Grades as Noisy Signals*

Brandon Joel Tan†

May 2020

Abstract

I study the consequences of letter grades serving as noisy measures of academic achievement using data from the National University of Singapore. I exploit a regression-discontinuity design with marks as the running variable and find that receiving a better grade in a single class results in $32 greater monthly earnings post-graduation. The effects are larger than expected from a corresponding cumulative grade point average increase via "employer-signaling", suggesting that future changes in behavior and outcomes may be important. I find that marginal students who receive a worse grade take significantly "easier" courses and earn lower grades in future semesters.

*I thank Emily Breza, Michael Droste, Edward Glaeser, Nathan Hendren, Ronak Jain, Lawrence Katz, Amanda Pallais, Jessica Pan, Yin Wei Soon, and workshop participants at Harvard University for their helpful comments and suggestions. Any errors are my own.

†Department of Economics, Harvard University, btan@g.harvard.edu

1 Introduction

Across all levels of education from elementary school to university, students are evaluated on letter grade scales. However, letter grades are noisy and blunt measures of academic achievement. Receiving a 'B' represents a wide range of points-based assignments and does not fully reflect the extent to which a particular student has (or has not) learned or performed in a given class. These letter grades nonetheless serve as important signals both to employers and to the student of his or her ability. I study the consequences of letter grades as noisy signals using administrative data from the National University of Singapore (NUS) which records both the letter grades as well as the precise marks (0-100) received for each course that a given student takes.

I exploit a regression discontinuity design – specifically, close to the letter grade cutoffs, individuals with very similar achievement will receive different letter grades. Almost all undergraduate modules at NUS have a sit-down final exam component and the grades are curved. In other words, a student's grade depends on his or her relative standing in the class, hence it is implausible that students can perfectly control their precise marks and hence letter grades. Additionally, once the grades are released, rare grade changes only happen under exceptional circumstances and must be justified under scrutiny at multiple levels (e.g. department, faculty, registration office etc). I find that observable characteristics such as race, gender and age, or measures of ability such as cumulative grade point average (CAP) prior to the class and student undergraduate admission scores (UAS) which determine one's entry into the university do not exhibit jumps at the letter grade thresholds. I also do not find significant bunching at the letter grade thresholds.

Receiving a better grade for students on the margin results in 45 dollars SGD (~\$32 USD) greater monthly earnings post-graduation. The direction of this result is unsurprising. Employers use grades as a signal of ability and as a result pay similar students different salaries based on this coarse measure ("employer-signaling"). However, the size of the effect is much larger than expected from a corresponding cumulative grade point average increase based on an estimated earnings function on grades and other observables. This suggests that future changes in behavior and outcomes may be important in explaining the large effects on earnings.

I find that future behavior and outcomes are affected by letter grades on the margin. Receiving a worse grade results in lower grades in future semesters. Additionally, these students take significantly "easier courses" in future semesters. Course levels range from 1 to 4, where level 1 courses are the most introductory and level 4 courses are the most advanced. Receiving a worse grade for students on the margin results in taking classes that are on average 0.2 standard deviations lower in course level. This suggests that students invest less in human capital accumulation as a result of receiving a noisy negative signal of ability. These results are consistent with a theory of "self-signalling" where students interpret better or worse grades as a signal of their own ability which affects their future behavior and outcomes. I find no evidence that the effects are driven by convexity in the earnings function, making it optimal for students to put in less effort or take easier courses in response to a bad grade – but do not rule it out definitively.

The point estimate returns in earnings from getting a better grade are the largest in year 1 ($69 SGD / $49 USD), followed by year 2 ($54 SGD / $39 USD), where there is greater margin (more future university years) for behavior change. Looking at each letter grade cut off, I find that the effect is largest at the A- cutoff ($82 SGD / $59 USD), followed by the A ($73 SGD / $52 USD) and B+ ($70 SGD / $50 USD) cutoffs. This suggests that the signal for being an "A" student is particularly important. There are null effects on wages for elective courses, which indicates that only signals for major-relevant courses may be important. However, my sample is not sufficiently large to reject homogeneity in effect sizes across these analyses.

Psychologists and sociologists have emphasized that the student self-conceptions about their ability can have important consequences for education (Brookover, Thomas, and Paterson 1964; Crocker et al. (2003); Shen and Pedulla 2000). B. R. Clark (1960) describes a “cooling out” process in which college students encounter obstacles which lead them to gradually lower their goals and aspirations. There is also a literature on discouragement in education within economics (Bettinger and Long 2009; Calcagno and Long 2008; Scott-Clayton and Rodriguez 2015). Martorell and McFarlin (2011) find evidence on the presence of discouragement or stigma effects from assignment to remediation. R. Stinebrickner and T. R. Stinebrickner (2014) show that students who perform poorly update beliefs about their ability to perform and link this learning to drop out.

This paper is directly related to the literature on the signaling value of education in the labor market, specifically building research that looks at credential effects with similar quasi experimental designs (Cameron and Heckman 1993; T. J. Kane et al. 1995; M. Kane, Crooks, and Cohen 1999; Tyler, Murnane, and Willett 2000; D. Clark and Martorell 2014; Feng and Graetz 2017). This literature reports small or null results from signaling in the labor market. My results show that the behavioral implications of receiving a bad grade on the margin may be more important. This paper is also related to Jacob and Lefgren (2009) and Manacorda (2012) which study the impact of receiving a failing grade in middle/secondary school, and find increased drop-out and lower educational attainment. There is also work exploring the effects of grades in specific settings (standardized testing, introductory courses etc.) on short-run educational attainment (Papay, Murnane, and Willett 2016; Smith, Hurwitz, and Avery 2017; Mcewan et al. 2019). To my knowledge, this paper is the first to study the effect of receiving higher letter grades for classes on labor market outcomes. The only other paper which includes earnings as an outcome is Diamond and Persson (2016) which explores the different but related topic of teacher discretionary grading in a unique pre-high school standardized testing context where schools compete with each other on scores and have incentives to inflate grades. My paper also uniquely studies grades and course taking at the university level which produce the most relevant skills and are the most relevant signals for the labor market.

This paper provides new empirical evidence that letter grades introduce distortions in wages and student behavior via coarse signalling. These distortions should be considered in policy discussions on whether educational institutions should move away from letter grade-based grading systems and towards point-based ones, or vice versa. Additionally, the results from this paper show that the behavioral responses to receiving a lower grade on the margin are significant, beyond the direct effects on wages via "employer-signalling". Interventions targeted at closing the gap between "unlucky" and "lucky" students at the margin may be valuable in remedying the distortions identified by this paper.

The rest of the paper proceeds as follows. The next section discusses the institutional background and the data. Section 3 outlines the estimation strategy. Section 4 examines the validity of the empirical design. Section 5 presents the results. I conclude in section 6.

2 Background and Data

My data is drawn from the administrative records of eight National University of Singapore admission cohorts (academic years 2005/2006 to 2013/2014) merged with information from employment surveys conducted by the university roughly six months after graduation.[1] The Graduate Employment Survey (GES) asks students about their salary, job search experience since graduation, and various information (e.g., sector, industry, full- or part-time) about their current job. All public universities in Singapore participate in the survey and have to ensure that the sampling is representative of the graduating population and that the response rate reaches at least 70 percent. The survey results are reported to the Singapore Ministry of Education and the aggregate information is released to inform the general public about the starting salaries for graduates from the public universities.[2] Comparing the means of key covariates shows that the merged survey sample is generally comparable to the full of administrative data (all NUS graduates). Importantly, the administrative data includes

[1] I thank Jessica Pan for language and informational details on National University of Singapore policies, administrative data, and the GES.

[2] More information about the survey and statistics for recent graduating cohorts can be found at: https://data.gov.sg/dataset/ graduate-employment-survey-ntu-nus-sit-smu-sutd

course-student level data on both letter grades and as well as precise marks (0-100).

Table 1 presents summary statistics for the full administrative data that includes all students admitted between 2005/2006 and 2013/2014 (columns 1 and 2), as well as the merged sample with matched employment records (columns 3 and 4). Comparing the means of key covariates such as student grades (e.g. Cumulative Point Average), university admission score, faculty, and standard demographic characteristics (gender, race, marital status, and year of birth), I find that the merged survey sample is generally comparable to the sample of admin data (all NUS graduates). The final merged sample – our main analysis sample – consists of 32,123 students. The largest school in my sample is Faculty of Arts and Social Science (28%), followed by Faculty of Science (19%). About 53% of the graduates are female, 90% Chinese, 4% Indian, and 4% Malays.

In 2018, Singapore's resident labor force was 2.3 million, in which approximately 36.7% held a university degree.[3] The graduating cohort in 2017 is about 16,160 and the vast majority are from six publicly funded universities in Singapore. Among these institutions, NUS, where our sample is drawn from, is the largest university and accounts for about 44% of the annual intake of undergraduates each year since the mid-2000s.[4] The next largest university, the Nanyang Technological University (NTU) accounts for about 41% of the annual intake.

Finally, in Singapore, students apply directly to their course (major) before starting university. In order to change one's course, one must re-apply for transfer and re-enroll which is uncommon. Additionally, classes do not have grade requirements for enrollment. Thus, grades do not directly affect the choice set of classes available to a student. Academic advising is hands-off. Each student is assigned a department advisor who serves as a resource, providing information on courses. Advisers do not approve or interfere in student course selection.

[3] http://www.mom.gov.sg/Documents/statistics-publications/manpower-supply/report-labour-2011/

[4] Education Statistics Digest 2011, Ministry of Education, Singapore. Prior to the mid-2000s, NUS accounted for more than 50% of the annual intake of undergraduate students each year. In the early 1990s and before, NUS accounted for more than 70% of all undergraduates.

3 Estimation Strategy

I exploit a regression discontinuity design (RDD) which compares students that receive marks greater than letter grade thresholds by a small margin with those that receive marks under by a small margin. The intuition is that students who score just above and below the letter grade threshold are similar in terms of their characteristics, course performance and learning; however, scoring just above the threshold results in receiving a higher letter grade. By comparing the future outcomes of students just above and just below the letter grade threshold, I can estimate the effect of receiving a higher letter grade. This design will distinguish the effect of receiving a different letter grade from other confounding factors as long as the determinants of outcomes are continuous at the letter grade cutoff. Under this assumption, any discontinuous jump in outcomes at the letter grade cutoff can be interpreted as the causal effect of receiving the higher letter grade.

As an illustration of our RD framework, consider the simple model:

$$y_i = \alpha_0 + \beta L_i + \eta X_i + \epsilon_i \tag{1}$$

where y_i is some outcome variable of interest for student i, L_i equals 1 if individual i receives an A- letter grade and 0 if individual i receives a B+ letter grade for some course, X_i is a vector of individual characteristics (i.e., student demographics, major, and undergraduate admission score), and ϵ_i is a i.i.d. error term. β measures the impact of receiving an A- relative to receiving a B+. The OLS estimate of equation (1) is biased if ϵ_i is correlated with L_i due to unobserved ability or individual characteristics. Thus, I use a regression discontinuity design specification:

$$y_i = \alpha_0 + \beta D(m_i > c) + f(m_i) + \eta X_i + \epsilon_i \tag{2}$$

where c is the A- letter grade cut-off, m_i is student i's class marks (0-100), $D(m_i > c)$ is a binary variable equal to 1 if student i's marks is greater than the A- letter grade cutoff (equivalent to L_i), and $f(m_i)$ is a flexible function. $f(m_i)$ represents the relationship between student i's precise marks and the outcome, thus the identification strategy hinges on the exogeneity of m_i at the threshold. The parameter β now identifies the local average treatment effect of receiving a A- versus a B+ on the treated, specifically those scoring near the letter grade threshold.

I will use a triangular kernel and local polynomial regressions in my main empirical specification, as suggested by Gelman and Imbens (2019). The optimal bandwidth, within which I include students above and below the letter grade threshold, used in my baseline RD estimates is based on the method developed by Calonico, Cattaneo, and Titiunik (2014). I report robust bias corrected standard errors. For a given letter grade and course, I define the letter grade cut-off, c, to be the minimum marks in the course for which a student gets that letter grade. I will use distance from the closest letter grade cut off in standard deviations as the running variable in all results in the next section to standardize across different courses.

4 Validity

A key identifying assumption is that, absent the treatment (receiving a higher letter grade), the average outcomes of the individuals would be smooth in the running variable (marks). As discussed by Lee (2008), it implies that individuals cannot fully manipulate the running variable, distance from a letter grade threshold. University institutional policies, as well as evidence that density and observable characteristics (race, gender, age, and undergraduate admission scores) are smooth across the letter grade thresholds suggest that manipulation is implausible.

4.1 Institutional Details and Policies

Almost all undergraduate modules at NUS have a sit-down final exam component and the grades are curved. In other words, a student's grade depends on his or her relative standing in the class, hence it is implausible that students can perfectly control their precise marks and hence letter grades. Moreover, all grades are released together by a centralized system after they have been scrutinized at multiple levels (e.g., department, faculty, registration office, etc). Faculty follow grade distribution guidelines set by their respective departments. The answer scripts are graded anonymously and are not returned to students. Once the grades are released, grade changes are extremely difficult and only happen under exceptional circumstances. In the rare case of re-grading, any changes must be justified by the module coordinator and approved by administrators. NUS considers the mishandling of examination matters to be serious misconduct.

4.2 Random Selection at the Letter Grade Thresholds

I find that observable demographics variables are smooth across the letter grade thresholds. Figure 1a plots the relationship between race and the distance from the closest grade threshold. Figure 1b plots the relationship between birth year and the distance from the closest grade threshold. Figure 1c plots the relationship between gender and the distance from the closest grade threshold. I find null regression discontinuity estimates for each demographic variable. Since I rely on the Graduate Employment Survey (GES) for earnings data, I also require an absence of a discontinuity in the probability of responding to the survey. Figure 1d plots the relationship between the probability of being matched to earnings data in the GES and the distance from the closest grade threshold. Again, I find null regression discontinuity estimates.

Most importantly, I show that measures of ability are smooth across the grade thresholds. Figure 2a plots the relationship between the student's University Admission Score (UAS) and the distance from the closest grade threshold. A student's University Admission Score is the score each student receives based on their pre-university grades for admissions and is a strong proxy for ability. Panel B in Table 3 presents the regression discontinuity estimates of Equation 2 with University Admission Score as y_i. I find that there is an insignificant relationship between University Admission Score and distance from the closest grade threshold. Figure 2b plots the relationship between the student's Cumulative Point Average[5] prior to enrollment (past CAP) in the class and the distance from the closest grade threshold. Panel C in Table 3 presents the regression discontinuity estimates of Equation 2 with past CAP as y_i. I find that there is an insignificant relationship between CAP prior to the class and distance from the closest grade threshold. Both results are robust to various demographic controls, major fixed effects, and using a local linear RD specification.

I also check the density of the running variable around the letter grade thresholds (McCrary 2008). I employ local polynomial density estimation using standard methods from Cattaneo, Jansson, and Ma (2018). I also plot the density of the running variable for visual evidence. Online Appendix Figure A1a plots the density of student marks for the largest course in the data where the red lines represent the B-, B, B+, A- and A grade cut offs. Online Appendix Figure A1b plots the density of all student marks in distances from the A Grade cut-off. Panel A in Table 3 reports the estimation results. Robust across various specifications, I find no evidence of significant bunching.

5 Results

5.1 Wages

First, I study the returns in post graduation wages from receiving a higher letter grade for students at the margin. Figure 3a plots the "reduced-form" relationship between earnings and distance to the closest letter grade cut off. On either side of the letter grade cut off, income has a flat relationship with distance to the cut off. This correlation suggests that letter grades as signals are important, while the precise marks that one receives may not be. In Panel A of Table 4, I present the regression discontinuity estimates of Equation (2) with monthly wages as y_i. I find that receiving a better grade for students on the margin results

[5] Same as Grade Point Average. 1-5 scale: 5 = A, 4.5 = A-, 4 = B+, 3.5 = B, 3 = B-, 2.5 = C+, 2 = C

in 32 US dollars ($45 SGD) greater monthly earnings post-graduation. This result is robust to including pre-university admission score and demographic controls, major fixed effects, and using a local linear RD specification.

There are two possible mechanisms for this result. First, employers may use grades as a signal of ability and as a result pay similar students different salaries based on this coarse measure. Second, students may interpret better or worse grades as a signal of their own ability which affects their future behavior and outcomes.

The size of the effects is much larger than might be expected based on an earnings function in CAP and other observables. In Figure 4a, I plot the relationship between post-graduation wage and final cumulative point average, controlling for demographics, major fixed effects and university admission score. I find that the earnings function is linear with respect to cumulative point average (1-5) above 3.3, and flat below 3.3. Note that less than 20% percent of students have a CAP below 3.3. In Online Appendix Table A1, I estimate that the return in monthly wages from a one point increase in CAP (above 3.3) is $712 SGD. The estimates remain quantitatively similar with and without various controls for demographics, major fixed effects, average course level, and university admission score. A single class on average counts for 3% of one's total credits. Thus, a letter grade increase of 0.5 in CAP for a single class, raises total CAP by 0.015.[6] This implies that getting a better grade on the margin should increase wages by $10.68 SGD[7] via "employer-signalling". This is far smaller than the regression discontinuity estimate – less than 60 percent the lower confidence interval. These back of envelope calculations imply that "employer-signalling" without consideration for future changes in behavior or outcomes can only partially account for large wage returns to receiving a higher grade on the margin.

[6] $0.5 * 0.03$

[7] $0.015 * 712 = 10.68$

5.2 Future Grades

I find that receiving worse grades for students on the margin results in slightly lower grades in future semesters. Figure 3b plots the "reduced-form" relationship between Cumulative Point Average (CAP)[8] in future semesters and distance to the closest letter grade cut off. The relationship is a linear positive slope except for the mass just below the grade cut off where future grades dip before jumping back up on the positive linear trend after the cut off. This suggests that receiving a lower letter grade for students on the margin negatively affects future grades. This could be explained by discouragement de-motivating the student, or that the student learns that he is of a "low ability type" then adjusts his future goals (e.g. that he is a "B' student') and effort (e.g. by studying less). In Panel B of Table 4, I present the regression results. Receiving a lower grade for students on the margin results in a 2 p.p. lower CAP in future semesters. I show that this result is robust to including pre-university admission score and demographic controls, major fixed effects, and using a local linear RD specification.

5.3 Future Course Difficulty

I find that receiving worse grades for students on the margin results in the student taking significantly lower level or easier courses in future semesters. Each course has a level from 1 to 4 where 1 represents the most introductory courses and 4 represents the most advanced courses. Figure 3c plots the "reduced-form" relationship between average course level in future semesters and distance to the closest letter grade cut off. The relationship is close to flat on either side of the cutoff except for the mass just before the grade cut off where future average course levels dip before jumping back up on the flat trend after the cut off. This suggests that receiving a lower letter grade for students on the margin negatively affects future average course levels. This indicates that students invest less in human capital accumulation as a result of receiving a noisy negative signal of ability. In Panel C of Table 4,

[8] Same as Grade Point Average. 1-5 scale: 5 = A, 4.5 = A-, 4 = B+, 3.5 = B, 3 = B-, 2.5 = C+, 2 = C

I present the regression results. Receiving a worse grade for students on the margin results in taking classes that are on average 0.2 standard deviations lower in course level. I show that this result is robust to including pre-university admission score and demographic controls, major fixed effects, and using a local linear RD specification.

The returns to taking higher level courses are large. Figure 4b plots the relationship between post-graduation wage and average course level across all enrolled years. In Table 5, I find that a standard deviation increase in average course level implies a $74.29 SGD increase in monthly earnings, controlling for sex, age, race, major fixed effects, final cumulative point average and university admission score.

5.4 Mechanisms

My findings are consistent with a theory of "self-signalling" where students interpret better or worse grades as a signal of their own ability affecting their future behavior and outcomes. Receiving a bad grade affects a student's perceptions of what he or she can achieve and can be discouraging. This may result in sub-optimal under investment in human capital development by taking fewer advanced classes, and in a reduction of effort leading to worse grades.

However, it may be possible that students are behaving optimally in response to "employer-signaling". If the earnings function is convex in cumulative point average, the return to getting a better grade is larger for students with higher CAPs. Thus, students may be rational in reducing effort in response to a getting a worse grade on the margin. However, in Figure 4a, I show that the earnings function is in fact linear with respective to cumulative point average above 3.3, and flat below 3.3 (the bottom 20 percent of the distribution). Thus, the returns to CAP are constant between 3.3 and 5.0. Based on such an earnings function, it is not optimal to lower effort in response to receiving a lower letter grade on the margin for students who have cumulative point averages sufficiently greater than 3.3 prior to the class. In Online Appendix Table A2, I restrict my sample to students with cumulative point averages greater than 3.5 prior to the class and find that effects on future grades remain significant and quantitatively similar. Another possibility is that the effects on future grades are driven by those at the very top of the distribution who are vying for the best jobs which are scarce. Thus, when one gets a lower grade and is no longer competitive for the best jobs, a student would rationally reduce their effort. In Online Appendix Table A2, I further restrict the sample to students with grades below 4.5 CAP, an A- average. Again, I find that the effects on future grades are robust and quantitatively similar.

Similarly, it may be optimal for students to take easier courses in response to receiving a worse grade on the margin if the returns to taking higher level courses are greater for those with higher CAPs. On the contrary, I find that the returns are greater for those with lower CAPs. In Table 5, I estimate returns to average course level interacted with CAP controlling for sex, race, age, major fixed effects and university admission score. A standard deviation increase in average course level is associated with a $140 increase in wages for students with CAPs below 3.7, $100 between 3.7 and 4.0, $50 for CAPs between 4.0 and 4.3, and $15 for CAPs above 4.3 prior to the class. Such an earnings function absent "self-signaling" would not predict that students optimally reduce future course difficulty in response to getting a worse grade on the margin. Again, as with effects on future grades, I rule out that the effects on future course taking are driven by the bottom or top of the distribution in Online Appendix Table A1.

5.5 Heterogeneity

In this subsection, I investigate heterogeneity in effects.

First, I show that the returns in earnings from receiving a better letter grade are driven by classes taken in years 1 and 2. In Table 6, I present the regression results by school year. The returns are highest in the first year at $49 USD ($69 SGD). In the second year, it is $39 USD ($54 SGD). I don't find significant effects for years 3 and 4. Year 3 and 4 are graduation years depending on one's major and students typically look for jobs in their final year with incomplete transcripts. The fact that the returns to getting a better grade are the largest for first year courses is consistent with "self-signalling", where first year students, not knowing where they fall in the school's ability distribution, use the first few grades they receive to decide where they fall and what they are able to achieve, thus shaping their future outcomes. First year students have the greatest margin (most future university years) for behavioral change. However, it might be possible that employers simply place more weight on first year letter grades. My sample is not sufficiently large to reject homogeneity in effect sizes across years.

Next, I find that these results only hold for courses that are required or relevant to the student's major. Table 7 presents the regression results by General Education electives versus Major requirements. There is a null effect for elective courses, while the returns to receiving a higher grade in required courses is significantly different from zero at $36 USD ($51 SGD). I find consistent results in future grades and course taking. There are null effects for elective courses on future grades, but significant effects for required courses. Effects on course taking are significant for both electives and required courses, though the point estimate is approximately 40 percent smaller for electives. These results suggest that signals from major-relevant courses may be more important than from elective courses. However, there are far fewer elective courses than required courses in my sample, thus I am less powered to detect effects on electives. I also cannot reject homogeneity in effect sizes.

In Table 8, I present the regression results by grade cutoff. I find that the return in earnings from getting a better grade on the margin is largest at the A- cutoff, followed by the A and B+ cutoffs. I find that the return to getting an A- versus an B+ is $59 USD ($82 SGD). The returns are $52 USD ($73 SGD) and $50 USD ($70 SGD) for the A versus A- and B+ versus B cut-offs respectively. The returns for the B versus B- and B- versus C+ cut offs are insignificant. However my sample is not sufficiently large to reject homogeneity in effect sizes across letters. The effects on behavior are similar in point estimates across grade cutoffs except for B- versus C where I find a null effect on future grades. This may suggest that an “A” signal for a given class is particularly important to employers on a transcript, but to students the signal from each grade cut-off is similar. Many employers in Singapore have attended NUS themselves and may look for grades, or A’s, for particular courses rather than focus on the aggregated CAP.

6 Conclusion

I study the consequences of noisy letter grades measures using administrative data from the National University of Singapore which records both the letter grades as well as the precise marks (0-100) received on each course that a given student takes. I exploit a regression discontinuity design – specifically, close to the letter grade cutoffs, individuals with very similar achievement will receive different letter grades. I find that receiving a better grade for students on the margin results in 32 US dollars greater monthly earnings post-graduation.

Employers use grades as a signal of ability and as a result pay similar students different salaries based on this coarse measure. However, the size of the effect is much larger than expected from a corresponding cumulative grade point average increase based on an estimated earnings function on grades and other observables. This suggests that future changes in behavior and outcomes may be important in explaining the large effects on earnings. I find that receiving a worse grade for students on the margin results in only slightly lower grades in future semesters, but these students take significantly “easier courses”. Receiving a worse grade for students on the margin results in taking classes that are on average 0.2 standard deviations lower in course level. This indicates that students invest less in human capital accumulation as a result of receiving a noisy negative signal of ability.

This paper has identified significant gaps between students who receive better and worse grades at letter grade thresholds in university classes. Policy makers should consider these distortions in decisions about grading regimes. Interventions to address these gaps in existing letter grading systems may be valuable.

References

Bettinger, Eric P. and Bridget Terry Long (July 2009). "Addressing the needs of under-prepared students in higher education: Does college remediation work?" In: *Journal of Human Resources* 44.3, pp. 736–771.

Brookover, Wilbur B., Shailer Thomas, and Ann Paterson (21 1964). "Self-Concept of Ability and School Achievement". In: *Sociology of Education* 37.3, p. 271.

Calcagno, Juan Carlos and Bridget Terry Long (2008). "The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance". In: *National Bureau of Economic Research Working Paper Series* No. 14194.April, p. 46.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (Nov. 2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". In: *Econometrica* 82.6, pp. 2295–2326.

Cameron, Stephen V. and James J. Heckman (Jan. 1993). "The Nonequivalence of High School Equivalents". In: *Journal of Labor Economics* 11.1, Part 1, pp. 1–47.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma (Nov. 2018). "Simple Local Polynomial Density Estimators". In: *Journal of the American Statistical Association*.

Clark, Burton R (1960). "The "Cooling-Out" Function in Higher Education". In: *American Journal of Sociology* 65.6, pp. 569–576.

Clark, Damon and Paco Martorell (2014). "The signaling value of a high school diploma". In: *Journal of Political Economy* 122.2, pp. 282–318.

Crocker, Jennifer et al. (Nov. 2003). *Contingencies of Self-Worth in College Students: Theory and Measurement*.

Diamond, Rebecca and Petra Persson (2016). "The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests". In: *National Bureau of Economic Research Working Paper Series* No. 22207.

Feng, Andy and Georg Graetz (Dec. 2017). “A question of degree: The effects of degree class on labor market outcomes”. In: *Economics of Education Review* 61, pp. 140–161.

Gelman, Andrew and Guido Imbens (July 2019). “Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs”. In: *Journal of Business and Economic Statistics* 37.3, pp. 447–456.

Jacob, Brian A. and Lars Lefgren (July 2009). “The Effect of Grade Retention on High School Completion”. In: *American Economic Journal: Applied Economics* 1.3, pp. 33–58.

Kane, Michael, Terence Crooks, and Allan Cohen (Oct. 1999). “Validating Measures of Performance”. In: *Educational Measurement: Issues and Practice* 18.2, pp. 5–17.

Kane, Thomas J et al. (1995). “Labor-Market Returns to Two- and Four-Year College”. In: *American Economic Review* 85.3, pp. 600–614.

Lee, David S (2008). “Randomized experiments from non-random selection in U.S. House elections $”. In: *Journal of Econometrics* 142, pp. 675–697.

Manacorda, Marco (2012). “The cost of grade retention”. In: *Review of Economics and Statistics* 94.2, pp. 596–606.

Martorell, Paco and Isaac McFarlin (2011). “Help or hindrance? The effects of college remediation on academic and labor market outcomes”. In: *The Review of Economics and Statistics* 93.2, pp. 436–454.

McCrary, Justin (Feb. 2008). “Manipulation of the running variable in the regression discontinuity design: A density test”. In: *Journal of Econometrics* 142.2, pp. 698–714.

Mcewan, Patrick J et al. (2019). *Grades and the Economics Major at a Women's College * We are grateful to.* Tech. rep.

Papay, John P., Richard J. Murnane, and John B. Willett (2016). *The Impact of Test Score Labels on Human-Capital Investment Decisions.* Tech. rep. 2, pp. 357–388.

Scott-Clayton, Judith and Olga Rodriguez (Jan. 2015). "Development, Discouragement, Or diversion? New evidence on the effects of college remediation policy". In: *Education Finance and Policy* 10.1, pp. 4–45.

Shen, Ce and Joseph J. Pedulla (2000). "The Relationship between students' achievement and their self-perception of competence and rigour of mathematics and science: A cross-national analysis". In: *Assessment in Education: Principles, Policy and Practice* 7.2, pp. 237–253.

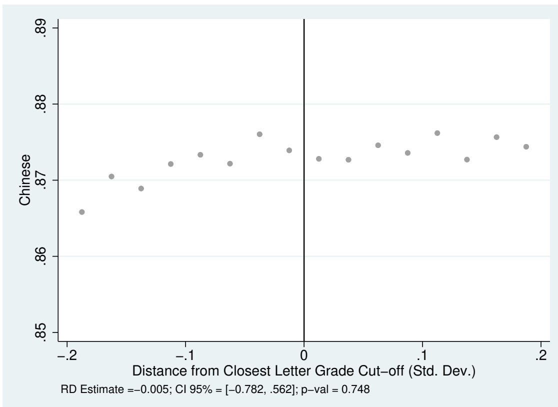
Smith, Jonathan, Michael Hurwitz, and Christopher Avery (2017). "Giving college credit where it is due: Advanced placement exam scores and college outcomes". In: *Journal of Labor Economics* 35.1, pp. 67–147.

Stinebrickner, Ralph and Todd R. Stinebrickner (Jan. 2014). "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout". In: *The Review of Economic Studies* 81.1, pp. 426–472.

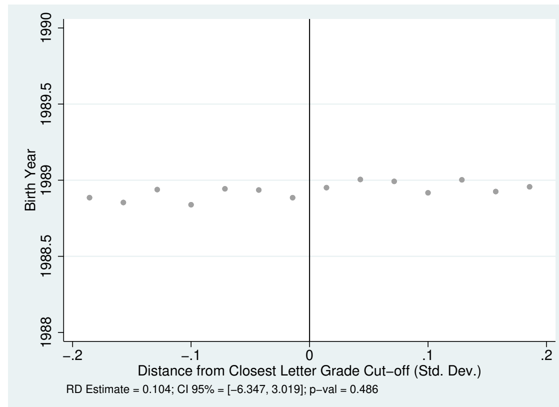
Tyler, John H., Richard J. Murnane, and John B. Willett (2000). "Estimating the labor market signaling value of the GED". In: *Quarterly Journal of Economics* 115.2, pp. 431–468.

Figures and Tables

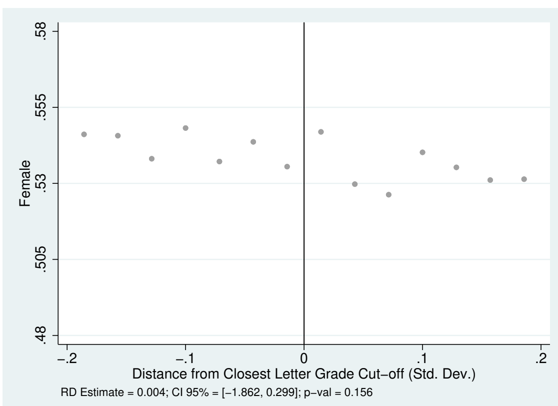
Figure 1: Continuity of Predetermined Variables



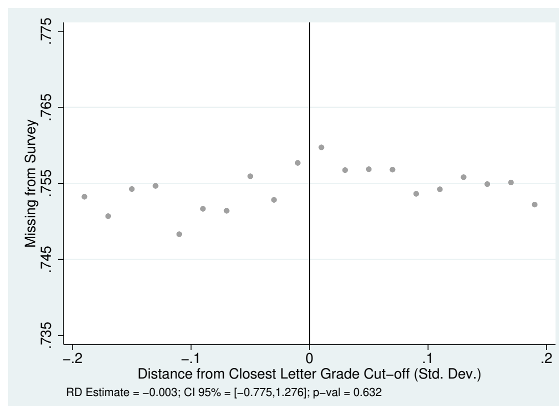
(a) Chinese



(b) Birth Year



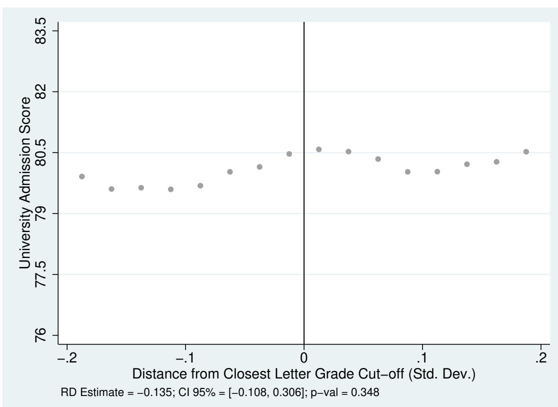
(c) Female



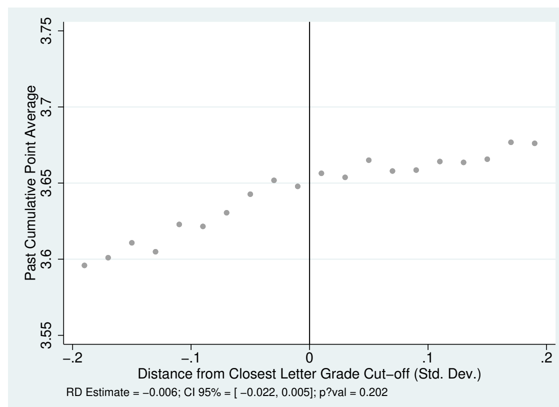
(d) Probability of Match to GES

Notes: These figure shows a regression discontinuity binned scatter plot of various pre-determined variables against the distance from a score's closest letter grade threshold in standard deviations. Each point shows the mean of the variable with distance from grade threshold in the given bin. Each figure reports estimates from a local polynomial Regression Discontinuity (RD) point estimator using methods developed in Calonico et al. (2014). The data set is a survey sample NUS students linked to administrative course data.

Figure 2: Continuity of Measures of Ability



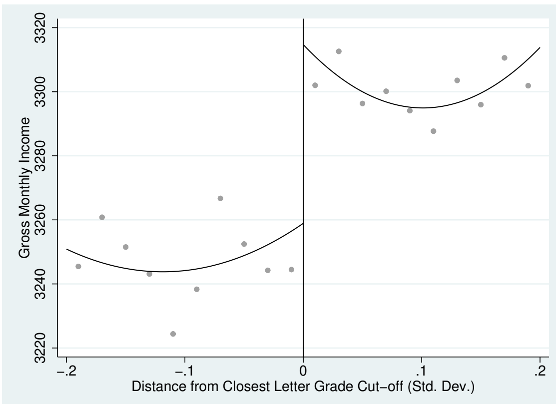
(a) University Admission Score



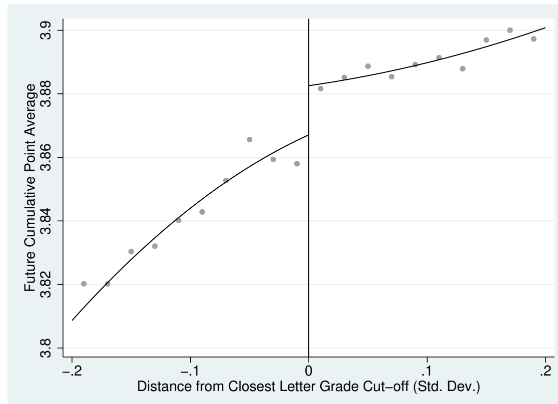
(b) Cumulative Average Point Prior to Class

Notes: These figure shows a regression discontinuity binned scatter plot of various measures of ability against the distance from a score's closest letter grade threshold in standard deviations. Each point shows the mean of the variable with distance from grade threshold in the given bin. Each figure reports estimates from a local polynomial Regression Discontinuity (RD) point estimator using methods developed in Calonico et al. (2014). The data set is a survey sample NUS students linked to administrative course data.

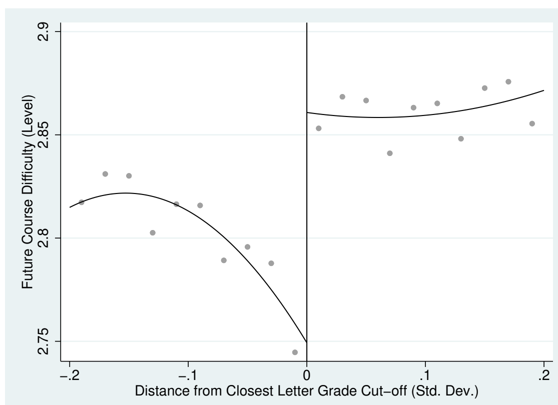
Figure 3: Impact of Receiving a Higher Letter Grade



(a) Monthly Gross Income



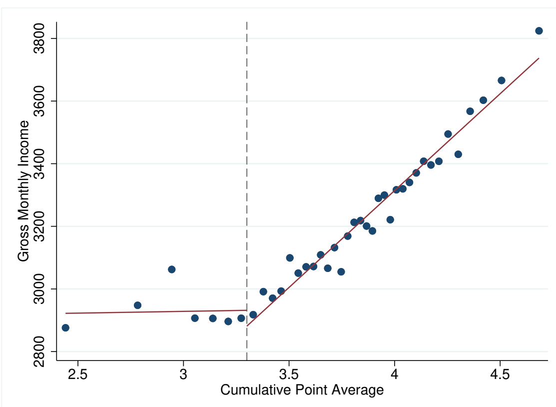
(b) Cumulative Point Avg. in Future Semesters



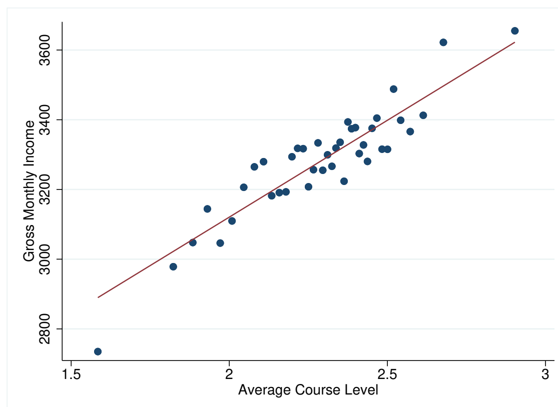
(c) Average Course Level in Future Semesters

Notes: These figures shows a regression discontinuity binned scatter plot of post-graduation gross monthly income, future cumulative point average (CAP), and future average course level against the distance from a score's closest letter grade threshold in standard deviations. Future CAP is the CAP for all future semesters. Future course level is the average course level from 1 to 4 for all future semesters. Each point shows the mean monthly income with distance from grade threshold in the given bin. The data set is a survey sample NUS students linked to administrative course data.

Figure 4: Earnings Function



(a) Cumulative Point Average



(b) Average Course Level

Notes: The Figure shows a binned scatter plot of post-graduation gross monthly income, against final cumulative point average in (a) and average course level across all years in (b). Each point shows the mean monthly income with cumulative point average or average course level in the given bin. Controls for gender, age, race, university admission score and major fixed effects are included. For both figures, observations are at the student level. The data set is a survey sample NUS students linked to administrative course data.

Table 1: Summary Statistics

	Full Sample		Survey Sample	
	Mean	SD	Mean	SD
University Admission Score	78.41	14.28	78.71	13.94
Cumulative Point Average	3.74	0.52	3.76	0.51
Female	0.55	0.50	0.53	0.50
Married	0.01	0.12	0.01	0.11
Chinese	0.90	0.30	0.90	0.30
Indian	0.04	0.20	0.04	0.19
Malay	0.03	0.18	0.04	0.18
Birth Year	1989.00	2.89	1989.15	2.84
Faculty of Arts and Social Science	0.29	0.45	0.28	0.45
Faculty of Business	0.10	0.30	0.10	0.30
School of Computing	0.05	0.21	0.05	0.21
School of Design and Environment	0.07	0.26	0.07	0.26
Faculty of Engineering	0.17	0.37	0.17	0.38
Faculty of Law	0.05	0.21	0.04	0.21
Faculty of Science	0.19	0.39	0.19	0.39
Observations	43406		32123	

Notes: The full sample includes all students admitted between 2005/2006 and 2013/2014 who are Singapore Residents. The survey sample is limited to those who had responded to the employment survey.

Table 2: Continuity of Measures of Ability

	(1)	(2)	(3)	(4)
Panel A: Continuity of University Admission Scores				
RD_Estimate	0.135	−0.025	0.109	0.060
	[0.093]	[0.096]	[0.061]	[0.066]
Robust 95% CI	[-.108 ; .306]	[-.275 ; .145]	[-.036 ; .241]	[-.12 ; .191]
Observations	1124489	1124489	1124437	1124489
Robust p-value	0.348	0.544	0.149	0.656
BW Loc. Poly. (h)	0.196	0.181	0.322	0.139
BW Bias (b)	0.325	0.313	0.515	0.258
Panel B: Continuity of Past Cumulative Point Average				
RD_Estimate	−0.006	−0.009	0.003	−0.008
	[0.006]	[0.006]	[0.006]	[0.005]
Robust 95% CI	[-.022 ; .005]	[-.024 ; .002]	[-.011 ; .013]	[-.022 ; .003]
Observations	757489	757489	757489	757489
Robust p-value	0.202	0.103	0.854	0.142
BW Loc. Poly. (h)	0.242	0.241	0.272	0.136
BW Bias (b)	0.499	0.494	0.544	0.252
Order Loc. Poly. (p)	2.000	2.000	2.000	1.000
Demo. Controls		X	X	X
Major FE			X	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The Table shows estimates from a local polynomial regression discontinuity (RD) estimator with asymptotic MSE-optimal bandwidth, triangular kernel, and robust bias-corrected confidence intervals developed in Calonico et al. (2014). In Panel A, the outcome variable is university admission score. In Panel B, the outcomes variables is cumulative point average prior to the class. The running variable is distance from the closest letter grade threshold in standard deviations. Demographic controls are included in columns 2, 3, and 4. Column 3 includes major fixed effects. Columns 1, 2, and 3 use a 2 degree local polynomial, and column 4 uses a linear local-polynomial. The data set is a survey sample NUS students linked to administrative course data.

Table 3: Manipulation Test of Distance from Letter Grade Threshold

	(1) Density	(2) Density	(3) Density	(4) Density
T - Statistic	1.3719	0.9339	0.7926	0.1538
P-Value	0.1701	0.3504	0.4280	0.8777
Observations	1703144	1703144	1703144	1703144
Order Loc. Poly. (p)	2.000	2.000	1.000	3.000
Order Bias (q)	3.000	3.000	2.000	4.000
Method	Conventional	Robust	Robust	Robust

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a polynomial density estimator with MSE-optimal bandwidth and triangular kernel as in Cattaneo et al. (2018). The running variable is distance from the closest letter grade threshold in standard deviations. Column 1 and 2 use a 2 degree polynomial specification. Column 3 uses a linear specification. Column 4 uses a 3 degree polynomial specification. Column 1 reports conventional test statistics. Columns 2 to 4 report robust robust bias-corrected statistics. The data set is a survey sample NUS students linked to administrative course data.

Table 4: Impact of Receiving a Higher Letter Grade

	(1)	(2)	(3)	(4)
Panel A: Monthly Wages				
RD_Estimate	47.808***	44.103***	43.459***	34.882***
	[12.304]	[14.129]	[14.163]	[10.822]
Robust 95% CI	[18.731 ; 72.525]	[14.253 ; 78.293]	[13.678 ; 77.879]	[13.665 ; 62.593]
Observations	525501	525449	525449	525449
Robust p-value	0.001	0.005	0.005	0.002
BW Loc. Poly. (h)	0.347	0.265	0.262	0.178
BW Bias (b)	0.623	0.412	0.408	0.384
Panel B: Future Cumulative Point Average				
RD_Estimate	0.017***	0.025***	0.019***	0.017***
	[0.004]	[0.005]	[0.005]	[0.004]
Robust 95% CI	[.006 ; .025]	[.013 ; .035]	[.007 ; .028]	[.009 ; .028]
Observations	974368	974324	974324	974368
Robust p-value	0.001	0.000	0.001	0.000
BW Loc. Poly. (h)	0.292	0.231	0.231	0.143
BW Bias (b)	0.586	0.418	0.440	0.284
Panel C: Future Average Course Level				
RD_Estimate	0.112***	0.101***	0.097***	0.106***
	[0.006]	[0.006]	[0.006]	[0.006]
Robust 95% CI	[.101 ; .127]	[.09 ; .116]	[.086 ; .112]	[.097 ; .119]
Observations	973518	973474	973474	973518
Robust p-value	0.000	0.000	0.000	0.000
BW Loc. Poly. (h)	0.152	0.141	0.145	0.080
BW Bias (b)	0.266	0.269	0.270	0.264
Order Loc. Poly. (p)	2.000	2.000	2.000	1.000
Order Bias (q)	3.000	3.000	3.000	2.000
Demo. Controls		X	X	X
Major FE			X	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a local polynomial regression discontinuity (RD) estimator with asymptotic MSE-optimal bandwidth, triangular kernel, and robust bias-corrected confidence intervals developed in Calonico et al. (2014). In Panel A, the outcome variable is monthly wage. In Panel B, the outcome variable is future cumulative point average (CAP). Future CAP is the CAP for all future semesters. In Panel C, the outcome variable is future average course level. Future course level is the average course level from 1 to 4 for all future semesters. The running variable is distance from the closest letter grade threshold in standard deviations. Columns 1-3 use a 2 degree local polynomial, and column 4 uses a linear local-polynomial. Demographic controls for race and gender are included in all specifications. Major fixed effects are included in columns 2-4. A control for pre-university admission score is included in columns 3-4. The data set is a survey sample NUS students linked to administrative course data.

Table 5: Returns to Average Course Level

	(1) Monthly Income	(2) Monthly Income	(3) Monthly Income	(4) Monthly Income	(5) Monthly Income
Average Level (Norm.)	288.6***	223.1***	207.2***	74.29***	138.0***
	(12.15)	(12.47)	(13.46)	(13.65)	(16.71)
Average Level (Norm.) * CAP 3.7 - 4.0					-41.58
					(30.52)
Average Level (Norm.) * CAP 4.0 - 4.3					-86.01**
					(35.58)
Average Level (Norm.) * CAP > 4.3					-121.0***
					(43.54)
Observations	20012	20012	20012	20012	20012
Major FE	X	X	X	X	X
Demo. Controls		X	X	X	X
UAS Control			X	X	X
CAP Control				X	
CAP Group FE					X

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a linear regression of gross monthly income on normalized average course level across all enrolled years. Major fixed effects are included in all columns. Demographic controls for sex, age and race are included in columns 2 to 5. Columns 3 to 5 control for University Admission Score. Column 4 controls for cumulative point average. Column includes interactions and fixed effects for students with CAPs below 3.7, between 3.7 and 4.0, between 4.0 and 4.3, and above 4.3. All observations are at the student level. The data set is a survey sample NUS students linked to administrative course data.

Table 6: Impact on Monthly Wages - By Year

	(1)	(2)	(3)	(4)
RD_Estimate	68.941***	53.948**	0.469	36.316
	[24.651]	[23.844]	[25.976]	[36.310]
Sample	Year 1	Year 2	Year 3	Year 4
Robust 95% CI	[10.782 ; 119.73]	[1.969 ; 108.436]	[-62.267 ; 50.886]	[-47.95 ; 110.55]
Observations	156717	153277	120889	86501
Robust p-value	0.019	0.042	0.844	0.439
BW Loc. Poly. (h)	0.269	0.320	0.326	0.383
BW Bias (b)	0.458	0.528	0.616	0.691

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a local 2 degree polynomial regression discontinuity (RD) estimator with asymptotic MSE-optimal bandwidth, triangular kernel, and robust bias-corrected confidence intervals developed in Calonico et al. (2014). The outcome variable is post-graduation monthly wage. The running variable is distance from the closest grade threshold in standard deviations. Column 1 includes all courses taken in year 1, column 2 includes all courses taken in year 2, column 3 includes all courses taken in year 3 and column 4 includes all courses taken in year 4. In Demographic controls for race and gender are included in all specifications. The data set is a survey sample NUS students linked to administrative course data.

Table 7: Impact of Receiving a Higher Letter Grade by General Education vs Major Requirement

	(1) Monthly Income	(2) Monthly Income	(3) Future CAP	(4) Future CAP	(4) Future Level	(5) Future Level
RD_Estimate	27.674	51.223***	0.004	0.018***	0.072***	0.114***
	[55.585]	[12.058]	[0.024]	[0.004]	[0.020]	[0.006]
Robust 95% CI	[-107.422 ; 143.214]	[23.583 ; 76.391]	[-.054 ; .053]	[.007 ; .026]	[.029 ; .118]	[.103 ; .129]
Observations	34042	491459	62301	912067	62249	911269
Robust p-value	0.780	0.000	0.988	0.001	0.001	0.000
BW Loc. Poly. (h)	0.268	0.385	0.221	0.304	0.261	0.152
BW Bias (b)	0.434	0.686	0.360	0.580	0.418	0.266
Elective	Yes	No	Yes	No	Yes	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a local 2 degree polynomial regression discontinuity (RD) estimator with asymptotic MSE-optimal bandwidth, triangular kernel, and robust bias-corrected confidence intervals developed in Calonico et al. (2014). The outcome variable is gross monthly wage in columns 1 and 2. The outcome variable is future cumulative point average (CAP) in columns 3 and 4. The outcome variable is future average course level in columns 4 and 6. The running variable is distance from the closest grade threshold in standard deviations. Future CAP is the CAP for all future semesters. Future course level is the average course level from 1 to 4 for all future semesters. Columns 1, 3 and 5 include all elective courses, and columns 2, 4 and 6 include all non-elective courses. Demographic controls for race and gender are included in all specifications. The data set is a survey sample NUS students linked to administrative course data.

Table 8: Impact of Receiving a Higher Grade - By Letter

	(1) All	(2) A	(3) A-	(4) B+	(5) B	(6) B-
Panel A: Monthly Wages						
RD_Estimate	47.808***	73.349***	82.061***	69.814***	21.251	49.275
	[12.304]	[25.617]	[13.680]	[15.326]	[19.192]	[46.958]
Robust 95% CI	[18.731 ; 72.525]	[27.751 ; 135.522]	[57.133 ; 115.187]	[41.429 ; 106.189]	[-24.85 ; 60.582]	[-61.886 ; 152.016]
Observations	525501	517398	521038	523817	517520	500234
Robust p-value	0.001	0.003	0.000	0.000	0.412	0.409
BW Loc. Poly. (h)	0.347	0.508	0.743	0.650	0.564	0.207
BW Bias (b)	0.623	0.938	1.315	1.157	0.871	0.320
Panel B: Future Cumulative Point Average						
RD_Estimate	0.017***	0.033***	0.027***	0.032***	0.030***	−0.003
	[0.004]	[0.007]	[0.005]	[0.005]	[0.006]	[0.014]
Robust 95% CI	[.006 ; .025]	[.022 ; .049]	[.018 ; .039]	[.023 ; .043]	[.017 ; .044]	[-.037 ; .027]
Observations	974368	963106	967284	972274	963500	943262
Robust p-value	0.001	0.000	0.000	0.000	0.000	0.738
BW Loc. Poly. (h)	0.292	0.481	0.535	0.597	0.511	0.248
BW Bias (b)	0.586	0.912	0.958	1.129	0.792	0.403
Panel C: Future Average Course Level						
RD_Estimate	0.112***	0.099***	0.080***	0.118***	0.093***	0.102***
	[0.006]	[0.008]	[0.006]	[0.008]	[0.008]	[0.017]
Robust 95% CI	[.101 ; .127]	[.086 ; .118]	[.07 ; .095]	[.105 ; .137]	[.078 ; .112]	[.061 ; .133]
Observations	973518	962315	966487	971458	962762	942601
Robust p-value	0.000	0.000	0.000	0.000	0.000	0.000
BW Loc. Poly. (h)	0.152	0.452	0.455	0.291	0.338	0.169
BW Bias (b)	0.266	0.974	0.881	0.593	0.536	0.289

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a local 2 degree polynomial regression discontinuity (RD) estimator with asymptotic MSE-optimal bandwidth, triangular kernel, and robust bias-corrected confidence intervals developed in Calonico et al. (2014). In Panel A, the outcome variable is post-graduation monthly wage. In Panel B, the outcome variable is future cumulative point average (CAP). Future CAP is the CAP for all future semesters. In Panel C, the outcome variable is future average course level. Future course level is the average course level from 1 to 4 for all future semesters. In column (1), the running variable is distance from the closest grade threshold in standard deviations. In columns (2) to (6), the running variable is distance from the A, A-, B+, B, and B- grade threshold in standard deviations respectively. Demographic controls for race and gender are included in all specifications. The data set is a survey sample NUS students linked to administrative course data.

Online Appendix

Table A1: Returns to Cumulative Point Average

	(1) Monthly Income	(2) Monthly Income	(3) Monthly Income	(4) Monthly Income
CAP ($\geq$ 3.3)	776.1***	706.5***	718.8***	711.9***
	(59.63)	(58.07)	(59.06)	(58.94)
CAP ($<$ 3.3)	20.34	-5.967	3.543	-9.592
	(50.32)	(50.40)	(49.67)	(49.76)
Observations	20012	20012	20012	20012
Major FE	X	X	X	X
Demo. Controls		X	X	X
UAS Control			X	X
Avg. Level Control				X

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a linear regression of gross monthly income on final cumulative point average (CAP). The slope for CAP greater than 3.3 is reported separately from CAP less than 3.3. Major fixed effects are included in all columns. Demographic controls for sex, age and race are included in columns 2 to 4. A control for University Admission Score is included in columns 3 and 4. A control for average course level is included in column 4. All observations are at the student level. The data set is a survey sample NUS students linked to administrative course data.

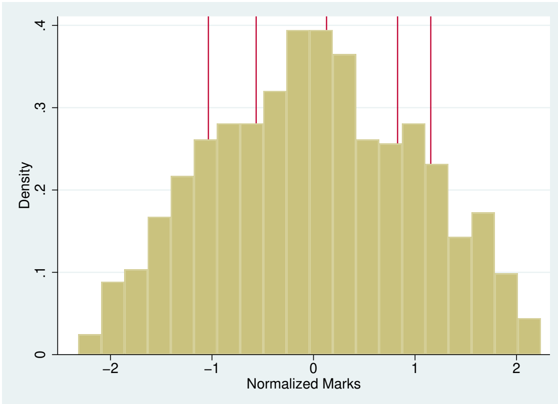
Table A2: Impact of Receiving a Higher Letter Grade on Future Behavior – Restricting by CAP Prior to Class

	(1) Future CAP	(2) Future CAP	(3) Future Level	(4) Future Level
RD_Estimate	0.026***	0.021***	0.087***	0.083***
	[0.007]	[0.007]	[0.009]	[0.009]
Robust 95% CI	[.011 ; .044]	[.007 ; .039]	[.07 ; .11]	[.066 ; .106]
Observations	418306	374131	418306	374131
BW Loc. Poly. (h)	0.210	0.209	0.160	0.166
BW Bias (b)	0.329	0.343	0.269	0.278
Past CAP ≥ 3.5	X	X	X	X
Past CAP ≤ 4.5		X		X

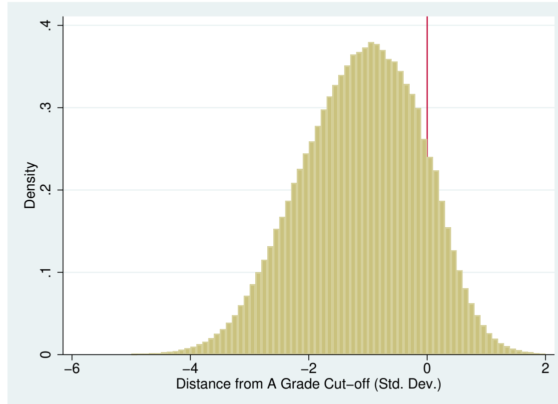
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Notes: The table shows estimates from a local 2 degree polynomial regression discontinuity (RD) estimator with asymptotic MSE-optimal bandwidth, triangular kernel, and robust bias-corrected confidence intervals developed in Calonico et al. (2014). In Columns 1 and 2, the outcome variable is future cumulative point average (CAP). Future CAP is the CAP for all future semesters. In Columns 3 and 4, the outcome variable is future average course level. Future course level is the average course level from 1 to 4 for all future semesters. The running variable is distance from the closest grade threshold in standard deviations. Demographic controls for race and gender are included in all specifications. Columns 1 and 3 include only students with a CAP greater than 3.5 prior to the class. Columns 2 and 4 include only students with a CAP greater than 3.5 and less than 4.5 prior to the class. Observations with no prior CAP (first year), are dropped. The data set is a survey sample NUS students linked to administrative course data.

Figure A1: Density Plots of the Running Variable



(a) All thresholds for a class



(b) All classes for a threshold

Notes: The histogram in (a) plots the density of marks for PC1432 in Fall 2006/2007. Marks are normalized to mean=0 and standard deviation=1. The histogram in (b) plots the density of distance in marks from the A grade cut-off for all courses. Marks are normalized to mean=0 and standard deviation=1 before computing the distance. The data set is a survey sample NUS students linked to administrative course data.