Empir Econ (2017) 52:411–446
DOI 10.1007/s00181-016-1078-4

Explaining university course grade gaps

Kevin P. Mongeon[1] · **Shawn W. Ulrick**[2] · **Michael P. Giannetto**[3]

Received: 8 August 2014 / Accepted: 7 February 2016 / Published online: 29 March 2016

Abstract This paper estimates the discrepancy in university mathematics and science course grades across races. Although there are significant Black–White and Hispanic–White grade discrepancies, or gaps, Black and Hispanic students who are equally prepared for university as White students do as well as White students. The grade gaps are explained after accounting for important factors such as a student's academic capabilities and socioeconomic status. Varying behaviors of university students relative to high school across races are ruled out as a possible source of the grade gaps.

Keywords Grade gaps · Education · Race

JEL Classification J15 · I21 · I24

The views herein do not necessarily reflect those of the Federal Trade Commission or any individual commissioner.

✉ Shawn W. Ulrick
sulrick@ftc.gov

Kevin P. Mongeon
kmongeon@brocku.ca

Michael P. Giannetto
Michael.Giannetto@asu.edu

1 Brock University, 500 Glenridge Ave., St Catharines, ON L2S 3A1, Canada

2 U.S. Federal Trade Commission, 600 Pennsylvania Ave., NW, Washington, DC 20580, USA

3 Arizona State University, P.O. Box 873201, Tempe, AZ 85287, USA

1 Introduction

This paper identifies disparities between course grades of both Black and Hispanic university students as compared to White university students and examines the sources that contribute to these grade gaps. Explaining the racial disparities of scholastic performance at all levels of education is of paramount importance. Namely, it allows for the development of policies aimed at reducing the socioeconomic status discrepancies across races, leading to a more egalitarian society. University achievement is important, as a college degree is a prerequisite for many higher paying jobs and more fulfilling careers.

Our regressions eliminate both the Black–White and Hispanic–White course grade gaps in general education university mathematics and science subjects when a student's academic capabilities and socioeconomic status are held constant. We emphasize that our results indicate that programs for academic development of minority students at the high school level or prior would likely be effective. This is because our results demonstrate that minority students who enter college after performing well in high school do just as well as their similarly accomplished White counterparts of similar social economic background. Minority students who performed poorly in high school do just as poorly as their White equivalents. This suggests that long-term policies specifically targeted at improving a minority student's high school—or prior—scholastic performance, coupled with reducing the effects caused by income disparities, should be considered as a means of improving a student's university grades and subsequent chances of finishing a university degree.

Many previous studies analyzed the gaps in student achievement in a variety of standard tests during a student's elementary and high school years across races (e.g., Clotfelter et al. 2009; Fryer and Levitt 2004; Hanushek and Rivkin 2009). Among papers that examined the gap in college, many focus on test scores (Jencks and Phillips 1998) or attendance, planned or actual (see Ginther et al. 2000; Solon et al. 2000 for summary). Course grades are a more direct measure of college-level scholastic performance than standardized tests or attendance because cumulative course grades contribute to advancement within an academic program. Course grades are also a strong predictor of finishing college (Tinto 1993). Finally, university students have self-selected themselves into an environment with a high incentive to succeed as well as a high cost to failing, which potentially suppresses superfluous factors that may contribute to the achievement gap in less heavily incentivized settings such as standardized tests. Findings of a substantial disparity in course grades across races might suggest that the issue of a grade gap is pervasive.

Prior studies examining the gap in college grades include Vars and Bowen (1998), who focus on a group of eleven highly selective universities, and Spenner et al. (2004), who examine the college grade gap at Duke University. Additionally, Clotfelter et al. (2015) study the gap across the University of North Carolina system.[1] Focusing on a single or a few colleges is a useful addition to research based on large survey datasets,

[1] The University of North Carolina system includes 16 colleges across a wide range of selectivity, including the state's flagship university, University of North Carolina at Chapel Hill, as well as several regional and historically black colleges.

because it reveals heterogeneity across schools and avoids errors related to aggregation. On the other hand, the results of any one such study cannot be applied nationally; as summarized by Vars and Bowen (1998, p. 460), “The strength of our database is also its weakness: Because we focus on a high-achieving population, our findings cannot be extrapolated to other institutional settings.” Our paper compliments Vars and Bowen (1998) and Spenner et al. (2004) by examining grades at a more moderately ranked university: Colorado State University—Pueblo (CSU-Pueblo). In that sense, our population is more similar to that in Clotfelter et al. (2015), though there are important differences between CSU-Pueblo and the University of North Carolina system, including higher racial diversity within a campus and students of more modest ability than found at the flagship campuses in the University of North Carolina system. Also, our dataset is somewhat unique in that it consists of actual student-level transcript data. This means that we avoid problems related to self-reporting. Our results cannot be extrapolated to all populations, but they are an observation that may be indicative of more common universities. We believe further research similar to ours is necessary for a more complete picture.

We utilize student-course grade data from CSU-Pueblo, a public four-year university with an enrollment of over 5100 students, in general education biology, chemistry, and mathematics courses held during the fall 2010 to spring 2012 semesters, to analyze Black–White and Hispanic–White grade gaps. University students have generated observable data concerning their prior intrinsic academic capabilities that are available for analytical purposes. Specifically, each student’s standardized test score [i.e., American College Testing (ACT) scores] and high school grade point average (GPA) provide measurable evidence of his previous academic performance.[2] We also include socioeconomic status (measured with zip-code income), high school quality, as well as additional student characteristics and other factors that can potentially contribute to explaining the grade gaps.

As in other papers, our data show that grade gaps exist across races, when not controlling for other characteristics. The following explanations have been offered for university achievement gaps across races: discrimination, varying behaviors, and unobservable differences in the characteristics of matriculating students (Breland 1978; Jencks and Phillips 1998; Spenner et al. 2004). Vars and Bowen (1998) dismiss the first explanation from anecdotal evidence. Concerning the latter point, the authors state (p. 472): “Because there are proportionately more white than black applicants with high SAT (scholastic aptitude test) scores and superior secondary school grades, colleges interested in a diverse school may choose a smaller, more ‘selective’ fraction of white applicants at the given level of SAT scores. As a result, white matriculates may be more likely to be exceptionally strong candidates not captured by SAT. Were this the case, SAT scores would under-predict White GPAs (grade point averages) but a more sophisticated academic bias would not.” After further empirical analysis that controls for academic and personal ratings, beyond SAT scores, they were unable to fully explain the grade gap. As a result, Jencks and Phillips (1998) as well as other

[2] The ACT tests are a standardized test for high school achievement and college admissions in the USA produced by ACT, Inc. The ACT consists of multiple choice subject tests in English, mathematics, reading, and science reasoning with scores ranging from 1 to 36.

researchers (Breland 1978; Spenner et al. 2004) conclude that behavior characteristics and college atmosphere that have adversely affected the academic performance of Black and non-White students may be the primary explanation for the grade gap.

We entirely remove the college grade gap with the high school performance variables and social economic indicators in our data. Through the regression analyses presented in this paper, varying behaviors of university students relative to high school across races are ruled out as a possible source of the grade gaps. This inference contradicts the inferences made in previous research which is based on standardized scores rather than grades (Jencks and Phillips 1998) or investigates highly selective schools (Spenner et al. 2004; Vars and Bowen 1998). It is, however, consistent with the Clotfelter et al. (2015) findings based on the University of North Carolina system.

The remainder of the paper is organized as follows: Section 2 provides a brief discussion of the previous literature. Section 3 describes and summarizes the data used in the analysis, addresses a number of potential selection bias issues, and provides evidence concerning the grade gaps. Section 4 presents and discusses the empirical evidence of the regression-adjusted grade gap estimates as well as the relationship between grades and a number of key covariates. Section 4 also presents analysis concerning the subject-specific grade gaps and explores the issue of varying coefficients across races. Section 5 discusses policy implications from our analysis and presents concluding remarks.

2 Previous literature

The government mandated Coleman report (Coleman et al. 1966) was one of the first studies which documented the existence of an achievement gap between students belonging to minority groups and White students. Since then, a thorough body of research has found supporting evidence identifying gaps in achievement across races and socioeconomic factors such as levels of family and school quality (e.g., Grissmer et al. 1998; Herrnstein and Murray 1996; Lee and Burkam 2002; Phillips et al. 1998).[3] Recently, McDonough (2015) found that Blacks are less upwardly mobile than Whites in achievement outcomes.

Since the socioeconomic status of minority groups is, on average, lower than Whites, the literature examines the portion of the gap that can be explained by socioeconomic factors such as income, parental education levels, and school quality (e.g., Duncan and Magnuson 2005; Yeung and Conley 2008).[4] For example, Cook and Evans (2000) find that both the 1970 and 1988 standardized test scores of Black students were 11.4 and 5.0 % lower than White students, respectively. This gap was specifically due to the differences in school qualities. Orr (2003) finds that family income accounts

[3] Also see Armor (1992), Fryer (2003), Hanushek et al. (2009), Jensen (1998), Krueger and Whitmore (2001), Tienda and Mitchell (2006), Wilson et al. (2006). Rouse et al. (2005) report that children below the poverty threshold are 1.3 times more likely to experience learning disabilities and developmental delays.

[4] In 1999 (2009), the US Census Bureau reports that the median income of Whites, Blacks, and Hispanics was $39,915 ($62,545), $21,423 ($38,409), and $23,431 ($39,730), respectively. All figures are reported in 2009 dollars. See Table 697 of http://www.census.gov/compendia/statab/cats/income_expenditures_poverty_wealth.html.

for approximately 15 % of the achievement gap between Black and White students. Phillips and Chin (2004) find a grade gap of 0.9 points (on a scale of 0–4) between high school students whose mother earned a B.A. compared to students whose mother did not graduate from high school.[5,6] Card and Rothstein (2007) find that the 1998–2002 SAT scores of students at all-black schools planning on attending a university are approximately 250 points lower than students at all-white schools, and Black students' relative achievement is lower in cities with more racially segregated schools. Studying educational attainment, Cameron and Heckman (2001) find that the education level of parents and family income account for the majority of the racial differences of whether or not a student will attend college. On the other hand, Fryer et al. (2015) implemented an experiment in which they rewarded parents when their elementary school children had good school attendance or completed homework. The authors found that the reward increased cognitive and non-cognitive outcomes for White and Hispanic but not for blacks, even when controlling for social economic status.

The National Assessment of Educational Progress (NAEP), a congressional mandated project that has tested the mathematics and reading ability of students in the fourth, eighth, and twelfth grades since 1970, allowed researchers to rigorously explore the achievement gaps between minorities and White students (e.g., Jencks and Phillips 1998; Jensen 1973; Kelly 2009; Perie et al. 2005; Reardon and Galindo 2009). Phillips and Chin (2004) summarize research that persistently finds gaps in the achievement level of minorities (Black, Hispanic, and Latino) relative to White Students, and find one set of the Black (Hispanic)-White grade gap estimates to be 0.70 (0.90) and 0.70 (0.83) standard deviations in fourth grade mathematics and reading tests, respectively.

A number of studies have also examined state-specific standardized tests. Analyzing standard test score data of students that remained in North Carolina's public schools from grades three to eight, Clotfelter et al. (2009) decompose the grade gaps into differences in covariates (e.g., gender, age, and parental education level), school effects (i.e., the type of school district), and an unexplained portion. Of the 0.78 (0.71) standard deviations reading (mathematics) Black–White grade gap, 0.23 (0.24), 0.02 (0.23), and 0.53 (0.45) are contributed to covariates, school effects, and unexplained portions, respectively. Similar results were found for the Hispanic–White grade gap. Using data from California elementary schools, Bali and Alvarez (2004) find both the mathematics and reading test score gaps to be approximately 0.50 and 0.22 standard deviations between Black–White and Hispanic–White students, respectively. Hanushek and Rivkin (2009) find that there is Black–White gap on the mathematics test in the amount of 0.50 standard deviations among Texas elementary school students, and Stiefel et al. (2007) find the mathematics (reading) Black–White gap among New York City public school students to be 0.84 (0.78) standard deviations.

Using data from the Early Childhood Longitudinal Study (ECLS), a nationally representative dataset of test scores and corresponding child information collected via

[5] Cook and Evans (2000) uses National Assessment of Educational Progress data.

[6] Orr (2003) use National Longitudinal Survey of Youth (NLSY79) data. The proportion of the gap explained by income was derived by finding the percentage change in the grade gap estimates from the regression models that did (1.82) and did not (1.54) contain the income variable from columns 2 and 3 of table 2 (p. 293).

interviews, Fryer and Levitt (2004) were the first researchers to explain the grade gap using control variables that serve as proxies for a (p. 451) "broad set of environmental and behavioral factors" in their regression analysis. Consequently, the researchers warn against causal interpretation of the coefficients. By examining the test scores of students entering school, they find the reading and mathematics Black (Hispanic)-White test score gap to be 0.40 and 0.64 (0.43 and 0.73) standard deviations. The regression-adjusted gaps were explained with inclusion of the aforementioned control variables such as the number of children's books in the child's home. Although the majority of the regression-adjusted grade gap can be explained using control variables in kindergarten, Fryer and Levitt (2011) find that the gap increases throughout school. Examining the ECLS data over time from the first to third grade, they substantially reduced the Black–White regression-adjusted grade gap to the point where Black students only score slightly lower (0.099 standard deviations) than White students at the beginning of school (i.e., fall kindergarten); however, the gap increases by almost 0.30 percentiles to 0.382 standard deviations by the spring of third grade.

The Fryer and Levitt studies (2004, 2011) also find that there are little differences in the average values of the school quality variables such as class size and the educational attainment levels of teachers across races and that these variables are not significant predictors of achievement. As a result, policies that alter these school quality variables for Black students in an attempt to reduce or eliminate the achievement gap should be reconsidered. Murnane et al. (2006) study a similar set of data as Fryer and Levitt to examine the extent to which their findings are sensitive to model specification and whether or not the information contained in more detailed data [National Institute of Child Health and Human Development (NICHD) Data] can explain the achievement gap patterns. The authors conclude that Fryer and Levitt's findings (p. 125) "stem from the narrow focus of the tests in the ECLS-K (kindergarten) data set," and a substantial Black–White grade gap exists at the beginning of kindergarten after controlling for similar characteristics as Fryer and Levitt and using the NICHD data.

Some research has examined the university grade gap from the perspective of cumulative GPAs. Betts and Darlene (1999) modeled GPAs of students at the University of California, San Diego, and found that, holding other factors constant including income, ethnic minorities had significantly lower GPAs than white students, but the grade gap became insignificant after controlling for their high school GPAs and SAT scores. Jencks and Phillips (1998, p. 464) examined data from 1989 at private universities and selective liberal art colleges and found that the Black–White grade gap was approximately 0.527 points, which decreased to 0.332 (37.0 %) when a student's SAT scores were included in the regression equations. In the end, they were unable to entirely explain the grade gap. In their regression analysis, the coefficient of SAT scores was 0.11. When race-specific regressions were analyzed, the coefficients of SAT scores were 0.12 and 0.08, respectively. Analyzing the university grade gap of 2001 freshman students attending Duke university, Spenner et al. (2004) found the Black–White grade gap to be approximately 0.39 points. After controlling for SAT scores, parental education, family income, and the availability of educational resources at home, the university grade gap was 0.23 points. Surveying approximately 4000 students from 30 colleges and universities and controlling for SAT scores and high school GPAs in their regression equation, Nettles et al. (1986) found the Black–White to be 0.258 points.

3 The data and potential sample selection issues

The core of our dataset comprises student-course grades and backgrounds, obtained from transcripts and admissions files at CSU-Pueblo. The 2011 edition of *US News & World Report's Best Colleges* labels this public university as a second-tier, national liberal arts college. "Second tier" means that the university ranks between 190 and 250 among national liberal arts colleges. (*US News* does not give further refinement of rankings in the second tier.) To place CSU-Pueblo in context, we note that the following, perhaps better known universities have the same first and third quartile of ACT scores: Auburn University at Montgomery, Western Illinois University, Indiana University South Bend, Old Dominion University, and University of Wisconsin Colleges.[7] CSU-Pueblo is also very diverse in terms of race. Approximately 42 % of the student body are minorities.[8] *US News* ranks CSU-Pueblo as the eighth most diverse national liberal arts school. Admission criteria are based on a combination of high school GPA, ACT or SAT scores.[9]

The CSU-Pueblo student data contain course grades from a number of biology (Human Physiology and Anatomy I and II), chemistry (General Chemistry I and II), and mathematics (Mathematical Explorations, College Algebra, and Introductory Statistics) courses from fall 2010 to spring 2012.[10] We supplement this data with two other variables: high school quality and zip-code income. High school quality is measured by the student's state-specific high school percentile rank, based on mathematics and reading test scores.[11] Zip-code income is obtained from the US Census Bureau.[12] For brevity, we will refer to zip-code income as "income," but this variable serves as a broad proxy for a student's socioeconomic status, not his family income.[13] Each

[7] Source: US Department of Education (https://nces.ed.gov/collegenavigator/?s=AL&l=93&tc=18&xc=20&id=101879#admsns) and collegesimplify.com. There are at least 45 other universities in the databases with the same first and third quartile ACT scores as CSU-Pueblo. The list presented in this paper reflects a handful we believe to be the recognizable to a broad audience. University of Wisconsin Colleges include thirteen smaller campuses in Wisconsin's state university system, some of which offer only two-year degrees; UW-Madison is not included (for a list, see https://www.uwc.edu/about/campuses).

[8] See the following link for CSU-Pueblo race composition: http://www.csupueblo.edu/Grants/currentuniversitydata/Pages/default.aspx.

[9] See the following link for CSU-Pueblo admission criteria: http://www.gocsupueblo.com/SiteCollectionDocuments/CCHEIndexChart.pdf.

[10] All students must pass at least one college-level mathematics course and complete two Natural and Physical science courses with laboratories to obtain their degrees. See p. 62–63 of the 2011/2012 CSU-Pueblo catalog, at http://www.csupueblo.edu/catalog/Pages/default.aspx.

[11] Schooldigger.com provides state-specific high school percentile ranks-based mathematics and reading test score data from their respective states' Departments of Education. We have each student's hometown zip code, and the Web site provides the distance of each high school from a zip code. The specific percentile ranks we use in our analysis are the five- year average percentile ranks of the two schools nearest to the student's hometown zip code.

[12] We use data from 1999 US Census, which is the most recent available.

[13] Previous literature suggests that zip-code income measures a variety of different factors such as peer effects, community factors, observable family characteristics, and parental education levels (e.g., Corcoran et al. 1992; Ginther et al. 2000; Jenks and Mayer 1990; Solon et al. 2000). Manski (1993) posits that the family and neighborhood factors are not separately identifiable, so we take zip-code income as a proxy for socioeconomic status.

student-course observation thus includes the following information: grade received in the course (i.e., A, B, C, D, and F), race, ACT score, high school GPA, high school GPA class percentile rank among her graduating class (henceforth called high school GPA percentile rank), zip-code income, age, gender, hometown state, admit type (e.g., freshman and transfer), university status (e.g., freshman, sophomore, junior, and senior), as well as whether or not the student was a varsity athlete or a first-generation university student.[14]

Our dataset contains 2220 student-course observations in which 1192; 232; 637; and 159 were generated by White, Black, Hispanic, and “Other” students, respectively.[15] The distribution of White, Black and Hispanic students are similar across biology, chemistry, and mathematics courses. Approximately 54, 11, and 26 % of the students across subjects are White, Black, and Hispanic, respectively.

Our dependent variable, grade, is only observed for students who chose to apply to CSU-Pueblo, met their admission criteria, and chose to attend the university. We note that our data omit students who did not attend the university because they did not meet the admission criteria as well as students who met the admission criteria and chose to attend a different school consisting of different characteristics (i.e., size, quality, and location) than CSU-Pueblo. Considering students who did not attend the university are unobservable for analytical purposes, we do not account for the selection criteria with an econometric technique.

In our sample, the respective mean (standard deviation) ACT composite, mathematics, and science scores are 20.9 (3.53), 21.1 (3.63), and 20.3 (3.69). Nationwide, according to the National Center for Education (NCES), in 2010, the means (standard deviations) of these scores were 21.0 (5.2), 21.0 (5.3), and 21.9 (5.1), respectively.[16] Furthermore, the NCES report the mean of ACT composite scores to be 22.3, 16.8, and 18.6 for White, Black, and Hispanic students, respectively. The race-specific mean ACT composite scores in our sample are 21.6, 18.6, and 20.4 for White, Black, and Hispanic students, respectively. Unfortunately, the NCES does not report race-specific ACT score standard deviations. As a result, we cannot determine whether or not the subject-specific sample means are significantly different from the population means reported by the NCES. The sample standard deviations are less than the population’s standard deviations of ACT composite, mathematics, and science scores. The sample means are less than the population means for ACT composite scores for Black and Hispanic students. Therefore, we may have the possibility of a sample selection issue

[14] The data on university course grades were obtained after all courses were complete. The analysis assumes that retests did not vary across races and that the sample is sufficiently large that transcript grades are an accurate representation of student performance. An example of high school GPA class percentile rank calculation is as follows: a student whose high school GPA ranked 67 out of a graduating class consisting of 250 students has a high school GPA percentile rank of $(67/250) \times 100 = 26.8$. High school percentile ranks range from 1 to 100 with 1 being the best and 100 being the worst rank possible. A first-generation student is defined as a student with neither parent having education past high school. See page iii of http://nces.ed.gov/pubs2001/2001153.

[15] There were a small number of students (159) in the data whose racial status is classified as “Other.” These include Asian (53), Indian (13), multi-ethnic (42), and unknown (51). While we use this information in the forthcoming regression analysis, we exclude Other race from the reported summary statistics.

[16] See the following link for NCES ACT score statistics: http://nces.ed.gov/programs/digest/d10/tables/dt10_155.asp. The NCES publishes summary statistics to one decimal place.

relative to the entire population who take the ACT. In interpreting these comparisons, it is important note that the NCES figures reflect the entire population of ACT takers, including those who ultimately did not attended college or who attended universities with different admission criteria, whereas our dataset (by definition) is limited to a group of students who chose to attend CSU-Pueblo. Nevertheless, it appears that the average academic ability of a student at CSU-Pueblo is at least similar to that nationwide.

In contrast, the first quartile of ACT composite scores at Duke university (the subject of Spenner et al.) is 30, nearly two standard deviations above the nationwide *mean* ACT score. Vars and Bowen (1998) also focus on an elite college population. Thus, our paper more closely reflects a moderate, average university. As we emphasize throughout, we believe additional studies using school-level data would paint a more complete picture.

Table 1 shows that at CSU-Pueblo there are approximately twice as many F's as A's. The average numerical grade is 1.82.[17] White students on average earn 0.47 (24.5 %) and 0.16 (8.5 %) grade points greater than Black and Hispanic students, respectively. The summary statistics of the covariates by race are presented in Table 2. There are substantial differences across races on many of these variables, including our key predictor variables: ACT scores, high school GPA, high school GPA percentile rank, and income. White students on average score 13.3 (5.0) and 12.7 (4.5) percent greater than Black (Hispanic) students on both the ACT science and mathematics sections. White students on average earn 12.1 (1.5) percent greater high school GPAs than Black (Hispanic) students. There is a negligible difference among high school GPA percentile rank of White and Black (Hispanic) students and incomes of White and Black students; however, White students on average have an 11.9 percent greater income than Hispanic students. White students on average attend high schools that are ranked 3.5 (6.4) percentile points greater than Black (Hispanic) students. There is a negligible discrepancy in the proportions of age-specific students among White, Black, and Hispanic students. Males comprise 40.6, 43.1, and 40.5 %, student-athletes comprise 9.7, 10.8, and 4.4 %, first-generation students comprise 23.4, 13.8, and 18.1 %, and out-of-state students comprise 0.06, 0.65, and 0.47 % of the White, Black, and Hispanic students in our sample, respectively.

4 Empirical models

4.1 The regression-adjusted grade gaps

We estimate the Black–White and Hispanic–White grade gaps through a series of regressions that include various sets of covariates nested within the following specification:

$$\mathrm{GRADE}_{y,t,c,i} = \mathrm{RACE}_i' \Gamma + \Lambda_c + \Delta_{y,t} + X'_{y,t,c,i} \Theta + F X_{y,t,c} + \varepsilon_{y,t,c,i} \quad (1)$$

[17] These courses have been identified by CSU-Pueblo as high-risk courses (i.e., a 30 % or greater "DFW" rate, where D, F, and W, stand for the letter grades D, F, and withdraw, respectively. CSU-Pueblo offers individual and group tutoring as well as supplemental instruction academic support programs to help students with these historically difficult courses).

Table 1 Grade distributions by race

Race	All	White	Black	Hispanic
Entire sample				
F (0 points)	21.97	19.63	31.44	24.00
D (1 points)	15.26	12.67	21.07	16.76
C (2 points)	29.36	31.36	26.09	27.68
B (3 points)	20.08	20.91	16.05	20.54
A (4 points)	13.32	15.43	5.35	11.03
Mean of numerical grade	1.88	2.00	1.43	1.78
Sample size	3341	1808	299	925
Sample that includes ACT scores and high school GPAs				
F (0 points)	22.70	20.64	30.60	23.86
D (1 points)	15.81	13.09	20.69	16.80
C (2 points)	29.68	31.46	25.86	27.94
B (3 points)	20.09	21.14	17.24	20.72
A (4 points)	11.71	13.67	5.60	10.68
Mean of numerical grade	1.82	1.94	1.47	1.78
Sample size	2220	1192	232	637
Sample that does not include ACT scores and high school GPAs				
F (0 points)	20.52	17.69	34.33	24.31
D (1 points)	14.18	11.85	22.39	16.67
C (2 points)	28.72	31.17	26.87	27.08
B (3 points)	20.07	20.45	11.94	20.14
A (4 points)	16.50	18.83	4.48	11.81
Mean of numerical grade	1.98	2.11	1.30	1.78
Sample size	1121	616	67	288

The p values from the likelihood ratio test of the null hypothesis that the race-specific distributions that does not include ACT scores and high school GPAs are not different than the distributions that do for the All, White, Black, and Hispanic students are 0.698, 0.728, 0.782, and 0.998, respectively

Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

where y, t, c, and i index the year, term (fall, summer, spring), course (Human Physiology and Anatomy I or II, General Chemistry I or II, Mathematical Explorations, Introductory Algebra, or Introductory Statistics), and student. The dependent variable, $\text{GRADE}_{y,t,c,i}$, takes two forms: numerical grades (whose letter grade point values were previously defined and are outlined in Table 1) and a pass–fail binary variable equaling one if the y, t, c, ith grade is an A, B, or C, and zero otherwise, with the latter dependent variable estimating a student's objective function of passing a course rather than earning higher grades. The matrix $RACE_i$ consists of a full set of race variables, with White students as the omitted category. The matrix Λ_c is a set of time-invariant course fixed effects, and the matrix $\Delta_{y,t}$ is a set of year–term fixed effects. The matrix $X_{y,t,c,i}$ denotes the student-specific covariates and $FX_{y,t,c}$ denotes other fixed effects included in the specification. The covariates that vary with time (i.e., age, student level, and

Table 2 Summary statistics by race—student characteristics and fixed effects

Race	All	White	Black	Hispanic
Dependent variable				
Numerical grade	1.8225	1.9413	1.4655	1.7755
	(1.3058)	(1.3087)	(1.2444)	(1.3073)
White	0.5496			
	(0.4977)			
Black	0.1070			
	(0.3091)			
Hispanic	0.2937			
	(0.4556)			
Other	0.0498			
	(0.2176)			
Student characteristics				
Science ACT	21.1222	21.8280	18.9353	20.7363
	(3.5435)	(3.3851)	(3.9554)	(3.3355)
Mathematics ACT	20.2946	20.9010	18.2543	19.9670
	(3.6749)	(3.6988)	(2.9906)	(3.5824)
H.S. GPA	3.2539	3.3244	2.9238	3.2747
	(0.5598)	(0.5567)	(0.4693)	(0.5498)
H.S. GPA rank	0.3598	0.3524	0.4102	0.3468
	(0.2134)	(0.2187)	(0.1796)	(0.2125)
Income	52,041.4343	53,959.8305	54,011.1509	47,523.5950
	(15,157.0302)	(16,574.6996)	(14,388.6875)	(10,733.9992)
H.S. quality	43.4077	45.4810	42.0057	39.0641
	(18.7359)	(18.8992)	(20.9018)	(16.1915)
Age	20.4103	20.4144	20.4914	20.4349
	(1.6422)	(1.5478)	(1.4324)	(1.8643)
Male	0.4099	0.4060	0.4310	0.4050
	(0.4919)	(0.4913)	(0.4963)	(0.4913)
Athlete	0.0802	0.0973	0.1078	0.0440
	(0.2717)	(0.2965)	(0.3107)	(0.2052)
First generation	0.2116	0.2341	0.1379	0.1805
	(0.4085)	(0.4236)	(0.3456)	(0.3849)
Out of state	0.0599	0.0646	0.0345	0.0471
	(0.2374)	(0.2459)	(0.1829)	(0.2120)
Fixed effects				
Admit: freshman	0.8755	0.8884	0.9095	0.8352
	(0.3302)	(0.3150)	(0.2875)	(0.3713)
Admit: transfer	0.0595	0.0587	0.0431	0.0754
	(0.2366)	(0.2352)	(0.2035)	(0.2642)

Table 2 continued

Race	All	White	Black	Hispanic
Admit: other	0.0650	0.0529	0.0474	0.0895
	(0.2466)	(0.2238)	(0.2130)	(0.2857)
Level: freshman	0.5445	0.5243	0.5431	0.5683
	(0.4981)	(0.4996)	(0.4992)	(0.4957)
Level: sophmore	0.3029	0.3062	0.3017	0.3093
	(0.4596)	(0.4611)	(0.4600)	(0.4626)
Level: junior	0.1130	0.1267	0.1207	0.0848
	(0.3166)	(0.3328)	(0.3265)	(0.2788)
Level: senior	0.0290	0.0327	0.0302	0.0235
	(0.1680)	(0.1780)	(0.1714)	(0.1518)
Level: other	0.0106	0.0101	0.0043	0.0141
	(0.1025)	(0.0999)	(0.0657)	(0.1181)
Biology I	0.1899	0.1871	0.1983	0.1947
	(0.3924)	(0.3901)	(0.3996)	(0.3963)
Biology II	0.0821	0.0822	0.0647	0.0958
	(0.2745)	(0.2748)	(0.2464)	(0.2945)
Chemistry I	0.1107	0.1216	0.1078	0.0958
	(0.3138)	(0.3270)	(0.3107)	(0.2945)
Chemistry II	0.0429	0.0453	0.0388	0.0424
	(0.2026)	(0.2081)	(0.1935)	(0.2016)
Mathematical explorations	0.0927	0.0898	0.1078	0.0895
	(0.2900)	(0.2860)	(0.3107)	(0.2857)
Intro. Algebra	0.2356	0.2181	0.3060	0.2370
	(0.4245)	(0.4131)	(0.4618)	(0.4256)
Intro. Statistics	0.2462	0.2559	0.1767	0.2449
	(0.4309)	(0.4365)	(0.3823)	(0.4304)
Fall 2010	0.2130	0.2366	0.2026	0.1837
	(0.4095)	(0.4252)	(0.4028)	(0.3875)
Fall 2011	0.2905	0.2785	0.2888	0.3061
	(0.4541)	(0.4485)	(0.4542)	(0.4612)
Spring 2011	0.2361	0.2383	0.2759	0.2245
	(0.4248)	(0.4262)	(0.4479)	(0.4176)
Summer 2011	0.0041	0.0034	0.0086	0.0031
	(0.0643)	(0.0579)	(0.0926)	(0.0560)
Spring 2012	0.2563	0.2433	0.2241	0.2826
	(0.4367)	(0.4292)	(0.4179)	(0.4506)
Observations	2200	1192	232	637

Standard deviations are in parentheses

Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

student-athlete) are relative to the end of the year–term–course. All continuous covariates are expressed as log difference relative to their respective means. The ACT science (mathematics) score variables are interacted with science (mathematics) course indicator variables equaling one if the *c*th course is science (mathematics) and zero otherwise.

Tables 3 and 4 present the estimation results from the various model specifications derived from the model with the numerical grade and pass–fail dependent variables, respectively. All estimations of the numerical grade model were done with a linear regression, though, to avoid a potential misspecification, categorical variables are modeled with dummy variables. In essence, rather than a linearly modeling categorical variables, we adopt what is effectively a semiparametric partially linear probability model. All models were also estimated with a tobit regression, to account for the inherent censoring of grades which are bounded between 0 and 4. The (unreported) results are similar to results of the linear regression model and support identical conclusions. The tobit results are available from the authors upon request. We use weighted least squares in the pass–fail model to account for the variance heterogeneity inherent in the binary process. In all estimations, the standard errors are clustered by year–term–course.

The respective specifications presented in the various columns of Table 3 and 4 are identical for both numerical grade and pass–fail models. Tables 3 and 4 present the coefficients of all of the variables used in the specification with the exception of the course, year–term, and instructor fixed effects. The estimation results presented in Tables 3 and 4 support similar conclusions concerning the explanation of the grade gaps. We explicitly discuss only the estimates of the grade gaps presented in Table 3, assuming all else constant.

Column 1 of Table 3 shows that the unconditional Black–White and Hispanic–White grade gaps are 0.47 and 0.16 points, respectively. Columns 2 adds the course-specific and year–term fixed effects to the specification in presented column 1, which have no significant effect on grade gaps. Column 3 adds the ACT science and mathematics score variables, which reduces the Black–White grade gap to 0.25 points (by 47 %, relative to column 1) and the Hispanic–White grade gap to 0.10 points (by 41 %).[18] Both the Black–White and Hispanic–White grade gaps are significant at the 0.01 level. Column 4 replaces the ACT science and mathematics variables with the high school GPA and high school GPA percentile rank variables, which reduces the Black–White grade gap to 0.16 points (by 67 %) and the Hispanic–White grade gap to 0.15 (by 11 %). The Black–White grade gap is significant at the 0.10 level, while the Hispanic–White grade gap remains significant at the 0.01 level. Columns 5 and 6 add the income variable to specifications presented in columns 3 and 4, respectively, which results in no substantial changes to the grade gaps. The specification presented in column 7 includes the ACT science, ACT mathematics, high school GPA, and high school GPA percentile rank variables, which reduces the Black–White grade gap to 0.03 points (by 93 %) and the Hispanic–White grade gap to 0.11 (by 35 %). The Black–White grade gap is not significant, while the Hispanic–White grade gap is significant at the 0.10 level. Column 8 adds the income variable to the specification presented in

[18] All grade gap percent changes presented in this paragraph are relative to the unconditional gap listed in column 1.

Table 3 The regression-adjusted estimates of the grade gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black	−0.4717***	−0.4712***	−0.2488**	−0.1553*	−0.2494**	−0.1473	−0.0349	−0.0371	−0.0161	0.0134
	(0.0940)	(0.0940)	(0.0917)	(0.0890)	(0.0923)	(0.0870)	(0.0895)	(0.0885)	(0.0876)	(0.0959)
Hispanic	−0.1617**	−0.1614**	−0.0969	−0.1446***	−0.0937	−0.0814*	−0.1056*	−0.0531	−0.0091	−0.0193
	(0.0706)	(0.0681)	(0.0711)	(0.0519)	(0.0681)	(0.0448)	(0.0551)	(0.0484)	(0.0483)	(0.0473)
Other	−0.3817***	−0.4048***	−0.3204**	−0.1941*	−0.3201**	−0.1775	−0.1537	−0.1417	−0.1026	−0.0423
	(0.1252)	(0.1322)	(0.1259)	(0.1107)	(0.1263)	(0.1134)	(0.1083)	(0.1116)	(0.1162)	(0.1127)
ln ACT science			1.4343***		1.4327***		0.8531***	0.7971***	0.9107***	0.8859***
			(0.1456)		(0.1471)		(0.1264)	(0.1320)	(0.1474)	(0.1665)
ln ACT mathematics			1.8223***		1.8182***		1.1296***	1.0254***	1.1465***	1.1137***
			(0.2854)		(0.2886)		(0.2693)	(0.2726)	(0.2990)	(0.3103)
ln H.S. GPA				2.0617***		2.1137***	2.0005***	2.0520***	2.1029***	2.0927***
				(0.3185)		(0.3136)	(0.3103)	(0.3049)	(0.3452)	(0.3470)
ln H.S. GPA rank				−0.3255***		−0.3494***	−0.2955***	−0.3188***	−0.2966***	−0.3004***
				(0.0502)		(0.0505)	(0.0496)	(0.0507)	(0.0536)	(0.0574)
ln income					0.0285	0.5520***		0.4843***	0.4567***	0.4919***
					(0.0944)	(0.1030)		(0.1038)	(0.1119)	(0.1085)
ln age									0.9039	1.0445
									(0.6718)	(0.6989)
Male									−0.0628	−0.0278
									(0.0853)	(0.0848)
Athlete									0.3417***	0.3595***
									(0.0919)	(0.0914)

Table 3 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
First generation									0.0248	0.0423
									(0.0457)	(0.0484)
Out of state									0.0054	−0.0228
									(0.0883)	(0.0820)
Admit: transfer									−0.1977*	−0.1788*
									(0.1007)	(0.0977)
Admit: other									−0.0309	−0.0215
									(0.0913)	(0.0860)
ln H.S. quality									0.0527	0.0355
									(0.0519)	(0.0505)
Level: sophmore									0.0928	0.0780
									(0.0761)	(0.0765)
Level: junior									0.1629*	0.1433
									(0.0942)	(0.0883)
Level: senior									0.2019	0.2037
									(0.1542)	(0.1516)
Level: other									0.4051*	0.4534**
									(0.2124)	(0.2175)
Constant	1.9372***	2.0568***	2.0206***	1.7828***	2.0208***	1.7704***	1.7782***	1.7680***	1.6652***	1.9381***
	(0.0582)	(0.1095)	(0.1071)	(0.1213)	(0.1071)	(0.1199)	(0.1181)	(0.1172)	(0.1201)	(0.0509)

Table 3 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Observations	2220	2220	2220	2220	2220	2220	2220	2220	2220	2220
Adj. R^2	0.0145	0.0438	0.0869	0.2363	0.0869	0.2473	0.2517	0.2601	0.2720	0.2980
Course FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Year–term FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Instructor FE	No	No	No	No	No	No	No	No	No	Yes

***, **, *, Significance at the 1, 5, and 10 % levels, respectively. Standard errors are in parentheses and are clustered by year–term–course
Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

Table 4 The regression-adjusted estimates of the probability of passing gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black	−0.1324***	−0.1443***	−0.1041***	−0.0766**	−0.1044***	−0.0747**	−0.0446	−0.0451	−0.0426	−0.0264
	(0.0332)	(0.0337)	(0.0355)	(0.0321)	(0.0354)	(0.0316)	(0.0334)	(0.0331)	(0.0306)	(0.0340)
Hispanic	−0.0385	−0.0424*	−0.0311	−0.0474**	−0.0287	−0.0286	−0.0378*	−0.0213	−0.0067	−0.0094
	(0.0232)	(0.0228)	(0.0231)	(0.0212)	(0.0225)	(0.0201)	(0.0211)	(0.0201)	(0.0208)	(0.0209)
Other	−0.1214**	−0.1341**	−0.1158**	−0.1023**	−0.1155**	−0.0987*	−0.0879*	−0.0853*	−0.0713	−0.0448
	(0.0489)	(0.0497)	(0.0502)	(0.0468)	(0.0506)	(0.0489)	(0.0462)	(0.0484)	(0.0481)	(0.0501)
ln ACT science			0.3643***		0.3632***		0.2921***	0.2780***	0.3124***	0.3015***
			(0.0692)		(0.0690)		(0.0619)	(0.0606)	(0.0744)	(0.0820)
ln ACT mathematics			0.3032**		0.2998**		0.2517**	0.2224*	0.2864**	0.2469*
			(0.1155)		(0.1168)		(0.1064)	(0.1088)	(0.1173)	(0.1228)
ln H.S. GPA				0.7910***		0.8123***	0.7862***	0.8065***	0.8085***	0.7618***
				(0.1204)		(0.1186)	(0.1199)	(0.1182)	(0.1347)	(0.1402)
ln H.S. GPA rank				−0.0478**		−0.0594**	−0.0436*	−0.0545**	−0.0465*	−0.0469*
				(0.0211)		(0.0215)	(0.0216)	(0.0221)	(0.0238)	(0.0259)
ln income					0.0211	0.1715***		0.1575***	0.1626***	0.1626***
					(0.0431)	(0.0438)		(0.0442)	(0.0505)	(0.0508)
ln age									0.1984	0.2337
									(0.2710)	(0.2636)
Male									−0.0169	−0.0054
									(0.0348)	(0.0345)
Athlete									0.1426***	0.1514***
									(0.0502)	(0.0486)
First generation									0.0110	0.0116
									(0.0205)	(0.0213)

Table 4 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Out of state									0.0090	−0.0059
									(0.0422)	(0.0418)
Admit: transfer									−0.0848*	−0.0738
									(0.0438)	(0.0455)
Admit: other									0.0190	0.0332
									(0.0373)	(0.0346)
ln H.S. quality									−0.0095	−0.0105
									(0.0248)	(0.0234)
Level: sophmore									0.0595**	0.0519*
									(0.0272)	(0.0262)
Level: junior									0.0842**	0.0791**
									(0.0403)	(0.0371)
Level: senior									0.0883	0.0941
									(0.0665)	(0.0645)
Level: other									0.0880	0.0856
									(0.0984)	(0.0889)
Constant	0.6105***	0.7085***	0.7150***	0.7118***	0.7153***	0.7125***	0.7172***	0.7180***	0.6621***	0.6911***
	(0.0204)	(0.0452)	(0.0417)	(0.0504)	(0.0417)	(0.0494)	(0.0466)	(0.0460)	(0.0487)	(0.0238)
Observations	2220	2220	2220	2220	2220	2220	2220	2220	2220	2220
R^2	0.0087	0.0291	0.0407	0.1137	0.0408	0.1209	0.1213	0.1274	0.1401	0.1556
Adj. R^2	0.0074	0.0233	0.0342	0.1076	0.0339	0.1145	0.1145	0.1202	0.1284	0.1358
Course FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

Table 4 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Year–term FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Instructor FE	No	No	No	No	No	No	No	No	No	Yes

***, **, *, Significance at the 1, 5, and 10 % levels, respectively. Standard errors are in parentheses and are clustered by year–term–course

Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

column 7, which has a negligible effect on the Black–White grade gap but reduces the Hispanic–White grade gap to 0.05 (by 67 %) where it is no longer significant. Column 9 adds the remaining student-specific covariates, the high school quality covariate as well as admit-type and student-level fixed effects (admit type and student-level freshman were omitted to avoid perfect multicollinearity). Column 10 replaces the course and year–term fixed effects with instructor fixed effects resulting in similar coefficients to that of column 9. Both the Black–White and Hispanic–White grade gaps are not significantly different from zero in the specifications presented in columns 9 and 10, which contain the full sets of covariates and fixed effects. The adjusted R^2 increases substantially in the specifications that include the high school GPA and high school GPA percentile rank variables relative to the specifications that do not.

In summary, the grade gaps are explained with ACT scores, high school GPA, high school percentile rank, and income variables. After controlling for a student's academic ability with his ACT score and high school GPA, income is a significant predictor of university grades and contributes to the explanation of the Hispanic–White grade gap. If the average income of Black students were lower than Whites, we might expect the income variable to be required in explaining the Black–White grade gap as well. The grade gaps are not significantly different from zero with the inclusion of the full set of covariates and fixed effects.[19]

4.2 The return, in terms of university grades, of the covariates

Given the heterogeneity of universities, we caution against extrapolating the magnitude of coefficients of the student covariates to explain student grades across all universities. However, our setting presents reasonable estimates of the factors that contribute to grades across a large number of academic institutions that have similar student and university characteristics as CSU-Pueblo.

The coefficients of the ACT scores, high school GPA, high school GPA percentile rank, and income variable change in magnitude across the various specifications presented in columns 2–7 of Table 3. However, the coefficients of these variables are similar across the more saturated model specifications presented in columns 8–10. The coefficients of the covariates beyond ACT scores, high school GPA, high school GPA percentile rank, and income in the specifications presented in columns 9 and 10 are also similar. We discuss the student-covariate coefficients in terms of grades presented in column 10 of Table 3 assuming all else constant.

At their respective means, a five-point increase in ACT science and mathematics scores results in a 0.07 and 0.15 grade point increase in science and mathematics courses, respectively. A 0.5 point increase in high school GPA results in a 0.22 grade point increase in a student's university grades, and a 0.10 point increase in high school

[19] The grade gap analysis holds constant gender and parental education level as factors that can potential influence student performance. The gender and first-generation indicator variable-specific grade gap results are similar to the results presented in Table 3. The grade gaps decreased in magnitude and were not significant when key covariates were included in the specification.

GPA percentile rank decreases a student's grade by 0.22 points.[20] A $25,000 increase in income increases a student's grade by 0.10 points. A student's age, gender or whether not they are an out-of-state student has no significant effect on university grades.

Whether or not a student is a first-generation student has no significant effect on university grades.[21] In 2009, CSU-Pueblo implemented a supplemental instruction (SI) program from a two-year, $100,000 Wal-Mart First Generation Grant Initiative. This SI program specifically targeted improving the performance of first-generation students. The SI program was institutionalized in 2011 and continues to operate. Consistent with previous research that find higher levels of parental education increases a student's academic achievement (Cameron and Heckman 2001; Phillips and Chin 2004), first-generation students are considered to be at risk of non-graduation. As a result, this SI program could potentially contribute to an explanation to what otherwise would be contradictory findings since first-generation students do not earn lower grades than their counterparts. To this end, the initiative could be viewed as a means of improving the grades of at-risk students beyond first-generation students, including minority students, low-performing high school students, and student from lower-income families.

Student-athletes earn 0.36 grade points better than non-student-athletes. Potentially contributing to this finding is the fact that student-athletes have an additional incentive compared to non-student-athletes to achieve high grades. Student-athletes are required to maintain their academic eligibility in accordance with NCAA and CSU-Pueblo compliance by maintaining a 2.0 cumulative GPA. Any student below a cumulative 2.0 GPA is placed on academic probation and can be suspended from competing in varsity competition if they do not raise their cumulative GPA to a least a 2.0.

4.3 The grade gap estimates by subject

Panels 1–3 of Table 5 present the biology, chemistry, and mathematics course grade distributions by race as well as the mean numerical grade and corresponding sample sizes. In biology courses, White students, on average, score 0.72 (40.0 %) and 0.33 (18.4 %) points greater than Black and Hispanic students, respectively. In chemistry courses, White students, on average, score 0.75 (41.5 %) and 0.06 (3.2 %) points greater than Black and Hispanic students, respectively. In mathematics courses, White students, on average, score 0.31 (15.2 %) and 0.12 (5.4 %) points greater than Black and Hispanic students, respectively.

There are only 340 chemistry course observations. Therefore, we empirically estimate the grade gaps across subjects (i.e., science and mathematics) rather than courses (i.e., biology, chemistry, and mathematics). Table 6 presents the estimated subject-specific grade gaps. Each subject, science and mathematics, is estimated with four regression models presented in columns (1–4) and (5–8), respectively. Columns 1 and

[20] Note that lower high school GPA percentile ranks indicate better academic performance than higher ranks.

[21] First-generation student also has no significant effect on university grades in models that did not include additional covariates.

Table 5 Subject-specific grade distributions by race

Races	All	White	Black	Hispanic
Biology courses				
F (0 points)	25.58	22.12	39.34	29.19
D (1 points)	18.98	16.51	29.51	17.30
C (2 points)	31.85	31.78	19.67	35.68
B (3 points)	17.00	20.25	8.20	14.59
A (4 points)	6.60	9.35	3.28	3.24
Mean of numerical grade	1.60	1.78	1.07	1.45
Sample size	606	321	61	185
Chemistry courses				
F (0 points)	25.00	23.12	38.24	23.86
D (1 points)	20.88	16.08	26.47	22.73
C (2 points)	27.06	31.16	26.47	20.45
B (3 points)	15.88	16.08	8.82	20.45
A (4 points)	11.18	13.57	0.00	12.50
Mean of numerical grade	1.67	1.81	1.06	1.75
Sample size	340	199	34	88
Mathematics courses				
F (0 points)	20.72	19.20	24.82	21.15
D (1 points)	12.95	10.57	15.33	15.11
C (2 points)	29.36	31.40	28.47	25.82
B (3 points)	22.68	23.07	23.36	23.90
A (4 points)	14.29	15.77	8.03	14.01
Mean of numerical grade	1.97	2.06	1.74	1.95
Sample size	1274	672	137	364

Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

5 present the coefficients from the specification that contains only race fixed effects. Columns 2 and 6 add the course and year–term fixed effects to the specification presented in columns 1 and 5. Columns 3 and 7 further add the ACT, high school GPA, high school GPA percentile rank, and income variables. Columns 4 and 8 further add the remaining covariates with the exception of the instructor fixed effects.

The Black–White and Hispanic–White science grades gaps are presented in column 1 of Table 6. The inclusion of the course and year–term fixed effects presented in column 2 has a negligible effect on either grade gaps; however, the inclusion of the ACT science score, high school GPA, high school GPA percentile rank, and income variables reduces the Black–White and Hispanic–White grade gaps to 0.25 (65 %) and 0.07 (71 %) points, respectively. The Black–White grade gap is significant at the 0.10 level and the Hispanic–White grade gap not significantly different from zero. Column 4 that contains the full set of covariates further reduces the Black–White grade gap to

Table 6 The regression-adjusted grade gaps by subject

Subject	Science courses				Mathematics courses			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	−0.7140***	−0.6959***	−0.2504*	−0.1972	−0.3169**	−0.3044**	0.1085	0.1333
	(0.1441)	(0.1438)	(0.1264)	(0.1231)	(0.1053)	(0.1056)	(0.1069)	(0.1053)
Hispanic	−0.2277***	−0.2137**	−0.0652	−0.0469	−0.1164	−0.1199	−0.0530	0.0232
	(0.0760)	(0.0772)	(0.0570)	(0.0526)	(0.1032)	(0.1044)	(0.0746)	(0.0742)
Other	−0.3486*	−0.2942	0.0595	0.1299	−0.4450**	−0.4487**	−0.2387	−0.1890
	(0.1838)	(0.2061)	(0.1699)	(0.1866)	(0.1564)	(0.1640)	(0.1330)	(0.1386)
ln ACT science			0.7553***	0.7860***				
			(0.1404)	(0.1711)				
ln ACT mathematics							1.0180***	1.2602***
							(0.2809)	(0.3065)
ln H.S. GPA			1.3601***	1.4115***			2.4534***	2.4702***
			(0.3283)	(0.3781)			(0.4266)	(0.4693)
ln H.S. GPA rank			−0.3845***	−0.3762***			−0.2972***	−0.2731***
			(0.0523)	(0.0612)			(0.0865)	(0.0857)
ln income			0.6430***	0.5128***			0.3761**	0.4120**
			(0.1025)	(0.1002)			(0.1615)	(0.1836)
ln age				0.5095				1.0353
				(0.9692)				(0.9836)
Male				0.1026				−0.1420
				(0.1715)				(0.0854)
Athlete				0.2768**				0.3978***
				(0.1202)				(0.1257)

Table 6 continued

Subject	Science courses				Mathematics courses			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First generation				0.0407				0.0221
				(0.0668)				(0.0548)
Out of state				−0.0499				0.0335
				(0.0985)				(0.1306)
Admit: transfer				0.0290				−0.3940**
				(0.1223)				(0.1367)
Admit: other				0.0791				−0.0550
				(0.1526)				(0.1034)
ln H.S. quality				0.1618**				0.0084
				(0.0687)				(0.0635)
Level: sophmore				0.0563				0.1261
				(0.0943)				(0.1137)
Level: junior				−0.0151				0.3491**
				(0.1197)				(0.1229)
Level: senior				−0.0496				0.4979**
				(0.1910)				(0.1923)
Level: other				0.6933				0.2711
				(0.4565)				(0.2346)
Constant	1.7772***	2.2066***	1.8857***	1.8003***	2.0614***	2.0264***	1.9054***	1.8306***
	(0.0940)	(0.1207)	(0.1527)	(0.1461)	(0.0478)	(0.1425)	(0.1156)	(0.1116)

Table 6 continued

Subject	Science courses				Mathematics courses			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Observations	946	946	946	946	1274	1274	1274	1274
Adj. R^2 0.0307	0.0602	0.2617	0.2730	0.0099	0.0125	0.2513	0.2749	
Course FE	No	Yes	Yes	No	No	Yes	Yes	No
Year–term FE	No	Yes	Yes	No	No	Yes	Yes	No
Instructor FE	No	No	No	Yes	No	No	No	Yes

***, **, *, Significance at the 1, 5, and 10 % levels, respectively. Standard errors are in parentheses and are clustered by year–term–course

Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

0.20, which is not significantly different from zero, and the Hispanic–White to 0.05 points.

The Black–White and Hispanic–White mathematics grade gaps are presented in column 5 of Table 6. The Hispanic–White grade gap is not significantly different from zero. Column 6 includes course and year–term fixed effects and has little effect on the grade gaps. The Black–White grade gap estimate presented in column 7 is not significantly different from zero. The grade gap estimates presented in column 8 are similar to those presented in column 7.

In both the science- and mathematics-specific regressions, the coefficients of the ACT scores, high school GPA percentile rank, and income variables are similar in magnitude and significant to their respective coefficients in the consolidated regression presented in Table 3. The coefficient of the high school GPA variable is 1.4 in the science regressions compared to 2.5 in mathematics, indicating that university mathematics course grades are more sensitive to high school GPA relative to science course grades. Most of the coefficients of student covariates are also similar to those presented in Table 3 with the exception of the high school quality variable in the science regression. The coefficient of the high school quality variable presented in column 4 of Table 6 is 0.16 and significant at the 0.05 level indicating that high school quality is a significant predictor of university science course grades, all else constant.

4.4 The coefficients of covariates by race

Table 7 explores the coefficients of the covariates across races. Columns (1–3), (4–6), and (7–9) present the coefficients of the covariates in different specifications for White, Black, and Hispanic students, respectively. Columns 1, 4, and 7 present the coefficients from the model that includes only ACT science and mathematics, high school GPA, high school GPA percentile rank, and income. Columns 2, 4, and 8 add course and year–term fixed effects. Columns 3, 6, and 9 add the remainder of the covariates with the exception of the instructor fixed effects. We explicitly discuss the coefficients from the most saturated model specifications presented in columns 3, 6, and 9 assuming all else constant.

High school GPA coefficients are 1.92, 2.28, and 2.48 in the White, Black, and Hispanic student-specific regressions, respectively. These coefficients provide strong evidence of the relevance of high school GPA when explaining university grades. The coefficients of high school GPA percentile rank are -0.36, -0.02, and -0.16 in the White, Black, and Hispanic-specific regressions, respectively. Only the coefficient in the White student-specific regression is significant. The t-values in the Hispanic regression is 1.44. The coefficients of the income variables are 0.55, 1.15, and 0.05 in the White, Black, and Hispanic student-specific regressions, respectively, with the White- and Black-specific regressions significant at the 0.01 and 0.05 levels, respectively.

ACT science and mathematics score coefficients of Black students compared to White and Hispanic students illuminate the most striking difference across races. Namely, White and Hispanic students have a positive and highly significant grade return to ACT scores, while Black students have a negligible grade return. Not only do Black students achieve lower results on standardized tests than White students (see

Table 7 The race-specific regression results of numerical grades conditioned on the covariates

Race	White			Black			Hispanic		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln ACT science	0.2155	0.7658***	0.9168***	0.3699	−0.2096	−0.1675	1.7990***	1.5953***	1.9714***
	(0.2616)	(0.2146)	(0.2127)	(0.4354)	(0.4199)	(0.4868)	(0.3311)	(0.3254)	(0.3442)
ln ACT mathematics	0.7928**	0.9714***	1.1009***	−1.1072	0.1836	−0.0290	0.7926*	1.2099***	1.5070***
	(0.3407)	(0.2961)	(0.3159)	(0.8207)	(0.7961)	(0.9200)	(0.3983)	(0.3880)	(0.4145)
ln H.S. GPA	1.7932***	2.0480***	1.9160***	1.7228**	1.9145**	2.2853***	1.7154***	1.9756***	2.4752***
	(0.4512)	(0.4470)	(0.4326)	(0.7212)	(0.8049)	(0.8166)	(0.4367)	(0.4366)	(0.5239)
ln H.S. GPA rank	−0.3757***	−0.3906***	−0.3697***	0.0304	−0.0959	0.0176	−0.2028*	−0.2043*	−0.1599
	(0.0748)	(0.0716)	(0.0705)	(0.2138)	(0.2160)	(0.2328)	(0.1032)	(0.1025)	(0.1107)
ln income	0.5117***	0.5431***	0.5503***	1.1250***	1.1508***	1.2501**	0.1815	0.2174	0.0516
	(0.1487)	(0.1377)	(0.1545)	(0.3769)	(0.3460)	(0.4660)	(0.1714)	(0.1629)	(0.1446)
ln age			−0.2315			0.0822			2.3477***
			(0.6925)			(1.4851)			(0.8030)
Male			−0.1317			0.0564			0.0274
			(0.1020)			(0.1007)			(0.1278)
Athlete			0.3789***			0.0181			0.5739***
			(0.1012)			(0.2139)			(0.1602)
First generation			−0.0650			−0.0219			0.0752
			(0.0470)			(0.1624)			(0.1277)
Out of state			−0.0603			0.0172			0.2209
			(0.1297)			(0.4296)			(0.2018)

Table 7 continued

Race	White			Black			Hispanic		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Admit: transfer			−0.0753			0.2036			−0.4366**
			(0.1592)			(0.4341)			(0.1694)
Admit: other			0.0253			−0.4912			0.0430
			(0.1295)			(0.3278)			(0.1934)
ln H.S. quality			−0.0144			−0.1015			0.2205*
			(0.0823)			(0.1299)			(0.1240)
Level: sophmore			0.0205			−0.0981			0.2457**
			(0.1068)			(0.1686)			(0.1177)
Level: junior			0.2121*			−0.1243			0.1188
			(0.1232)			(0.2307)			(0.1769)
Level: senior			0.3271*			−0.7272*			0.3446
			(0.1818)			(0.3891)			(0.2864)
Level: other			0.2848			1.9443***			0.4335
			(0.2836)			(0.6167)			(0.3799)
Constant	1.8164***	1.4067***	1.4215***	1.6191***	1.8749***	2.0454***	1.7830***	1.4242***	1.1195***
	(0.0777)	(0.1450)	(0.1567)	(0.1659)	(0.2872)	(0.3479)	(0.0932)	(0.2203)	(0.2738)
Observations	1192	1192	1192	232	232	232	637	637	637
Adj. R^2	0.2323	0.2914	0.3039	0.0914	0.2426	0.2644	0.1653	0.2168	0.2609
Course FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year–Term FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

***, **, *, Significance at the 1, 5, and 10 % levels, respectively. Standard errors are in parentheses and are clustered by year–term–course
Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

previous literature section), but also a Black student's score has a statistically insignificant predictive value in terms of university grades. If the experiences/knowledge/skills (henceforth referred to as ACT EKS) that are assessed through the ACT contribute to university grades, we would expect the coefficients of the ACT scores for Black students to be positive, rather than zero. ACT EKS contributes to university grades for both White and Hispanic students. Specifically, the coefficients of ACT science and mathematics variables in the White (Hispanic) students specific regressions are 0.92 and 1.10 (1.97 and 1.51), respectively.

A Black student's ACT EKS might not be fully utilized when earning grades at a university. Fleming (2002) summarizes evidence that environmental factors contribute to the magnitude of the correlation between standardized test scores (i.e., SAT) and university cumulative GPA's for Black students. Namely, the predictive validity (i.e., amount of variance accounted for) between Black students' SAT scores and their cumulative GPAs was 11.5 percent greater in Black colleges compared to White colleges.[22]

Another interpretation that the coefficients of ACT scores for Black students are not significantly different from zero centers on the issue that standardized tests are designed in a racially biased manner that favor White students compared to minorities (Delpit 2006; Ferguson 2003; Rodgers and Spriggs 1996). If this were the case, Black students could potentially possess the ACT EKS, although it was not reflected in their ACT scores. Consequently, the ACT variables in the Black student-specific regressions would not be significant predictors of university course grades. The fact that the coefficients of ACT scores for Black students are not significantly different from zero could support the notion that the ACT is designed in a racially biased manner. This notion would be strengthened if a similar trend was found among Hispanic students as well.

Differences in coefficients across races concern both the student-athlete and high school quality variables. White and Hispanic student-athletes earn 0.38 and 0.57 higher grade points than White and Hispanic non-student-athletes, respectively. Black student-athletes do not earn significantly higher or lower grades than Black non-student-athletes. There could be a number of reasons for this finding including heterogeneity in the sports that students participate in across races. The coefficient of the high school quality variable is only significant in the Hispanic-specific regression. Specifically, it is 0.22 and significant at the 0.10 level. Discrepancies in the coefficients across races of the student-athlete and high school quality variable across races have no impact on whether or not the grade gaps are fully explained.

4.5 Robustness of the grade gap estimates

Table 8, which presents only the grade gap estimates, uses more flexible regression specifications compared to the specification presented in Table 3 by allowing the ACT scores, high school GPA, high school GPA percentile rank, and income variables, to

22 Breland (1978) and Morgan (1990) find a strong positive correlation between SAT scores and GPAs, and Boyd (1977) found no relationship between SAT scores and GPA. Also see Fleming and Garcia (1998) for a summary of the literature pertaining to the relationship between SAT scores and GPA's.

Table 8 The regression-adjusted estimates of the grade gaps, conditioned on race indicator variables and key covariate interactions

	(1)	(2)	(3)	(4)
Black	−0.1987	−0.1732	−0.1694	−0.1492
	(0.1459)	(0.1411)	(0.1389)	(0.1443)
Hispanic	−0.0348	−0.0095	0.0108	−0.0093
	(0.0623)	(0.0569)	(0.0594)	(0.0549)
Other	−0.0518	−0.1776	−0.1683	−0.1640
	(0.1495)	(0.1574)	(0.1534)	(0.1467)
Constant	1.8178***	1.9488***	1.8785***	1.5848***
	(0.0785)	(0.0931)	(0.0925)	(0.0453)
Observations	2220	2220	2220	2220
Adj. R^2	0.2074	0.2644	0.2699	0.2699
Race–key covariate interactions	Yes	Yes	Yes	Yes
Course FE	No	Yes	Yes	No
Year–term FE	No	Yes	Yes	No
Full covariates and other FE:	No	No	Yes	Yes
Instructor FE	No	No	No	Yes

***, **, *, Significance at the 1, 5, and 10 % levels, respectively. Standard errors are in parentheses and are clustered by year–term–course. Race–key covariate interactions are White, Black, Hispanic, and Other race indicator variables each interacted the ACT science, ACT mathematics, high school GPA, high school GPA percentile rank, and income variables

Source CSU-Pueblo transcript data, fall 2010 to spring 2012; Schooldigger.com; and US Census Bureau, 1999

vary across races as a means of checking the robustness of the grade gap estimates. We include the more flexible specifications because the coefficient of ACT scores was not significant in the Black-specific regressions presented in columns (4–6) of Table 7. If Black's ACT return is truly zero, then a model that treats Blacks and Whites as equal has the potential to misstate the returns to ACT for both Blacks and Whites. Allowing the interaction terms allays this concern.

In addition to the Black, Hispanic, and Other race indicator variables, the specification presented in column 1 includes White, Black, Hispanic, and Other race indicator variables each interacted the ACT science, ACT mathematics, high school GPA, high school GPA percentile rank, and income variables. Columns 2 adds the course-specific and year–term fixed effects to the specification in presented column 1. Column 3 adds the remainder of the covariates with the exception of the instructor fixed effects. Column 4 replaces the course-specific and year–term fixed effects with the instructor fixed effects while keeping all other covariates in the specification presented in column 3. None of the grade gap estimates presented in Table 8 are significantly different from zero. Across the four specifications, the Black–White grade gap ranges from 0.21 to 0.28 points. Each of the four Hispanic–White grade gaps are close to zero.

Propensity score matching that accounts for differences in covariates across races supports the linear model's conclusions. Following the work of Rosenbaum and Rubin

Table 9 Treatment model results

	(1)	(2)	(3)
Base student group	Non-white	Black	Hispanic
White effect	0.0227	0.1797	0.0532
	(0.0644)	(0.1700)	(0.0716)
Propensity score results based on a logistic regression of a White on covariates			
ln ACT science	3.0548***	5.7180***	2.6837***
	(0.4129)	(0.6962)	(0.4996)
ln ACT mathematics	1.6331***	3.9677***	1.2745***
	(0.3567)	(0.6574)	(0.4078)
ln H.S. GPA	2.7420***	6.7183***	1.3373***
	(0.4327)	(0.7439)	(0.4893)
ln H.S. GPA rank	0.4277***	0.9512***	0.3163***
	(0.0902)	(0.1599)	(0.0986)
Constant	0.3364***	2.4029***	0.7273***
	(0.0532)	(0.1209)	(0.059)
Observations	2220	1424	1829

***, **, *, Significance at the 1, 5, and 10 % levels, respectively. Standard errors are in parentheses

(1983), the average treatment effect was estimated by imputing the outcomes of similar students from different races based on the probability of being White. Logit regression models that account for ACT scores, high school GPA, and high school GPA percentile rank were used to estimate the probability that the student was White compared to non-White (entire sample), White compared to Black (White and Black student sample), and White compare to Hispanic (White and Hispanic student sample). Pairs of White and non-White individuals were then matched based on the estimated probabilities (i.e., propensity scores). These matched pairs composed the sample for comparing grade outcomes. Table 9 presents the treatment model results. The top panel presents the average White effects, and the bottom panel presents the propensity score logit regression results. The resulting White effect, in terms of higher grades, is approximately 0.02 compared to non-White students, 0.18 compared to Black students, and 0.05 compared to Hispanic students, with none of the estimates significant at the 0.10 level. All of the estimates in the logit regression are significant at the 0.01 level. Higher (better) ACT scores and high school grade performance increase the likelihood that a student is White across student treatment groups.

Matching on the propensity score in the Black and White sample results in a reasonably well-balanced sample. The differences of the covariates in the matched sample are less 0.25 for all covariates. Figure 1 presents the balance density plots for the White–Black comparison, for both the raw and matched data.[23] The raw comparison shows that there are few White students with a high propensity to be Black. The right-hand side of the figure shows that the pool control units resulted in a well-balanced

[23] This approach is similar to that in Chapter 10 of Gelman and Hill (2006) figure 10.7.

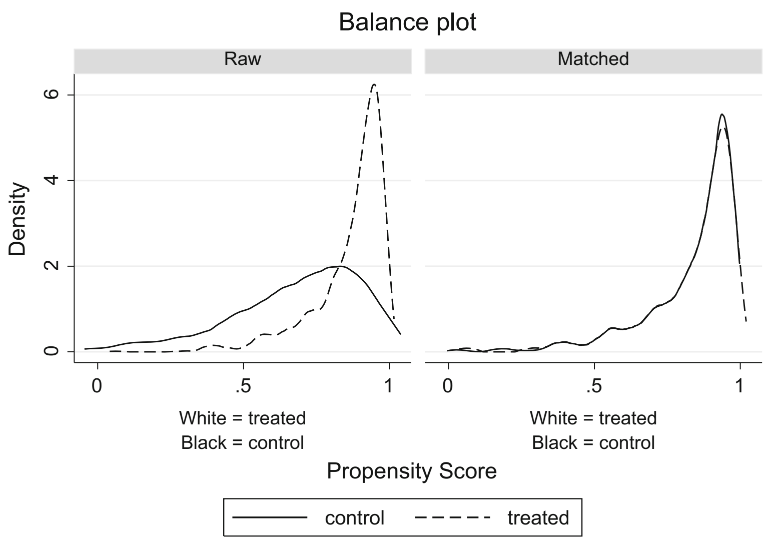

Fig. 1 Black–White balance plots

sample and provided good matches, except perhaps at the tails.[24] A larger sample of Black students would provide more robust diagnostics.

As with the Black–White sample, matching on propensity scores for White and Hispanic and White and non-White create balanced samples. In the White and Hispanic sample, the standardized difference in covariates in the matched data is less than 0.03 for all covariates. In the White and non-White sample, these differences are all less than 0.06. Figure 2 depicts the balance density plots for the White–Hispanic comparison; the figure shows that the matched sample is well balanced with good overlap. The matched propensity densities are also almost identical for the White versus non-White sample, so we omit the figure for brevity.

Finally, we examined whether repeated student-course observations impact the grade gap results. Assuming students retake courses during their current or subsequent academic year, repeated course observations during the 2010 academic year are unobservable. There were 1748 observations generated from fall 2011 to spring 2012. One thousand five hundred and twenty-seven observations (812 White, 155 Black, 445 Hispanic, and 115 Other) were generated from students during their first course attempt, and 221 (98 White, 30 Black, 75 Hispanic, and 18 Other) were generated by students during their second or third time completing a course. The likelihood ratio test does not reject the null hypothesis at the 0.05 level (p-value =0.088) that the race distribution across first and multiple course student observations are the same. There are also no significant grade gaps based on only the 1748 single student-course observations and the specifications presented in columns (7–9) of Table 3. Based

[24] We repeated the Black–White analysis using a stricter tolerance on the overlap assumption by removing propensity scores within 0.025 and 0.10 of zero or one. Trimming with these tolerances removes 110 and 693 observations, respectively, and results in even smaller Black–White differences than with no trimming (though on different subsets of the population).

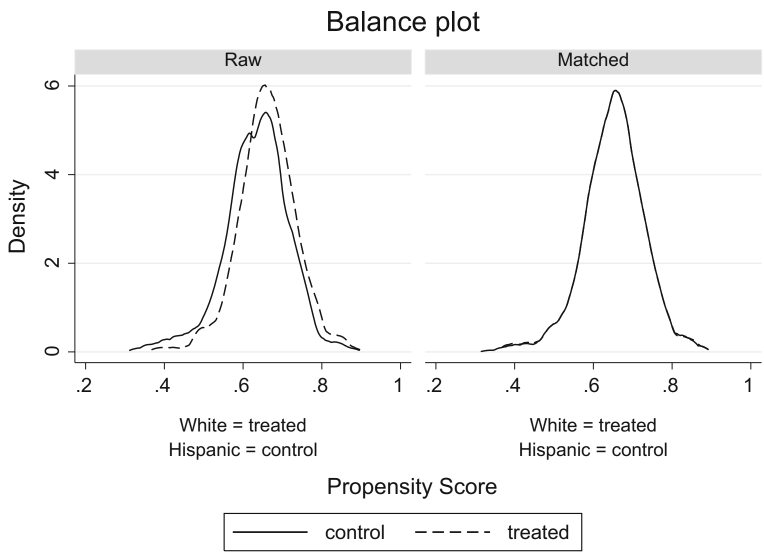

Fig. 2 Hispanic–White balance plots

on the specification presented in column (7), the grade estimates (standard errors) for Black, Hispanic, and Other students are −0.0853(0.0626), −0.001(0.0971), and −0.1430(0.1257).

5 Discussion and conclusion

In this paper, we examine the grade discrepancies of Black and Hispanic students compared to White students at a moderately ranked university. This setting for analysis is relatively unique, compared to most of the previous literature that examines race achievement gaps using standardized tests of elementary or high school students. University students have self-selected themselves into an environment that has a high incentive to succeed and a high cost of failure, creating an atmosphere that would suppress any superfluous factors that contribute to the grade gap in less heavily incentivized settings.

To the extent previous literature has looked at the college achievement gap, it has been primarily the likelihood of attending college or other self-reported measures of college success. Transcript grades are a more accurate predictor of college success than test scores or attendance (planned or actual), and our actual transcript data are not prone to errors innate in self-reported data. Moreover, there may be heterogeneity across schools not captured in national data. With respect to the other studies that do measure college grade gaps, we note that they are at highly selective schools. (The exception is Clotfelter et al. (2015), who also study modestly ranked universities, and whose results are the most similar to ours.) The student populations at more elite schools are different than the more moderately ranked university in this paper. We suspect that other, unmeasured factors such as study skills might differ between minorities and Whites who attend highly selective universities. Moderately ranked schools like the one study we are relevant to substantially more individuals, because many more students attend schools similar to CSU-Pueblo than elite universities.

Although there are significant Black–White and Hispanic–White grade gaps, it appears that Black and Hispanic students that are equally prepared for college as Whites do as well. Since minorities with similar academic capabilities to Whites perform equally well, our results suggest that earlier intervention, designed to increase high school performance, would help minority students close the achievement gap in college. Placed in context of Fryer and Levitt (2011, showing achievement gap among young children growing over time), this intervention should perhaps start early and continue.

While the majority of the grade gaps can be accounted for with a student's academic capabilities, socioeconomic status remains an important factor for explaining the gaps. Fully explaining the university grade gaps with a student's previous academic knowledge and socioeconomic status provides evidence to the contrary that varying behaviors of university students relative to high school students across races as a possible source of the grade gaps.

We view our results as an informative and important supplement to existing research that focuses on survey data or highly ranked universities. Because our results rely on one university, we view it necessary that future research examines students at other moderately ranked universities.

Acknowledgements The authors would like to thank Armando Rodriguez, Robert Rosenman, Hajime Hadeishi, Seth Sacher, Luke Olson, and Christopher Lutz for helpful suggestions.

References

Armor D (1992) Why is Black educational achievement rising? Public Interest 108:65–80
Bali V, Alvarez R (2004) The race gap in student achievement scores: longitudinal evidence from a racially diverse school district. Policy Stud J 32(3):393–415
Betts JR, Darlene M (1999) The determinants of undergraduate grade point average: the relative importance of family background, high school resources, and peer group effects. J Hum Resour 34(2):268–293
Boyd W (1977) SATs and minorities: the dangers of underprediction. Change 77
Breland H (1978) Population validity and college entrance measures. Educ Test Serv
Cameron S, Heckman J (2001) The dynamics of educational attainment for black, hispanic, and white males. J Political Econ 109(3):455–499
Card D, Rothstein J (2007) Racial segregation and the Black–White test score gap. J Public Econ 91(11):2158–2184
Clotfelter C, Ladd H, Vigdor J (2009) The academic achievement gap in grades 3 to 8. Rev Econ Stat 91(2):398–419
Clotfelter C, Ladd H, Vigdor J (2015) Public universities, equal opportunity, and the legacy of Jim Crow: evidence from North Carolina. NBER Working paper 21577
Coleman J, Campbell E, Hobson C, McPartland J, Mood A, Weinfeld F, York R, Gardner J (1966) Equality of education. Opportunity. US Government Printing Office, Washington, DC
Cook M, Evans W (2000) Families or schools? Explaining the convergence in White and Black academic performance. J Labor Econ 18(4):729–754
Corcoran M, Gordon R, Laren D, Solon G (1992) The association between men's economic status and their family and community origins. J Hum Resour 27(4):575–601
Delpit L (2006) Other people's children: cultural conflict in the classroom. The New Press, New York City

Duncan G, Magnuson K (2005) Can family socioeconomic resources account for racial and ethnic test score gaps? Future Child 15(1):35–54
Ferguson R (2003) Teachers' perceptions and expectations and the Black–White test score gap. Urban Educ 38(4):460–507
Fleming J (2002) Who will succeed in college? When the sat predicts Black students' performance. Rev High Educ 25(3):281–296
Fleming J, Garcia N (1998) Are standardized tests fair to African Americans? Predictive validity of the SAT in black and white institutions. J High Educ 471–495
Fryer R (2003) An economic approach to cultural capital. Manuscript, University of Chicago
Fryer R, Levitt S (2004) Understanding the Black–White test score gap in the first two years of school. Rev Econ Stat 86(2):447–464
Fryer R, Levitt S (2011) The Black–White test score gap through third grade. Am Law Econ Rev 8(2):249–281
Fryer R, Levitt S, List J (2015) Parental incentives and early childhood achievement: a field experiment in chicago heights. NBER Working Paper No 21477
Gelman A, Hill J (2006) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York
Ginther D, Haveman R, Wolfe B (2000) Neighborhood characteristics as determinants of children's outcomes: How robust are the results? J Hum Resour 35(4):603–642
Grissmer D, Flanagan A, Williamson S (1998) Why did the black–white score gap narrow in the 1970s and 1980s. The Black–White test score gap, pp 182–228
Hanushek E, Rivkin S (2009) Harming the best: How schools affect the Black–White achievement gap. J Policy Anal Manag 28(3):366–393
Hanushek E, Kain J, Rivkin S (2009) New evidence about brown vs. board of education: the complex effects of school racial composition on achievement. J Labor Econ 27(3):349–383
Herrnstein R, Murray C (1996) The bell curve: intelligence and class structure in American life. Free Press, New York
Jencks C, Phillips M (1998) The Black–White test score gap. Brookings Inst Press, Washington, DC
Jenks C, Mayer SE (1990) Consequences of growing up in a poor neighborhood. In: Lynn L, McGeary M (eds) Inner-city poverty in the United States. National Academy Press, Washington, DC
Jensen A (1973) Educability and group differences. Free Press, New York
Jensen A (1998) The g factor: the science of mental ability. Praeger Westport, CT
Kelly S (2009) The Black–White gap in mathematics course taking. Sociol Educ 82(1):47–69
Krueger A, Whitmore D (2001) Would smaller classes help close the black–white achievement gap? 451, Industrial Relations Section, Princeton University
Lee V, Burkam D (2002) Inequality at the starting gate: social background differences in achievement as children begin school. Economic Policy Institute, Washington, DC
Manski CF (1993) Identification of endogenous social effects: the reflection problem. Rev Econ Stud 60:531–542
McDonough I (2015) Dynamics of the black–white gap in academic achievement. Econ Educ Rev 47:17–33
Morgan R (1990) Predictive validity within categorizations of college students: 1978, 1981, and 1985
Murnane R, Willett J, Bub K, McCartney K (2006) Understanding Trends in the black–white achievement gaps during the first years of school. Brookings-Wharton Papers on Urban Affairs pp 97–135
Nettles M, Thoeny A, Gosman E (1986) Comparative and predictive analyses of Black and White students' college achievement and experiences. J High Educ 289–318
Orr A (2003) Black–White differences in achievement: the importance of wealth. Sociol Educ 281–304
Perie M, Moran R, Lutkus A (2005) NAEP 2004 trends in academic progress: three decades of student performance in reading and mathematics. National Center for Education Statistics, US Department of Education, Institute of Education Sciences
Phillips M, Chin T (2004) School inequality: what do we know. Social Inequal 467–519
Phillips M, Crouse J, Ralph J (1998) Does the Black–White test score gap widen after children enter school?. The Booking Institute, Washington, DC
Reardon S, Galindo C (2009) The Hispanic–White achievement gap in math and reading in the elementary grades. Am Educ Res J 46(3):853–891
Rodgers W, Spriggs W (1996) What does the AFQT really measure: race, wages, schooling and the AFQT score. Rev Black Political Econ 24(4):13–46

Rosenbaum P, Rubin D (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70(1):41–55

Rouse C, Brooks-Gunn J, McLanahan S (2005) Introducing the issue. Future Child 15(1):5–14

Solon G, Page ME, Duncan GJ (2000) Correlations between neighboring children in their subsequent educational attainment. Rev Econ Stat 82(3):383–392

Spenner K, Buchmann C, Landerman L (2004) The Black–White achievement gap in the first college year: evidence from a new longitudinal case study. Res Soc Stratif Mobil 22:187–216

Stiefel L, Schwartz A, Gould Ellen I (2007) Disentangling the racial test score gap: probing the evidence in a large urban school district. J Policy Anal Manag 26(1):7–30

Tienda M, Mitchell F (2006) Multiple origins, uncertain destinies: Hispanics and the american future. National Academy Press, Washington, Dc

Tinto V (1993) Leaving college: rethinking the causes and cures of student attrition, 2nd edn. University of Chicago Press, Chicago

Vars F, Bowen W (1998) Scholastic aptitude test scores, race, and academic performance in selective colleges and universities. The Black–White test score gap pp 457–79

Wilson D, Burgess S, Briggs A (2006) The dynamics of school attainment of England's ethnic minorities

Yeung W, Conley D (2008) Black–White achievement gap and family wealth. Child Dev 79(2):303–324