Math Achievement, College Success and Financial Wellness: Findings from the National College Databases

Jack Marley-Payne, Andrew Davidson, Philip Dituri

College attendance can confer many advantages within the US when it comes to financial wellness. However, the benefits are not uniform and attending college also comes with significant costs. One aspect of this is the relationship between the academic achievements of incoming students and their experiences after attending college. Combining data from the US Department of Education Integrated Postsecondary Education Data System (IPEDS), and the US Department of Education College Scorecard (Scorecard) allows us to investigate this relationship. We find that there is a significant positive relationship between math level and a range of measures of college success: future earnings, graduation rate, and loan repay rate. This suggests that when looking to college attendance as a tool for improving financial wellness, we must pay careful attention to prior math achievement.

Introduction

College attendance can confer many advantages within the US when it comes to financial wellness. However, the benefits are not uniform and attending college also comes with significant costs. Understanding the various factors that determine how the benefits of college are distributed is of vital importance. It will allow incoming students to have a better sense of what to expect, given their circumstances, and also may aid policy makers and financial counsellors in making sure future students are as well prepared as possible for college success.

One aspect of this is the relationship between the academic achievements of *incoming* students and their experiences after attending college. Combining data from the US Department of Education Integrated Postsecondary Education Data System (*IPEDS*), and the US Department of Education College Scorecard (*Scorecard*) allows us to investigate this relationship.

Previous research has shown that both math achievement and math education level are correlated with a range of positive outcomes, including further educational achievement, future salary, and a range of beneficial financial behaviors. This study uses the aforementioned data sets to further explore this issue, creating a model that predicts the impact of math level on college success.

We find that there is a significant positive relationship between math level and a range of measures of college success: future earnings, graduation rate, and loan repay rate. This suggests that when looking to college attendance as a tool for improving financial wellness, we must pay careful attention to prior math achievement.

Background

In the US, those with college degrees earn significantly more, on average, than those without. A report from the Urban Institute states that "Even after accounting for paying higher taxes (and for paying for college), postsecondary education pays off for most people." (Baum 2014, p. 9) The report calculates that the average difference in annual earnings between college graduates and high school graduates is around $20,000, in line with the general research consensus.[1] This has led to policy proposals aimed at increasing college attendance with the goal of reducing poverty and increasing mobility for future generations in the US (Brookings 2019). For young people, this means that one of key pieces of financial advice they receive is to complete a college degree in order to ensure future financial stability

However, a range of research has cast doubt on the idea that college attendance is an equally good financial option for everyone. As the Urban Institute report goes on to note "there is considerable variation in outcomes and not every college graduate earns more than every high

[1] See also, for example, Abel (2014).

school graduate." (Baum 2014, p. 9) A key issue is that many college attendees fail to complete their degree and so lack the desired qualification for improved employment (Bound et. al. 2010). Hardy and Marcotte (2018) find that college attendees from low income families are particularly at risk of failing to complete their degree. Further, Morgan and Steinbaum (2018) show that many college attendees are significantly burdened by their student debt.

A key point, stemming from these considerations, is that outcomes of college attendance vary dramatically. Despite the higher average, many college attendees end up worse off financially than the average high school graduate, after factoring in the various costs of college attendance. For any prospective college attendee, there is huge uncertainty as to the outcome of college attendance. Building on this point, Balakrishnan and Cynamon (2018) demonstrate that the value of college is significantly lower than usually stated once one factors in risk aversion.

Because of this, it's important to look at what factors influence the likelihood of successful results of college attendance for potential students. In what follows, we'll focus on how prior mathematics achievement affects outcomes. Research demonstrates that participating and succeeding in mathematics courses is one of the key predictors of financial success in the US today. As Arcidiacono puts it:

> Students who choose natural science majors earn substantially more than humanities majors. In fact, economists have reported that differences in returns to majors are much larger than differences in returns to college quality… "while sending your child to Harvard appears to be a good investment, sending [them] to your local state university to major in Engineering, to take lots of math, and preferably to attain a high GPA, is an even better private investment." (Arcidiacono 2004 p. 252).

To take advantage of this connection, students must come out of high school with a strong foundation in mathematics if they are to successfully complete the kind of math intensive major that tends to lead to higher future earnings. Additionally, research from Goodman (2019) suggests that simply taking additional math courses in high school leads to increased financial benefits, especially for disadvantaged demographic groups.

Further, Cole, Paulson and Shastry (2016) have found that taking additional mathematics courses increases a range of positive financial outcomes, not limited to increased income – for example, greater financial market participation, higher investment income, and better credit management. As well as college success specifically, it's well established that higher levels of math ability are correlated with great financial stability (Banks & Oldfield 2007, Lusardi 2012, Marley-Payne, Dituri & Davidson ms).

Purpose

The purpose of our study is to build on this research by examining the relationship between math score and college success across all colleges with selective admissions (henceforth "selective colleges") nationwide in recent years. In particular, it looks specifically at how successful *colleges* are, in terms of the success rates of their students across a range of measures of success.

Hypothesis

Our hypothesis is that there is a significant positive relationship between the math level of incoming students at a college, and the success rates for students at the college. However, as mentioned above, research shows that there is a significant relationship between family income and college success (Hardy and Marcotte 2018; see also Restuccia & Urrutia 2002). In addition, we know that there is a significant relationship between family income and math achievement (Butcher 2017). One might argue, then, that once we adjust for family income, any significant relationship between math achievement and college success disappears. We will have to respond to this line of argument in order to uphold our hypothesis.

Data and Methodology

The data we will use comes from two sources: The Integrated Postsecondary Education Data System (IPEDS) and College Scorecard. IPEDS is a collection of data on post-secondary education institutions collected by the National Center for Education Statistics. College Scorecard is also a source of data on post-secondary institutions, primarily intended for use by students evaluating potential enrollment options. We make use of the Urban Institute's Education Data Portal to integrate these datasets effectively.[2]

Together they provide a wealth of information on colleges and their attendees – though all information is aggregated at the college level annually, rather than providing individual student data.

We will look at three distinct measures of college success:

1) Median earnings
2) Graduation rate
3) Repayment rate.

Earnings is the usual measure of how worthwhile college is, so we include a measure of it here. The datasets provide information on median earnings for attendees of college in a given year a certain number of years out from their enrollment. For example, it might say that the median earnings for students at Big City University who enrolled in 2008 was $40,000 in 2014, six years

[2] https://educationdata.urban.org/

after enrollment. Data was available for median earnings at 6, 8 and 10 years from enrollment, and for cohort years 2001-2008.

As discussed above, there is risk involved in college attendance which makes median earnings alone an inadequate measure of college success. Failing to graduate college is a significant risk, since one then lacks the qualification likely to lead to financial security – so probability of graduation is an important statistic for incoming students. The measure we use here is the percentage of students in a year group at a college who graduate in 150% of the recommended time – i.e. completing a four-year degree in six years, or a two-year degree in three years. Data here is available for cohort years 2001-2014.

Finally, since cost of living varies based on location and lifestyle, repayment rate offers additional insight on whether college has provided students with sufficient financial security – if a graduate is failing to make their loan payments, then their post college income is not sufficient. This shows why probability of meeting loan repayments is a significant measure. The dataset provides information on the repayment rate of a college cohort at a certain number of years after they have entered into repayment. The years after repayment for which data is collected are 1, 3, 5 and 7. The cohort years for which data is available are 2006-2013.

Our goal is to look at the influence of math ability on these outcomes, and the datasets allow us to measure this as they provide information on the SAT and ACT scores of incoming students. In particular, they break the results down by category – math score and reading score. For the reasons discussed above, we will be focused primarily with math score, but we will also add the reading score to our model for the purposes of comparison. As independent variables, we use 25th percentile SAT/ACT math score and SAT/ACT reading score for a college's annual cohort (the data set only provides 25th and 75th percentile score, not median). Only colleges with selective admissions provide information on SAT/ACT scores of incoming students. This means we are only looking at about 20% of colleges in the US in our model. In addition, not all colleges provided data on different outcome variables so those included in the regression analysis varies between models.

In building our models, we aim to control for other significant factors relevant to our outcomes. As discussed above, one factor that has a major influence on future success is family income, which we must account for in our model. The dataset includes median family income for a college's annual cohort, so we include this as an independent variable in our model. We also include dummy variables for key demographic factors – see appendices 1-3 for a complete list.

We create separate models with SAT and ACT scores as the independent variables for academic achievement. Since graduation and repayment rates must be between 0 and 1, we use logistic regression for modelling these outcomes, while we use linear regression for the model of average earnings. For, the earnings and repayment models, we add dummy variables for years since enrollment. This gives us the following:

1. $Median\ Earnings = \beta(Math\ Score,\ Reading\ Score,\ Years\ Out,\ Controls)$
2. $Graduation\ Rate = \sigma(\beta(Math\ Score,\ Reading\ Score,\ Controls))$
3. $Repayment\ Rate = \sigma(\beta(Math\ Score,\ Reading\ Score,\ Years\ Out,\ Controls))$

Here β is a vector of linear coefficients and σ is the logistic function. Note that each of the three models has two version, one using SAT score and one using ACT scores. We can now state our hypothesis precisely: in each of our three models, there will be a positive statistically significant coefficient associated with Math Score.

Results

The key results of our models, displayed in table 1, show that there is a statistically significant, positive relationship between incoming math score (both SAT and ACT) and each of the three outcome variables, in line with our hypothesis – in each case, the coefficient is significant at 1%. Full regression results can be found in appendices 1-3.

	SAT Scores	ACT Scores
Median Earnings		
Math Score	6572.80 ***	5250.12 ***
Reading Score	-3023.63 ***	-1542.83 ***
Median Family Income	3022.65 ***	2181.89 ***
N	17311	13056
Graduation Rate		
Math Score	0.39 ***	0.20 ***
Reading Score	0.10 ***	0.23 ***
Median Family Income	0.29 ***	0.28 ***
N	13464	10806
Repayment Rate		
Math Score	0.16 ***	0.12 ***
Reading Score	0.04	0.04
Median Family Income	0.31 ***	0.31 ***
N	32859	28320

All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 1

It's also worth noting that for all models, apart from graduation rate with ACT score, the positive effect for math score is larger than for reading score. Indeed, for repayment rate there is no statistically significant relationship with reading score, while for earnings, the relationship is negative. As expected, there is also a statistically significant positive correlation between family income and college success displayed in all of the models. However, even with this relationship included in the model, the significant effect of math score remains, contrary to the potential objection to our hypothesis considered above.

The numerical regression results are complemented by *figures 1-3*, which provide a graphical representation of the relationship between math score and each of the outcome variables. The shape of the data in the scatter plots suggests that the regression methods used in our models are appropriate.

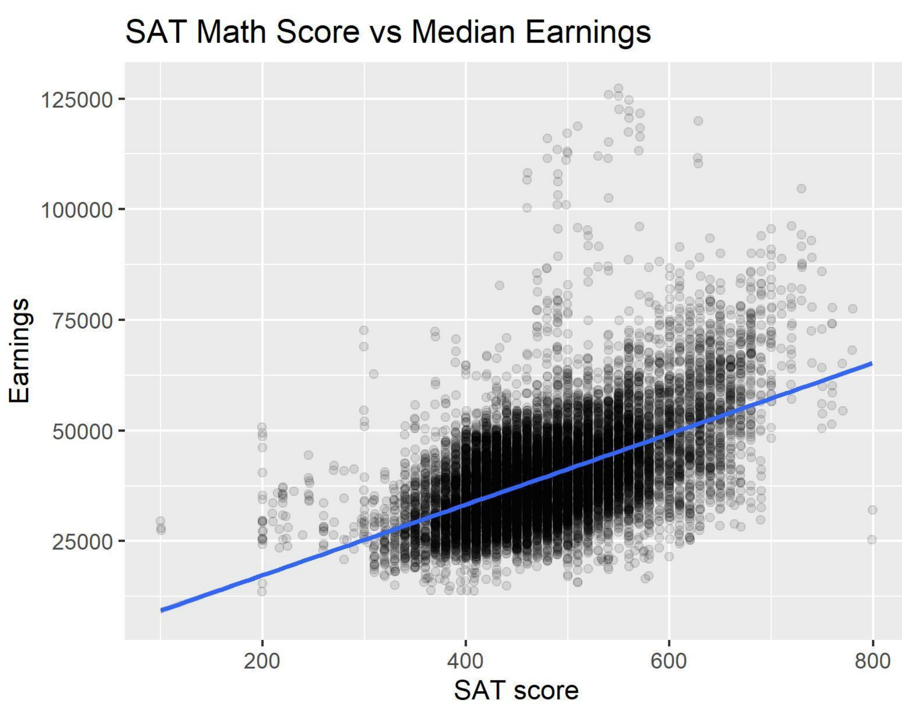

figure 1

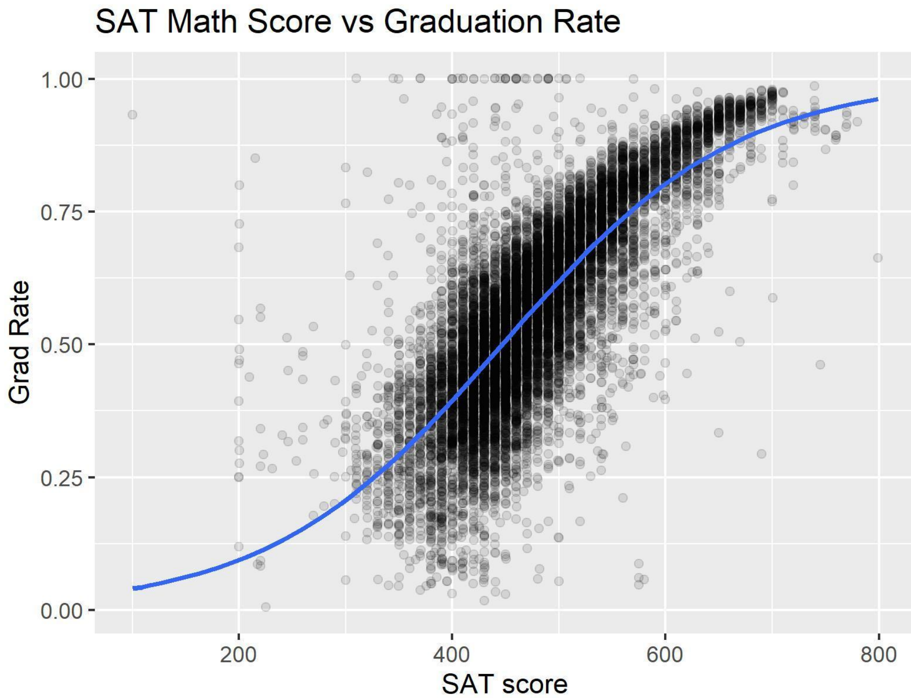

figure 2

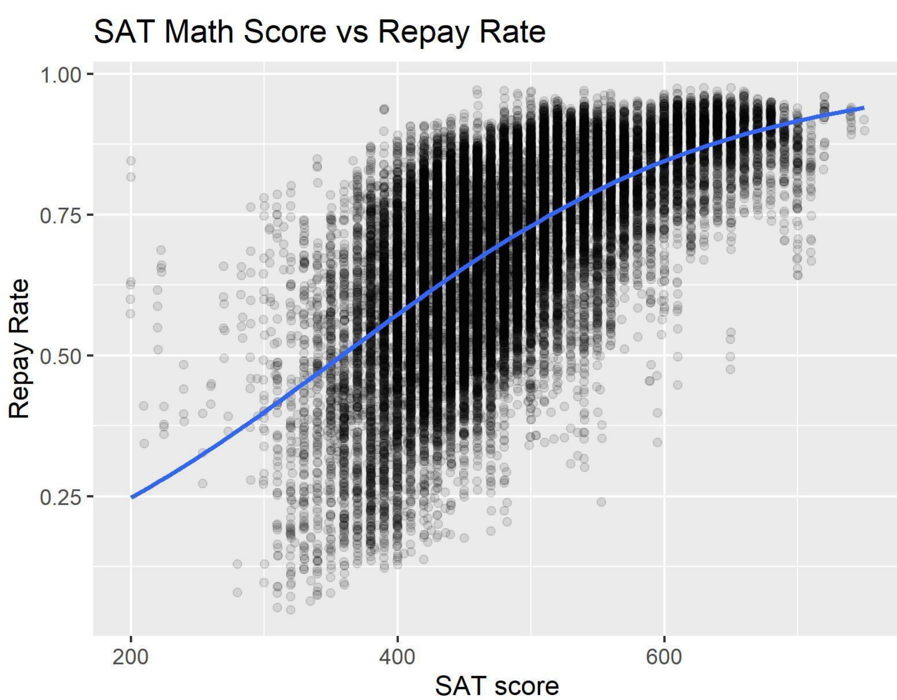

figure 3

We can also plug in some numbers to get a sense of the strength of these results. This is most straight forward for earnings as it is a linear model. This tells us that when the math scores of an incoming college cohort, specifically the cohort's 25th percentile, increase by one standard deviation, the median future income increases by around $6500. In other words, the difference in expected earnings per year for an attendee of a mathematically average college is $6500 less than for an attendee of a very good college.

For graduation and repay rate, it's not quite so straightforward since we are dealing with a logistic regression. We need to plug in numbers for all variables to get a prediction from the model. To do this, we'll take the mean value for all continuous values and set dummy values to zero then look at what happens when we vary the math score. For a college with an average SAT math score, graduation rate is predicted to be 57.2%, with a score one standard deviation above the mean, it's predicted to be 66.4%, and with a score one standard deviation below the mean, it's predicted to be 47.5%. Meanwhile, for a college with an average score, *repay* rate is predicted to be 64.8%, with a score one standard deviation above the mean, it's predicted to be 68.4%, and with a score one standard deviation below the mean, it's predicted to be 61.1%.

A final point worth examining is how the selective colleges in used in our regression analysis compare with the overall set of colleges in the US. The table below compares mean values for each of the outcome variables we have been considering for selective colleges and all colleges. As can be seen in *table 2*, in each case, the average value is higher for selective colleges.

	Selective	All
Earnings	$39,600	$30,800
Graduation Rate	56.2%	53.2%
Repay Rate	68.7%	49.0%

Table 2

Discussion

The results clearly show the importance of math achievement for college success and future financial well-being. Unpacking the consequences of this leads to a range of different issues when looking at college attendance as tool for improving financial wellness. A central point is that examining the math level of students at a given college provides additional information on the prospects of success for potential attendees. In discussing these results, though, great care must be taken in how we frame prospects for different kinds of students considering attending college – we are not aiming to discourage anyone from college, just presenting more nuanced information about the possible outcomes.

These findings also inform policy considerations. In particular, proposals to expand college attendance without changing pre-college achievement levels appear risky, given that additional college attendees are likely to be disproportionately math low achievers. It suggests the need to improve math achievement in earlier years as a first priority when looking to use education as a means to improve future financial wellness.

Another point worth discussing is the *negative* coefficient for reading score in the model for earnings – in other words a higher reading score predicts lower future earnings. One might plausibly speculate that students with higher reading scores are more likely to major in subjects such as the humanities associated with lower future earnings, and this is certainly a phenomenon worthy of further investigation.

It's also important to note the limitations of the current study. Most significantly, it only uses college level information. This means it is impossible to separate the contributions of individual ability and college quality when looking at their impact on outcomes. We don't know the extent which an individual's prospects would change if, holding their math abilities fixed, they switched from a low-math college to a high math college. To understand this issue better requires a study that connects college level data with individual longitudinal data. The evidence here suggests that this is an important investigation.

References

Abel, J. R., & Deitz, R. (2014). Do the benefits of college still outweigh the costs? *Current Issues in Economics and Finance*, *20*(3).

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, *121*(1-2), 343-375.

Balakrishnan, S., & Cynamon, B. Z. (2018). Borrowing Arrangements and Returns to College Education. *Available at SSRN 3263639*.

Banks, J., & Oldfield, Z. (2007). Understanding pensions: Cognitive function, numerical ability and retirement saving. *Fiscal Studies*, *28*(2), 143-170. doi:10.1111/j.1475-5890.2007.00052.x

Baum, S. (2014). Higher education earnings premium: Value, variation, and trends. *Urban Institute*.

Bound, J., Lovenheim, M. F., & Turner, S. (2010). Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. *American Economic Journal: Applied Economics*, *2*(3), 129-57.

Brookings Institute (2019) OPPORTUNITY, RESPONSIBILITY, AND SECURITY: A CONSENSUS PLAN FOR REDUCING POVERTY AND RESTORING THE AMERICAN DREAM. AEI/Brookings Working Group on Poverty and Opportunity

Butcher, K. F. (2017). Assessing the long-run benefits of transfers to low-income families. In *Hutchins Center on Fiscal and Monetary Policy at Brookings' conference "From Bridges to Education: Best Bets for Public Investment," held at Brookings*

Cole, S., Paulson, A., & Shastry, G. K. (2016). High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of Human Resources*, *51*(3), 656-698. doi:10.3368/jhr.51.3.0113-5410r1

Goodman, J. (2019). The labor of division: Returns to compulsory high school math coursework. *Journal of Labor Economics*, *37*(4), 1141-1182.

Hardy, B. L., & Marcotte, D. E. (2018). Ties that bind? Family income dynamics and children's post‐secondary enrollment and persistence.

Lusardi, A. (2012). *Numeracy, financial literacy, and financial decision-making* (No. w17821). National Bureau of Economic Research.

Marley-Payne, J., Dituri, P., Davidson, A. (ms) The Combined Effects of Financial Education and Mathematical Confidence: Findings from the 2018 National Financial Capability Survey. *Financial Life Cycle Education*.

Morgan, J., & Steinbaum, M. (2018). The Student Debt Crisis, Labor Market Credentialization, and Racial Inequality. *Roosevelt Institute*.

Restuccia, D., & Urrutia, C. (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review*, *94*(5), 1354-1378.

Appendix 1: Earnings Regression Results

	SAT Scores	ACT Scores
(Intercept)	35276.46 ***	33640.32 ***
	(88.41)	(86.09)
Math Score	6572.80 ***	5250.12 ***
	(144.28)	(135.17)
Reading Score	-3023.63 ***	-1542.83 ***
	(137.15)	(128.47)
Median Family Income	3022.65 ***	2181.89 ***
	(82.99)	(82.41)
Female (%)	-5.71	-235.39 ***
	(57.47)	(57.92)
Married (%)	1813.71 ***	1647.03 ***
	(69.08)	(67.29)
Dependent (%)	-332.95 ***	-336.50 ***
	(59.82)	(59.36)
First Generation Student (%)	127.65 *	-9.75
	(62.05)	(60.43)
Race		
(Ref: Non-Stated)		
White (%)	-1052.47	-806.59
	(645.82)	(595.13)
Black (%)	-621.29	-805.80
	(526.92)	(513.96)

Hispanic (%)	-30.98	340.39
	(260.43)	(219.81)
Asian (%)	2367.75 ***	1505.54 ***
	(184.70)	(153.69)
Other (Race) (%)	213.81	390.77
	(331.25)	(301.20)
Years Since Enrollment		
(Ref: 6)		
8 Years	5361.09 ***	4786.76 ***
	(131.80)	(129.62)
10 Years	10082.32 ***	8594.64 ***
	(144.19)	(146.64)
N	17311	13056
R2	0.53	0.53

All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Appendix 2: Graduation Rate Regression Results

	SAT Scores	ACT Scores
(Intercept)	0.29 ***	0.14 ***
	(0.02)	(0.02)
Math Score	0.39 ***	0.20 ***
	(0.04)	(0.05)
Reading Score	0.10 **	0.23 ***
	(0.03)	(0.05)
Median Family Income	0.29 ***	0.28 ***
	(0.03)	(0.03)
Female (%)	-0.04	-0.06 *
	(0.02)	(0.02)
Married (%)	0.02	0.01
	(0.02)	(0.03)
Dependent (%)	-0.01	-0.02
	(0.02)	(0.03)
First Generation Student (%)	-0.06 *	-0.08 **
	(0.02)	(0.02)
Race		
(Ref: Non-Stated)		
White (%)	-0.00	-0.08
	(0.22)	(0.25)
Black (%)	0.01	-0.07

	(0.17)	(0.21)
Hispanic (%)	0.00	-0.00
	(0.11)	(0.09)
Asian (%)	0.10	0.07
	(0.06)	(0.06)
Other (Race) (%)	0.02	-0.02
	(0.11)	(0.12)
N	13464	10806
Pseudo R2	0.38	0.37

All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Appendix 3: Repay Regression Results

	SAT Scores	ACT Scores
(Intercept)	0.61 ***	0.54 ***
	(0.02)	(0.02)
Math Score	0.16 ***	0.12 ***
	(0.02)	(0.03)
Reading Score	0.04	0.04
	(0.02)	(0.03)
Median Family Income	0.31 ***	0.31 ***
	(0.02)	(0.02)
Female (%)	-0.01	-0.01
	(0.01)	(0.01)
Married (%)	-0.02	-0.02
	(0.02)	(0.02)
Dependent (%)	-0.00	0.00
	(0.02)	(0.02)
First Generation Student (%)	-0.02	-0.03
	(0.01)	(0.01)
Race		
(Ref: Non-Stated)		
White (%)	0.04	0.04
	(0.27)	(0.27)
Black (%)	-0.19	-0.21
	(0.22)	(0.24)

Hispanic (%)	-0.02	-0.02
	(0.13)	(0.11)
Asian (%)	0.10	0.10
	(0.07)	(0.07)
Other (Race) (%)	-0.01	-0.02
	(0.13)	(0.13)
Years Since Entering Repay		
(Ref: Year 1)		
Year 3	0.18 ***	0.17 ***
	(0.03)	(0.03)
Year 5	0.41 ***	0.41 ***
	(0.03)	(0.04)
Year 7	0.61 ***	0.61 ***
	(0.04)	(0.04)
N	32859	28320
Pseudo R2	0.27	0.27

All continuous predictors are mean-centered and scaled by 1 standard deviation. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.