education policy analysis archives

A peer-reviewed, independent, open access, multilingual journal

epaa aape

Arizona State University

Volume 29 Number 57 May 3, 2021 ISSN 1068-2341

Taxonomy of Grading Practices in the University of Puerto Rico at Bayamón, 1995-96 to 2015-16

In loving memory of my son, Horacio Matos-De Jesús (March 12, 1983 – December 20, 2009), who passed too early and whom I sorrowfully miss, each second of each day.

Horacio Matos-Díaz
University of Puerto Rico at Bayamón
Puerto Rico

Citation: Matos-Díaz, H. (2021). Taxonomy of grading practices in the University of Puerto Rico at Bayamón, 1995-96 to 2015-16. *Education Policy Analysis Archives, 29*(57). https://doi.org/10.14507/epaa.29.5845

Abstract: Faculty members and their corresponding academic fields at the University of Puerto Rico at Bayamón are classified with regard to grading practices over time. Based on the effects on the intercept of the equations that predict the GPA and the proportion of student withdrawals observed in each of the 39,337 courses offered during 41 consecutive terms, faculty members and academic fields are scaled from the easiest to the most difficult. Evidence points to the conclusion that the courses of the most difficult academic fields are offered primarily by the hardest grading faculty members and attended by the most academically able students, while the courses of the easiest academic fields are offered primarily by the easiest grading faculty members and attended by less academically able students. The conclusion of such self-sorting processes is reinforced by evidence from maximum likelihood models demonstrating that the probability that a randomly selected faculty member behaves like a high-grader or a low-grader is highly and significantly related to the cluster of academic fields to which the faculty member belongs. Such a probability is also strongly and significantly

Journal website: http://epaa.asu.edu/ojs/
Facebook: /EPAAA
Twitter: @epaa_aape

Manuscript received: 7/24/2020
Revisions received: 11/24/2020
Accepted: 1/12/2021

influenced by the heterogeneity of student academic ability distribution. Hence, faculty members are very responsive to signals sent by their students' characteristics. This empirical result deserves further detailed analysis given that it implies a scenario in which faculty members and students engage in a shopping-around process in which both parties free-ride from each other, altering institutional norms and academic standards.
Keywords: grading taxonomy; multinomial logistic models; easy and tough grading; Puerto Rico

Taxonomía de las prácticas de evaluación en la Universidad de Puerto Rico en Bayamón, 1995-96 al 2015-16

Resumen: Los profesores de la Universidad de Puerto Rico en Bayamón y sus correspondientes programas académicos son clasificados en términos de sus prácticas de evaluación a través del tiempo. A base de su efecto sobre el intercepto de las ecuaciones que predicen el GPA y la proporción de bajas observadas en cada uno de los 39,337 cursos ofrecidos durante 41 semestres consecutivos, se les ordena desde los más fáciles hasta los más difíciles. La evidencia apunta a la conclusión de que los cursos de los programas académicos más difíciles son ofrecidos principalmente por los profesores más exigentes, y están constituidos por los estudiantes académicamente más aventajados; mientras que los cursos de los programas académicos más fáciles son ofrecidos principalmente por los profesores menos exigentes, y están constituidos por los estudiantes académicamente más desventajados. La conclusión sobre estos procesos de auto selección es reforzada por evidencia proveniente de modelos de verosimilitud máxima que demuestran que la probabilidad de que un profesor aleatoriamente seleccionado se comporte como un evaluador exigente y riguroso o como uno lenitivo está significativamente determinada por el grupo de programas académicos al que pertenezca. Tal probabilidad está significativa y fuertemente influida por la heterogeneidad de la distribución estudiantil de habilidades académicas. Por tanto, los profesores son muy responsivos a las señales enviadas por las características de sus estudiantes. Este resultado merece análisis más detallado pues implica un escenario donde profesores y estudiantes participan en un proceso (*van de compras*) donde cada parte pretende tomar ventajas de la otra, impactando adversamente las normas y los estándares académicos institucionales.
Palabras clave: taxonomía de calificaciones; modelos logísticos multinomiales; evaluación lenitiva y exigente; Puerto Rico

Taxonomia de práticas avaliativas na Universidade de Puerto Rico-Bayamón, 1995-96 até 2015-2016

Resumo: O quadro docente e seus campos acadêmicos correspondentes na Universidade de Porto Rico-Bayamón são classificados de acordo com suas práticas de avaliação no decorrer do tempo. Baseado em efeitos de interceptação de equações que prevêem o GPA (média de pontos por nota) e pela proporção de abandonamentos de estudantes observados em cada uma das 39.337 matérias oferecidas durante 41 semestres consecutivos, docentes e campos acadêmicos são escalonados do mais fácil ao mais difícil. A evidência aponta para a conclusão que as matérias dos campos acadêmicos mais difíceis são ofertadas principalmente por docentes que avaliam mais rigorosamente e por alunos mais academicamente habilidosos, enquanto que matérias em campos acadêmicos mais fáceis são ofertadas principalmente por docentes que avaliam de maneira mais branda e composta por alunos menos academicamente habilidosos. A conclusão desses processos auto-regulatórios é reinforçado por evidências de modelos de máxima probabilidade os quais demonstram que a probabilidade de selecionar um docente

acadêmico randomicamente e que se comporta como um avaliador-brando ou avaliador-rigoroso é alta e significantemente relacionado a um cluster de campos acadêmicos do qual o docente do quadro de professores pertence. Tal probabilidade é também forte e significantemente influenciada pela heterogeneidade da distribuição da habilidade acadêmica do estudante. Portanto, os docentes são bastante responsivos aos sinais característicos enviados por seus alunos. Os resultados empíricos merecem análises futuras detalhadas uma vez que se implica um cenário no qual docentes e discentes se engajam num processo de "busca no mercado" no qual as duas partes se encontram gratuitamente, alterando as normas institucionais e os padrões acadêmicos.
Palavras-chave: taxonomia avaliativa; modelos de logística multinominais; avaliação branda e rigorosa; Porto Rico

Introduction

This paper aims to develop a taxonomy for classifying faculty members and their corresponding academic fields (AFs) at the University of Puerto Rico at Bayamón (UPR-Bayamón) according to grading practices over time. To this end, a rich and detailed panel data comprising all 39,337 courses offered during 41 consecutive terms (1995-96 to 2015-16) is used. Thus, the unit of analysis will be individual courses. The grade point average (GPA) and the proportion of student withdrawals (PWs) observed in each course will provide the underpinnings to scale faculty members and AFs from the easiest to the most difficult. Both parameters, GPA and PWs, will be defined conceptually and operationally later on.

Once faculty members and AFs are properly classified according to the defined categories, the paper focuses on shedding light on several relevant academic issues. For instance, it is important to determine whether faculty members and AFs are consistent regarding their grading practices over time. Professors will behave consistently to the extent that they can be classified as easy or difficult with respect to both GPA and PWs. On the other hand, an inconsistent behavior will be documented to the extent that the professor can be classified as difficult with respect to one parameter and easy with respect to the other. The same criteria will be applied to the AFs.

Once the consistency (or inconsistency) of the grading practices of faculty members and AFs is determined, the taxonomy will allow the analysis of the distribution of easy and difficult faculty members by easy and difficult AFs. Research questions considered and discussed in the text are the following: Is there a symmetric distribution of faculty members by AFs in terms of difficulty levels? That is, do difficult (easy) faculty members belong to difficult (easy) AFs or are there crossovers or asymmetries? Which are the AFs with the greatest proportion of difficult and easy faculty members? Which are the characteristics of the students enrolled in the most difficult (easy) AFs? Are there self-sorting processes matching difficult (easy) faculty members with difficult (easy) AFs with the academically ablest (least able) students? If such is the case, what are their academic consequences?

It should be emphasized that university norms and academic standards share the two fundamental properties that distinguish public goods: nonexclusive and nonrival. Given these properties, public goods are subject to free-riding behavior, i.e., consumers will not reveal their willingness to pay for the good preferring to rely on others to pay for it. What kind of behaviors will free-riding induce on faculty members and students, and what will the academic consequences be?

The final stage of the paper explores the analytical insights coming from the classification of each faculty member by categories through the estimation of a multinomial logistic model using maximum likelihood methods. The estimated model sheds light on the determinants of the probability that a randomly selected faculty member belongs to a determinate category and to determinate AFs.

To the best of my knowledge, this is the first paper proposing a classification of the grading practices of both faculty members and their corresponding AFs at the university level that is solidly based on empirical evidence gathered by using different econometric techniques and a detailed panel data. Although the reported results are not generalizable, the methodology adopted allows replicating the estimation process in other institutions using their own data.

The remainder of the paper focuses on reviewing the scant literature gathered, describing the data underlying the present study; contextualizing the relationship between GPA and course withdrawals; motivating the empirical formulation of the models; justifying and motivating the taxonomy; presenting the empirical results, and summarizing the central findings of the study.

Unsuccessful Quest of Relevant Literature: A Note

According to Johnson (2003), the origins of the discussion on the relationship between student evaluations of teaching (SET) and grade inflation dates back to studies conducted by Remmers (1928, 1930) in the late 1920s. Since then, much has been written on the subject. Even in AFs far from the scope of economics, such as medical imaging and radiation sciences, the subject of grade inflation is passionately debated nowadays (Watts & Winters, 2016). However, with the exceptions of Sabot & Wakeman-Linn (1991), Achen & Courant (2009), as well as Butcher et al. (2014), there has been little published work aimed at classifying faculty members and their corresponding AFs using a scale from low-grading to high-grading. Although these three papers are devoted to the study of the phenomenon of grade inflation at three U.S. universities (Williams College, The University of Michigan, and Wellesley College, respectively), their authors classified the AFs of each institution from low- to high-grading in terms of their average GPA through time, which is consistent with some of the procedures adopted by the taxonomy proposed in this paper.

According to Sabot & Wakeman-Linn (1991), Mathematics (2.53), Chemistry (2.66), and Economics (2.81) exhibit the lowest mean grades among all the AFs. Exactly the same order is reported by Achen & Courant (2009) using data from The University of Michigan: Mathematics (2.53), Chemistry (2.66), and Economics (2.81). Although the data reported by Butcher et al., (2014) for AFs at Wellesley College change the order, Chemistry, Economics, and Mathematics are among the lowest-grading fields. All three papers report that liberal arts AFs are at the top of the grades distribution, which is consistent with this paper's results.

Several papers related to the results reported in this study will be discussed whenever it is necessary throughout the text. They address important academic issues; particularly, the relationship between teachers' quality and students' academic achievements through time (Aaronson et al., 2007; Keng, 2018; Koedel, 2008, 2011; Rivkin et al., 2005; Rockoff, 2004; Weiss & Rasmussen, 1960), as well as the relationship at college level between grading practices and earnings differences by AFs and their effects on students' majors selection (Ahn et al., 2019), grade inflation, students' academic efforts, and human capital accumulation through time (Babcock, 2010; Stinebrickner & Stinebrickner, 2008), as well as grade inflation, enrollment in higher education, and earnings (Nordin et al., 2019).

Data and Empirical Models

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 Commission on Higher Education (MSCHE), the institution offers associate and bachelor's degrees, as well as articulated transfer programs to the Rio Piedras, Mayagüez, and Medical Sciences campuses. In the fall of 2019, total enrollment at UPR-Bayamón was 3,937, including 3,444 full-time students. All enrolled students were Hispanic, with 31% first-generation college students. The fact that 70% of UPR-Bayamón students receive Pell Grants reflects a mostly low-income status population. The 2018 graduating class had 524 graduates. Applications and admissions for 2018-19 present a behavior similar to recent years, with 4,430 applications received, 1,117 admitted students, and 978 enrolled students. Even though numbers can change from year to year, approximately 51% of the student population is female, and 53% comes from public high schools.

In the UPR system, the admission decisions made on each campus and by its academic programs are based exclusively on each applicant's GAI (General Admission Index) score, 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 (CEEB) test. Every year, each of the UPR's eleven campuses establishes the minimum GAI required by its different academic programs in response to trends in enrollment demand and the programs' capacity. For instance, during the academic year 2018-19 at UPR-Bayamón, the minimum required GAI ranged from 320 (B.S. in Electrical and Mechanical Engineering transfer programs) to 250 (B.A. in Preschool, Elementary, and Adapted Physical Education, as well as Accounting). However, during the academic year 2014-15, these minimum GAI were 340 (Mechanical Engineering transfer program), 270 (Preschool and Elementary Education), 265 (Adapted Physical Education), and 290 (Accounting). It should be mentioned that at other UPR campuses the same academic programs could require different minimum GAI from year to year.

The fact that the GAI required for each program is made public every year has led from its inception, to a self-inclusion/exclusion process by which students themselves decide whether to apply to the UPR (and a particular UPR program) based on their GAI and the minimum established by the program. Hence, the GAI plays a critical role, not only for admission to the different UPR campuses, but also for admittance to particular programs. Thus, it is used as the best available proxy of student academic ability (student quality).

Data Description

For each one of the 39,337 courses offered at UPR-Bayamón from 1995-96 to 2015-16, the following variables are available: enrollment, instructor who taught the course, letter grade distribution (As, Bs, Cs, Ds, Fs, and Ws), GPA, and the variance of the GPA distribution. A total of 21 AFs that offer affined courses were defined using dummy variables.[1] The mean and variance of the following variables are used as proxies to account for student academic ability at course level: high school graduation GPA (HSGPA), GAI, and the score on each of the five sections of the standardized admission test administered by the CEEB.[2] Furthermore, for each course offered, the proportions of students by gender and type of high school (public or private) are available. Dummies control for academic schedule (weekdays and hours) and for summer terms. For each faculty member in the sample, the following time-varying variables are available: age, academic rank, degree, and tenure status. Dummies control for instructor's gender and whenever students evaluate

[1] For specific details related to the distribution of AFs by academic departments, as well as the characteristics of the offering of each academic department, refer to Table Appendix 2.

[2] The CEEB test includes five sections: verbal and mathematical aptitude, and achievement in Spanish, English, and mathematics.

the course using the SET instrument. A set of 41 dummies identifying term/year captures time effects.[3] Table Appendix 1 describes the variables used.

The Relationship between GPA and Course Withdrawals (Ws)

According to Matos-Díaz (2018), academically lagging students, who belong to the bottom of the grade distribution, pull down the expected course GPA. Thus, Ws allow the cleaning of each course by separating students who will withdraw (leavers) from those who will remain (stayers). If so, leavers will not experience the expected decrease in their GPA, and the course GPA of the stayers will increase. Several studies invoke arguments very close to this line of reasoning in order to explain grade inflation at particular U.S. universities (Hoyt & Reed, 1976; McSpirit & Jones, 1999; Oglive & Jelavic, 2013; Rosovsky & Hartley, 2002). Hence, the liberalization of Ws policies should increase the GPA of both leavers and stayers, other things being equal. Such a prediction would be correct to the extent that the GPA of the stayers was greater than the observed one if leavers would have remained in the course, which sounds reasonable. However, an observability problem makes this prediction untestable. By the end of a term, one can observe the effect of students' decision only on the GPA of the stayers, but one cannot know how the leavers would have performed if they had not done so.[4]

To shed light on this issue, consider the following expressions (Matos-Díaz, 2018):

$$\left(GPA_i \middle| Ws\right) = \text{ GPA observed in course } i \text{ given that some students withdrew} \tag{1}$$

$$E\left(GPA_i \middle| \underline{Ws}\right) = \text{ GPA expected in course } i \text{ if withdrawn students would have stayed} \tag{2}$$

According to the previous argument, it should be expected that (1) > (2). If so, the net benefit arising from Ws could be measured using expression (3)

$$\left(GPA_i \middle| Ws\right) - E\left(GPA_i \middle| \underline{Ws}\right) \tag{3}$$

However, only the first term of (3) is observable. Unless the second term was computed, it would not be possible to uncover the relationship between GPA and Ws. With the exception of Matos-Díaz (2018), none of the published studies until now has undertaken such a task.

Contrariwise, it is reasonable to posit that Ws directly vary with the course's inherent difficulty level, according to students' criteria.[5] When course difficulty approaches zero, the average GPA approaches its maximum (implying grade compression), and Ws will tend toward zero. Conversely, in courses with greater difficulty, Ws will increase, and the GPA observed by the end of the term will be significantly lower than that observed in less difficult courses. If so, Ws and GPA should move in opposite directions. Evidence points to the conclusion that such is the case prevailing at UPR-Bayamón over time.

---

[3] Courses offered during summers are counted as part of fall sessions.

[4] Even though each course withdrawal is included into the student's official transcript, they have no effect on the estimated GPA. The deadline for course withdrawals is scheduled by the end of the term. Thus, by that time students have a clear idea about their academic success probabilities in each enrolled course.

[5] It could be possible that some students decide to withdraw for reasons not directly related to the course's inherent difficulty, such as economic adversities, personal problems, illnesses, etc. Otherwise, why would someone be willing to waste time and money withdrawing from a relatively easy but expensive course?

Matos-Díaz (2018) shows that estimates of the average GPA from three different samples over the 41 terms studied allow the isolation of this inverse relationship. Figure Appendix 1 plots the three series. Each exhibits an increasing tendency over time (grade inflation). However, sub-sample two $(\mathrm{Ws}=0)$ exhibits the highest GPA by term, ranging from 2.97 to 3.24. Conversely, sub-sample one $(\mathrm{Ws}\geq 1)$ exhibits the lowest GPA by term, ranging from 2.42 to 2.66. Right in the middle lies the full sample, in which case GPA by term ranges from 2.54 to 2.84. Therefore, Ws and GPA move in opposite directions over time. Figure Appendix 2 uses color bars to represent triplets of GPA distributed by AFs, succinctly depicting this inverse relationship. Furthermore, this relationship also prevails between and within AFs (Matos-Díaz, 2018).

Modeling GPA_{ij} and PWs_{ij} at Course Level

Let *As*, *Bs*, *Cs*, *Ds*, and *Fs* be the letter grades assigned by instructor *j* at the end of the term to the students enrolled in course *i*. Let Ws_{ij} be the number of students who withdraw from course *i* after the deadline to add or drop a course. Thus, total enrollment in course *i* offered by instructor *j* is proportional to

$$N_{ij}=\left(As+Bs+Cs+Ds+Fs+Ws\right) \tag{4}$$

Hence, the grade point average observed in the course by the end of the term $\left(GPA_{ij}\right)$ will be

$$GPA_{ij}=\left\{\frac{As\cdot 4+Bs\cdot 3+Cs\cdot 2+Ds\cdot 1+Fs\cdot 0}{As+Bs+Cs+Ds+Fs}\right\} \tag{5}$$

Likewise, the proportion of withdrawals observed in such a course $\left(PWs_{ij}\right)$ will be

$$PWs_{ij}=\frac{Ws_{ij}}{N_{ij}} \tag{6}$$

A caveat is in order here. It might be expected that academically ablest students sort themselves into the most difficult content courses offered by the more prestigious and high-paying academic fields. If so, a selectivity issue might arise.[6] However, to replicate the suggested method at student-course level, the dimensions of the data matrix should be increased substantially in order to accommodate the academic records of the students enrolled in each one of the 39,337 courses analyzed, as well as the covariates defining their characteristics. A panel at such specificity levels is unavailable. Thus, this study uses the course as the unit of analysis. Nevertheless, suppose that student-course is the unit of analysis instead of the course itself. By the end of a term, only the integers {0, 1, 2, 3, 4} corresponding to letter grades {*Fs*, *Ds*, *Cs*, *Bs*, and *As*}, for each course completed by student *h*, or {*Ws*} for each one who withdrew (missing), will be observed. Thus, GPA is not defined and maximum likelihood methods, such as ordered probit or ordered logit, should be used to model the probability of each letter grade. Correcting for selectivity in such cases might be a challenge (Heckman, 1979).

[6] This point was brought to my attention by one anonymous referee to whom I am grateful.

The preceding discussion justifies the specification of the following two equations in order to model both variables at course level.

$$Y_{ij\ell}^{GPA} = \alpha_{GPA} + \gamma_j^{GPA} + \sum_{\ell=2}^{21} \phi_\ell^{GPA} AF_\ell + \mathbf{X}'_{ij\ell}\boldsymbol{\beta} + \varepsilon_{ij\ell}^{GPA} \tag{7}$$

$$Y_{ij\ell}^{PWs} = \alpha_{PWs} + \gamma_j^{PWs} + \sum_{\ell=2}^{21} \phi_\ell^{PWs} AF_\ell + \mathbf{X}'_{ij\ell}\boldsymbol{\beta} + \eta_{ij\ell}^{PWs} \tag{8}$$

The parameters α_{GPA} and α_{PWs} are the overall constants (intercepts) in the models. The duplets $\left(\gamma_j^{GPA},\ \gamma_j^{PWS}\right)$ and $\left(\phi_\ell^{GPA},\ \phi_\ell^{PWs}\right)$ account for unobservable faculty heterogeneity (UFH) and AFs' effects, respectively. $\mathbf{X}'_{ij}$ is a k-vector of regressors and $\boldsymbol{\beta}$ is a vector of parameters to be estimated. Finally, ε_{ijl}^{GPA} and $\eta_{ij\ell}^{PWs}$ are the error terms, $i \in [1,\ 39{,}337]$, $j \in [1,\ 987]$, and $\ell \in [1,\ 21]$.

The vector $\mathbf{X}'_{ij\ell}$ includes the following semi-continuous regressors: instructor's age, GAI, GAI variance, proportion of students from private high schools, and proportion of female students. The vector also includes several dummies: AFs, faculty academic ranks, degree, and tenure status, class size, time and days of the courses, summer sessions, the use of SET in the course, as well as term/years.

Standardizing and evaluating the continuous regressors at their means, and evaluating the dummies at their reference groups, allow isolating the effects of UFH $\left(\gamma_j^{GPA},\ \gamma_j^{PWS}\right)$ and AFs $\left(\phi_\ell^{GPA},\ \phi_\ell^{PWs}\right)$ on the intercept of the estimated versions of expressions (7) and (8), which will collapse into expressions (9) and (10)

$$\hat{Y}_{j\ell}^{GPA} = \hat{\alpha}_{GPA} + \hat{\gamma}_j^{GPA} + \sum_{\ell=2}^{21} \hat{\phi}_\ell^{GPA} AF_\ell \tag{9}$$

$$\hat{Y}_{j\ell}^{PWs} = \hat{\alpha}_{PWs} + \hat{\gamma}_j^{PWs} + \sum_{\ell=2}^{21} \hat{\phi}_\ell^{PWs} AF_\ell \tag{10}$$

Given that GPA and PWs move in opposite directions at course level, low- and high-grading instructors would exert different effects on the intercept of the estimated equations. For example, $\hat{\gamma}_j^{GPA} < 0$ and $\hat{\gamma}_j^{PWs} > 0$ will imply that, other things being equal, the presence of instructor *j* in course *i* has the effect of shifting downward the intercept of the GPA and shifting upward the intercept of the PWs. Thus, the instructor is a low-grader. The instructor will be lenient (high-grader) whenever the signs of the estimated coefficients reverse. On the other hand, $\hat{\phi}_\ell^{GPA} < 0$ and $\hat{\phi}_\ell^{PWs} > 0$ will imply that, compared to the reference group, AF_ℓ has the effect of shifting downward the intercept of the GPA and shifting upward the intercept of the PWs, implying a low-grader AF. Whenever signs reverse, the AF will be reclassified as a high-grader one. Thus, the coefficients $\hat{\alpha}_{GPA}$, $\hat{\alpha}_{PWs}$, $\hat{\gamma}_j^{GPA}$, $\hat{\gamma}_j^{PWs}$, $\hat{\phi}_\ell^{GPA}$ and $\hat{\phi}_\ell^{PWs}$ allow classifying faculty members and AFs by categories according to the following principles.

Taxonomy Underpinning

The academic output of the 15 academic departments of UPR-Bayamón is classified into 21 AFs that offer affined courses. In order to properly classify faculty members and AFs in terms of how easy or difficult they are, the first stage of the study consists of estimating the models specified in expressions (7) and (8). Instructor *j* classifies as high-grader; high-grader (*HH*) whenever the respective intercept of the GPA and PWs equations shifts upward and downward, other things being equal. In the respective opposite case, instructor *j* classifies as low-grader; low-grader (*LL*). There are two other asymmetric cases (cross-categories): high-grader; low-grader (*HL*), and low-grader; high-grader (*LH*). Hence, there are four mutually exclusive and collectively exhaustive categories: *LL*, *HL*, *LH*, and *LL*. The order of the terms signals the effect of the instructor *j* and the AF_ℓ on the intercept of the GPA and PWs equations, respectively. Each category implies a different pattern of signs: *HH* = $(+,-)$, *HL* = $(+,+)$, *LH* = $(-,-)$, and *LL* = $(-,+)$.

Thus, *HL* implies that instructor $j\ (\text{or } AF_\ell)$ shifts the intercept of both equations upward. If, related to the GPA variable, instructor *j* is a high-grader, why do so many students drop from her courses? Maybe, several students are confident that there are other faculty members easier than she is, and they are willing to pay the price (Ws) of better grades. The relationship among GPA, Ws and student academic ability proxies that will be discussed later, justifies conjecture one (C1). C1 sustains that the incidence of *HL* instructors should be greater in AFs characterized by higher proportions of *HH* instructors, higher GPA and lower PWs, less academically able students, as well as lower admission requirements, and less difficult academic content courses. Hereafter and for concreteness, less competitive AFs.

Contrariwise, *LH* implies that instructor $j\ (\text{or } AF_\ell)$ shifts the intercept of both equations downward. Then, why would instructor *j* be a low-grader related to GPA but a high-grader with respect to Ws in this case? One plausible explanation is that many faculty members are very concerned about the possibility that only the best and hardworking students can access the highest grades (A and B). However, they might be willing to adjust their grading curves allowing failed students to obtain a letter grade of D rather than the deserved F. Thus, students that under normal circumstances would withdraw, now have an incentive to complete the course. Hence, Ws should decrease, while D grades would increase pulling down the overall course GPA. Then, conjecture two (C2) is in order. C2 establishes that the incidence of *LH* instructors should be greater in AFs characterized by higher proportions of *LL* instructors, lower GPA and higher PWs, academically ablest students, as well as higher admission requirements and more difficult academic content courses. Hereafter, more competitive AFs.

Each faculty member who has offered courses at UPR-Bayamón from 1995-96 to 2015-16 is classified into one of these categories, as well as into one of the 21 AFs defined. Therefore, the taxonomy will also allow the classification of each AF into one of these categories based on its own estimated effects $\left(\hat{\phi}_\ell^{GPA},\ \hat{\phi}_\ell^{PWs}\right)$ as well as on its members' mean estimated effects $\left(\bar{\hat{\gamma}}_\ell^{GPA},\ \bar{\hat{\gamma}}_\ell^{PWs}\right)$. Finally, the estimates of a multinomial logistic model allow classifying faculty members by categories.

The taxonomy sheds light on several academic issues that have important policy implications. For instance, it provides insights to explain why the distribution of high- and low-grading instructors significantly varies by AFs. There is strong empirical evidence confirming that teachers' quality exerts profound and significant effects on students' academic achievements and dropout outcomes through time (Aaronson et al., 2007; Keng, 2018; Koedel, 2008; Rivkin et al., 2005; Rockoff, 2004). Likewise, Babcock (2010) and Stinebrickner & Stinebrickner (2008) show that grade inflation in college reduces student effort, which in turn reduces human capital accumulation. On the contrary, Nordin et al., (2019) using data from Sweden, argue that grade inflation at the upper secondary education level benefits students through the increases in the probability of being accepted to a university of higher quality or to a high-paying field of education. Thus, according to the authors, it seems that what really matters for higher earnings are the academic credentials rather than the stock of human capital embodied in graduates. However, such an issue is not addressed in the paper.

The taxonomy allows analyzing the academic, economic, and social policy implications of two extreme plausible academic scenarios that could exist at UPR-Bayamón. The first one considers the possibility that, attracted by the convincing incentive of higher grades at a very low price (academic effort), academically weakest students sort themselves into those AFs offering the less difficult academic content courses taught by the highest-grading instructors. The second one considers the opposite extreme, in which case academically ablest students sort themselves into the AFs offering the more difficult academic content courses taught by the lowest-grading instructors. If so, it would have significant and profound consequences on the nature, relevance, and pertinence of the human capital embodied in students through time. Such an issue deserves urgent research.[7]

The second scenario discussed in the previous paragraph has been analyzed in a recent working paper. Using data from the University of Kentucky, Ahn et al. (2019) study, among other academic issues, the relationship among faculty grading practices, student workloads and efforts, and students' demand for courses and majors. They show that high earnings majors are the lower-grading and require a higher workload and effort from students.

Modeling the Probability that Instructor j Belongs to a Determinate Category

The final stage of the study explores the analytical insights coming from the classification of each faculty member by categories through the estimation of a multinomial logistic model.[8] For analytical purposes, the *HL* and *LH* cross-categories represent asymmetric behavior from faculty members. Thus, they were consolidated into a single one denoted as ambiguous (hereafter, *AA*), which was used as the reference group to estimate the model, and the estimated model allowed classifying each faculty member into one of these three categories: *HH*, *LL*, or *AA*. The categories are coded as *AA* = 0, *HH* = 1 and *LL* = 2.

Following Hosmer & Lemeshow (2000, pp. 260-262), let

$$g_1(\mathbf{X}) = ln\left[\frac{P(Y=1|\mathbf{X})}{P(Y=0|\mathbf{X})}\right] = \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \cdots + \beta_{1p}X_p = \mathbf{X}'\boldsymbol{\beta}_1 \tag{11}$$

$$g_2(\mathbf{X}) = ln\left[\frac{P(Y=2|\mathbf{X})}{P(Y=0|\mathbf{X})}\right] = \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \cdots + \beta_{2p}X_p = \mathbf{X}'\boldsymbol{\beta}_2 \tag{12}$$

After some algebraic manipulations, it can be shown that the conditional probabilities of each outcome category given the covariate vector are computed as follows:

[7] To this regard, refer to Arcidiacono (2004), who analyzes the causes of the ability sorting across majors.
[8] For details, refer to Greene (2012, p. 764), as well as Hosmer & Lemeshow (2000).

$$P\left(Y=0|\mathbf{X}\right)=\frac{1}{1+exp\left(g_1\left(\mathbf{X}\right)\right)+exp\left(g_2\left(\mathbf{X}\right)\right)} \tag{13}$$

$$P\left(Y=1|\mathbf{X}\right)=\frac{exp\left(g_1\left(\mathbf{X}\right)\right)}{1+exp\left(g_1\left(\mathbf{X}\right)\right)+exp\left(g_2\left(\mathbf{X}\right)\right)} \tag{14}$$

$$P\left(Y=2|\mathbf{X}\right)=\frac{exp\left(g_2\left(\mathbf{X}\right)\right)}{1+exp\left(g_1\left(\mathbf{X}\right)\right)+exp\left(g_2\left(\mathbf{X}\right)\right)} \tag{15}$$

Results and Discussion

Classifying Faculty Members by Categories

As already mentioned, the first stage of the study consisted of estimating the equations of GPA_{ij} and PWs_{ij} described in expressions (7) and (8), using the panel data of 39,337 courses offered in UPR-Bayamón during 41 consecutive terms. Both estimated equations exhibited an excellent statistical fit.[9] In order to isolate the effects of UFH $\left(\gamma_j^{GPA},\ \gamma_j^{PWs}\right)$ and the AFs $\left(\phi_\ell^{GPA},\ \phi_\ell^{PWs}\right)$ on the intercept of each equation, the continuous regressors were standardized and evaluated at their mean values, while dummies were evaluated at their reference groups. After completing such procedures, the expressions (9) and (10) were computed, allowing classification of the 987 total faculty members and the 21 AFs by categories.

Instructors who did not teach at least four three credit courses and either retired during the first term or began to teach during the last term of the study, were excluded, rendering a final sample of 700 instructors. Table 1 reports the distribution of this final sample by categories. The results discussed in the remainder of the paper are based on that sample.

Table 1

Faculty Members Distributed by Categories

	Categories				
	HH	*HL*	*LH*	*LL*	Total
Frequencies:	265	100	110	225	700
(%)	(37.86%)	(14.29%)	(15.71%)	(32.14%)	(100%)

Notes: *HH* = (high-grader; high-grader), *HL* = (high-grader; low-grader), *LH* = (low-grader; high-grader), *LL* = (low-grader; low-grader).

[9] Their coefficients are not reported, but they are available upon request.

The greatest proportion of faculty members belongs to the category of *HH* (265/700 = 37.86%). A total of 225 members is classified as *LL* (32.14%) and 210 (30%) are classified as *AA* (*HL* or *LH*). Tables 2 and 3 report the distribution of instructors by categories and AFs. According to Table 2, a great proportion of the *HH* instructors is concentrated in the AFs at the top of the distribution, while a great proportion of the *LL* ones is concentrated in the AFs at the bottom. For instance, the proportion of *HH* instructors in the first seven AFs of the distribution runs from 40% in Humanities to 76% and 77% in English and Management, respectively. Conversely, the five AFs at the bottom contain the highest proportions of *LL* instructors and the lowest proportions of *HH*. For example, the proportion of *LL* is 46% in Biology and Economics & Statistics (hereafter, Econ/Stat), 54% in Engineering Technologies, 57% in Mathematics, and 100% in Chemistry. Several AFs in the middle of the distribution also exhibit high proportions of *LL* instructors: Computer Sciences (55%), Office Systems (58%), Engineering Transfers (67%), and Electronics (67%). Therefore, the distribution of *HH* and *LL* instructors by AFs does not follow a random process.

Table 2

Faculty Distributed by Categories and Academic Fields

Academic Fields	*HH*	*HL*	*LH*	*LL*	Total
Physical Education	18 (53%)	15 (44%)	0	1 (3%)	34
Education	34 (59%)	16 (28%)	5 (9%)	3 (5%)	58
Marketing	5 (71%)	0	2 (29%)	0	7
Social Sciences	13 (29%)	45	9 (13%)	10 (15%)	67
English	2 (67%)	3	0	4 (5%)	75
Management	20 (77%)	0	5 (19%)	1 (4%)	26
Humanities	18 (40%)	12 (27%)	2 (4%)		
Engineering Transfers (A)	0	0	1 (33%)		
Finance	10 (43%)	1 (4%)	7 (30%)	5 (22%)	23
Spanish	14 (32%)	3 (7%)	10 (23%)	17 (39%)	44
Electronics	3 (11%)	2 (7%)	4 (15%)	18 (67%)	27
Computer Sciences	11 (26%)	1 (2%)	7 (17%)	23 (55%)	42
Office Systems	0	14 (42%)	0	19 (58%)	33
Accounting	13 (41%)	2 (6%)	8 (25%)	9 (28%)	32
Physics	5 (22%)	2 (9%)	10 (43%)	6 (26%)	23
Materials Management (B)	1 (20%)	0	3 (60%)	1 (20%)	5
Engineering Technologies (C)	4 (17%)	1 (4%)	6 (25%)	13 (54%)	24
Biology	0	0	25 (54%)	21 (46%)	46
Econ/Stat	5 (38%)	0	2 (15%)	6 (46%)	13
Chemistry	0	0	0	26 (100%)	26
Mathematics	12 (26%)	4 (9%)	4 (9%)	27 (57%)	47
	265	100	110	225	700
Engineering (A + B + C)	5 (16%)	1 (3%)	10 (31%)	16 (50%)	32

Why does the incidence of *HH* and *LL* instructors significantly vary by AFs? To shed light on this issue, Table 3 reports the GPA and PWs observed in all the courses offered by each AF during the 41 terms analyzed and the means of the mathematical aptitude (MA) and verbal aptitude (VA) tests of the students enrolled. The GPA column reports in parentheses the total courses offered by each AF. Thus, the AFs at the top of the distribution, which exhibit the highest proportions of *HH* instructors, also exhibit the highest GPA and the lowest PWs. Contrariwise, the AFs at the bottom, which exhibit the highest proportions of *LL* instructors, also exhibit the lowest GPA and the highest PWs.

Table 3

Faculty Taxonomy and Students' Characteristics

	Students characteristics				Categories				
Academic Fields	GPA	MA	VA	PWs (%)	*HH*	*HL*	*LH*	*LL*	Total
Physical Education	3.35 (1,970)	524	506	4.62	18	15	0	1	34
Education	3.21 (2,714)	523	527	6.41	34	16	5	3	58
Marketing	3.19 (738)	571	546	2.66	5	0	2	0	7
Social Sciences	2.97 (2,551)	560	550	5.97	35	13	9	10	67
English	2.95 (3,922)	567	546	7.57	57	14	0	4	75
Management	2.95 (1,423)	573	546	5.41	20	0	5	1	26
Humanities	2.93 (2,903)	562	548	7.08	18	12	2	13	45
Engineering Transfers (A)	2.92 (405)	671	605	10.6	0	0	1	2	3
Finance	2.89 (791)	589	552	6.98	10	1	7	5	23
Spanish	2.78 (2,948)	560	542	5.19	14	3	10	17	44
Electronics	2.76 (2,471)	590	535	13.4	3	2	4	18	27
Computer Sciences	2.76 (2,496)	603	564	11.6	11	1	7	23	42
Office Systems	2.75 (1,780)	497	505	8.34	0	14	0	19	33
Accounting	2.61 (1,843)	592	555	14.3	13	2	8	9	32
Physics	2.58 (1,321)	602	562	11.4	5	2	10	6	23
Materials Management (B)	2.56 (309)	567	535	8.01	1	0	3	1	5
Engineering Technologies (C)	2.48 (1,203)	547	502	11.2	4	1	6	13	24
Biology	2.45 (1,806)	571	562	11.8	0	0	25	21	46
Econ/Stat	2.28 (875)	592	559	17.4	5	0	2	6	13
Chemistry	2.21 (1,084)	597	569	15.1	0	0	0	26	26
Mathematics	1.73 (3,784)	585	556	29.2	12	4	4	27	47
Total	2.72 (39,337)	568	544	10.65	265	100	110	225	700
Engineering (A + B + C)	2.59 (1,917)	577	531	10.3	5	1	10	16	32

Notes: MA = mathematical aptitude; VA = verbal aptitude; PWs = proportion of course withdrawals; *HH* = high-high; *HL* = high-low; *LH* = low-high; *LL* = low-low. Values in parentheses in the first column (GPA) represent the total courses offered by the program during the 41 terms studied.

There are other empirical findings related to student academic ability deserving discussion. The Physical Education and Education AFs exhibit the highest GPA (3.35 and 3.21) and relatively lower PWs (4.62% and 6.41%), respectively. However, the students enrolled in their courses are the less academically able ones, exhibiting mean MA scores of only 524 and 523 points, respectively.[10] Conversely, Econ/Stat, Chemistry, and Mathematics, which exhibit the respective lowest GPA (2.28, 2.21 and 1.73) and highest PWs (17.4%, 15.1% and 29.2%), are constituted by more academically able students given that their mean scores on such a test are 592, 597 and 585, respectively. Therefore, the AFs at the bottom of the distribution behave like the more competitive. Nonetheless, among the AFs at the top, the opposite occurs. Apparently, there is a self-sorting process matching academically ablest students to low-grading and more difficult academic content courses and less academically able students to high-grading and less difficult content ones. Evidence also points to the conclusion that the distributions of GPA and Ws prevailing in each AF guide the incentive mechanism process. Thus, for students, the attractiveness of grades is a powerful and convincing incentive that AFs are using as an input into their processes of recruiting, retaining, and graduating students (Matos-Díaz, 2012).

To contextualize these results, it would be convenient to analyze the relationship among student academic ability proxies, GPA, and Ws through time. The best available student academic ability proxies come from the mean score of the mathematical and verbal aptitude tests computed over the 41 terms. Figures Appendixes 3 and 4 clearly depict the growth-path of the student academic ability proxies, as well as the GPA and PWs. Four proxies are used to measure student academic ability at the course level: GAI (see footnote 2), high school GPA (HSGPA), as well as mathematics and verbal aptitude. Although GAI exhibits an increasing tendency over time, it should be mentioned that such a tendency is pushed by the self-sustained growth-path of HSGPA, which increases over time (implying grade inflation). However, according to mathematics and verbal aptitude figures, student academic ability decreases over time.[11] Therefore, it is expected that GPA goes down and PWs goes up. However, quite the opposite occurs. The GPA increasing tendency without a concomitant increase in student quality points to grade inflation. Meanwhile, the decreasing tendency of PWs points to the existence of a potentially even greater problem: diminishing academic standards through time.

The distribution of instructors by cross-categories (*HL* and *LH*) behaves as hypothesized. According to Table 2, the majority (70%) of the *HL* instructors is concentrated in five of the first seven AFs at the top of the distribution (Physical Education (15), Education (16), Social Sciences (13), English (14), and Humanities (12)). The respective numbers of *HH* instructors of these AFs (18, 34, 35, 57, and 18) totalize 162, which represent 61% (162/265) of the total *HH* instructors. Thus, these AFs behave like the less competitive. Hence, not only are the academic departments using the attractiveness of grades as mechanisms for recruiting students but also the faculty members are doing the same within their AFs. Thus, students are free to choose among the easier

---

[10] These results are in accordance with the earliest findings reported by Weiss & Rasmussen (1960), and with more recent ones from U.S. public universities (Koedel, 2011). Using data from Indiana University-Bloomington, Miami University-Ohio, and the University of Missouri-Columbia, Koedel (2011) shows that, compared to the departments of Mathematics, Sciences, and Economics; Humanities; and Social Sciences, the Departments of Education exhibit significantly greater GPAs. The respective GPA gaps range from 0.61, 0.59, and 0.49 grade points in Bloomington; 0.82, 0.53, and 0.68 grade points in Ohio, as well as 0.81, 0.64, and 0.68 grade points in Columbia.

[11] The decreasing tendency in the mathematical aptitude test scores reported in this study has also been documented at the international level, particularly in evidence from Norway. For details, refer to Bratsberg & Rogeberg (2018) and the references cited therein.

instructors when looking for better grades, even though the price to pay would be an increase in Ws. Therefore, conjecture one (C1) cannot be rejected.

On the other hand, the majority $(58/110 \approx 53\%)$ of the *LH* instructors is concentrated in the eight AFs at the bottom of the distribution in Table 2. However, only 21% (23/110) of them belongs to the seven AFs at the top, while the remaining 26% (29/110) is concentrated in the AFs at the center. As discussed, AFs at the bottom behave like the more competitive. Moreover, many of them (Physics, Materials Management, Engineering Technologies, Biology, and Chemistry) include laboratories in their curriculum as prerequisites. Furthermore, the Mathematics curriculum requires a series of prerequisite courses scheduled in a particular order. Precalculus 1 and 2, Calculus 1, 2 and 3, and Differential Equations are examples. Failure at the beginning of the sequences implies high opportunity costs for students in terms of time and money. Thus, given the inherently difficult contents of the courses offered by these AFs, the possibility of an unofficial tradeoff between F and D letter grades might be a good deal for so many students. So, conjecture two (C2) cannot be rejected.

It is important to note that such structure of mechanisms could induce significant and distorted social and economic results. UPR-Bayamón has eight academic departments that grant bachelor's degrees and seven academic departments that offer service courses.[12] To survive academically, degree AFs should compete among themselves to recruit, retain, and graduate students during the allotted time. It is expected that the best students apply to those AFs offering the more competitive, prestigious, and remunerated careers in the labor market and/or to those that increase the likelihood of gaining admission to graduate school. Such AFs offer the more difficult content courses (more competitive), as Tables 3 and 4 report. On the other hand, the less academically able students should have to apply to the academic departments offering the courses that belong to the less competitive AFs in order to improve their possibilities of admission.

Therefore, the education production function prevailing at the institution will generate at least two differentiated educational outputs. The first one will consist of students whose human capital stock has been enhanced with relevant, pertinent, and state-of-the-art knowledge and skills highly valued by the labor market and the graduate schools. On the contrary, the education received by students of the second output lacks such attributes, implying that their members' human capital stock would be doomed to fast obsolescence rates over time. To the extent that the labor market and the graduate schools would be able to distinguish between both groups of students, the members of the first group will be successful, and those belonging to the second one will not. For instance, Babcock (2010) shows that grade inflation in college reduces student effort, which in turn reduces human capital accumulation.[13] Therefore, a university education would be enhancing the inequality of the income distribution rather than improving it, as might be expected by enrolled students, graduates, and society.

The previous examples of the Physical Education and Education AFs illustrate this potential problem very well. Both AFs are responsible for providing the academic skills to the future teachers who in turn, will be responsible for the education of the next generations of students who will apply to UPR-Bayamón and other private and public institutions. However, both AFs would have ended up recruiting the less academically able students through the convincing incentives of higher grades at a very low price (academic effort) provided by a high-grading faculty who offers less difficult academic content courses. As mentioned, empirical evidence strongly confirms that teachers' quality exerts profound and significant effects on students' academic achievements and dropout outcomes

---

12 Refer to Table Appendix 2, for specific details.

13 Interested readers should also consult Stinebrickner & Stinebrickner (2008).

(Aaronson et al., 2007; Keng, 2018; Koedel, 2008; Rivkin et al., 2005; Rockoff, 2004).[14] Hence, the result is a vicious circle, which perpetuates the poverty traps surrounding the educational environment.

Classifying Academic Fields by Categories

The set of their own estimated effects and their members' mean estimate effects allows classifying AFs by categories.[15] Table 4 and Figures 1 and 2 report the relevant information. Nine (43%) AFs' estimated effects on GPA $\left(\hat{\phi}_{\ell}^{GPA}\right)$ are statistically significant. Six of them are positive and three are negative (English, Mathematics, and Econ/Stat). On the other hand, six (28.6%) AFs' estimated effects on PWs $\left(\hat{\phi}_{\ell}^{PWs}\right)$ are significant. Out of those, two are negative and four are positive (Engineering Transfers, Accounting, Mathematics, and Econ/Stat). It should be noted that for Chemistry, both estimated effects were insignificant and exhibited a pattern of signs consistent with an *HH* category: $\left(\hat{\phi}_{\ell}^{GPA} = 0.32\right)$ and $\left(\hat{\phi}_{\ell}^{PWs} = -5.71\%\right)$. Hence, based on their AF's estimated effects, only Mathematics and Econ/Stat consistently classify as *LL* AFs. The estimated effects of the remainder AFs are statistically insignificant, or the AFs behave like *AA*.

For illustration, consider the case of Mathematics. According to Table 4, the mean UFH effect on GPA is negative $\left(\bar{\hat{\gamma}}_{Math}^{GPA} = -0.13\right)$, and the respective mean effect on PWs is positive $\left(\bar{\hat{\gamma}}_{Math}^{PWs} = 2.98\right)$. Thus, based on their faculty members' mean effects, Mathematics is a low-grading AF $\left(LL = (-.+)\right)$. The AF's own effects reinforce such a faculty mean effect since the estimated coefficients are negative $\left(\hat{\phi}_{Math}^{GPA} = -0.79\right)$ and positive $\left(\hat{\phi}_{Math}^{PWs} = 14.81\right)$, and both are highly significant. Hence, after accounting for AF and faculty mean effects, the GPA $\left(\hat{Y}_{Math}^{GPA}\right)$ and PWs $\left(\hat{Y}_{Math}^{PWs}\right)$ expected in a randomly selected mathematics course are:

$$\hat{Y}_{Math}^{GPA} = 2.92 - 0.13 - 0.79 = 2.0, \text{ and } \hat{Y}_{Math}^{PWs} = 11.56 + 2.89 + 14.81 = 29.26\% .$$

Over the 41 terms analyzed, the means of the GPA and PWs of mathematics courses are 1.73 and 29.2%, respectively. Therefore, the predicted values are quite good approximations to the observed ones. A very similar result will emerge if the example uses the Econ/Stat figures rather than those of Mathematics. The replication of this procedure allows studying the other AFs.

---

[14] For technical details, refer to the papers and the references cited therein.

[15] For example, the estimated mean effect on GPA and Ws of the faculty members of AF_{ℓ} are defined as follows: $\bar{\hat{\gamma}}_{\ell}^{GPA} = \frac{1}{n}\sum \hat{\gamma}_{\ell}^{GPA}$ and $\bar{\hat{\gamma}}_{\ell}^{PWs} = \frac{1}{n}\sum \hat{\gamma}_{\ell}^{PWs}$.

Table 4

Uncovering the Effects of Unobservable Faculty Heterogeneity and Academic Fields on GPA and PWs

Intercepts:		$\hat{\alpha}_{GPA} = 2.92$			$\hat{\alpha}_{PWs} = 11.56$	
Academic Fields	GPA	$\bar{\hat{\gamma}}_{\ell}^{GPA}$	$\hat{\phi}_{\ell}^{GPA}$	PWs	$\bar{\hat{\gamma}}_{\ell}^{PWs}$	$\hat{\phi}_{\ell}^{PWs}$
Physical Education	3.35	0.55	0.03	4.62	0.04	-4.49*
Education	3.21	0.46	-0.04	6.41	-1.12	-1.57
Marketing	3.19	0.13	0.37***	2.66	-6.96	-0.33
Social Sciences	2.97	0.26	RG	5.97	-1.44	RG
Management	2.95	0.24	0.1	5.41	-6.08	1.82
English	2.95	0.66	-0.38***	7.57	-3.19	0.56
Humanities	2.93	0.18	0.03	7.08	1.42	-3.37
Engineering Transfers (A)	2.92	-0.63	0.17*	10.6	0.13	4.13*
Finance	2.89	-0.002	0.26***	6.98	-3.46	0.56
Spanish	2.78	-0.15	0.2283	5.19	-1.1	-4.47
Electronics	2.76	-0.22	0.38***	13.4	4.48	-4.13
Computer Sciences	2.76	-0.12	0.18	11.6	2.42	-2.07
Office Systems	2.75	-0.04	0.1	8.34	8.78	-12.09***
Accounting	2.61	-0.03	-0.03	14.3	-2.01	6.43***
Physics	2.58	-0.18	-0.01	11.4	-1.23	2.96
Materials Management (B)	2.56	-0.27	0.41***	8.01	0.25	-2.62
Engineering Technologies (C)	2.48	-0.22	0.37***	11.2	-2.24	-1.5
Biology	2.45	-0.66	0.15	11.8	-0.43	4.19
Econ/Stat	2.28	-0.02	-0.23***	17.4	1.99	3.08*
Chemistry	2.21	-1.05	0.32	15.1	12.63	-5.71
Mathematics	1.73	-0.13	-0.79***	29.2	2.89	14.81***
Engineering (A + B + C)	2.59	-0.36	N/A	10.3	-0.45	N/A
Sample size: the whole panel data of 39,337 courses offered along 41 consecutive terms						

Notes: $\bar{\hat{\phi}}_{\ell}^{GPA}$ and $\bar{\hat{\gamma}}_{\ell}^{PWs}$ are the averages of the unobservable faculty heterogeneity on the equations predicting GPA and PWs, respectively; $\hat{\phi}_{\ell}^{GPA}$ and $\hat{\phi}_{\ell}^{PWs}$ are the estimated AF's effects on the intercept of the corresponding equation; *******, ******, ***** = statistically significant at 1%, 5% and 10% level, respectively; RG = reference group; N/A = do not apply.

Figure 1 adds perspective to the discussion by showing the mean UFH effects on GPA by AFs, from smallest to largest. AFs to the left of the abscissa zero point are low-grading, while those to the right of it are high-grading. There are 12 low-grading AFs led by Chemistry (-1.05) and nine high-grading ones led by English (0.66). On the other hand, Figure 2 plots the mean UFH effects on PWs by AF. To be consistent (*LL* or *HH*), each AF in Figure 1 should rotate 180 degrees to its right or to its left. That is, the bars' direction in Figure 2 should be the opposite of those of Figure 1. Notwithstanding, such is not the case. Table 5 reports the distribution of AFs by categories, based on their faculty members' mean effects. There are 11 consistent AFs. Six are *LL* (Chemistry, Mathematics, Econ/Stat, Computer Sciences, Office Systems, and Electronics) and five are *HH* (Education, Marketing, Management, Social Sciences, and English). On the other hand, eight AFs behave as *AA* (Physics, Biology, Engineering, Accounting, Finance, Humanities, Spanish, and Education).

Figure 1

Unobservable faculty heterogeneity effects on GPA

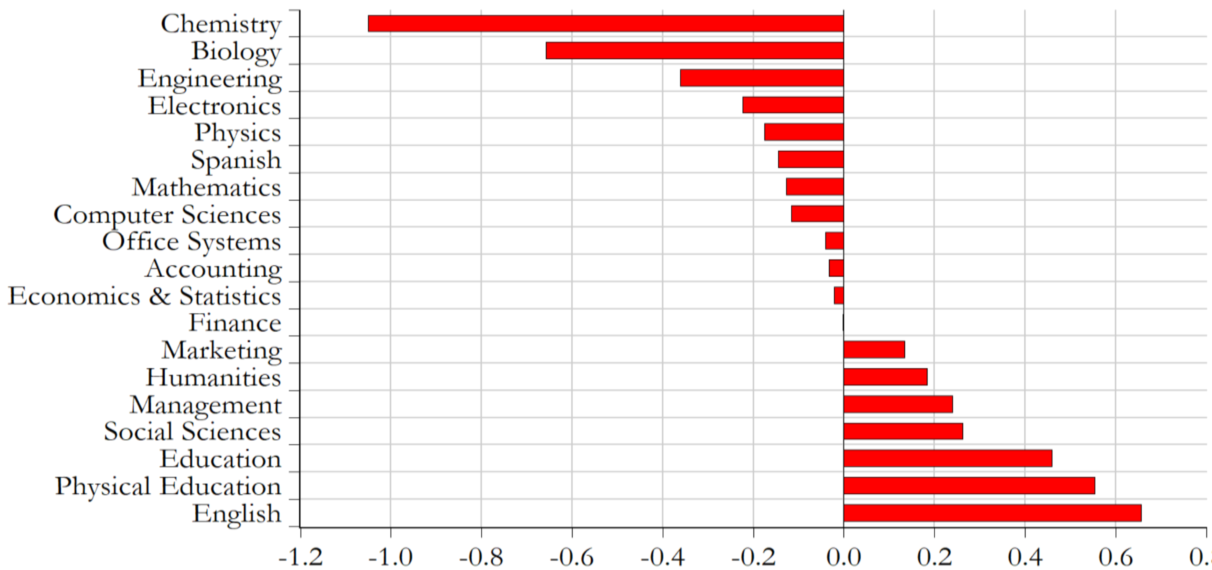


Figure 2

Unobservable faculty heterogeneity effects on PWs

Table 5

Academic Fields Classified by Categories

AA (*LH* or *HL*)	*HH*	*LL*
Physics (*LH*)	Marketing	Mathematics
Biology (*LH*)	Management	Chemistry
Engineering (*LH*)	Social Sciences	Economics & Statistics
Accounting (*LH*)	English	Electronics
Finance (*LH*)	Education	Computer Sciences
Spanish (*LH*)		Office Systems
Humanities (*HL*)		
Physical Education (*HL*)		

Thus, based on its faculty members' mean effects on GPA and PWs, Econ/Stat occupies the third place among the *LL* AFs, preceded only by Mathematics and Chemistry. As previously mentioned, based on their own estimated effects $\left(\hat{\phi}_{\ell}^{GPA},\ \hat{\phi}_{\ell}^{PWs}\right)$, only the AFs of Mathematics and Econ/Stat behave as *LL*. Moreover, among all the 21 AFs, the mean GPA (2.28) and PWs (17.4%) of Econ/Stat are the third lowest and the second highest ones, respectively. Thus, independently of the criteria used, Econ/Stat ranks with the lowest-grading AFs, even lower than Physics, Biology, and Engineering Transfers.

Predicting the Probability that Instructor j Belongs to a Determinate Category

Table 6 reports the results of the estimated multinomial model described in expressions (13) – (15), while Tables 7 and 8 report their estimated marginal effects. The model includes 10 dummies and 5 standardized semi-continuous regressors. Among the 32 (including the constant terms) estimated coefficients reported in Table 6, 10 (31.25%) are statistically significant. For instance, the ln-odds between the categories *HH* and *AA* is 1.78 points higher for faculty members from Cluster 3, and such a coefficient is highly significant. The problem is that the magnitudes of these kinds of coefficients are difficult to interpret.[16] What really matters is the direction (sign) and the magnitudes of the relationship between the dependent variable and the regressors. However, neither the sign of the estimated regression coefficients nor their values shed light on such an issue. In order to circumvent this limitation, the attention centers on their estimated marginal effects, which Table 7 reports.

Partial derivatives from expressions (13) - (15) are cumbersome. After some algebraic manipulations, they simplify to[17]

$$\frac{\partial P\left(Y=0\right)}{\partial X_i}=P\left(Y=0\right)\left\{0-\beta_{1i}\cdot P\left(Y=1\right)-\beta_{2i}\cdot P\left(Y=2\right)\right\} \tag{16}$$

16 For details, refer to Greene (2012, p. 764).
17 For specific details, refer to Greene (2012, pp. 763-766).

$$\frac{\partial P(Y=1)}{\partial X_i} = P(Y=1)\left\{\beta_{1i} - \beta_{1i}\cdot P(Y=1) - \beta_{2i}\cdot P(Y=2)\right\} \tag{17}$$

$$\frac{\partial P(Y=2)}{\partial X_i} = P(Y=2)\left\{\beta_{2i} - \beta_{1i}\cdot P(Y=1) - \beta_{2i}\cdot P(Y=2)\right\} \tag{18}$$

The evaluation of the derivatives occurs at the means of the independent variables, which normally include several dummies. By definition, a partial derivative measures the effect on the dependent variable of an infinitesimal change in a determinate regressor, other things being equal. However, such a procedure is unfeasible in the case of dummies. According to Greene (2012, p. 690), the appropriate marginal effect for a binary independent variable d would be

$$P\left(Y=1 \mid \mathbf{X}, d=1\right) - P\left(Y=1 \mid \mathbf{X}, d=0\right) \tag{19}$$

However, NLOGIT 5 (2012) computes all partial derivatives according to expressions (16) - (18), including those for the dummies. On the other hand, this study aims to measure the effect on the probability of a determinate outcome of changes of one standard deviation above or below the mean of a regressor. Changes of such magnitude are in no way infinitesimals. Thus, Table 8 reports marginal effects for dummies computed according to expression (19), as well as for the semi-continuous standardized regressors. In both instances, the evaluation of the base line model outcomes probabilities occurs at the mean values and at the reference groups. Hence, the base line model includes only the constant terms of the equations reported in Table 6. The estimation of marginal effects uses the difference between the base line model outcomes probabilities and the respective ones when the regressor of interest enters into expressions (13) – (15), one at a time. The discussion that follows rests on results from Table 8, while those of Table 7 will serve as a basis for comparison.

For analytical purposes, it is convenient to classify AFs into four clusters: C1, C2, C3, and C4. C1 consists of Mathematics, Chemistry, Physics, and Econ/Stat AFs. C2 includes the AFs that offer bachelor's degrees: Accounting, Management, Finance, Marketing, Education, Physical Education, Biology, Electronics, Computer Sciences, Materials Management, and Office Systems. C3 consists of liberal arts AFs such as Spanish, English, Humanities, and Social Sciences. Finally, C4, which is the reference group, includes the AFs of Engineering Technologies (except Materials Management) and Engineering Transfers.

Several of the dummies' marginal effects reported in Table 8 are very similar to their Table 7 counterpart, and only four reverse signs. Compared to the reference group (C4), the probability of *AA* decreases, while the probability of *HH* increases if instructor *j* belongs to whichever other three clusters. The fact that the probability of *LL* increases only if instructor *j* belongs to Mathematics, Chemistry, Physics, or Econ/Stat (C1) is interesting. In such a case, the probability increases by 13.41 percentage points (pp). On the other hand, other things being equal, liberal arts (C3) and bachelor's degree AFs (C2) diminish that probability by 13.77 pp and 6.6 pp, respectively.

Female and doctorate covariates exhibit very small marginal effects, which is consistent with results reported in Table 6, where both dummies were insignificant. Compared to instructors, the probability of *AA* and *HH* inversely varies, while the probability of *LL* directly varies with the other academic ranks (assistant, associate, and full professor). Finally, compared to part-time faculty members, the probability of *AA* inversely varies, while the probability of *HH* and *LL* directly vary whenever instructor *j* has a tenure-track or tenured academic status.

Table 6

Multinomial Logit Estimated Probabilities

Regressors	$ln\left\{\frac{P(Y=1)}{P(Y=0)}\right\}$	p-values	$ln\left\{\frac{P(Y=2)}{P(Y=0)}\right\}$	p-values
Constant	-1.0587	0.1196	-0.6087	0.2553
Dummy variables				
Cluster 1	1.1643	0.11	1.1216**	0.0442
Cluster 2	1.0782	0.1135	-0.0562	0.9135
Cluster 3	1.7789***	0.0094	-0.0085	0.9876
Female	-0.1451	0.4995	0.0241	0.9158
Asst Professor	-0.8152*	0.0645	0.3946	0.3374
Assoc Professor	-0.1383	0.8099	0.4882	0.4098
Professor	-0.9273	0.1242	0.4186	0.4904
Doctorate	0.1334	0.6692	0.0401	0.9015
Probation	1.0597**	0.0195	0.831*	0.0613
Tenured	0.3066	0.5424	0.3221	0.513
Semi-continuous standardized regressors				
Age	0.4905***	0.0000	-0.3207**	0.0172
GAI	-0.2579*	0.0636	-0.1521	0.2982
GAI Variance	0.1084	0.2028	-1.2374***	0.0036
PHSSP	0.108	0.3527	-0.0512	0.6591
FSP	-0.1365	0.2732	-0.3112**	0.0107
Log likelihood function: -675.85				
McFadden Pseudo R-squared: 0.1172				
Sample size: 700				

Notes: ***, **, * = statistically significant at 1%, 5%, and 10% level, respectively; $Y = 0 = AA$; $Y = 1 = HH$; $Y = 2 = LL$; PHSSP = private high school students proportion; FSP = female students proportion; Cluster 1 includes the following AFs: Mathematics, Chemistry, Physics, as well as Econ/Stat. Cluster 2 includes the AFs that offer bachelor's degrees: Accounting, Management, Finance, Marketing, Education, Physical Education, Biology, Electronics, Computer Sciences, Materials Management, and Office Systems. Cluster 3 consists of liberal arts AFs such as Spanish, English, Humanities, and Social Sciences. Finally, Cluster 4 consists of Engineering Technologies (except Materials Management), and Engineering Transfers AFs.

Among the semi-continuous standardized regressors, only the private high school students' proportion (PHSSP) was insignificant in both estimated equations reported in Table 6. Consistent with this result, their marginal effects are in the neighborhood of zero in Tables 7 and 8. Conversely, the other four regressors were significant at least in one of the equations. The Age covariate, which is highly significant in both equations, exhibits an interesting behavior around its mean, consistent with relative extrema. Changes of one standard deviation to the right or to the left of its mean tend to decrease the probability of *LL* by 8.64 pp or 3.46 pp, respectively. Meanwhile, the probability of *HH* will increase by 10.54 pp or 22.58 pp because of those changes. On the other hand, the probability of *LL* inversely varies with the female students proportion (FSP). Increases or decreases of one standard deviation in FSP tend to decrease or increase the probability of such a category by 5.34 pp or 5.93 pp, respectively. However, their marginal effects on the probability of *HH* are practically zero.

Table 7

Marginal Effects

Regressors	$P(Y=0=AA)$	$P(Y=1=HH)$	$P(Y=2=LL)$	Total
		Dummy variables		
Cluster 1	-0.2547**	0.1534	0.1013	0
Cluster 2	-0.1324	0.261*	-0.1286	0
Cluster 3	-0.2264	0.4214***	-0.195*	0
Female	0.0163	-0.0369	0.0207	0
Assistant Professor	0.0669	-0.2356***	0.1687**	0
Associate Professor	-0.0284	-0.0857	0.1142	0
Professor	0.0789	-0.2647**	0.1858*	0
Doctorate	-0.0208	0.0272	-0.0064	0
Probation	-0.2139**	0.1602*	0.0536	0
Tenured	-0.0696	0.0375	0.0321	0
		Semi-continuous standardized regressors		
Age	-.0324	0.1508***	-0.1184***	0
GAI	0.0473*	-0.0444	-0.0029	0
GAI Variance	0.103**	0.1601***	-0.2631***	0
PHSSP	-0.009	0.0311	-0.0221	0
FSP	0.0468*	0.0015	-0.0484**	0

Notes: ***, **, * = statistically significant at 1%, 5%, and 10% level, respectively. Marginal effects were estimated by NLOGIT 5.

Considering the student academic ability proxies, the most relevant one is GAI because it constitutes the institutional admission policy criterion. Therefore, it is expected that both GAI and GAI variance exert a significant effect on the probabilities of the categories under analysis. According to the marginal effects reported in Table 7, GAI is significantly related only to the probability of *AA*, which estimated coefficient is approximately 0.05. This value is equal to those reported in Table 8 for that category (*AA*) in response to changes of one standard deviation around the GAI mean.

On the other hand, the GAI variance regressor is highly significant in the second equation reported in Table 6 but insignificant in the first one. All marginal effects reported in Table 7 (by NLOGIT 5) for this covariate are highly significant. Therefore, the pattern of signs exhibited by their estimated coefficients needs further explanation. The heterogeneity of student academic ability, proxied by this covariate, might have different effects on the probability of the categories depending on the instructor's attitude toward risk (Matos-Díaz & Ragan, 2010). For instance, faced with courses of highly heterogeneous students, a risk-averse instructor would relax the academic standards to allow students belonging to the lower bound of the academic ability distribution to exceed threshold GPA values that induce them to remain in the course. Thus, relaxing academic standards would increase GPA and decrease Ws, improving the distribution of course grades. Under such a scenario, it is expected that the probability of *HH* increases while the probability of *LL* decreases whenever GAI variance increases. Evidence points to the conclusion that this is the case prevailing at UPR-Bayamón. According to Table 8, an increase of one standard deviation in GAI variance increases the probability of *HH* by 6.68 pp and decreases the probability of *LL* by 18.55 pp. Furthermore, the diminishing of one standard deviation in such a regressor decreases and increases the probability of *HH* and *LL* by 8.58 pp and 30.08 pp, respectively. Therefore, faculty members are very responsive to the signals sent by students' characteristics. That is, faculty members and students free-ride from each other altering institutional norms and academic standards. This empirical result has profound academic policy implications given that, to the extent that student academic ability heterogeneity increases in different AFs, academic standards will tend to diminish significantly, which in turn implies the provision of an irrelevant education. If such is the case prevailing in AFs like Education and Physical Education discussed earlier, their academic and economic policy implications would be of considerable social concerns.

Table 8

Re-estimating Marginal Effects Through Simulations

	Estimated probabilities			Changes in estimated probabilities		
Variables	$P(Y=0)$	$P(Y=1)$	$P(Y=2)$	$P(Y=0)$	$P(Y=1)$	$P(Y=2)$
Dummy variables						
Cluster 1	0.2738	0.3043	0.4219	-0.255	0.1211	0.1341
Cluster 2	0.3853	0.3929	0.2218	-0.1435	0.2094	-0.066
Cluster 3	0.2782	0.5717	0.1501	-0.2506	0.3882	-0.1377
Female	0.5384	0.1615	0.3	0.0096	-0.022	-0.0001
Assistant Professor	0.51	0.0783	0.4117	-0.0188	-0.1052	0.124
Associate Professor	0.4569	0.138	0.405	-0.0719	-0.0455	0.1173
Professor	0.5288	0.1835	0.2877	-0.0197	-0.1136	0.1333
Doctorate	0.5095	0.202	0.2885	-0.0193	0.0185	0.0008
Probation	0.3077	0.308	0.3843	-0.2211	0.1245	0.0966
Tenured	0.45	0.2121	0.3379	-0.0788	0.0286	0.0502
Semi-continuous standardized variables						
Age						
$Z=1$	0.5099	0.2889	0.2013	-0.0189	0.1054	-0.0864
$Z=-1$	0.376	0.4093	0.3531	-0.1912	0.2258	-0.0346
GAI						
$Z=1$	0.5763	0.1545	0.2639	0.0475	-0.029	-0.0184
$Z=-1$	0.4802	0.2156	0.3042	-0.0486	0.0321	0.0165
GAI Variance						
$Z=1$	0.6475	0.2503	0.1022	0.1187	0.0668	-0.1855
$Z=-1$	0.3138	0.0977	0.5885	-0.215	-0.0858	0.3008
PHSSP = private high school students proportion						
$Z=1$	0.5254	0.203	0.2716	-0.0034	0.0195	-0.0161
$Z=-1$	0.5308	0.1653	0.3039	0.002	-0.0182	0.0162
FSP = female students proportion						
$Z=1$	0.5878	0.1779	0.2343	0.059	-0.0056	-0.0534
$Z=-1$	0.4672	0.1858	0.347	-0.0616	0.0023	0.0593
Base line model	0.5288	0.1835	0.2877	N/A	N/A	N/A

Notes: The base line model probabilities are estimated at the mean of the continuous variables and at the dummies reference group. Changes in estimated probabilities are computed with respect to the base line model values. N/A = do not apply. Per row sums of the estimated probabilities and their estimated changes are equal to one and zero, respectively. Z = standardized random variable, implying that $Z=\left(X_i-\bar{X}\right)/\sigma \Rightarrow E(Z)=0$ and $\sigma^2(Z)=1$.

Policy Implications on Institutional Norms and Academic Standards as Public Goods

The grading practices adopted by faculty members in their corresponding AFs over time have profound and significant impact on institutional norms and academic standards. On the other hand, norms and academic standards provide invaluable information with regard to reputation, scholastic achievements, and prestige reached by the university. According to Marks (2002) and Johnes (2004), university norms and academic standards share the two fundamental characteristics that distinguish public goods: nonexclusive and nonrival.

A good is nonexclusive if it is impossible, or extremely costly, to exclude individuals from benefiting from its consumption. National defense and herd immunity coming from the inoculation against disease programs are examples. A good is nonrival if the social marginal cost of its provision to an additional consumer is zero. The use of a highway during a period of low traffic volume, one more viewer tuned in to a public television channel, as well as the use of a lighthouse by a ship are commonly used textbook examples. Given these two properties, public goods are subject to free-riding behavior, i.e., consumers will not reveal their willingness to pay for the good preferring to rely on others to pay for it.

Both faculty members and students have an incentive to cheat out of institutional norms and academic standards. For instance, even if they belong to a strict grading department, individual faculty members seeking better SET ratings and/or more convenient teaching schedules have incentives for grading leniently, free-riding on the grading norms that they expect their peers will support (Johnes, 2004). It is also possible that other faculty members free-ride through benefiting from the prestige and recognition of the academic unit where they work without contributing to (paying for) it through their research and teaching achievements. On the other hand, seeking or accepting undeserved grades and committing plagiarism are clear examples of students' free-riding behavior.

A cautionary note is in order here. Suppose an academic environment is characterized by continuous decline in the growth of student enrollments, external accountability pressures, as well as an increasing tendency for defining and measuring 'good teaching' through SET ratings (McKenzie & Staaf, 1974). If so, it can be hypothesized (Correa, 2001, Matos-Díaz, 2012) that faculty members (academic departments) will compete not only for favorable SET ratings but also for enrollment in the courses they teach. The reasons might vary. They could be interested in recruiting, retaining, and graduating more students with certain traits, characteristics, and academic abilities from their programs, or they might be motivated to obtain convenient teaching schedules, or simply because more enrolled students implies greater budget assignments. Additionally, suppose that there are only two academic departments that compete for attracting students using the strategies of "increase grade distribution" or "maintain grade distribution." Using a simple game-theoretic model, McKenzie & Staaf (1974) convincingly show that both academic departments will be trapped into the prisoner's dilemma, i.e., they end up in the worst cell of the payoff matrix since the dominant strategy will be to "increase grade distribution." Therefore, each department increases grades distribution, but neither increases its enrollment. To the extent that the increase in grades will not be accompanied by a concomitant increase in students' academic ability, the process will give rise to the phenomenon of grade inflation. In such an academic environment, a strict grader professor (who does not inflate grades) can free-ride and take advantage of his colleagues' grade inflation, benefitting from the department's increased budget (McKenzie & Staaf, 1974, p. 61). However, in this particular situation the professor's free-riding behavior benefits the university academic standards and the society. Therefore, faculty members and university administrators do not live in an ivory tower free of market pressures. They respond to economic incentives that can induce them to modify the established norms and academic standards. Thus, the nature of the institutional norms and academic standards as "good" or "bad" should be clearly specified and subject to continuous revisions.

It has been established that free-riding behavior by students and faculty members will depend on consent or complicity with each other (Marks, 2002). For instance, students reward high-grading faculty members through high SET ratings; faculty members reward non-hardworking students by awarding them with undeserved high grades. It should be pointed out that free-riding behavior is most easily identified and prevented when groups are smaller. However, the smaller the course, the greater its cost, making the solution of the problem very costly to the university.

Therefore, free-riding behavior is at odds with university norms and academic standards. It should be emphasized that the establishment of norms and academic standards at universities is a hard and time-intensive production process; however, it is very fragile and easy to destroy. Their preservation through time is in the best interests of the institution.

Summary and Conclusions

The paper proposes and develops a taxonomy to analyze the grading practices of the faculty members and the AFs defined at UPR-Bayamón during 41 consecutive terms, from fall 1995-96 to fall 2015-16. To this end, the estimation of two models using the same set of regressors and specifications allows the accurate prediction of the GPA and PWs observed in each of the 39,337 total courses offered in the institution during 41 consecutive terms. The adoption of different econometric techniques that capture their effects on the intercept of these equations permits the classification of faculty members and their AFs into four mutually exclusive and collectively exhaustive categories scaled from low- to high-grading.

Evidence points to the existence of a self-sorting process that matches academically ablest students to low-grading and more difficult content AFs and less academically able students to high-grading and less difficult content courses. The incentive mechanism guiding such a process depends on the distributions of GPA and PWs prevailing in each AF.

Evidence points to the conclusion that the students enrolled in those AFs portrayed as less competitive, recruited through the attractive incentive of higher grades at a very low price, effectively obtain significantly greater GPA and lower PWs.[18] On the contrary, those AFs portrayed as more competitive and which recruited the ablest students, ended up exhibiting the lowest GPA and the highest PWs. These contradictory results might be plausibly explained to the extent that academic standards have decreased in the first group of AFs and have remained approximately constant or have even increased in the second. If so, the human capital stock of the students from the first group will suffer from fast obsolescence rates across time, while the respective stock of the students from the second group will be enhanced through the addition of relevant, pertinent, and updated skills and knowledge. To the extent that the labor market and the graduate schools can distinguish among the students of each group of AFs, the possibility for the university to play its role as a promoter of social mobility and greater equity in the social income distribution significantly diminishes, at least for the students of the first group.

---

[18] Therefore, the *LL* category is likely even more rigorous than the results suggest. With a high number of withdrawals, the students left are among the strongest within a difficult discipline. Hence, low-grading in this cohort is much tougher, relatively speaking, that it is for the three other groupings. This point was brought to my attention by Dennis L. Weisman, to whom I am very grateful.

Acknowledgements

For helpful comments and suggestions, I am indebted to Panagiotis (Peter) Tsigaris, Nélida Matos-Díaz, Dennis L. Weisman, Manuel Frau-Ramos, Dwight García-Vázquez, Suphan Andic, Manuel Junco de la Fuente, Brenda Domínguez-Rosado, Luis A. Avilés, Myrna L. Torres-Pérez, José La Luz-Concepción, Arturo Avilés-González, Javier Zavala-Quiñones, Gabriel Rodríguez-Matos, Jorge F. Rovira-Álvarez, and Nellie J. Sieller. I am also deeply grateful to two anonymous referees for their comments and suggestions, as well as to Molly Ott and Stephanie McBride-Schreiner for their guidance and support throughout the publishing process. Any remaining errors are my sole responsibility.

References

Aaronson D., Barrow L., & Sander W. (2007). Teachers and students achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95-135. https://www.jstor.org/stable/10.1086/508733

Achen, A. C., & Courant, P. N. (2009). What are grades made of? *Journal of Economic Perspectives*, *23*(3), 77-92. https://doi.org/10.1257/jep.23.3.77

Ahn, T., Arcidiacono, P., Hopson, A., & Thomas, J.R. (2019). *Equilibrium grade inflation with implications for female interest on STEM major*. [Working paper 26556]. National Bureau of Economic Research. http://www.nber.org/papers/w26556

Arcidiacono, P. (2004). Ability sorting and returns to college major *The Journal of Econometrics*, *121*(1-2), 343-375. https://doi.org/10.1016/j.jeconom.2003.10.010

Babcock, P. (2010). Real costs of nominal grade inflation? New evidence from student course evaluations. *Economic Inquiry*, *48*, 983-996. https://doi.org/10.1111/j.1465-7295.2009.00245.x

Brastberg, B., & Rogeberg, O. (2018). Flinn effect and its reversal are both environmentally caused. *PNAS*, *115*(26), 6674-6678. https://doi.org/10.1073/pnas.17187931155

Butcher, K. F., McEwan, P. J., & Weerapana, A. (2014). The effects of an anti-grade-inflation policy at Wellesley College. *Journal of Economic Perspectives*, *28*(3), 189-204. http://dx.doi.org/10.1257/jep.28.3.189

Correa, H. (2001). A game theoretic analysis of faculty competition and academic standards. *Higher Education Policy*, *14*, 175-182.

Greene, W. H. (2012). *Econometric analysis*. Prentice Hall.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*, 153-161.

Hosmer, D. W., & Lemeshow, S. (2000) *Applied logistic regression*. John Wiley and Sons.

Hoyt, D., & Reed J. (1976). *Grade inflation at KSU: Student and faculty perspectives*. Research Report No. 36. June 1976. (ERIC Document Reproduction Service No. ED160018).

Johnes, G. (2004). Standards and grade inflation. In G. Johnes & J. Johnes (Eds.), *International handbook on the economics of education* (pp. 462-483). Edward Elgar.

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

Keng, S. H. (2018). Tenure system and its impact on grading leniency, teaching effectiveness and student effort. *Empirical Economics*, *55*(3), 1207-1227. https://doi.org/10.1007/s00181-017-1313-7

Koedel, C. (2008). Teacher quality and dropout outcomes in a large, urban school district, *Journal of Urban Economics*, *64*, 560-572. https://doi.org/10.1016/j.jue.2008.06.004

Koedel, C. (2011). Grading standards in education departments at universities. *Education Policy*

Analysis Archives, 19, 1-23. https://doi.org/10.14507/epaa.v19n23.2011

McKenzie, R. B., & Staaf. R. J. (1974). *An economic theory of learning: Students sovereignty and academic freedom*. [Report]. Virginia Polytechnic Institute and State University.

Marks, D. (2002). Academic standards as public goods and varieties of free-rider behaviour. *Education Economics*, *10*(2), 145-163.

Matos-Díaz, H. (2018). Estimating the effects of partial withdrawals on GPA through time: Evidence from the University of Puerto Rico. *EconomiA*, *19*, 90-104. https://doi.org/10.1016/j.econ.2017.10.004

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*(3), 296-309. https://www.jstor.org/stable/23271209

Matos-Díaz, H., & Ragan, J. (2010). Do student evaluations of teaching depend on the distribution of expected grade? *Education Economics*, *18*(3), 317–330. https://doi.org/10.1080/09645290903109444

McSpirit, S., & Jones, K. E. (1999). Grade inflation rates among different ability students, controlling for other factors. *Education Policy Analysis Archives*, *7*, 1-16. https://doi.org/10.14507/epaa.v7n30.1999

NLOGIT 5. (2012). *NLOGIT 5*. [Computer software]. Econometric Software, Inc.

Nordin, M., Heckley, G., & Gerdtham, Ulf-G. (2019). *The impact of grade inflation on higher education enrolment and earnings* [Working paper 2019:1]. Department of Economics, Lund University, Sweden.

Oglive, K., & Jelavic, M. (2013). Grade inflation in the US higher educational environment: A faculty perception study. *International Journal of Management in Education*, *7*, 406-416.

Remmers, H. H. (1928). The relationship between students' marks and student attitude toward instructor. *School and Society*, 759-760.

Remmers, H. H. (1930). To what extent do grades influence student ratings of instructors? *Journal of Educational Research*, *21*, 314-316.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247-252. https://www.aeaweb.org/articles?id=10.1257/0002828041302244

Rosovsky H., & Hartley, M. (2002). *Evaluation and the academy: Are we doing the right thing? Grade inflation and letters of recommendation*. American Academy of Arts and Sciences.

Sabot, R., & Wakeman-Linn, J. (1991). Grade inflation and course choice. *Journal of Economic Perspectives*, *5*(1), 159-170. https://www.aeaweb.org/articles?id=10.1257/jep.5.1.159

Stinebrickner, R., & Stinebrickner, T. R. (2008). The causal effect of studying on academic Performance. *The B. E. Journal of Economic Analysis and Policy*, *8*, 1-53. https://doi.org/10.2202/1935-1682.1868

Watts, L. K., & Winters, R. (2016). Examining grade inflation and considerations for Radiologic Sciences: A literature review. *Journal of Medical Imaging and Radiation Sciences*, *48*(1), 95-102. https://doi.org/10.1016/j.jmir.2016.08.002

Weiss, R. M., & Rasmussen, G. (1960). Grading practices in undergraduate education courses: Are the standards too low? *The Journal of Higher Education*, *31*, 143-149. https://doi.org/10.1080/00221546.1960.11777549

Appendix

Table A1

Descriptive Statistics (panel data)

Dummy variables					
Variable	Mean	Variable	Mean	Variable	Mean
Accounting	0.0539 (0.2259)	Marketing	0.0125 (0.111)	Probation	0.0854 (0.2794)
Biology	0.0515 (0.221)	Materials Management	0.0075 (0.0863)	Tenure	0.6155 (0.4865)
Chemistry	0.0349 (0.1836)	Mathematics	0.1301 (0.3364)	Class size 1	0.0598 (0.2372)
Computer Sciences	0.0633 (0.2436)	Physical Education	0.0375 (0.19)	Class size 3	0.2773 (0.4477)
Econ/Stat	0.0277 (0.164)	Physics	0.0294 (0.1689)	Morning	0.541 (0.4983)
Education	0.0595 (0.2366)	Office Systems	0.0381 (0.1914)	Night	0.0927 (0.29)
Electronic	0.0587 (0.2351)	Social Sciences	0.0615 (0.2403)	Summer	0.0156 (0.1241)
Engineering Technologies	0.0285 (0.1663)	Spanish	0.0682 (0.2521)	SET	0.1465 (0.3536)
Engineering Transfers	0.01 (0.0994)	Female faculty	0.5285 (0.4992)	Monday to Friday	0.017 (0.1292)
English	0.0994 (0.2992)	Doctorate	0.2545 (0.4356)	Monday & Wednesday	0.3181 (0.4658)
Finance	0.0183 (0.1341)	Assistant	0.2256 (0.418)	Tuesday & Thursday	0.3671 (0.482)
Humanities	0.0758 (0.2647)	Associate	0.2149 (0.4107)	M/T/W	0.1625 (0.3689)
Management	0.0337 (0.1804)	Professor	0.2269 (0.4188)	M/T/W/Th	0.01 (0.0777)

Semi-continuous variables					
Variable	Description	Mean	Std. Dev.	Min	Max
PWs	proportion of course withdrawals	10.89	12.72	0	92.86
GPA	grade point average	2.72	0.7169	0	4.00
Age	professor's age (years)	47.63	9.64	23	76
GAI	General Application Index	285	23.03	174	372
GAI Variance	GAI distribution variance	795.2	686	0	20,031
FSP	female students proportion	0.5329	0.2662	0	1
PHSSP	private HS proportion	0.4722	0.1622	0	1

Note: For all dummies, standard deviations are reported in parentheses, Max = 1 and Min. = 0.

Table A2

Academic Fields by Academic Departments at UPR-Bayamón

Academic Departments	Commentaries…	Academic Fields
1) Business Administration	Accredited by the Accreditation Council for Business Schools and Programs (ACBSP) since 2009, the departments grants bachelor's degree (BBA) in Accounting, Finance, Management, and Marketing. Furthermore, it provides service courses for other academic departments. The department offers two sequencies of two courses of Economics and Statistics that are required for all admitted students. These four courses comprise the cluster referred to as Econ/Stat.	1) Accounting 2) Finance 3) Management 4) Marketing 5) Econ/Stat
2) Adapted Physical Education	Accredited by the National Council for Accreditation of Teacher Education (NCATE), the department grants the bachelor's degree in Special and Elementary Physical Education.	6) Physical Education
3) Pedagogy	Accredited by the *NCATE*, the department grants the bachelor's degree in Preschool and Elementary Education.	7) Education
4) Electronics	Accredited by the Engineering Technology Accreditation Commission (ABET), the department grants bachelor's and associated degrees in Electronics.	8) Electronics
5) Computer Sciences	Accredited by the Computing Accreditation Commission of ABET, the department grants the bachelor's degree in Computer Sciences.	9) Computer Sciences
6) Biology	The department grants the bachelor's degree in General Biology, as well as in Biology with a Human Approach.	10) Biology
7) Office Systems	Accredited by the *Accreditation Council for Business Schools and Programs, the department grants the bachelor's degree in Office Systems.*	11) Office Systems

Table A2 (cont'd.)

Academic Fields by Academic Departments at UPR-Bayamón

Academic Departments	Commentaries…	Academic Fields
8) Engineering and Engineering Technologies	Accredited by the Engineering Technology Accreditation Commission (ABET), the department grants the bachelor's degree in Materials Management, as well as associate degrees in the engineering technologies. Furthermore, it offers an Articulated Transfer Program to the College of Engineering at UPR-Mayagüez. New entrant students are admitted to the following academic programs that offer associate degrees: Civil Engineering Technology, Construction, Surveying and Highway Engineering Technology, and Industrial Engineering Technology.	12) Materials Management 13) Articulated Engineering Transfer Program 14) Engineering Technologies
9) Social Sciences	The department offers transfer programs to other UPR Campuses, as well as service courses to all academic departments (AD) at UPR-Bayamón.	15) Social Sciences
10) Humanities	The department offers transfer programs to other UPR campuses, as well as service courses to all AD at UPR-Bayamón.	16) Humanities
11) English	The department offers service courses to all AD at UPR-Bayamón.	17) English
12) Spanish	The department offers service courses to all AD at UPR-Bayamón.	18) Spanish
13) Chemistry	The department offers two articulated transfer programs to the campuses of UPR-Cayey (bachelor's degree in Chemistry), and UPR-Medical Sciences (bachelor's degree in Nursing). Furthermore, it offers service courses to all AD at UPR-Bayamón.	19) Chemistry
14) Physics	The department offers service courses to all AD at UPR-Bayamón.	20) Physics
15) Mathematics	The department offers service courses to all AD at UPR-Bayamón.	21) Mathematics

Figure A1

GPA through time under different Ws patterns at UPR-Bayamón

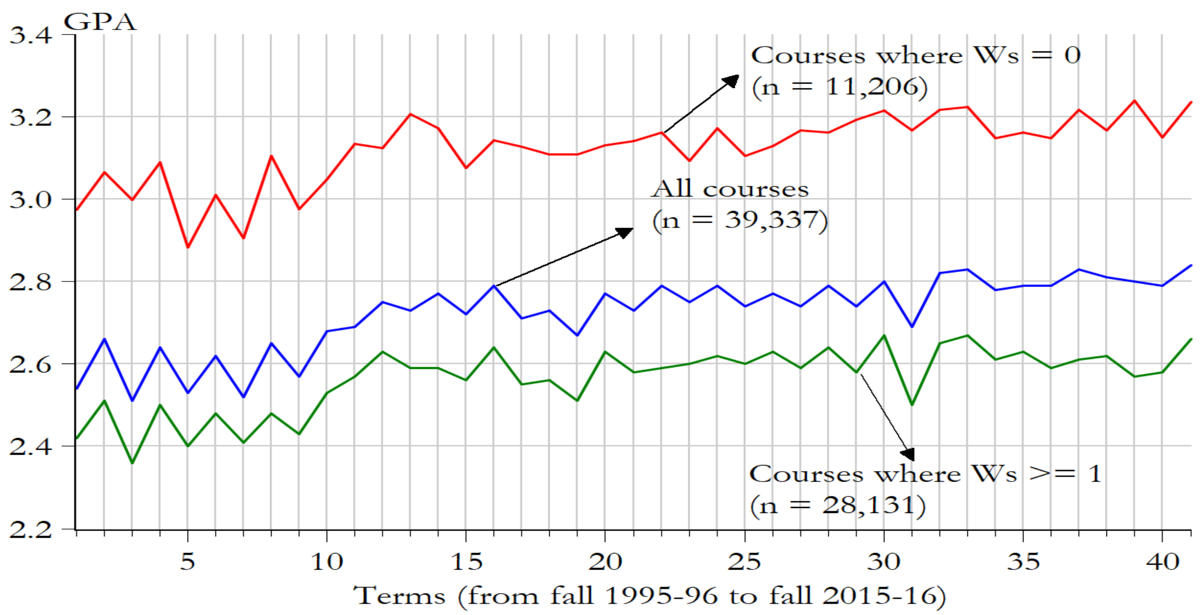


Figure A2

Triplets of GPA by academic fields

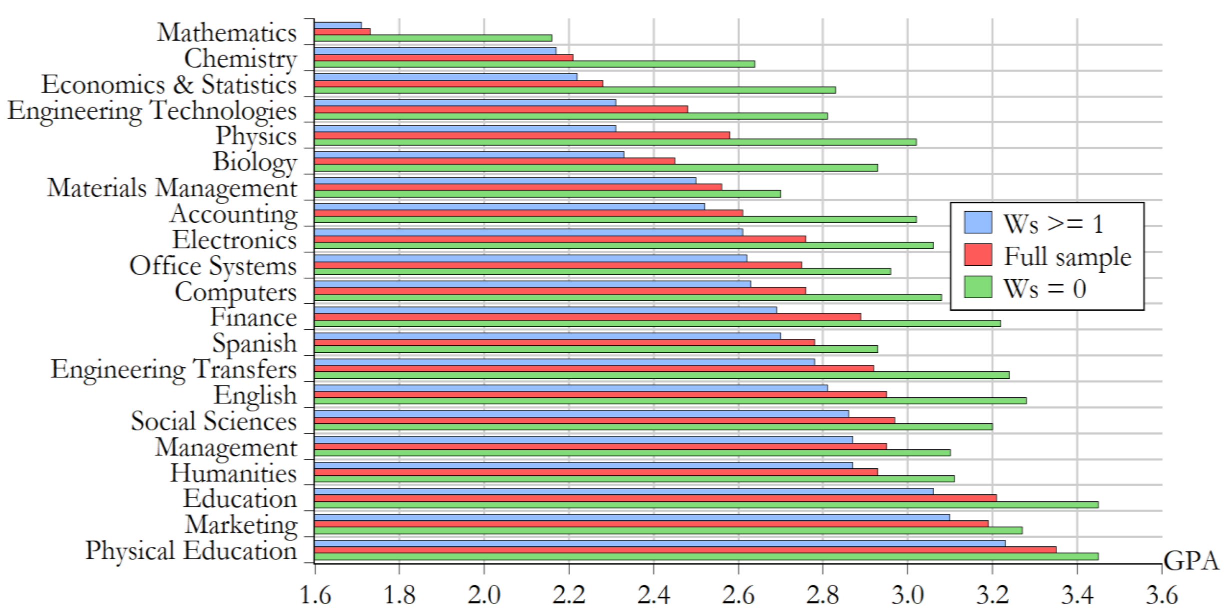


Figure A3

Student academic ability proxies through time

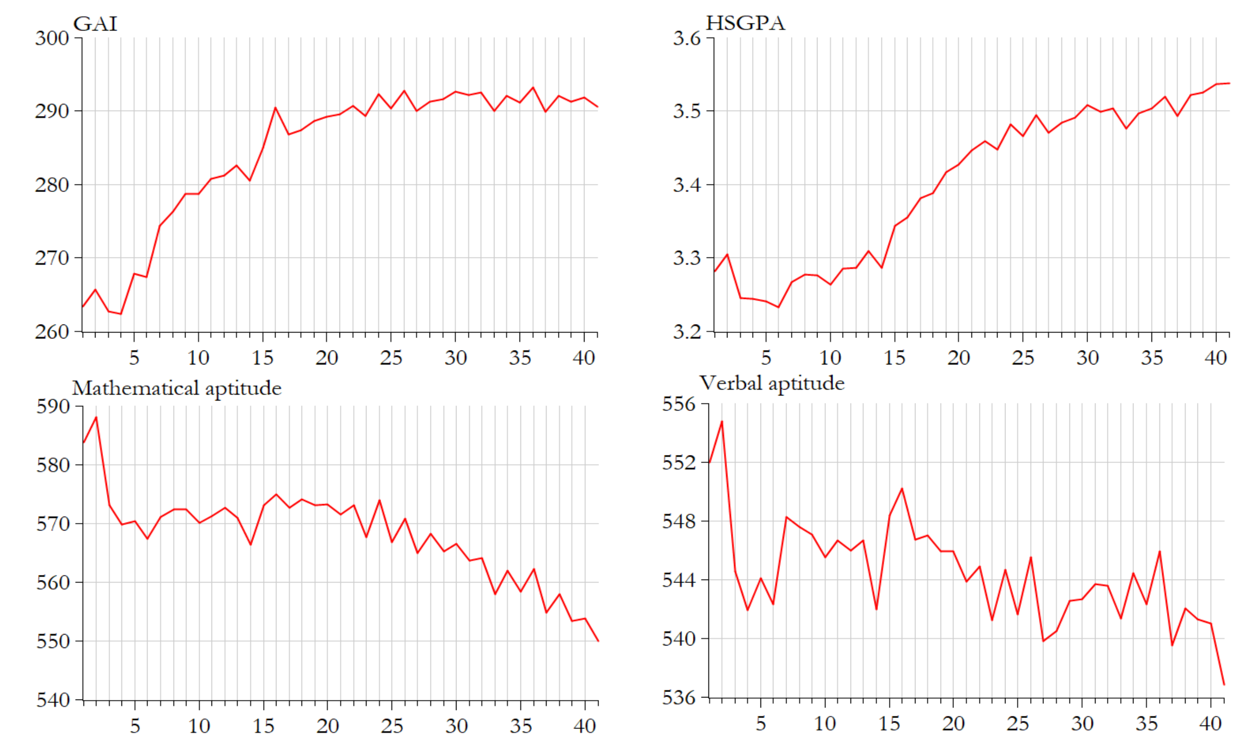
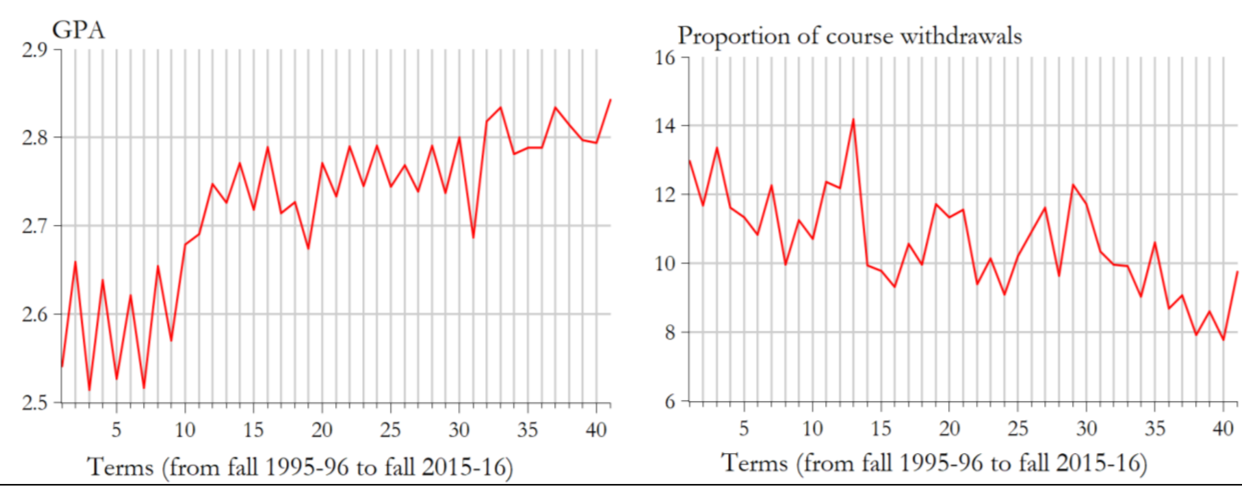


Figure A4

GPA and PWs through time



About the Author

Horacio Matos-Díaz
University of Puerto Rico at Bayamón
horacio.matos1@upr.edu
Horacio Matos-Díaz is a Professor of Economics and Statistics at the University of Puerto Rico at Bayamón, where he has taught and conducted research during the last 43 years. He obtained the B.A. (1976) and M.A. (1978) degrees in Economics from the University of Puerto Rico at Río Piedras Campus, while his Ph.D. (2000) in Economics was granted by Kansas State University. His research has been published in journals such as *Economics of Education Review, Education Economics, Eastern Economic Journal, Cogent Economics & Finance, EconomiA*, this journal (EPAA), as well as in several Spanish journals from Puerto Rico and abroad.

education policy analysis archives

Volume 29 Number 57 May 3, 2021 ISSN 1068-2341