such, in many cases, department and administration leadership were contacted to understand data changes. Further, much of the faculty information was obtained by internet searches or contacting instructors directly. These sources of information were particularly used for early years and non-tenure track instructors.

### ***Variable Descriptions and Summary Statistics***

Based on the literature, a theoretical model of the three sets of factors influencing grades is presented in Figure 1. Institutional, instructor, and student factors come together to determine the class mean GPA and standard deviation. Empirical variables used as proxies for the theoretical model, their descriptions, and summary statistics are presented in Table 1 and Figure 1. Most of the variables are self-explanatory. Class averages for variables that are based on student information, such as SAT scores, are based on all students in the class that had such information reported. TAMU, COALS, and departments early in the data set did not require students to provide SAT scores or their rank in high school for admission to the University. As such, the class average may be based on less students than in the class, especially in the early years. To account for this non-reporting, a qualitative 0–1 variable for the years 1985–1988 is included with one indicating the years 1985–1988. This variable is then interacted with SAT and High School rank. Variables associated with Association of American Universities (AAU) and foreign universities are included as a proxy for potential differences in instructor background (not quality). AAU is comprised of some of the leading universities in the U.S. and Canada.

Qualitative variables' (listed as 0–1 variables in Table 1) mean values can be interpreted as the share of class in the category. As to institutional variables, most classes meet either in the morning or afternoon, meet twice a week, and are upper division classes. Average class size is 51 students but ranges from five to 349 students. Approximately 78 percent of the instructors are male. Few instructors' last degree is from a foreign institution. At the time of their graduation, slightly more instructors received their degree from non-AAU-affiliated institutions than from AAU-affiliated universities. Average percent of male students in class is slightly higher than females. SAT scores are somewhat above the national average (College Board, 2015). High school rank average is approximately the 79 percentile (or top 21% of the graduating class in the high school). Students took on average 14 credits per semester with only a small percentage of students dropping or receiving a no grade in a class.

In Figure 2, mean GPA and standard deviation are graphed by semester for all classes. Mean weighted GPA overall years is 3.08. In fall 1985, mean weighted GPAs is 2.88. By fall 2019, GPA

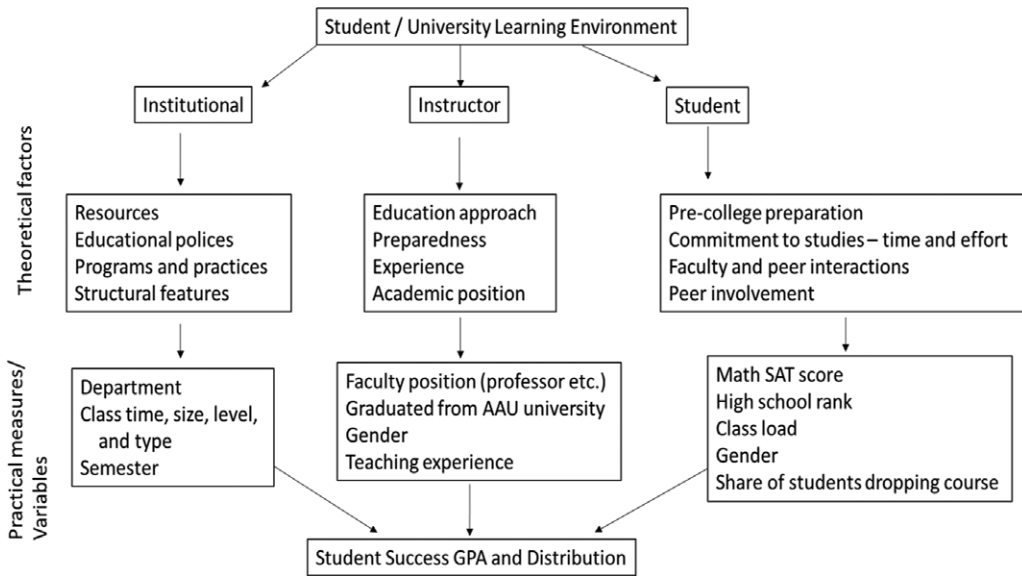


Figure 1. Institutional, instructor, and student variables that may affect a students' grade.

has increased to 3.28. Standard deviation declined from 0.85 to 0.73. Examining the percentage of letter grades helps explain the above increase in GPA and decrease in standard deviation (Figure 3). The percentage of A's increases from about 32% to over 52%. All other letter grades percentages have decreased over the time frame with the percentage of C's showing the largest decrease (from approximately 22% to a little more than 13%). Mean GPAs and standard deviations by department show considerable variation, with mean GPAs, ranging from 2.79 for the Department of Renewable and Natural Resources to 3.46 for the Department of Agricultural Leadership Education and Communication (Figure 4). Standard deviation and GPA show an inverse relationship caused by the increase in A's and decrease in F's. Such departmental differences provide support for the inclusion of departments as a level in the model.

**Estimation Methodology**

When discussing observations that are a part of one group or a cluster, there is risk of drawing inferences about the groups based on simple aggregation of the individual data. Diez-Roux (1998) argues the relationships observed at the individual level may not hold between the same variables observed in a group, which will result in inferential errors. At the group level, there may be specific group properties that will affect the variables in the same way, regardless of the individual characteristics. In the scope of this study, there may be department-specific characteristics that might follow the same pattern within the same department. This means that data within the same department might be correlated, whereas data from different departments may be considered independent. A similar argument holds for instructors.

Multilevel mixed modeling accounts for both within-individual and between-individual variations (Douglas, 2004; Wu, 2019; Shinn and Rapkin, 2000). Following Kokkelenberg, Dillon, and Christy (2008), Beenstock and Feldman (2016), and Hernández-Julián and Looney (2016), a mixed effect model is used to incorporate the variations of grading between the departments and individual instructors, as well as to account for the internal correlation while controlling for other factors.

**Table 1.** Variable description and summary statistics

Variable Name	Variable Description	Mean	Standard Deviation	Minimum Value	Maximum Value
<i>Dependent Variable</i>					
GPA	Unweighted class mean GPA	3.216	0.490	1.2	4
StDev	Standard deviation of the grades received in the class	0.774	0.303	0	1.952
<i>Grade Inflation Variable</i>					
Ln trend	Natural logarithm of trend as given by semester	3.559	0.652	0	4.263
<i>Institutional Variables</i>					
Department	Name of the department used as level, 14 departments				
Morning	Equals 1 if class starts before 12:01, 0 otherwise (dropped to avoid perfect collinearity)	0.623	0.485	0	1
Afternoon	Equals 1 if class starts between 12:01 and 15:59, 0 otherwise	0.356	0.479	0	1
Evening	Equals 1 if class starts at 16:00 or later, 0 otherwise	0.021	0.144	0	1
Meet 1	Equals 1 if the class meets once per week—usually class duration is 2.5 hours for a three-credit class, 0 otherwise	0.288	0.452	0	1
Meet 2	Equals 1 if the class meets twice per week—usually class duration is 75 minutes for a three-credit class, 0 otherwise	0.487	0.500	0	1
Meet 3	Equals 1 if the class meets three times per week—usually class duration is 50 minutes for a three-credit class, 0 otherwise (dropped to avoid perfect collinearity)	0.226	0.418	0	1
Lower division	Equals 1 if the class is listed as a 100 or 200 level class, 0 otherwise (dropped to avoid perfect collinearity)	0.232	0.422	0	1
Upper division	Equals 1 if the class is listed as a 300 or 400 level class, 0 otherwise	0.768	0.422	0	1
Total Students	Number of students receiving a grade A–F and no grades (see share below) in the class	51.084	50.334	5	349
Low_credit	Equals 1 if the class is 1 or 2 credit hours, 0 otherwise—very few classes are 2 credits	0.257	0.437	0	1
High_credit	Equals 1 if the class is 3 credit hours or more, 0 otherwise very few classes have more than 3 credits	0.742	0.437	0	1
Semester	Equals 1 for classes held in the fall and 0 for spring classes	0.514	0.500	0	1
Y85-88	Equals 1 if the year is 1985, 1986 or 1987, 0 otherwise	0.033	0.179	0	1
Y85-88*SAT	Interaction term between Y85-88 and SAT—defined under student variables	36.001	194.366	0	1359.5
Y85-88*HS	Interaction term between Y85-88 and High School—defined under student variables	2.713	14.857	0	99.434
<i>Instructor Variables</i>					
Instructor	Instructor name used as a level, 1377 instructors				
Gender	Gender of the instructor, male = 1 and female = 0	0.780	0.414	0	1

(Continued)



Table 1. (Continued)

Variable Name	Variable Description	Mean	Standard Deviation	Minimum Value	Maximum Value
Prof	Equals 1 if the position at the time of instruction was professor, 0 otherwise (dropped to avoid perfect collinearity)	0.414	0.493	0	1
Assoc prof	Equals 1 if the position at the time of instruction was associate professor, 0 otherwise	0.212	0.408	0	1
Assist prof	Equals 1 if the position at the time of instruction was assistant professor, 0 otherwise	0.134	0.341	0	1
Lec Grad	Equals 1 if the position at the time of instruction was graduate student, 0 otherwise	0.101	0.301	0	1
Other	Equals 1 if the position at the time of instruction was other lecturer, 0 otherwise (includes visiting faculty, lecturers, non-graduate instructors)	0.140	0.347	0	1
AAU	Equals 1 if the university was AAU member at the time of the instructor's graduation (includes Canadian universities), 0 otherwise (dropped to avoid collinearity)	0.365	0.481	0	1
Non-AAU	Equals 1 if the university was not AAU member at the time of the instructor's graduation (includes Canadian universities), 0 otherwise	0.527	0.499	0	1
Foreign	Equals 1 if the instructor's terminal degree was from a non-US or non-Canadian university, 0 otherwise	0.019	0.137	0	1
<i>Student Variables</i>					
Percent Male	Percentage of male students in the class	0.497	0.208	0	1
SAT	Class average of students' combined SAT scores	1101.751	71.280	721.667	1472.500
Load	Average number of credits students in the class are enrolled	13.959	0.762	2.828	20.667
High School rank	The average high school rank of students in the class, calculated as the percentile of students in the school that rank below the given student	79.424	8.375	12.903	99.857
Share	Share of students who enrolled in the class but did not receive an A-F grade for the class. Includes students who dropped beyond the initial drop date, received an incomplete grade, took the class pass/fail, or was dropped from the class by the dean's office divided by total students	0.032	0.050	0	0.982

The model is a three-stage hierarchical model with two levels, where the first stage measures the fixed effect or within-individual variation, while the second and third stages measure the random effect or the between-individual variations for instructors and for departments. The estimation consists of separate sets of regressions nested within each other (Greene, 2012). Combining the stages, one obtains

$$y_{ijk} = \beta_0 + X_k\beta_k + u_i + u_{ij} + \epsilon_{ijk} \tag{1}$$

where

$k$  indicates an individual class;

$j = 1, \dots, 1359$  instructors nested within  $i = 1, \dots, 14$  departments;

$y_{ijk}$ —is the individual class mean GPA or class GPA standard deviation;

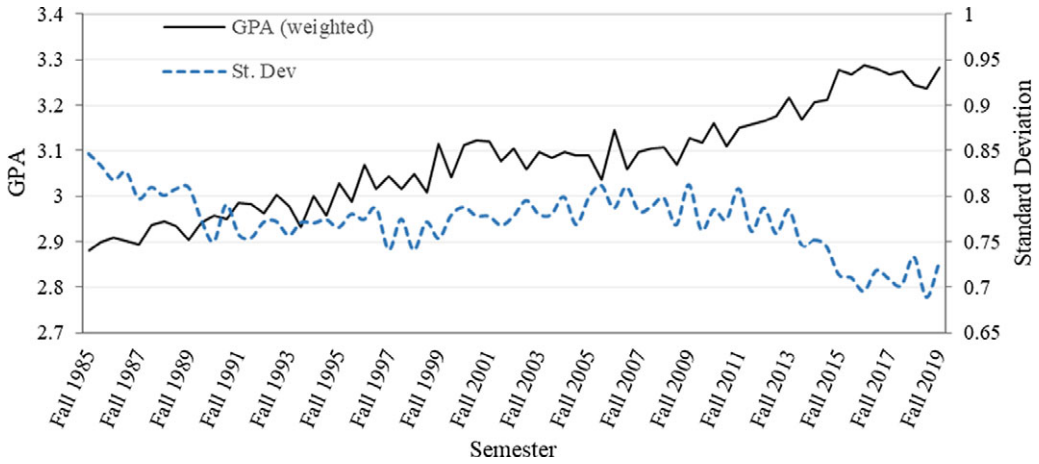


Figure 2. Weighted (by class size) average COALS GPA and standard deviation by semester from fall 1985 to fall 2019.

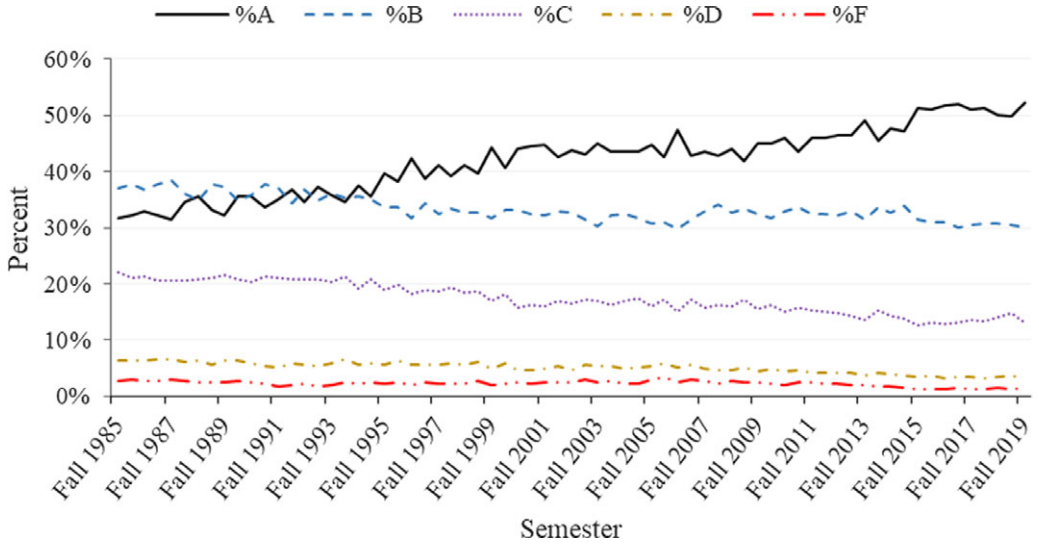


Figure 3. Percentage of letter grades in COALS in all courses from fall 1985 to fall 2019.

$\beta_0$ —is the intercept to be estimated;  
 $\beta_k$ —is a vector of coefficients to be estimated associated with the fixed part of the model;  
 $X_k$ —is the matrix of independent variables described in Table 1;  
 $u_i$ —is the particular department GPA (standard deviation) variation from the mean GPA (standard deviation) in the fixed part of the model;  
 $u_{ij}$ —is the variation of the instructor mean GPA (standard deviation) from the department mean GPA (standard deviation) in the random part of the model; and  
 $\epsilon_{ijk}$ —is the variation of each class GPA (standard deviation) from the instructor mean GPA (standard deviation) in the random part of the model.

The model is estimated using mixed command in Stata 15 (StataCorp, 2017).

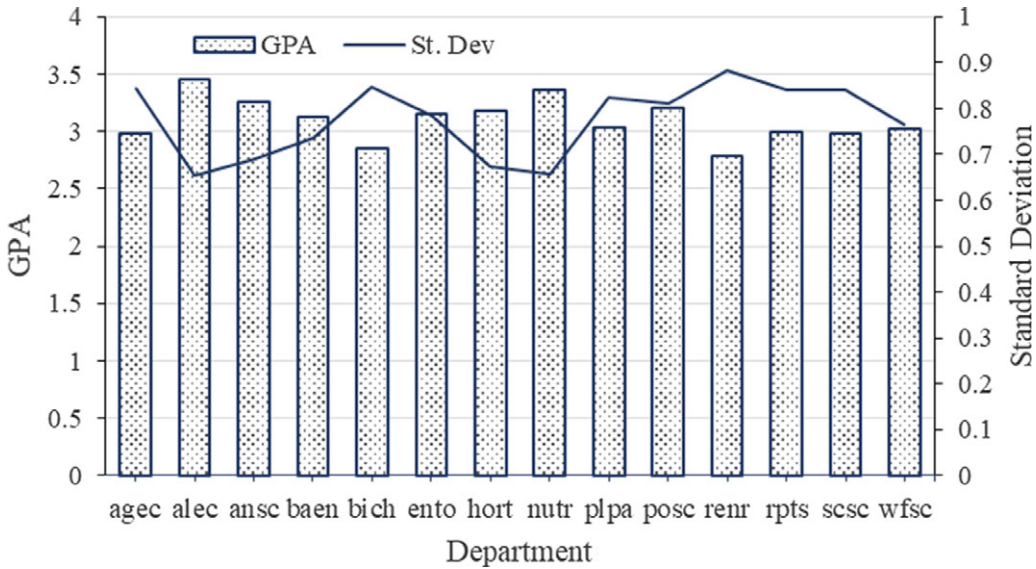


Figure 4. Overall weighted by student mean GPA and standard deviation by department in COALS for the fall 1985 to fall 2019.

**Results and Discussions**

Estimation results for both the GPA and standard deviation (StDev) models are presented in Table 2. Wald statistics for both models indicate jointly all variables except the intercept differ from zero at p-values less than 0.01. Likelihood ratio tests reject the null hypotheses; the models including random effects are equivalent to a linear regression model. Evidence of clusterization of the data is tested using interclass correlation coefficients, which are the ratio of one variance to the sum of the other two variances. The coefficient for level two, for example, is the ratio of level two estimated variance to the sum of the level one and level two estimated variances. Because this measure is greater than 0 (0.14 for the department and 0.55 for the instructor given department), the conclusion clustering is indicated (Park and Lake, 2005). The selection of both levels (random effects) is also supported by model selection information criteria, and minimum values were attained using both levels when compared to one level or no level. Finally, both random effects parameters differ from zero in the two models. As is common practice, the coefficients associated with the levels are not presented given the number of coefficients, rather just over all significance is discussed. Overall, model statistics provide evidence for the appropriateness of the mixed effects model.

**Trend Variables**

In the GPA model, the trend variable is significant and positive, indicating over the period 1985–2019, there is evidence of systematic grade increases in COALS (Table 2). Without the department and instructor levels included in the model, trend remains positive, significant, and is larger in magnitude. Department and instructor levels appear to partially account for the increases in GPA seen in Figure 2. In the StDev model, the trend variable is significant and positive. Without the department and instructor levels, the coefficient on trend in StDev model is not significant.

**Table 2.** Estimated coefficients of the GPA and StDev models

Variable Name	GPA			StDev		
	Coefficient	Std. Error	P-value	Coefficient	Std. Error	P-value
Constant	1.552	0.093	0.000	1.862	0.066	0.000
<i>Grade Inflation Variables</i>						
Ln trend	0.026	0.007	0.000	0.019	0.005	0.000
<i>Institutional Variables</i>						
Afternoon	0.051	0.006	0.000	-0.045	0.005	0.000
Evening	0.047	0.018	0.010	-0.077	0.014	0.000
Meet 1	0.091	0.011	0.000	-0.035	0.009	0.000
Meet 2	0.003	0.008	0.653	0.003	0.006	0.654
Upper division	0.004	0.009	0.683	-0.055	0.007	0.000
Total Students	-0.002	0.0001	0.000	0.001	0.0001	0.000
High credit	-0.261	0.010	0.000	0.116	0.008	0.000
Semester	0.008	0.005	0.105	-0.006	0.004	0.129
Y85-88	0.727	0.232	0.002	-0.745	0.178	0.000
Y85-88*SAT	-0.0004	0.0002	0.065	0.001	0.0002	0.001
Y85-88*HS	-0.004	0.001	0.000	0.002	0.001	0.004
<i>Instructor Variables</i>						
Gender	-0.013	0.019	0.500	-0.016	0.011	0.167
Assoc prof	0.021	0.010	0.028	0.009	0.007	0.201
Assist prof	0.015	0.012	0.227	0.024	0.009	0.006
Lec Grad	0.038	0.019	0.041	0.016	0.012	0.205
Other	0.100	0.022	0.000	-0.032	0.014	0.026
Non-AAU	-0.022	0.011	0.044	0.009	0.008	0.242
Foreign	-0.092	0.037	0.013	0.044	0.025	0.081
<i>Student Variables</i>						
Percent Male	-0.181	0.019	0.000	0.137	0.015	0.000
SAT	0.001	0.0001	0.000	-0.001	0.0001	0.000
Load	0.012	0.003	0.000	-0.016	0.003	0.000
High School rank	0.004	0.0004	0.000	-0.003	0.0003	0.000
Share	-1.046	0.049	0.000	0.638	0.038	0.000
<i>Random -Effects Parameters</i>						
Department	0.026	0.010		0.007	0.003	
Instructor	0.074	0.004		0.020	0.001	
Residual	0.082	0.001		0.051	0.001	
<i>Overall Model Fit</i>						
Wald Test	$\chi^2 = 7577.47$		P-value = 0.000	$\chi^2 = 3957.16$		P-value = 0.000
Likelihood Ratio vs. Linear Model	$\chi^2 = 9046.42$		P-value = 0.000	$\chi^2 = 4272.74$		P-value = 0.000

### ***Institutional-Level Data***

Estimated institutional random effects parameters are significant in both the GPA and StDev models indicating common grading patterns within a department. These results are in line with previous studies such as Diette and Raghav (2015), Hernández-Julián and Looney (2016), and O'Connor (2020). Variables associated with Y85-88 indicate university reporting requirements may help explain GPA. All other institutional variables in the GPA model except semester, upper division classes, and meeting twice a week are significant at p-values of 0.01 or less indicating institutional variables affect student learning as measured by GPA. Semester being insignificant indicates class GPA does not vary by spring and fall semesters, a pleasing result for COALS. GPAs for classes held twice a week do not significantly differ from classes held three times a week. However, GPAs for classes meeting once a week are higher than those meeting three times a week. Classes that meet more often per week are generally shorter in duration (Table 1). Contrary to Henebry (2010), longer classes appear to result in higher GPAs. This result may be because longer duration classes tend to more specialized attracting interested and motivated students.

Classes held during the morning hours result in lower GPA than classes that are held in the afternoon and in the evening. Based on a Wald test, grades for evening and afternoon classes do not significantly differ from each other. One possible explanation for morning classes having lower GPAs is students in morning classes may not have had enough sleep, which affects their performance during the class (Wahlstrom, Hendrix, and Frederickson, 1998; Wheaton, Chapman, and Croft, 2016). Further, evening classes tend to be more specialized courses that may be contributing to the larger GPAs. There is no significant difference between grades in upper and lower division classes. Classes with three or more credit hours tend to have lower GPAs than classes with one or two credit hours.

GPA is negatively correlated with the class size, which is consistent with previous findings (Kokkelenberg, Dillon, and Christy, 2008). Higher grades for smaller classes may be a result of better learning outcomes (Nye, Hedges, and Konstantopoulos, 2000), because in smaller classes instructors are able to provide more time to individual students both during class and office hours (Jewell, McPherson, and Tieslau, 2013).

All institutional variables' coefficients in the StDev model are significant at p-values of 0.01 or less except those for semester and classes that meet twice a week. The signs of the significant coefficients in the StDev model tend to be opposite of the signs in the GPA model. Generally, variables that increase GPA tend to decrease standard deviation or variability of grades. Most likely cause of this inverse relationship is GPAs having an upper bound of four.

### ***Instructor Specific Variables***

As in the case with departments, the random part of the model indicates significant variability in grading patterns of individual instructors. Coefficients associated with associate professor, lecturer graduate student, and lecturer other are significant and positive indicating classes taught by these instructors on average have higher GPAs. The increasing trend of using graduate students and other lecturers accounts for some of the increase seen in GPAs overtime. The questions are do students learn more from these non-professorial rank instructors, which would not be grade inflation, or do these instructors grade easier, which is grade inflation? Unfortunately, the data set cannot answer these questions. Anecdotal evidence provided by some graduate students suggest their incentive is to get good teaching evaluations to help in obtaining the first faculty position. Easy grading resulting in higher GPAs may contribute to higher evaluations (Peace, 2017; Howard and Maxwell, 1980; Eiszler, 2002). Kezim, Pariseau, and Quinn (2005) and Sonner (2000) also find visiting and adjunct faculty grades are higher. They suggest that the reason for the higher grades may be the expectation of higher student evaluations because these instructors are often hired on the term-by-term basis and higher student evaluations are more likely to result in their contract

being extended. Classes with instructors who graduated from an AAU university have higher GPAs than classes with non-AAU or foreign universities' instructors. Instructor gender does not significantly influence GPAs. This result differs from some previous studies (Moore and Trahan, 1998; Kezim, Pariseau, and Quinn, 2005) and is in line with others (Figlio, Schapiro, and Soter, 2015; Ronco and Cahill, 2004). The insignificance may arise from controlling for individual instructor variability.

In the StDev model compared to professors, assistant professors tend to have more variable GPAs, while other faculty members show smaller variability in grades. Assistant professors having a larger variability may be because of differences in experience and the number of times an instructor has taught a particular class.

### ***Student Specific Variables***

All student variables are significant. Coefficients associated with variables measuring student motivation and ability, SAT, class load, and high school rank are positive. These results are supported by previous studies, such as Kobrin et al. (2008) and Westrick et al. (2015) who find high school performance and innate ability of students are good indicators of university performance. Student motivation to achieve and less time to procrastinate may explain why larger class loads results in slightly higher GPAs. As the percent of no grades increases, the class GPA decreases, which may be because student tend to drop harder classes or classes with more work to maintain their GPA. As the percent of females in the class increases (percent male decreases), GPAs increase. Kuh and Hu (1999), and Voyer and Voyer (2014) find females tend to have higher grades than male students.

Both the number of female students in COALS and the students average SAT have been increasing over time (Figure 5). These increases help explain the increasing grades in COALS. Increasing SAT would not cause grade inflation, because better students are expected to obtain better grades. The reason females tend to have better grades determines if the increased percent of females is associated with grade inflation. If instructors simply grant higher grades to females, then grade inflation is present; however, if females are more motivated and study harder, it is not grade inflation. Unfortunately, the dataset cannot clarify this issue. Dampening the GPA increase is student load and high school rank, both of which decrease slightly overtime.

All student variables are significant in the StDev model. There is higher variability in grades for classes with more male students than female students. The higher the student load, SAT score, and high school rank, the lower the variability in grades.

#### *Excluding Earlier Years*

Grade inflation is considered an empirical issue occurring over decades, as such the analysis used all years in which data are available. To examine if problems with data consistency in the years 1985–1988 impact the results, the models are re-estimated excluding these years. In removing the years, 1,735 observations are lost. All coefficients' signs are unchanged. Both trend and lecture grades coefficients are less significant but still significant at the six percent or less. As such, there is no real change in inferences between including or excluding the early years.

### **Discussion**

As discussed earlier, grade inflation is important to universities, instructors, and students. Universities competing for reputation, students and resources may permit grade inflation to occur (Jewell, McPherson, and Tieslau, 2013). Instructors may inflate grades for personal gain such as favorably biasing student evaluations (Jaschik, 2008; Peace, 2017), whereas students may pressure academics for higher grades (Peace, 2017). Grade inflation, however, not only affects universities, instructors, and students, but also spills over into the employment sector where the recruitment of a good candidate is often based on university GPA (Adams, 2015; Fossati, Wilson, and Bonoli,

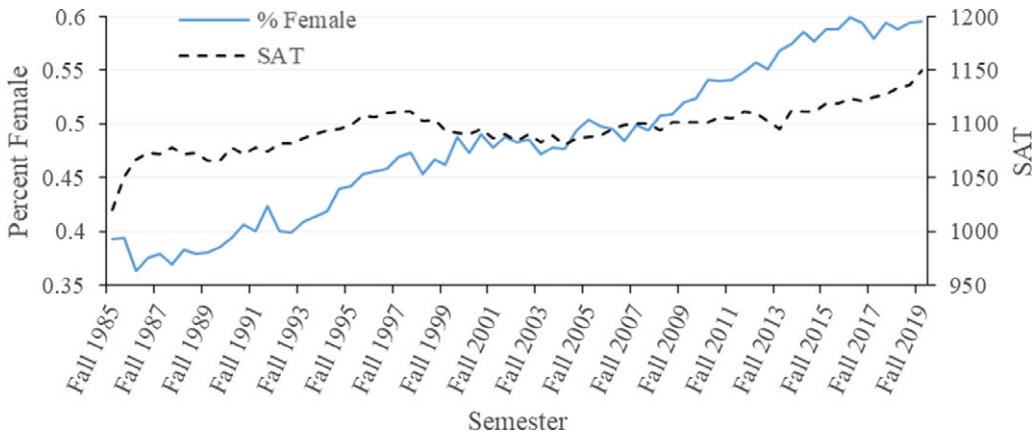


Figure 5. Average SAT score for all students in COALS and percentage of female students by year from fall 1985 to fall 2019.

2020). A high GPA may suggest a job candidate has high aptitude and motivation, but on the other hand, it could just be the result of inflated grading. It is clear comparing student GPAs across the time in the presence of grade inflation is inappropriate and may result in poor hires by companies. Whether grade increases are grade inflation or caused by improvements in teaching and students' preparedness is an empirical issue. As such, using data from Texas A&M College of Agriculture and Life Sciences (COALS) three interrelated objectives are to examine: 1) has grade inflation occurred, 2) has the distribution of grades as measured by the standard deviation changed,; and 3) how various factors influence class mean GPA and grade distributions. Even though the study is conducted based on data from a specific institution, the results may be generalized to other land grant universities, because the results support findings of earlier studies that did not explicitly focus on colleges of agriculture.

Utilizing data from 14 departments in COALS for the years 1985–2019, this study found evidence of systematic increases in mean class GPAs and standard deviation of grades after controlling for the influence of institution, instructors, and student factors. The results are suggestive of grade inflation, but with the caveat not all factors potentially influencing grades could be included. All three groups of factors are important in explaining class GPAs and standard deviations. Overall, the findings and results are quite robust to changes in model assumptions. If the model is simplified to one level (department) or to no levels, majority of conclusions about the inferences of the various characteristics on mean class GPAs and standard deviation hold, supporting the validity of model.

Findings have implications on recruitment policies and requirements. With respect to instructors, a positive signal to administration is seen in that there are no differences in GPAs by instructor gender. The practice of increasingly hiring graduate students and other instructors is associated with increasing grades, but is it inflation or do these instructors provide a better learning environment? Other instructors, for example, tend to have less of a research role, which leaves more time to allocate to class preparation and lecturing. If inflation, the trend of hiring non-tenure track instructors may be particularly worrisome as universities move towards increasing use of non-permanent faculty. Addressing this increase in grades, however, is not easy given an environment of decreasing budgets and increasing accountability as measured by evaluations of instructors' performance. Studies have suggested grades and evaluations are positively correlated (Marsh and Roche, 2000). If this is the case, then other measures of instructor effectiveness may have to be used. Hiring instructors with AAU backgrounds may result in higher GPAs but the reason is not clear. More research is necessary here; is it for example better preparedness by instructors or more focus on research that results in different class GPAs. Assuming GPAs are

good measures of learning, administrators may choose to schedule their core, degree-specific classes later during the day to maximize student learning. Further, smaller class sizes appear to be beneficial for learning. Universities should stride for smaller classes instead of larger classes. Budgetary considerations, however, are driving both class size and use of graduate students and other instructors. Some larger classes, however, may serve to weed-out students from the major, instructors in these classes may give lower grades (Weston et al., 2019). This observation complicates the issue.

From an instructor standpoint, implications are few. Given the trend of using student evaluations and assuming higher class GPAs are associated with better evaluations, instructors may want to ask to teach smaller class sections of specialized courses. Another consideration is class duration. Longer classes yield higher GPAs, but this is most likely because longer and less frequent classes tend to be more specialized.

With the goal of potentially higher GPA, when selecting between different sections of a class, students could check the instructor background and should try to register for the sections that have fewer students. In addition, given a chance, they may consider taking classes offered later during the day. Students may take classes taught by graduate students and other instructors in anticipation of higher grades, but should be aware of the potential opportunity cost of missing on insights when taught by more experienced professorial track faculty members.

### ***Suggestions for Future Research and Limitations***

The results suggest further research on separating the effects of grade improvement and grade inflation is necessary. Because variations are seen by department, examining departments separately or by general focus such as STEM or social is seen as a valuable contribution. Improved understanding of this phenomena and identifying source of grade increases is essential to determine if grade inflation is occurring. Research on grading patterns directly before and after instructor's promotion would provide insights on how the promotion process impacts grades. Whether instructors shift their time and effort towards publications and away from teaching or if they grade easy to improve evaluations needs exploring. Although different colleges of agriculture have somewhat different employment areas, they do compete against each other for hiring of their graduates. Further, how graduate schools and employers use/compare students' GPAs from different university should be explored, given the empirical nature of grading. Can or should grades from different universities be compared? Being an empirical issue, how do grades in colleges of agriculture vary by university. Such research would benefit both graduate colleges and employers. How has COVID-19 change the grading environment? If changes occurred, are they permanent changes?

As with all studies, limitations exist. Being a case study, all limitations of such studies apply here. The primary limitation is the inability to include all characteristics influencing class GPAs. As mentioned by a reviewer, instructor level may not account for all influence of an instructor, such as changing teaching styles. Unfortunately, such information is not available going back to 1985. An attempt was made to include infrastructure (building, classroom, lab modernization, electronic equipment available, etc.) in each classroom and how it evolved over time. Such details, however, are not available for the classrooms. Including interaction terms may be appropriate, but with the number of variables in this study the question becomes which interactions to include because including all interactions creates an unwieldy model. Interactions are not included to avoid a data mining exercise.

**Data availability statement.** Data are available from the authors.

**Author contributions.** Conceptualization, J.W.M., K.K.L., A.Y.; Methodology, J.W.M., A.Y.; Formal Analysis, J.W.M., A.Y.; Data Curation, J.W.M., K.K.L., A.Y.; Writing—Original Draft, J.W.M., K.K.L., A.Y. Writing—Review and Editing, J.W.M., K.K.L., A.Y.; Supervision, J.W.M.; Funding Acquisition, NA.



**Financial support.** This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

**Conflict of interest.** Anna Yeritsyan, James Mjelde, and Kerry Litzenberg declare none.

## References

- Adams, S. "Do Employers Care About College Grades." *Forbes*, 2015. Internet site: <https://www.forbes.com/sites/susanadams/2015/07/08/do-employers-care-about-college-grades/?sh=586e2d3a4b07> (Accessed November 9, 2021).
- Adelman, C. *Principal Indicators of Student Academic Histories in Postsecondary Education, 1972-2000*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2004. Internet site: <https://eric.ed.gov/?id=ED483154> (Accessed September 7, 2018).
- Adelman, C. "Undergraduate Grades: A More Complex Story than Inflation." *Grade Inflation: Academic Standards in Higher Education*. L.H. Hunt, ed. Albany: State University of New York Press, 2008.
- Bachan, R. "The Drivers of Degree Classification." Research Report, University of Brighton, 2018. Internet site: <https://research.brighton.ac.uk/en/publications/the-drivers-of-degree-classification> (Accessed September 21, 2021).
- Baird, A.F., J.S. Carter, and J.M. Roos "Seeking Evidence of Grade Inflation at For-profit Colleges and Universities." *Sociological Focus* 52,4(2019):343-58.
- Beenstock, M., and D. Feldman "Decomposing University Grades: A Longitudinal Study of Students and Their Instructors." *Studies in Higher Education* 43,1(2016):114-33.
- Bejar, L., and E. Blew "Grade Inflation and Validity of the Scholastic Aptitude Test." *American Educational Research Journal* 18,2(1981):143-56.
- Bergtold, J., E. Yeager, and T. Griffin "Academic, Demographic and Spatial Factors in the Classroom Affecting Student Performance in Principles of Agricultural Economics Courses." Paper presented at the 2016 AAEA Annual Meeting, Boston, MA, 2016. Internet site: <https://econopapers.repec.org/papers/agsaaea16/235930html> (Accessed January 21, 2022).
- Brighouse, H. "Grade Inflation and Grade Variation: What's All the Fuss About?" *Grade Inflation: Academic Standards in Higher Education*. L.H. Hunt, ed. Albany: State University of New York Press, 2008, pp. 73-91. ISBN: 978-0-7914-7497-6.
- Bond, T., and K. Mumford *The Causes and Consequences of Purdue Grade Inflation*, 2019. Internet site: <https://www.purdue.edu/senate/documents/meetings/Causes-and-Consequences-of-Purdue-Grade-Inflation.pdf> (Accessed November 7, 2021).
- Boretz, E. "Grade Inflation and the Myth of Student Consumerism." *College Teaching* 52,2(2004):42-6.
- Chen, K., Z. Hansen, and S. Lowe "Why Do We Inflate Grades? The Effect of Adjunct Faculty Employment on Instructor Grading Standards." *Journal of Human Resources* 56,3(2021):878-921.
- College Board. *SAT Total Group Profile Report*, 2015. Internet site: <https://secure-media.collegeboard.org/digitalServices/pdf/sat/total-group-2015.pdf> (Accessed October 16, 2020).
- Cuseo, J. "The Empirical Case Against Large Class Size: Adverse Effects on the Teaching, Learning, and Retention of First-Year Students." *Journal of Faculty Development* 21,1(2007):5-21.
- Danilowicz-Gösele, K., K. Lerche, J. Meya, and R. Schwager "Determinants of Students' Success at University." *Education Economics* 25,5(2017):513-32.
- DATAUSA. *Agriculture*, 2020. Internet site: <https://datausa.io/profile/cip/agriculture> (Accessed September 21, 2021).
- Diette, T., and M. Raghav "Class Size Matters: Heterogeneous Effects of Larger Classes on College Student Learning." *Eastern Economic Journal*; Basingstoke 41,2(2015):273-83.
- Diez-Roux, A. "Bringing Context Back into Epidemiology: Variables and Fallacies in Multilevel Analysis." *American Journal of Public Health* 88,2(1998):216-22.
- Douglas, L. *Multilevel Modeling (Quantitative Applications in the Social Sciences) Second Multilevel Edition*. Thousand Oaks, CA: Sage Publishing, 2004.
- Eiszler, C. "College Student's Evaluations of Teaching and Grade Inflation." *Research in Higher Education* 43,4(2002):483-501.
- Figlio, D., M. Schapiro, and K. Soter "Are Tenure Track Professors Better Teachers?" *The Review of Economics and Statistics* 97,4(2015):715-24.
- Flynn, J. "Massive IQ Gains in 14 Nations: What IQ Tests Really Measure." *Psychological Bulletin* 101,2(1987):171-91.
- Fossati, F., A. Wilson, and G. Bonoli "What Signals Do Employers Use When Hiring? Evidence from a Survey Experiment in the Apprenticeship Market." *European Sociological Review* 36,5(2020):60-779.
- Greene, W. *Econometric Analysis*. 7th ed. London: Pearson, 2012.
- Henebry, K. "The Impact of Class Schedule on Student Performance in a Financial Management Course." *Journal of Education for Business* 73,2(2010):114-20.
- Hernández-Julían, J., and A. Looney "Measuring Inflation in Grades: An Application of Price Indexing to Undergraduate Grades." *Economics of Education Review* 55,C(2016):220-32.
- Hoffmann, F., and P. Oreopoulos "A Professor Like Me: The Influence of Instructor Gender on College Achievement." *Journal of Human Resources* 44,2(2009):479-94.

- Howard, G., and S. Maxwell** "Correlation between Student Satisfaction and Grades: A Case of Mistaken Causation?" *Journal of Educational Psychology* 72,6(1980):810–20.
- Jaschik, S.** "Students Fail and Professor Loses Job." Inside Higher Ed, May 14 2008. Internet site: <https://www.insidehighered.com/news/2008/05/14/students-fail-and-professor-loses-job> (Accessed September 13, 2018).
- Jephcote, C., E. Medland, and S. Lygo-Baker** "Grade Inflation versus Grade Improvement: Are Our Students Getting More Intelligent?" *Assessment & Evaluation in Higher Education* 46,4(2021):547–71.
- Jewell, R., M. McPherson, and M. Tieslau** "Whose Fault is It? Assigning Blame for Grade Inflation in Higher Education." *Applied Economics* 45,9(2013):1185–200.
- Joyce, T., S. Crockett, D. Jaeger, O. Altindag, and S. O'Connell** "Does Classroom Time Matter? A Randomized Field Experiment of Hybrid and Traditional Lecture Formats in Economics." NBER Working Paper No. 20006, 2014. Internet site: <http://www.nber.org/papers/w20006.ack> (Accessed September 10, 2018).
- Juola, A.** "Grade Inflation in Higher Education: What Can or Should be Done?" Paper presented at the Annual Meeting of National Council on Measurement in Education, San Francisco, California, April 1979. Internet site: <https://eric.ed.gov/?id=ED129917> (Accessed September 7, 2018).
- Kapitanoff, S., and C. Pandey** "Stereotype Threat, Anxiety, Instructor Gender, and Underperformance in Women." *Learning in Higher Education* 18,3(2017):213–29.
- Kamber, R., and M. Biggs** "Grade Inflation: Metaphor and Reality." *Journal of Education* 184,1(2004):31–37.
- Kezim, B., S. Pariseau, and F. Quinn** "Is Grade Inflation Related to Faculty Status?" *Journal of Education for Business* 80,6(2005):358–64.
- Kokkelenberg, E., M. Dillon, and S. Christy** "The Effects of Class Size on Student Grades at a Public University." *Economics of Education Review* 27,2(2008):221–3.
- Kobrin, J., B. Patterson, E. Shaw, K. Mattern, and S. Barbuti** "Validity of the SAT for Predicting First-Year College Grade Point Average." Research Report No. 2008-5 College Board, 2008. Internet site: <https://eric.ed.gov/?id=ED563202> (Accessed September 11, 2018).
- Kohn, A.** "The Dangerous Myth of Grade Inflation." *The Chronicle of Higher Education* 49,11(2002):B7–9.
- Kolevzon, M.** "Grade Inflation in Higher Education: A Comparative Study." *Research in Higher Education* 15,3(1981):195–212.
- Kostal, J., N. Kuncel, and P. Sackett** "Grade Inflation Marches on: Grade Increases from the 1990s to 2000s." *Educational Measurement: Issues and Practice* 35,1(2016):11–20.
- Kuh, G., and S. Hu** "Unraveling the Complexity of the Increase in College Grades from the Mid-1980s to the Mid-1990s." *Educational Evaluation and Policy Analysis* 21,3(1999):297–320.
- Kuh, G., J. Kinzie, J. Buckley, B. Bridges, and J. Hayek** "What Matters to Student Success: A Review of the Literature." Commissioned Report for the National Symposium on Postsecondary Student Success, 2006. Internet site: [https://www.ue.ucsc.edu/sites/default/files/WhatMattersStudentSuccess\(Kuh,July2006\).pdf](https://www.ue.ucsc.edu/sites/default/files/WhatMattersStudentSuccess(Kuh,July2006).pdf) (Accessed September 11, 2018).
- Marini, J., E. Shaw, L. Young, and M. Ewing** "Getting to Know your Criteria: Examining College Course Grades and GPAs over Time." Research Report College Board, 2018. Internet site: <https://files.eric.ed.gov/fulltext/ED582569.pdf> (Accessed September 21, 2021).
- Marsh, H., and L. Roche** "Effects of Grading Leniency and Low Workload on Students' Evaluations of Teaching: Popular Myth, Bias, Validity, or Innocent Bystanders?" *Journal of Educational Psychology* 92,1(2000):202–38.
- McAllister, C., X. Jiang, and F. Aghazadeh** "Analysis of Engineering Discipline Grade Trends." *Assessment and Evaluation in Higher Education* 33,2(2008):167–78.
- McElroy, T., and L. Mosteller** "The Influence of Circadian Type, Time of Day and Class Difficulty on Students' Grades." *Electronic Journal of Research in Educational Psychology* 4,3(2006):611–22.
- Moore, M., and R. Trahan** "Tenure Status and Grading Practices." *Pacific Sociological Association* 41,4(1998):775–781.
- Mostrom, A., and P. Blumberg** "Does Learning-Centered Teaching Promote Grade Improvement?" *Innovative Higher Education* 37,5(2012):397–405.
- Nye, B., L. Hedges, and S. Konstantopoulos** "The Effects of Small Classes on Academic Achievement: The Results of the Tennessee Class Size Experiment." *American Educational Research Journal* 37,1(2000):123–51.
- O'Connor, L.** "The Decline and Fall of Grade Deflation." The Daily Princetonian, January 2020. Internet site: <https://www.dailyprincetonian.com/article/2020/01/the-decline-and-fall-of-grade-deflation> (Assessed September 30, 2020).
- Park, S., and E. Lake** "Multilevel Modeling of a Clustered Continuous Outcome: Nurses' Work Hours and Burnout." *Nursing Research* 54,6(2005):406–13.
- Pattison, E., E. Grodsky, and C. Muller** "Is the Sky Falling? Grade Inflation and the Signaling Power of Grades." *Educational Researcher* 42,5(2013):259–65.
- Peace, J., II** "How Employers Can Stanch the Hemorrhaging of Collegiate GPA Credibility." *Business Horizons* 60,1(2017):35–43.
- Rojstaczer, S., and C. Healy** Grading in American Colleges and Universities, 2010. Internet site: <http://gradeinflation.com/tcr2010grading.pdf> (Accessed September 7, 2018).

- Rojstaczer, S., and C. Healy "Where A Is Ordinary: The Evolution of American College and University Grading, 1940 – 2009." *Teachers College Record* **114**,7(2012):1–23.
- Ronco, S., and J. Cahill "Does It Matter Who's in the Classroom? Effect of Instructor Type on Student Retention, Achievement and Satisfaction." Paper presented at the 44th Annual Forum of the Association for Institutional Research, Boston, Massachusetts, 2004. Internet site: <https://eric.ed.gov/?id=ED491000> (Accessed September 10, 2018).
- Rothstein, J. "College Performance Predictions and the SAT." *Journal of Econometrics* **121**,1-2(2004):297–317.
- Schutz, K., B. Drake, J. Lessner, and G. Hughes "A Comparison of Community College Full-Time and Adjunct Faculties' Perceptions of Factors Associated with Grade Inflation." *The Journal of Continuing Higher Education* **63**,3(2015):180–92.
- Shinn, M., and B. Rapkin "Cross-level Research without Cross-ups in Community Psychology." *Handbook of Community Psychology*. Shinn, M., and B. Rapkin, eds. Boston, MA: Kluwer Academic Publishers, 2000.
- Skinner, N. "University Grades and Time of Day of Instruction." *Bulletin of the Psychonomic Society* **23**,1(1985):67.
- Solanki, S., and D. Xu "Looking Beyond Academic Performance: The Influence of Instructor Gender on Student Motivation in STEM Fields." *American Educational Research Journal* **55**,4(2018):801–35.
- Sonner, B. "A is for 'Adjunct': Examining Grade Inflation in Higher Education." *Journal of Education for Business* **76**,1(2000):5–8.
- StataCorp. Stata Multilevel Mixed-Effects Reference Manual release 15. ISBN-10: 1-59718-244-3, 2017. Internet site: <https://www.stata.com/manuals15/me.pdf>. (Assessed November 8, 2020).
- Stratton, R., S. Myers, and R. King "Faculty Behavior, Grades, and Student Evaluations." *Journal of Economic Education* **25**,1(1994):5–15.
- Suslow, S. A Report on an Interinstitutional Survey of Undergraduate Scholastic Grading 1960s to 1970s, 1976. Internet site: <https://eric.ed.gov/?id=ED129187> (Accessed September 7, 2018).
- Trahan, L., K. Stuebing, M. Hicoek, and J. Fletcher "The Flynn Effect: A Meta-Analysis." *Psychology Bulletin* **14**,5(2014):1332–60.
- Voyer, D., and S. Voyer "Gender Differences in Scholastic Achievement: A Meta-Analysis." *Psychology Bulletin* **140**,4(2014):1174–204.
- Wahlstrom, K., V. Hendrix, and J. Frederickson School Start Time Study. Technical Report, Volume II: Analysis of Student Survey Data. Minneapolis, MN: University of Minnesota, Center for Applied Research and Educational Improvement, 1998. Internet site: <https://conservancy.umn.edu/bitstream/handle/11299/4249/SST-1998.Vol.II-Student%20Survey%20results%20%281%29.pdf?sequence=7&isAllowed=y> (Assessed October 17, 2020).
- Wang, M., and J. Degol "School Climate: A Review of the Construct, Measurement, and Impact on Student Outcomes." *Educational Psychology Review* **28**,2(2016):315–52.
- Wajtrakul, B. "Factors Affecting Students' Intentions to Study at Universities Adopting the 'Student-as-Customer' Concept." *International Journal of Educational Management* **28**,6(2014):676–93.
- Weston, T., E. Seymour, A. Koch, and B. Drake "Weed-Out Classes and Their Consequences." *Talking about Leaving Revisited*. E. Seymour and A.-B. Hunter, eds. Cham: Springer, 2019, pp. 197–243.
- Westrick, P., H. Le, S. Robbins, J. Radunzel, and F. Schmidt "College Performance and Retention: A Meta-Analysis of the Predictive Validities of ACT Scores, High School Grades, and SES." *Educational Assessment* **20**,1(2015):23–45.
- Wheaton, A., D. Chapman, and J. Croft "School Start Times, Sleep, Behavioral, Health, and Academic Outcomes: A Review of the Literature." *Journal of School Health* **86**,5(2016):363–81.
- Wu, L. Mixed Effect Models for Complex Data. Chapman and Hall/CRC, London, 2019. ISBN: 9780367384913 (44-45).