\[ \text{GPA}_{i,j,T} = (\alpha + \varphi_t) + \rho \text{Female} + \beta \text{PS} + \lambda_1 \text{RC} + \lambda_2 \text{RC}^2 + \psi \text{GAI}_t + \sum_{j=1}^{12} \omega_j \text{AP}_j \\
+ \sum_{t=2}^{16} \left( \gamma_t S_t + \phi_t (S_t \cdot \text{GAI}_t) \right) + \sum_{T=2}^{22} \theta_T C_T + \sum_{T=2}^{22} \sum_{j=1}^{12} \delta_{T,j} (C_T \cdot \text{AP}_j) + e_{i,j,T} \]

\[ \frac{\partial (\text{GPA}_{i,j,T})}{\partial (C_T)} = \hat{\theta}_T + \sum_{j=1}^{12} \hat{\delta}_{j,T} (\text{AP}_j) \quad \text{and} \quad \frac{\partial (\text{GPA}_{i,j,T})}{\partial (\text{GAI}_t)} = \frac{1}{\sigma_{\text{GAI}}} \left\{ \psi + \sum_{t=2}^{16} \phi_t S_t \right\} \]

\[ \frac{\partial (\text{GPA}_{i,j,T})}{\partial (S_t)} = \gamma_t + \phi_t \left( \frac{\text{GAI}_t - \overline{\text{GAI}}}{\sigma_{\text{GAI}}} \right) = \gamma_t + \phi_t \cdot Z_{\text{GAI}} \]

\[ \text{GAI}_t^* > \left( \frac{\hat{\gamma}_t \cdot \sigma_{\text{GAI}} + \overline{\text{GAI}}}{\phi_t} \right) \Rightarrow Z_{\text{GAI}}^* > \frac{\hat{\gamma}_t}{\phi_t}, \quad \forall \ t = 2, \ldots, 16. \]

**RESEARCH ARTICLE**

Measuring grade inflation and grade divergence accounting for student quality

Horacio Matos-Díaz

*Cogent Economics & Finance* (2014), 2: 915756
Measuring grade inflation and grade divergence accounting for student quality

Horacio Matos-Díaz

**Abstract:** This study uses a rich longitudinal data-set of 13,202 full-time students belonging to 11 cohorts over 22 consecutive semesters (Fall 1995 to Spring 2006) to model the determinants of the grade inflation rates prevailing at the University of Puerto Rico at Bayamón. The following new interesting findings are reported: (1) Estimated rates vary significantly among and within the academic programs, implying grade divergence, depending on the time reference used: cohort time dummies or semesters since admission to the institution. (2) The rates are significantly related to the proportions of female students, students who switch from their original academic programs, and students from private schools. (3) Results suggest that, under determinate circumstances, average and low-quality students consider higher grades as normal goods; conversely, high-quality students consider them as inferior goods.

**Keywords:** grade inflation index, GPA, student quality, random and fixed effects models

**JEL classifications:** C23, I20, I21

Received: 06 November 2013
Accepted: 11 April 2014
Published: 03 June 2014

© 2014 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license.
1. Introduction

The phenomenon of grade inflation—the tendency of student grade point average (GPA) to increase without a concomitant increase in achievement—has been observed and widely studied since the 1970s (Hu, 2005; Johnes, 2004; Johnson, 2003). Rosovsky and Hartley (2002) discuss several explanations offered for grade inflation, including student evaluation of teaching (SET) ratings. Students’ grades serve several purposes (Sabot & Wakeman-Linn, 1991). They send signals to the labor market and to graduate schools about the quality of university graduates and provide information to the students about their achievement levels, progress, and academic success. To the extent that it deviates from these traditional purposes, the university grading system lessens the value society attributes to higher education.

According to the leniency hypothesis (Gump, 2007), faculty members can buy higher SET ratings, recruit more students, or even become more popular, by relaxing their academic standards through leniency grading. McKenzie and Staaf (1974), McKenzie (1975), and Lichty, Vose, and Peterson (1978) have developed economic models that rationalize this conjecture and its relationship to grade inflation. Recent models in this tradition include Kanagaretnam, Mathieu, and Thevaranjan (2003) and Love and Kotchen (2010). Following the general guidelines of those models, a series of papers have studied the relationship between SET ratings and students’ known or expected grades (EG) using different methods, including OLS (Dilts, 1983), 2SLS (Krautmann & Sander, 1999; Nelson & Lynch, 1984; Seiver, 1983), 3SLS (Zangenehzadeh, 1988), as well as fixed-effects models (Isely & Singh, 2005; McPherson, 2006). The majority of these studies do not properly model the phenomenon of grade inflation. Normally, authors present statistical evidence in favor of (or against) the interdependence between the SET and EG that leads them to infer or conclude that there is or there is not grade inflation. It should be emphasized that a relationship between EG and SET would provide a plausible explanation for high grades, but not for grade inflation which refers to an upward time trend in grades.

There are several exceptions to this pattern. For example, Jewel, McPherson, and Tieslau (2013), using data from the University of North Texas, report significant grade inflation rates over two decades. They attribute these variations to characteristics of academic departments, university-level factors or instructor-specific characteristics. Grove and Wasserman (2004) reported significant grade inflation rates over the life-cycle pattern of collegiate GPA of five consecutive cohorts at a large USA private university. Sabot and Wakeman-Linn (1991) analyze the problems of grade inflation as well as of grade divergence. They present comparative evidence on the average academic grades given from 1962/1963 to 1985/1986 by the different departments of seven USA universities and find that grades tend to increase, but with marked differences among the departments, which implies grade divergence. The issue of grade divergence is analyzed in detail by Freeman (1999) and Achen and Courant (2009).

Matos-Díaz (2012) analyzes the issues of reductions in academic standards and lenient grading as well as their relationship with the SET process at the University of Puerto Rico at Bayamón (UPR-Bayamón). Results suggest that faculty members might be able to increase enrollment in their courses, get better SET ratings, and/or improve their teaching schedules by (1) adjusting their academic standards in order to increase students’ EG and (2) promoting the academic conditions that would transform pessimistic relative EG into optimistic ones. The author suggests the possibility that such strategies probably lead to a decrease in academic standards and grade inflation at UPR-Bayamón.

This study uses the student as the proper unit of analysis in order to model the phenomena of grade inflation and grade divergence over two different time references: cohort time dummies and semesters since admission to UPR-Bayamón (hereafter “time reference 1 and 2”, respectively). To estimate grade inflation accurately, it is necessary to account for student quality. Because student quality is proxied by the general admission index (GAI), which is a time-invariant variable, the grade inflation index (GII) cannot be defined at the student level. Thus, its estimation will require
transforming panel data into highly aggregated historical data. The aggregation of data over each time reference generates two sets (1 and 2) of 22 and 16 sub-samples, respectively. The partition of each set by academic programs produces two different restricted panels (panel 1 and 2) of 286 and 179 sub-samples, respectively.

This paper contributes to the literature in several ways. First, it uses a rich panel data of 13,202 full-time students (12 or more credits per semester) belonging to 11 consecutive cohorts that entered UPR-Bayamón from 1995–1996 to 2005–2006. Students were classified into 13 different academic programs and were tracked every semester until each one left the institution. Second, this is the first paper using two different time references in order to accurately estimate grade inflation and grade divergence. Third, the GII, accounting for student quality, is defined and its behavior through academic programs over both time references is analyzed using two-way fixed-effects models. Both models account for academic programs' unobserved heterogeneity while controlling for other student characteristics. Grade divergence and grade inflation are captured in the estimated cross-sections and period fixed-effects coefficients, respectively. Thus, this is the first paper in the field able to directly measure the determinants of the GII at such econometric accuracy and specificity levels. Finally, results from the unrestricted panel are consistent with and tend to confirm the historical growth paths of the GPA, GAI, and GII variables over both time references.

The remainder of the paper is organized as follows: Section 2 deals with the data and the models to be estimated, while Section 3 presents the results. A summary providing some conclusions closes the paper.

2. Data-set and Models

2.1. UPR-Bayamón and Its Admission Criterion

The UPR-Bayamón is an autonomous unit of the University of Puerto Rico system (UPR). Accredited by the Middle States Association of Colleges and Secondary Schools, the institution offers associate and bachelor’s degrees, as well as articulated transfer programs to the Río Piedras, Mayagüez and Medical Sciences campuses. In the fall of 2006, total enrollment at UPR-Bayamón was 4,565, including 3,737 full-time students. The admission policy at the UPR is based on the GAI of each applicant, which is the weighted mean of the high school GPA (50%) and the scores in the verbal aptitude (25%) and mathematical aptitude (25%) sections of the College Entrance Examination Board test. The GAI plays a critical role, not only for admission to the different campuses of the UPR, but also for admittance to particular programs. Thus, it is used as the best available proxy of student quality.

2.2. The Unrestricted Panel

This study uses a rich longitudinal data-set that describes in detail every one of the 13,202 students belonging to the 11 cohorts that entered UPR-Bayamón from 1995–1996 to 2005–2006. The following variables are available for each student: GAI; school of origin code; gender; age; academic program to which the student is admitted; registered semester credit-hours (RC); and GPA. A set of 22 time dummies \( (C_T, 1 \leq T \leq 22) \), which are included to capture time trend effects, identify the entry cohort to which each student belongs, as well as the corresponding semester. Hence, odd numbers designate fall semesters, while even ones denote spring semesters. For example, \( C_8 \) designates all spring semesters for students in cohort 4, while \( C_7 \) designates all fall semesters.

To track each student throughout time reference 2, a second set of 16 time dummies was defined, indicating the number of semesters since admission to UPR-Bayamón. For a few of the students in the sample this number is greater than 16, so they were grouped into the following single category: 16 or more semesters. The variable is defined as \( S_t \), where \( 1 \leq t \leq 16 \). Dummies are used to control for academic programs (13), high school of origin and student’s gender. Table 1 describes all the variables used in the study.
2.3. The Restricted Panels

From each sub-sample of sets 1 and 2, the arithmetic means of the GPA, GAI, and GII variables are computed. Means are indexed to identify their time reference and, in the case of the restricted panels 1 and 2, their academic program. To document grade inflation, it is necessary to present evidence showing a self-sustained increase in GPA over time without a concomitant increase in student achievement or academic ability. However, neither achievement nor academic ability is observable. To overcome this data limitation, the GAI is used as the best available proxy of student’s academic ability, achievement or quality.

Given that GPA and GAI variables are expressed in different metrics, it was necessary to normalize both of them. This was done by dividing each value in the series by its first one, generating their respective time reference growth rates. It should be expected that GPA vary directly and strongly with GAI. In the limit case, where they vary directly proportional, the series would mimic each other and their ratio would always be equal to one. If GPA tends to grow over time while GAI remains constant, then it could be conjectured that academic standards have diminished, providing the bases for grade inflation. Thus, current students receive the same grades as past students, but with less

| Table 1. Descriptive Statistics of the Unrestricted Sample |
|---------------------------------------------------------|
| Dummy variables                                        |
| Var. | Mean | Var. | Mean | Var. | Mean | Var. | Mean |
|-------|------|------|------|------|------|------|------|
| Accounting | .0696 (.2544) | S_1 | .1854 (.3886) | S_16 | .0025 (.0496) | C_15 | .0471 (.2119) |
| Business | .1585 (.3652) | S_2 | .1576 (.3643) | C_1 | .0603 (.238) | C_16 | .0371 (.1889) |
| Computer Sciences | .0869 (.2817) | S_3 | .147 (.3541) | C_2 | .0502 (.2184) | C_17 | .0391 (.1938) |
| Education | .1046 (.306) | S_4 | .1217 (.3269) | C_3 | .0652 (.2469) | C_18 | .03 (.1706) |
| Electronics | .0821 (.2745) | S_5 | .0973 (.2963) | C_4 | .0549 (.2278) | C_19 | .0339 (.18) |
| Engineering Technologies | .0351 (.1841) | S_6 | .078 (.2682) | C_5 | .0576 (.2329) | C_20 | .0231 (.1502) |
| Engineering Transfers | .0446 (.2065) | S_7 | .0661 (.2485) | C_6 | .0489 (.2157) | C_21 | .0266 (.1609) |
| Humanities | .0312 (.2065) | S_8 | .0547 (.2273) | C_7 | .0661 (.2485) | C_22 | .0129 (.1129) |
| Materials Management | .038 (.1911) | S_9 | .0402 (.1963) | C_8 | .0547 (.2273) | |
| Natural Sciences | .1422 (.3492) | S_10 | .0234 (.1512) | C_9 | .0402 (.1963) | |
| Office Systems | .0846 (.2783) | S_11 | .0121 (.1093) | C_10 | .0488 (.2155) | |
| Physical Education | .0662 (.2486) | S_12 | .007 (.0835) | C_11 | .055 (.2281) | |
| Social Sciences | .0565 (.2309) | S_13 | .0037 (.0037) | C_12 | .0461 (.2098) | |
| Female | .5793 (.4937) | S_14 | .0021 (.0021) | C_13 | .0517 (.2213) | |
| PS | .4546 (.4979) | S_15 | .0014 (.0374) | C_14 | .0422 (.201) | |

| Continuous variables |
|----------------------|
| Var. | Mean | SD | Max | Min |
|-------|------|----|-----|-----|
| GPA | 2.74 | .5956 | 4.00 | .00 |
| GAI | 285.66 | 35.5 | 385 | 137 |
| RC | 15.14 | 2.09 | 21 | 12 |
| RC-square | 233.68 | 64.15 | 641 | 144 |

Notes: For all dummies, standard deviations (SD) are reported in parentheses; Max = 1 and Min = 0; Var. = variables.
academic requirements, a lesser amount of content, and significantly less learning. The problem of grade inflation would be even greater if GPA tends to increase while GAI tends to decrease over time, because in such a case grading did not decrease in order to account for less qualified students. These considerations give rise to the following operational definition:

\[
\text{GII} = \frac{\text{NGPA}}{\text{NGAI}}
\]  

The “N” in front of the terms stands for normalized. If GII > 1, then evidence will point to grade inflation which could be consistent with diminishing academic standards over time. Grade deflation will be documented whenever GII < 1; that is, when GPA tends to decrease even when GAI is constant or increasing over time. Finally, GII = 1 would imply neither grade inflation nor grade deflation.

2.4. Models to be Estimated

The following two-way fixed-effects models will be estimated using data from the restricted panels 1 and 2, respectively.

\[
\text{GII}_{j,T} = (w + a_j + d_T) + b_1 \Pi_F + b_2 \Pi_{PS} + b_3 \Pi_{\Delta AP} + e_{j,T}
\]  

\[
\text{GII}_{j,t} = (c + g_j + h_t) + z_1 \Pi_F + z_2 \Pi_{PS} + z_3 \Pi_{\Delta AP} + v_{j,t}
\]

Subscript \( j \) \((j = 1, \ldots, 13)\) indexes the academic programs, while subscripts \( T \) and \( t \) index the cohort time dummies and semesters since admission to UPR-Bayamón, respectively \((T = 1, \ldots, 22 \text{ and } t = 1, \ldots, 16)\). Finally, \( w \) and \( c \) are the overall intercepts, while \( e_{j,T} \) and \( v_{j,t} \) are the composite error terms.

Both models share the following covariates: proportion of female students \( (\Pi_F) \), proportion of students from private schools \( (\Pi_{PS}) \), and proportion of students who switched from their original academic programs \( (\Pi_{\Delta AP}) \). Some data facts related to these three proportions should be emphasized. For example, in those academic programs with low mathematics requirements such as Education, Physical Education, Humanities, Social Sciences, and Office Systems, the proportions of female students and private school students are 80 and 39%, respectively. On the other hand, in those academic programs with higher mathematics requirements and possibly greater inherent difficulty of content, such as Engineering Transfers, Electronics, and Computer Sciences, the respective proportions are 18 and 43%. Finally, for the whole sample, the proportion of students who switched from their original academic program is only 4.4%. However, 10% of the students enrolled in Social Sciences came from other programs by means of transfers.

Given these data, it is reasonable to assume that if students self-sorting by academic programs in such a way that \( \Pi_F \) is greater in programs of low mathematics and difficulty contents, then it should be expected that \( b_1 \) and \( z_1 \) be positive and significant. On the other hand, if \( \Pi_{PS} \) is greater in programs of high mathematics and difficulty contents, then both estimated coefficients \( b_2 \) and \( z_2 \) should be negative and significant. Lastly, if students transfer from more demanding programs to less demanding ones looking for better grades with less academic effort, the estimated coefficients \( b_3 \) and \( z_3 \) should be positive and significant. Thus, it should be expected that the GII would respond significantly to changes in those three proportions across both time references.

In order for grade divergence to hold, it will be necessary to observe significant differences in the GII across academic programs. If so, grade divergence can be captured in the estimated
cross-section fixed-effects coefficients ($\hat{d}_f$ and $\hat{h}_t$). On the other hand, grade inflation across both time references can be captured in differences in the estimated period fixed-effects coefficients ($\hat{d}_f$ and $\hat{h}_t$).

Model 4 is specified using GPA as the dependent variable and GAI as a covariate because it is a time-invariant variable at student level. The model controls for student’s gender (female = 1), school of origin (private = 1), RC and academic programs (Social Sciences = reference group). Unobservable student heterogeneity ($\varphi_i$) is modeled as random-effects.²

$$GPA_{i,t,j} = (\alpha + \varphi_j) + \beta\text{Female} + \beta PS + \lambda_1 RC + \lambda_2 RC^2 + \psi \text{GAI}, \quad (4)$$

$$+ \sum_{j=1}^{12} \omega_j AP_j + \sum_{t=2}^{16} \{ \gamma_1 S_t, + \phi_1 (S_t \cdot \text{GAI}_j) \} + \sum_{t=2}^{22} \theta_{1,t} C_t + \sum_{t=2}^{22} \sum_{j=1}^{12} \delta_{j,t} (C_t \cdot AP_j) + \epsilon_{i,t,j}$$

3. Results and Discussion

3.1. Grade Inflation and Grade Divergence: Historical Data Reports

Average time growth rates of GPA, GAI, and GII normalized variables, estimated from sub-sample sets 1 and 2, are reported in Figures 1 and 2. Compared to the base year, GAI exhibits an increasing path over time reference 1 (except for cohort 2),¹ reaching its maximum growth rate (5.95%) at $C_{20}$. However, GAI behaves quite differently throughout time reference 2. It increases until $S_8$, driven by the attrition of low-quality students during their first semesters at the institution. Between $S_8$ and $S_{14}$, GAI decreases because many high-quality students transfer to other UPR campuses or obtain their associate degrees, although it remains above the unitary threshold. Finally, GAI decreases after $S_{14}$ and goes below the unitary threshold because the best students complete their bachelor’s degree and leave the institution. Only the lower quality students remain in campus. This behavior explains the “life-cycle” pattern of GPA over time reference 2, which in turn will explain the inverted U-shape exhibited by the variable over time reference 1.

Along time reference 1, GPA is always over the unitary threshold (except at $C_3$) until $C_{14}$, reaching its maximum growth rate (7.3%) at $C_{14}$. Then, it decreases and is underneath the unitary threshold at three times ($C_{15}$, $C_{21}$, and $C_{22}$). That is, GPA increases, reaches its absolute maximum, and then decreases. To explain such an inverted U-shape, it is necessary first to analyze the behavior of GPA over time reference 2. Over it, GPA exhibits a similar inverted U-shape. It increases, reaching its maximum growth rate (13.8%) at $S_{14}$, then it decreases and reaches its minimum growth rate (5.3%) at $S_{14}$. However, it is always above the unitary threshold. To the right of $S_{14}$, GPA is decreasing because only the lower quality students of each cohort remain in the sample.

Students of cohorts 8, 9, 10, and 11 remain in the sample at the most 4, 3, 2, and 1 years, respectively. Thus, the maximum number of semesters since admission ($S_j$) of cohort 8 students is 8 ($S_8 \leq S_j$), 6 for cohort 9 students ($S_8 \leq S_j$), 4 for cohort 10 students ($S_9 \leq S_j$), and only 2 for cohort 11 students ($S_{11} \leq S_j$). Therefore, students of cohorts 8–11 should be to the left of $S_8$, $S_9$, $S_10$, and $S_{11}$, respectively, and exhibit lower GPAs along time reference 2, which is equivalent to reversing movements from $S_j$ to the origin. Hence, the GPA decreasing pattern to the right of $S_j$ over time reference 2 explains its decreasing pattern to the right of $C_{14}$ through time reference 1. That is to say, movements to the right of $C_{14}$ over time reference 1 correspond to movements to the origin from the left of $S_j$ over time reference 2. In such a direction, GPA decreases over time. Consequently, to the right of $C_{14}$, GPA tends to diminish because of insufficient time reference 2 sub-sample observations. Thus, the life-cycle pattern of GPA over time reference 2 explains the inverted U-shape of GPA over time reference 1, particularly, its decreasing portion to the right of $C_{14}$. 
Figures 1 and 2 describe the behavior of the GII along both time references. The index is always over the unitary threshold until \( C_{14} \) and reaches its maximum rate (5.32%) at \( C_{4} \). Based on the points discussed in the prior paragraph, it is misleading to analyze the behavior of the index to the right of \( C_{14} \). That is, if cohorts 8–11 would have been in the sample during longer periods, the behavior of GPA would be different from the exhibited pattern. Therefore, evidence shows that the 14 sub-samples drawn from the first seven cohorts admitted to UPR-Bayamón (1995–1996 to 2001–2002) exhibit grade inflation. The eight sub-samples drawn from the last four cohorts (2002–2003 to 2005–2006), do not have enough time reference 2 observations to accurately estimate the index.

The behavior of the GII is quite different along time reference 2. It increases and reaches its maximum rate (13.3%) at \( S_{8} \). Then, it decreases and reaches its minimum rate (8.7%) twice: at \( S_{12} \) and \( S_{13} \). Therefore, irrespective of the time reference used, the evidence points to the same conclusion: the existence of grade inflation. However, the magnitude of the phenomenon will depend on the time reference used. It is not argued here that one reference is better than the other; both complement each other. Hence, the growth pattern of GPA along time reference 1 cannot be explained without knowing the corresponding pattern over time reference 2. This last point has seemingly been ignored in the mainstream literature.

Using data from restricted panels 1 and 2, the GII was computed by academic program and plotted graphically over time reference 1 and 2. There are marked differences in the index’s value among academic programs, implying grade divergence, as well as along both time references, implying grade inflation. The pattern displayed is very similar to the one depicted in Figure 1.5

3.2. The Two-way Fixed-effects Models
The estimated regression coefficients of Equation 2 are reported in Table 2. The three coefficients \((\hat{b}_1, \hat{b}_2, \text{ and } \hat{b}_3)\) behave as conjectured. Each one exhibits the expected sign and is significant. For example, other things being equal, increments of 10 percentage points (pp) in \( \Pi_F \) will increase the GII by 1.17 pp. However, if such increase were in \( \Pi_P \), the GII will decrease by .71 pp. Likewise, if it
were in \( \Pi_{\Delta AP} \), then the gii will increase by 2.06 pp. Table 2 also reports the estimated cross-section and period fixed-effects. They were submitted to three different statistical tests to determine the significance of the cross-section, the period and all of the effects jointly. The three null hypotheses of statistical insignificance were rejected through the \( F \) and chi-square tests (\( p \)-value = .0000).

The distribution of signs and values of the estimated cross-section effects (\( \hat{a}_j \)) induces changes in the overall intercept of Equation 2, which in turn implies that the grade inflation rates will vary significantly among the academic programs. On the other hand, 13 out of the first 14 estimated period effects (\( \hat{d}_T \)) are positive, while seven out of the last eight are negative. Given that the intercept of the equation (\( \hat{w} \approx .96 \)) is 4 pp beneath the unitary threshold, such a result is in accordance with the
growth pattern exhibited by the GII as reported in Figure 1: it increases from the origin until $C_{14}$ and then decreases. These results show that the GII varies significantly across academic programs, implying grade divergence. In the same way, it varies over time reference 1, implying grade inflation.

The estimated regression coefficients, cross-section and period fixed-effects of Equation 3 are reported in Table 3. The coefficients $\hat{z}_1$, $\hat{z}_2$, and $\hat{z}_3$ exhibit the expected sign, but only $\hat{z}_1$ is significant. Other things being equal, increments of 10 pp in $\Pi_F$ will increase the GII by 1.2 pp. The null hypotheses about the effects’ insignificance were rejected through the $F$ and chi-square tests ($p$-value = .0000). The distribution of signs and values of the estimated cross-section effects ($\hat{g}_j$) implies that the rates vary across the academic programs, which is consistent with grade divergence. The overall intercept of Equation 3 ($\hat{c} \approx 1.04$) is 4 pp above the unitary threshold. So, the distribution of signs and values of the estimated period effects ($\hat{h}_t$) imparts an inverted U-shape to the GII over time reference 2, similar to the curvature depicted in Figure 2. Accordingly, the statistical results are consistent with and provide analytical support to the story told by the historical data reported in Figures 1 and 2. Hence, in order to have a good grasp of the magnitude and scope of the grade inflation and grade divergence phenomena, their analysis should have to consider not only the distribution of grades among academic programs, but also their path over both time references.

### Table 3. Determinants of the GII over Time Reference 2

| Variable                          | Coefficient  | Std. Error | t-Statistic | Prob.  |
|----------------------------------|--------------|------------|-------------|--------|
| Constant                         | 1.036490     | .028400    | 36.49591    | .0000  |
| Female proportion                | .120986      | .040928    | 2.956095    | .0036  |
| Private school proportion        | -.048557     | .032106    | -1.512395   | .1326  |
| Program changes proportion       | .015553      | .120995    | .128547     | .8979  |

Effects specification

Cross-section fixed (dummy variables)

Period fixed (dummy variables)

| Programs (cross-section) | Effects | Semesters (period) | Effects |
|--------------------------|---------|--------------------|---------|
| P1                       | -.065928| 1                  | -.080349|
| P2                       | -.005617| 2                  | -.029051|
| P3                       | .075258 | 3                  | -.027754|
| P4                       | -.027072| 4                  | -.003901|
| P5                       | .054562 | 5                  | -.003031|
| P6                       | .080556 | 6                  | .013712 |
| P7                       | -.029746| 7                  | .005020 |
| P8                       | -.059835| 8                  | .030095 |
| P9                       | .066833 | 9                  | .020489 |
| P10                      | .009494 | 10                 | .023374 |
| P11                      | -.073785| 11                 | -.007577|
| P12                      | .039965 | 12                 | .007047 |
| P13                      | -.064663| 13                 | -.008977|
|                          |         | 14                 | -.006116|
|                          |         | 15                 | .024200 |
|                          |         | 16                 | .034586 |

**Notes:** The standard errors of the fixed-effects coefficients are not reported because, for estimation purposes, EViews treats them as nuisance parameters. Periods included: 16; cross-sections included: 13; total panel (unbalanced) observations: 179; R-square = .74.
3.3. The Unrestricted Panel

The estimated regression coefficients of Equation 4 are reported in Table 4. The statistical fit of the model is excellent. The impact of the academic programs on GPA is measured by their estimated main and time dummies interaction effects coefficients: $\hat{\omega}_j$ and $\hat{\delta}_{jT}$, respectively. Their respective proportions of significant coefficients are 42 and 68%. Females and students from private schools have advantages over their counterparts. The GPA expected by a female is .22 points greater than the expected by a male; while the expected by students from private schools is .019 points greater than the expected by students from public schools. There is a significant and concave relationship ($\hat{\lambda}_1 > 0$ and $\hat{\lambda}_2 < 0$) between GPA and the Rc variable, which is standardized. GPA reaches its absolute maximum at Rc= 18 credit-hours.

In order to detect grade inflation and grade divergence, the attention is centered on the estimated main effects coefficients of the time dummies ($\hat{\theta}_T$) and their interactions effects with the academic programs ($\hat{\delta}_{jT}$), which are captured into the following 21 partial derivatives ($T=2, \ldots, 22$):

$$
\frac{\partial (\text{GPA}_{i,j,t})}{\partial (\text{C}_T)} = \hat{\theta}_T + \sum_{j=1}^{12} \hat{\delta}_{jT}(\text{AP}_j)
$$

(5)

Sixty seven percent of $\hat{\theta}_T$ coefficients are significant. The first 15, which belong to the sub-samples drawn from the first eight entering cohorts are positive, implying GPAs greater than the GPA of the reference group ($\text{C}_i =$ fall semesters of cohort 1). On the other hand, three of the last six, which belong to the last three entering cohorts, are negative, implying GPAs lesser than the GPA of the reference group. The pattern of signs of these main effects is in accordance with the historical data reported in Figure 1: grade inflation for almost all sub-samples drawn from the first seven cohorts. Sixty eight percent of the interaction effects ($\hat{\delta}_{jT}$) are significant. They allow for the possibility of capturing the instances in which GPA changes among and within the academic programs throughout time reference 1. Thus, the estimated interaction coefficients ($\hat{\delta}_{jT}$) allow to measure grade divergence across different academic programs along time reference 1.

Figure 2. Growth path of GPA, GAI and GII over time reference 2.
Let $\sigma_{GAI}$ be the standard deviation of GAI distribution, then the following two partial derivatives warrant some explanation:

$$
\frac{\partial (GPA_{i,t,T})}{\partial (\sigma_{GAI})} = \frac{1}{\sigma_{GAI}} \left\{ \psi + \sum_{t=2}^{16} \phi_i \cdot S_t \right\} 
$$

(6)

$$
\frac{\partial (GPA_{i,t,T})}{\partial (S_t)} = \gamma_t + \phi_t \left\{ \frac{GAI - \overline{GAI}}{\sigma_{GAI}} \right\} = \gamma_t + \phi_t \cdot Z_{GAI}
$$

(7)

Expression 6 describes the relationship between GPA and GAI (student quality) over time reference 2. The variables move in the same direction since the estimated main effect coefficient ($\psi$) is positive.

---

**Table 4. Predicting GPA at UPR-Bayamón: 1995–1996 to 2005–2006**

| IV          | Coefficient   | IV          | Coefficient   | IV          | Coefficient   |
|-------------|---------------|-------------|---------------|-------------|---------------|
| Constant    | 2.36** (.0346)| $S_2$       | .0554** (.0143)| $GAI \times S_5$ | $-0.296** (.004)$ |
| $AP_1$      | .1229** (.0399)| $S_3$       | .0226** (.0036)| $GAI \times S_6$ | $-0.0341** (.0044)$ |
| $AP_2$      | .1068** (.0363)| $S_4$       | .0523** (.0144)| $GAI \times S_7$ | $-0.0352** (.0045)$ |
| $AP_3$      | .004 (.0406)  | $S_5$       | .0504** (.004) | $GAI \times S_8$ | $-0.0365** (.0049)$ |
| $AP_4$      | .0389 (.0402) | $S_6$       | .0747** (.0145)| $GAI \times S_9$ | $-0.0482** (.0055)$ |
| $AP_5$      | $-0.0092 (.0439)$ | $S_7$     | .0822** (.0045)| $GAI \times S_{10}$ | $-0.054** (.0068)$ |
| $AP_6$      | .0421 (.049)  | $S_8$       | .1106** (.0147)| $GAI \times S_{11}$ | $-0.0558** (.0086)$ |
| $AP_7$      | .0287 (.0461) | $S_9$       | .112** (.0055)| $GAI \times S_{12}$ | $-0.0614** (.0113)$ |
| $AP_8$      | .3235** (.0555)| $S_{10}$    | .1289** (.0154)| $GAI \times S_{13}$ | $-0.0838** (.0157)$ |
| $AP_9$      | .0635 (.0493) | $S_{11}$    | .1128** (.0091)| $GAI \times S_{14}$ | $-0.0883** (.0221)$ |
| $AP_{10}$   | .0008 (.0348) | $S_{12}$    | .1273** (.0181)| $GAI \times S_{15}$ | $-0.0568** (.0245)$ |
| $AP_{11}$   | .1885** (.0452)| $S_{13}$    | .1500** (.0161)| $GAI \times S_{16}$ | $-0.0704** (.019)$ |
| $AP_{12}$   | .2243** (.046) | $S_{14}$    | .1325** (.0254)| $C_2$ | .0381 (.0312) |
| Female      | .1518** (.0101)| $S_{15}$    | .1155** (.026) | $C_3$ | .1714** (.0454) |
| PS          | .0189* (.0093)| $S_{16}$    | .1364** (.023) | $C_4$ | .1877** (.0479) |
| RC          | .0141** (.0013)| $GAI \times S_2$ | $-0.0217** (.0039)$ | $C_5$ | .1201** (.0458) |
| $(RC)^2$    | $-0.0024** (.0009)$ | $GAI \times S_3$ | $-0.0211** (.0035)$ | $C_6$ | .1212** (.048) |
| GAI         | .3229** (.0057)| $GAI \times S_5$ | $-0.0275** (.0004)$ | $C_7$ | .1042** (.0426) |

Notes: Panel robust standard errors are in parentheses. The equation also controls for 252 interactions between academic programs and time dummies. IV = independent variables. $R^2 = .17$; cross-sections included: 13,066; total panel (unbalanced) observations: 70,085.

* Statistically significant at the 5% level.

** Statistically significant at the 1% level.
and significant. However, all the estimated interaction effects coefficients ($\hat{\phi}_t$) are negative and significant. This result implies that the official admission criterion used by the UPR-Bayamón predicts academic success better early on in the education, which is what would be expected if the contents of university education gradually shift from those of high school. This result is in accordance with predictive validity studies, which often find that the correlation between the admission policy and the outcome variable is higher the earlier in the education the outcome is measured.  

Expression 7 describes the behavior of GPA over time reference 2. All the estimated main effects coefficients ($\hat{\gamma}_t$) are positive and significant, but all the estimated interaction effects coefficients ($\hat{\phi}_t$) are negative and significant. So, the 15 partial derivatives have positive intercept and negative slope. For average-quality students ($\overline{GAI}$), the expression collapses to its intercept because the GAI variable is standardized. Hence, the slopes become positive and steeper, increasing the value of the derivatives in all cases in which student quality is less than average. Contrary to what would be expected, this result implies that the GPA of low-quality students ($GAI < \overline{GAI}$) tends to increase faster than the GPA of high-quality ones ($GAI > \overline{GAI}$) over time reference 2. How could this result be explained?

It has been argued that high-quality students face lower marginal costs because they learn faster than low-quality students and consequently incur in lower opportunity costs (Winkler, 1987). If so, the greater the student quality the lower the cost of acquiring higher grades. Subsequently, if there has been a reduction in academic standards and professors have become more lenient in grading, then the price of higher grades tends to decrease, inducing the two well-known substitution and income effects. It seems that, among low-quality students, the substitution effect dominates, providing them a greater incentive to move up along the distribution of grades at UPR-Bayamón. Thus, for average and low-quality students, this result implies grade inflation over time reference 2, which is consistent with the growth pattern of the GII in Figure 2.

The implications of such a result are quite different in the case of high-quality students. Each partial derivative will become negative, implying grade deflation over time reference 2, whenever:

$$GAI_t^* > \left\{ \frac{\hat{\gamma}_t \cdot \sigma_{GAI}}{\hat{\phi}_t} + \overline{GAI} \right \} \Rightarrow Z_{GAI} > \frac{\hat{\gamma}_t}{\hat{\phi}_t}, \forall t = 2, \ldots, 16. \tag{8}$$

The estimated threshold lies in the following range: $1.07 \leq Z_{GAI} \leq 3.03$. That is, for those high-quality students whose GAI is at least $1.071\sigma_{GAI}$ greater than the mean, income effect dominates, inducing them to substitute higher grades for other activities such as more leisure hours, part- or full-time jobs, or simply, enjoying life doing whatever they want to do. Under such conditions, high-quality students would have been considering higher grades as inferior goods, since their consumption tends to diminish to the extent that faculty members decrease their market price through lenient grading. The conjecture of higher grades as inferior goods was formally deduced long ago by Lichty et al. (1978). It should be rigorously analyzed in future research projects using new data-sets from different universities.

4. Summary

This study uses a rich longitudinal data-set to analyze the relationship between student quality and grade inflation at UPR-Bayamón from 1995–1996 until 2005–2006. All full-time students belonging to the 11 entering cohorts are tracked during each semester until they leave the institution. Two different time references are used: cohort time dummies (time reference 1) and semesters since admission to UPR-Bayamón (time reference 2). The path of the normalized GPA, GAI, and GII variables is analyzed over both references. Along time reference 1, evidence shows grade inflation for almost all the 14 sub-samples drawn from the first seven cohorts admitted to UPR-Bayamón (1995–1996 to 2001–2002); while the eight sub-samples drawn from the last four ones (2002–2003 to 2005–2006), do not have sufficient information to accurately estimate the index. Grade inflation is
documented throughout time reference 2. The maximum and minimum rates are 13.3 and 8.7%, respectively. Therefore, conclusions about the presence and magnitude of grade inflation depend on the time reference used, a point that has been ignored by the mainstream economic literature.

Given that GAI (proxy of student quality) is a time-invariant variable, it is not possible to define the GII at student level. Notwithstanding, a statistical model was specified to estimate the determinants of GPA using GAI as a covariate. Controlling for unobservable student heterogeneity (random-effects), student's gender, high school of origin, enrolled credit-hours, student’s quality and academic programs, the evidence is in accordance with the growth pattern exhibited by the GPA, GAI, and GII variables over the time references reported in Figures 1 and 2.

To analyze the behavior of the grade inflation rates among academic programs, sub-samples sets 1 and 2 were partitioned generating two restricted panels of 286 and 179 sub-samples, respectively. Evidence from two-way fixed-effects models shows that over both time references the estimated grade inflation rates are directly related to the proportions of female students and students who shift their original academic programs and indirectly related to the proportion of private school students. Furthermore, over time reference 1, the estimated coefficients of the three proportions are statistically significant. However, over time reference 2, only the coefficient of the female proportion is significant. Finally, the rates vary significantly among and within academic programs over both time references.

This relationship among students’ characteristics and the phenomena of grade inflation and grade divergence could have serious academic and institutional policy implications. An overview of the data shows that there are six academic programs with higher GAI’s: Engineering Transfers (335), Natural Sciences (310), Accounting (309), Computer Science (300), Business Administration (288), and Electronics (287). A second cluster comprises the seven programs with the lower GAI’s: Social Sciences (279), Education (277), Materials Management (272), Humanities (264), Office Systems (260), Engineering Technologies (257), and Physical Education (250). The average GAI (student quality) of the whole sample is 286. Therefore, on average, the programs of the first (second) cluster comprise students whose quality is greater than (less than) the mean. On the other hand, the proportions of female and private high school students in the programs of the first and second clusters are (47 and 49%) and (70 and 40%), respectively.

Thus, the seven programs in the second cluster are predominantly constituted by below average-quality students, females (70%), and public school students (60%). The last two proportions are directly and significantly related to the GII over time reference 1. On the other hand, according to expression 7, evidence points to grade inflation along time reference 2 only among average and below average-quality students. So, on average, the incidence of grade inflation is concentrated among and within the seven programs of the second cluster. It seems that the faculty members and administrators of these programs have become aware of the powerful and persuasive incentive of grades and are using them as a mechanism in order to recruit, retain, and graduate more students. The impact of increasing lenient grading on the quality, relevance, and pertinence of the education received by students needs to be analyzed in future studies.

Acknowledgments
I would like to recognize my deep debt of gratitude to Dr James F. Ragan, Jr., my mentor and loved friend, with whom I had the benefit of discussing these ideas on innumerable occasions. The paper was originally conceived as a joint research project, but was postponed several times while he finished other projects. Professor Ragan, who passed away on 13 October 2009, will always be missed by his students and friends. I am also deeply grateful to Gilberto Calderón for data compilation and, for their helpful comments and suggestions, to Dennis L. Weisman, Richard W. Lichty, Suphan Andic, Larry D. Singell, Paul G. Hare, Jaesung Kang, Abdullahi Abdulkadri, Nélida Matos-Díaz, Dwight García-Vázquez, Wendy Stock, Nellie J. Sieller, Gabriel Rodríguez-Matos, Arturo Avilés-González and one anonymous referee. Any remaining errors are of my sole responsibility.
Achen, A. C., & Courant, P. N. (2009). What are grades made of? Journal of Economic Perspectives, 23, 77–92.

Davies, S., & Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. Social Forces, 75, 1417–1438.

Dilts, D. A. (1983). A statistical interpretation of student evaluation feedback. The Journal of Economic Education, 71, 10–15.

EvViews. (2009). User guide. Irvine, CA: Quantitative Micro Software.

Freeman, D. G. (1999). Grade divergence as a market outcome. The Journal of Economic Education, 30, 344–351. http://dx.doi.org/10.1080/00220489909596091

Greene, W. H. (2012). Econometric analysis. New York, NY: Prentice Hall.

Groove, W. A., & Wasserman, T. (2004). The life-cycle pattern of collegiate GPA: Longitudinal cohort analysis and grade inflation. The Journal of Economic Education, 35, 162–174. http://dx.doi.org/10.3200/JECE.35.2.162-174

Gump, S. E. (2007). Student evaluation of teaching effectiveness and the leniency hypothesis: A literature review. Educational Research Quarterly, 30, 55–68.

Hu, S. (2005). Beyond grade inflation: Grading problems in higher education (ASHE Higher Education Report. Vol. 30(6)). San Francisco: Jossey-Bass, Wiley.

Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? The Journal of Economic Education, 36, 29–42. http://dx.doi.org/10.3200/JECE.36.1.29-42

Jewell, R. T., McPherson, M. A., & Tieslau, M. A. (2013). Whose fault is it? Assigning blame for grade inflation in higher education. Applied Economics, 45, 1185–1200.

Johnes, G. (2004). Standards and grade inflation. In G. Johnes & J. Johnes (Eds.), International handbook on the economics of education (pp. 462–483). Northampton, MA: Edward Elgar Publishing.

Johnson, V. E. (2003). Grade inflation: A crisis in college education. New York, NY: Springer-Verlag.

Kanagaretnam, K., Mathieu, R., & Thevaranjan, A. (2003). An economic analysis of the use of student evaluations: Implications for universities. Managerial and Decision Economics, 24, 1–13. http://dx.doi.org/10.1002/20011099.1468

Krautmann, A. C., & Sander, W. (1999). Grades and student evaluations of teachers. Economics of Education Review, 18, 59–63. http://dx.doi.org/10.1016/S0222-7757(98)00044-1

Lichy, R. W., Vose, D. A., & Peterson, J. M. (1978). The economic effects of grade inflation on instructor evaluation: An alternative interpretation. The Journal of Economic Education, 10, 3–11. http://dx.doi.org/10.2307/118216O

Love, D. A., & Kotchen, Matthew J. (2010). Grades, course evaluations, and academic incentives. Eastern Economic Journal, 36, 151–163. http://dx.doi.org/10.1057/eej.2009.6

Matos-Díaz, H. (2012). Student evaluation of teaching, formulation of grade expectations, and instructor choice: Explorations with random-effects ordered probability models. Eastern Economic Journal, 38, 296–309.

McKenzie, R. B. (1975). The economic effects of grade inflation on instructor evaluations: A theoretical approach. The Journal of Economic Education, 6, 99–106. http://dx.doi.org/10.2307/1182459

McKenzie, R. B., & Staaf, R. J. (1974). An economic theory of learning: Students’ sovereignty and academic freedom. Blacksburg, VA: University Publications.

McPherson, M. A. (2006). Determinants of how students evaluate teachers. The Journal of Economic Education, 37, 3–20. http://dx.doi.org/10.3200/JECE.37.1.3-20

Nelson, J. P., & Lynch, K. A. (1984). Grade inflation, real income, simultaneity, and teaching evaluations. The Journal of Economic Education, 15, 21–39.

Rosovsky, H., & Hartley, M. (2002). Evaluation and the academy: Are we doing the right thing? Grade inflation and Letters of Recommendation (pp. 1–25). Cambridge, MA: American Academy of Arts and Sciences.

Sabol, R., & Wakeman-Linn, J. (1991). Grade inflation and course choice. Journal of Economic Perspectives, 5, 159–170. http://dx.doi.org/10.1257/jep.5.1.159

Seiver, D. A. (1983). Evaluations and grades: A simultaneous framework. The Journal of Economic Education, 14, 32–38. http://dx.doi.org/10.2307/1182856

Szafran, R. F. (2001). The effect of academic load on success for new college students: Is lighter better? Research in Higher Education, 42, 27–50. http://dx.doi.org/10.1023/A:1018712527023

Winkler, D. R. (1987). Screening models and education. In G. Psacharopoulos (Ed.), Economics of education: Research and studies (pp. 287–291). New York, NY: Pergamon Press.

Zangenehzadeh, H. (1988). Grade inflation: A way out. The Journal of Economic Education, 19, 217–223. http://dx.doi.org/10.3200/JECE.19.2.217-223.
