

IZA DP No. 6107

The Returns to Four-Year College for Academically Marginal Students

Seth Zimmerman

November 2011

The Returns to Four-Year College for Academically Marginal Students

Seth Zimmerman

*Yale University
and IZA*

Discussion Paper No. 6107
November 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Returns to Four-Year College for Academically Marginal Students^{*}

I combine a regression discontinuity design with rich data on academic and labor market outcomes for a large sample of Florida students to identify the returns to four-year college for students on the academic margin of college admission. In addition, I develop a theoretical model of college choice with and without credit constraints that allows for intuitive tests of the importance of credit constraints within this population. I find that students who obtain high school grades just above the threshold value for admissions eligibility at a large public university in Florida are much more likely to attend a four-year college and much less likely to attend a community college than students with grades just below the threshold. The earnings returns to a year of four-year college for affected students are 8.7 percent, nearly identical to returns to college for the population of Florida high school students. Consistent with the credit constraints hypothesis, poorer students who are more likely to be credit constrained work more while in college and realize higher post-college returns.

JEL Classification: I20, J30

Keywords: returns to college, credit constraints, community college

Corresponding author:

Seth Zimmerman
Department of Economics
Yale University
37 Hillhouse Ave.
New Haven, CT 06520
USA
E-mail: seth.zimmerman@yale.edu

^{*} I thank Joseph Altonji, Martin Hackmann, Fabian Lange, Costas Meghir, and Christopher Neilson for their helpful comments. I thank Jamie Champion, Nancy Copa, and Tammy Duncan at the Florida Department of Education Data Warehouse for helping me access to student data, and Sherby Jean-Leger at the College Board for providing access to SAT score data. All remaining errors are my own. SAT scores are derived from data provided by the College Board. Copyright 1996-2004 The College Board. www.collegeboard.com.

1 Motivation

The college wage premium has risen dramatically over the past 30 years. In 1980, college graduates earned roughly 50 percent more than high school graduates; by 2008, they earned 97 percent more.¹ A series of influential papers (e.g., Katz and Murphy (1992), Goldin and Katz (2008), Acemoglu and Autor (2011)) show that this change is at least in part the product of rapidly rising demand for skilled labor with slower increases in the supply of skilled labor. For instance, Goldin et al. (2007) estimate that between 1980 and 2005, the demand for college graduates increased by about 3.5 percent per year, while the relative supply of college graduates increased by only 2 percent per year. The net result was growth in the college wage premium at the rate of 0.9 percent per year.

Why has supply not kept pace with demand? One possibility is that the returns for students on the margin of college attendance are much lower than the average returns to college. This is consistent with the large body of evidence suggesting that many US primary and secondary schools do a poor job of preparing their students for college (Roderick et al. 2009). Students who receive low quality education through secondary school may reap few benefits from attending a four year college, particularly if less costly educational alternatives like community college or vocational training are also available. The principle competing hypothesis is that many students would like to attend a four year college but cannot because they face short-term credit constraints.² It is certainly true that the real costs of college have risen rapidly since 1980, and that a large and increasing fraction of college students receive some type of financial aid.³ What is unclear is whether these trends result from credit constraints or merely a functioning financial aid system in the context of rising costs. Within the economics literature, evidence on the importance of credit constraints is mixed (see, e.g., Belley and Lochner (2007), Stinebrickner and Stinebrickner (2008), Cameron and Taber (2004)).

¹Source: Acemoglu and Autor (2011). Estimates adjust for changes in demographic composition.

²Following Carneiro and Heckman (2002), children's inability to purchase better early-life inputs can be thought of as a long term credit constraint. In this paper, I focus exclusively on short-term liquidity constraints that affect students' ability to pay for college.

³Between the 1979-1980 and 2009-2010 school years, the mean annual direct cost of a four-year college (including tuition and room and board, in constant 2008-2009 dollars) rose by 240 percent (NCES Digest of Education Statistics 2010, Table 345). The percentage of students at public four year institutions receiving some type of financial aid rose from 71.3 in 2000-2001 to 78.9 in 2008-2009. Over the same period the percentage of students at private non-profit four-year institutions receiving aid rose from 82.9 to 87.2 (NCES Digest of Education Statistics 2010, Table 350).

Because they lead to very different policy recommendations, distinguishing between these two lines of reasoning is of critical importance for higher education policy. If credit constraints are the dominant story, then policies aimed at expanding access to college by reducing these constraints will be enough to increase the supply of college graduates. If low marginal returns are the dominant story, then policies aimed at improving primary and secondary education so that students emerge better-prepared for college are more appropriate. The key question here is whether students who are only marginally prepared for college are able to realize economic returns large enough to justify the investment of time and money.

This paper combines a rich and novel dataset on high school, college, and labor market outcomes for a large sample of Florida high school students with a regression discontinuity design around a state-level GPA cutoff for admission to the Florida State University System (SUS) to estimate the returns to education for academically marginal applicants at a specific four-year institution, Florida International University (FIU). Students just above the grade threshold are 17 percent more likely to be admitted to FIU and 9 percent more likely to enroll in any four-year university than those just below the threshold. 58 percent of the students induced to attend a university would otherwise have enrolled in community college. 63 percent of policy compliers obtain a BA within 6 years, and each additional year of university enrollment raises their earnings by 8.7 percent. This estimate is nearly identical to OLS estimates of the returns to university attendance for all students, which suggests that concerns that academically marginal students lack the preparation necessary to benefit from college unfounded, at least for students who choose to attend college when given the option. Importantly, students do not appear to manipulate their grades in order to surpass the admissions threshold, likely due to opaque and institution-specific grade weighting procedures.

To better understand the role credit constraints play in the decision to attend college for this policy-critical group of students, I develop a model of college choice in which credit constraints are defined as a higher rate of interest while in college. In the spirit of Cameron and Taber (2004), the model predicts that, compared to unconstrained students, credit constrained students should spend more time in the labor market while in college and realize higher returns to college on average. Continuing to exploit the regression discontinuity at the admissions threshold, I test these claims by looking for differences in in-college labor force participation and post-college earnings returns for students who did and did not receive free lunch in high school. The idea is that free lunch students are more likely to be credit constrained than other students. The data

bear out both of the model's predictions; of particular interest is the fact that returns to college are 4.3 percentage points higher for free lunch students. My findings indicate that expanding need-based financial aid or reducing the qualifications needed to qualify for existing Florida merit aid programs would help students make investments with large private returns.

Finally, an important feature of this student population is that the majority (58 percent) of students induced to attend college by threshold-crossing would otherwise have attended community college within three years after high school graduation. Threshold-crossing has no effect on students' probability of receiving an associate's degree or vocational certificate, indicating that students forced to attend a community college by rejection from a four-year college rarely graduate. Coupled with the observation that the high observed returns to four year college reflect returns net of any earnings benefits that accrue to below-threshold community college attendance, this finding suggests that students on the margin of attending a four-year college do not benefit much from going to a community college, even though they on average have stronger academic backgrounds than other community college students. One explanation that can reconcile my results with prior findings of relatively high average returns to community college attendance (Kane and Rouse (1995), Jacobson et al. (2005)) is that community college works best for students who are not academically prepared to succeed at a four-year college; i.e., its benefits are primarily remedial. This is not implausible given that the majority of community college students engage in some type of remedial coursework (Attewell et al. (2006)).

The paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the model of college choice under credit constraints. Section 4 describes the regression discontinuity design. Section 5 describes the data, and section 6 describes the SUS and how its admissions policies map into the RD framework. Section 7 presents the paper's main findings. Section 8 addresses the robustness of the empirical design, and section 9 concludes.

2 Prior work

This paper builds on two related strands of work. The first strand is the instrumental variables literature on heterogeneous returns to schooling. The canonical challenge facing researchers trying to recover the returns to education is how to obtain estimates that

are purged of bias resulting from the tendency of more able individuals to obtain higher levels of schooling. One strategy is to instrument for education using a variable that is plausibly uncorrelated with ability. As discussed in Card (1999) and Meghir and Rivkin (2010), IV estimation with a valid instrument under standard monotonicity conditions yields an unbiased estimate of the effect of schooling on earnings for individuals whom the instrument induces to change their educational choices. The individuals who switch education levels are known as 'compliers', and the effect itself is termed a local average treatment effect, or LATE.

Because the effects of education on earnings are heterogeneous (see, e.g., Carneiro et al. (2010)), the estimated LATE for individuals who comply with one policy cannot in general be extrapolated to those who comply with a different policy. For instance, returns for admitted students who attend college only if offered financial aid may not be the same as those for students who attend college only if pulled off of an admissions wait-list. Credible use of LATE estimates for policy evaluation therefore depends on finding an instrument that shifts students across the same margin as the proposed policy. The instrument here is grade threshold-crossing for students with grades close to the cutoff value. This instrument focuses tightly on academically marginal students and offers the answer to a concrete policy question: what are the returns to college for students who attend if we slightly decrease the threshold from its current value?

The second strand of literature addresses the role of credit constraints in determining educational attainment. Evidence on the role credit constraints play in college attendance decisions is mixed. A number of quasi-experimental studies find strong effects of grant aid on college attendance and college graduation (e.g., Kane (2003), Dynarski (2002), Dynarski (2004), Dynarski (2008); see Deming and Dynarski (2009) or Kane (2006) for reviews of this literature). One limitation of these studies is that grant aid represents a subsidy, not increased access to credit. Structural models of the choice to attend college typically find that subsidies increase college attendance, but loosening borrowing constraints has little impact (Keane and Wolpin (2001), Johnson (2010)). An alternate strategy, pursued in Cameron and Taber (2004), henceforth CT, uses the intuition that credit constrained students face higher interest rates while in college, and therefore require higher marginal returns to college to make attendance worthwhile. CT develop several empirical tests for credit constraints, focusing on the differential impacts of college cost-shifters on constrained and unconstrained students. They find no evidence of credit constraints.

I draw out two implications from a model similar to that used in CT. First, constrained students who face higher interest rates will supply more labor while in college, since they effectively face a higher price of labor. This prediction extends CT, who do not consider labor supply while in school. Second, the model yields the CT prediction that constrained students who attend college should realize higher returns than unconstrained students. I test these predictions by examining heterogeneity in the responses of free lunch and non-free lunch students to the quasi-random variation in college access generated by threshold crossing. The idea is that free lunch students are poorer and more likely to be credit constrained. Unlike CT, I find that both empirical tests support the credit constraints hypothesis. This may be due to differences in the population of interest or the instrumental variable used.⁴ Regardless, the key contributions here are that a) credit constraints appear to be operative in the policy-relevant group of academically marginal college applicants, and b) the CT methodology is consistent with an important role for credit constraints.

My empirical strategy draws on several papers that investigate the effect of crossing test score or grade thresholds on educational attainment and labor market outcomes. Kane (2003) uses cutoff rules in income, assets, and grades to estimate the effects of financial aid eligibility on college enrollment for California students. Van der Klaauw (2002) takes a similar approach, estimating the changes in enrollment induced by sharp cutoffs for aid eligibility in an index of academic performance at an unnamed college on the East Coast. In the most closely related paper, Hoekstra (2009) estimates the returns to college for students who attend a flagship state university only if their SAT score is just above a cutoff value for admissions. This paper differs from Hoekstra in that it identifies returns for a very different group of students. Presumably, many students who are not admitted to the flagship state university attend other four-year colleges (though Hoekstra cannot verify this with the data he has), while in the current application, threshold-crossing results in a large change in the probability of enrollment at any university. From the perspective of access to tertiary education, this is the policy-critical group, and one about which little is known. This paper presents the first quasi-experimental evidence on returns for these policy-critical students.

⁴CT focus on instruments based on local labor market conditions and the presence or absence of local colleges

3 Theoretical strategy

To assess the importance of credit constraints in the decision-making process, I develop a stylized model of the choice to attend college, labor supply while in college, and the earnings returns to college. The model draws on CT in that a) it defines credit constraints as higher rates of interest for borrowers, and b) it yields empirical tests for credit constraints based on the differential responses of constrained and unconstrained groups of students to instruments that shift college attendance. It builds upon CT by developing an additional test for credit constraints based on heterogeneity in labor supply while in college. The model makes two concrete predictions:

1. Constrained individuals who choose to attend college will realize higher returns than unconstrained individuals who make the same choice.
2. Constrained individuals will supply more labor while in college than unconstrained individuals.

I now develop these predictions formally. Let agents live for three periods, $t = 0, 1, 2$. In period zero, agents choose whether or not to attend college. The direct cost of attending college is τ , paid in period one. For simplicity, I set initial asset levels to zero. In period one, agents who choose to attend college do so. Whether or not an agent chooses to attend college, he chooses how much wage labor l to supply at wage rate w , and how much to consume, c_1 . Attending college has a time cost h , which, when added to l , determines total labor supply. In period two, agents' earnings depend on their level of education, and earnings for college attendees depend in addition on their own type, θ . Specifically, period two earnings are given by $Y_C(\theta)$ for college attendees and $Y_{HS}(\theta) = Y_{HS}$ for non-attendees. Agents choose only how much to consume, c_2 .

Let $s \in C, HS$ denote a given educational choice, and define $d = 1(s = C)$. Then utility from choice s is given by

$$V_s(\theta, R_s) = \max_{c_1, c_2, l} u(c_1) + \beta U(c_2) - v(l + dh) \quad (1)$$

subject to the budget constraint

$$c_1 + \frac{1}{1 + R_s} c_2 + d\tau = wl + \frac{1}{1 + R_s} Y_s(\theta). \quad (2)$$

The interest rate R_s depends on the agent's educational choices. Credit constraints correspond to higher values of R_C . To facilitate the analysis I make a number of functional form assumptions. First, $u', U' > 0$ and $u'', U'' < 0$. The utility functions in periods one and two are strictly increasing and concave. They need not be the same function. Similarly, the labor cost function v satisfies $v' > 0, v'' > 0$ so that costs are strictly increasing and convex. Returns to college are increasing in θ : $Y'_C(\theta) > 0$. Based on empirical observation, further assume that students who attend college are net borrowers in period one and students who do not are net savers. Formally, let l^s and c_t^s be the optimized labor supply and period t consumption for agents who choose schooling option s . I assume that $c_1^C + \tau > wl^C$ and $c_1^{HS} < wl^{HS}$.

Let R_{HS} be fixed in the population, while R_C varies, and let θ be distributed according to the continuous density function $F(\theta|R_C)$. The following proposition demonstrates that credit constrained students who choose to attend college realize larger returns than unconstrained students making the same choice.

Proposition 1. *If given any value of R_C at least some agents attend college, then for each R_C there exists $\bar{\theta}(R_C)$ such that all agents for whom $\theta > \bar{\theta}(R_C)$ will choose to go to college and all agents for whom $\theta < \bar{\theta}(R_C)$ will choose not to. Further, $\frac{d\bar{\theta}(R_C)}{dR_C} > 0$.*

This implies that

Corollary 1.

$$\frac{dE[Y_C(\theta) - Y_{HS}|R_C, s = C]}{dR_C} > 0 \quad (3)$$

Proof. See Appendix C. □

Corollary 1 corresponds to model prediction one above. The intuition for the proof, which follows directly from the envelope theorem, is as follows. Credit constrained students face higher interest rates if they attend college. Since college attendees are net debtors in period one, the higher interest rates reduce their net income. This additional cost of college means that credit constrained students must realize larger returns to college than unconstrained students to make attending worthwhile.

The second prediction is obtained using straightforward comparative statics:

Proposition 2. *For fixed θ ,*

$$\frac{dl^C}{R_C} > 0 \tag{4}$$

Proof. See Appendix C. □

The intuition here is that raising interest rates affects labor supply in two ways. Raising R_C raises the value of labor while in college, since saving money (or borrowing less) has a higher rate of return. Raising R_C also makes net debtors poorer, lowering consumption and raising the marginal utility of additional labor supply. Note that Proposition 2 does not formally resolve the question of how average labor supply for college attendees will differ across constrained and unconstrained students, since the optimized value of l^C depends on θ , and the distribution of θ conditional on college attendance will be different for constrained and unconstrained students. In general, the effect of changes in θ on $\frac{dl^C}{dR_C}$ are ambiguous and depend on the shape of the third derivatives of u and U . I assume that the third derivatives are such that cross-group compositional differences do not drive differences in average labor supply.

Testing these predictions empirically requires credible identification of heterogeneity in in-college labor supply and the returns to college across constrained and unconstrained students. Ideally, this could be achieved through randomization of both college attendance and constraints. In practice, it is difficult to obtain this type of double randomization. Instead, I will test these predictions using interactions between exogenous variation in college access generated by a regression discontinuity in admissions policy and students' free lunch status. The idea is that students who receive free lunch are more likely to be credit constrained than students from wealthier backgrounds. Comparing how free-lunch and non-free-lunch students respond to exogenous offers of admission then allows for tests of heterogeneity in work while in college and returns to college.

In this framework, the prediction from equation 3 can be tested by comparing the returns to college for free lunch and non-free lunch students whom threshold-crossing induces to attend college. The students who attend college if given the option are precisely the group we're interested in here. The prediction from equation 4 can be tested by comparing the labor force effects of college attendance for free lunch and non-free lunch students.

The main drawback of this approach is that, because credit constraints are non-randomized, I cannot rule out the possibility that it is in fact some other correlate of free lunch status that is driving the observed effects. I discuss this issue further in section 9.

4 Empirical strategy

I recover estimates of the returns to college using a fuzzy regression discontinuity (FRD) design that compares outcomes for students with grades just below the grade cutoff for FIU admission to outcomes for students with grades just above the cutoff. The intuition is that students with grades very close to the cutoff on either side are comparable in terms of the observable and unobservable determinants of wages, but that those just above the cutoff are more likely to attend college. In FRD designs, threshold crossing causes a discontinuous jump in the probability of treatment, but this jump is not from zero to one. The idea here is that some students with grades below the cutoff are admitted and attend college, and some students with grades above the cutoff are not admitted or are admitted but choose not to attend. Because the students who respond to threshold crossing may differ from other students with similar grades, the estimates I obtain should be interpreted as a local average treatment effect for students at academic margin of college attendance who attend only if their grades surpass the threshold value. One way to think of this group is as the group of ‘compliers’ with the admissions policy (Angrist and Imbens 1995).

Following Imbens and Lemieux (2008), I implement the FRD design using local linear instrumental variables regressions. Restricting the sample to students with grades within 0.3 grade points above or below the cutoff value, I estimate the equation

$$\ln Y_{it} = X_i' \delta_1 + f(t) + \gamma_1^1 g_i + \gamma_2^1 Z_i g_i + \rho S_i + u_{1i} \quad (5)$$

where Y_{it} is real earnings for individual i at experience level t , X_i is a set of demographic variables including race dummies, cohort dummies, district dummies, SAT score dummies, and free lunch status, $g_i = grade_i - c_i$ is the difference between the GPA for individual i and individual-specific grade cutoff c_i (see next section for a description of how grade cutoffs are determined), and S_i is years of SUS attendance. $Z_i = 1[g_i \geq 0]$ is an indicator variable for threshold crossing; I allow for distinct slopes in grades above and

below the threshold value to reflect the fact that the standard college admissions process (effective above the cutoff) may differ from the non-standard admissions process (effective below the cutoff) in terms of the relationship between grades and the determinants of earnings. The coefficient of interest is ρ , which measures returns to a year of SUS attendance. Because S_i is likely correlated with unobservable wage determinants u_{1i} , I instrument for S_i using the first stage equation

$$S_i = Z_i\pi + X_i'\delta_2 + \gamma_1^2 g_i + \gamma_2^2 Z_i g_i + u_{2i} \quad (6)$$

To test for heterogeneous effects by free lunch status, I interact S_i with a free lunch dummy in equation 5, and instrument using the interaction between that dummy and Z_i in equation 6. Note that the free lunch dummy itself is a component of X_i .

For this analysis to produce consistent and interpretable results, several conditions must hold. First, the interpretation of ρ as a mean effect for compliers requires the monotonicity condition that there are no individuals who attend SUS for fewer years if they have above-cutoff grades than if they have below-cutoff grades (Imbens and Angrist 1995). This condition seems plausible. Second, threshold-crossing variable Z_i must be conditionally uncorrelated with unobservable earnings determinants u_{1i} when g_i is within some narrow window around zero. As discussed in Lee and Lemieux (2009), this restriction will typically hold if a) applicants do not attempt to manipulate grades so as to just surpass the cutoff score, or b) applicants do attempt grade manipulation, but manipulation is imprecise because students can only choose their grades up to some continuous disturbance. In either case, earnings determinants other than college attendance will be change smoothly near the cutoff value, and the discontinuity will reflect only the desired treatment effect.

As I discuss in the section 6, the use of an idiosyncratic and non-transparent set of weights in the computation of the GPAs used for admissions makes precise grade manipulation much less likely. In section 8 I test the claim empirically and find no evidence of discontinuous grade manipulation.

5 Data

5.1 Data description

I use population data for seven cohorts of 12th-grade students from fifteen Florida counties.⁵ The counties were chosen to be diverse in size and location (though not necessarily representative of the student population of the state). The cohorts represented are defined by graduating years 1996, 1997, 1999, 2000, 2001, 2002, and 2004, though the sample is not limited to students who graduate. For each student, I have access to basic demographic data, high school transcript data, applications and admissions data for the Florida State University System (SUS), SUS attendance, graduation, and transcript data, and community college attendance and graduation data. Graduation data are available up to the 2008-2009 school year, while attendance records are available through 2008-2009 for the SUS and 2006-2007 for community colleges. I link the education data to labor market outcomes from Florida Unemployment Insurance records. These records provide quarterly earnings data from in-state employment from 1995 through the first quarter of 2010. If students are employed outside of Florida, I do not observe their earnings. See Appendix A for a description of how key variables were constructed.

5.2 Descriptive statistics

The first three columns of Table 1 present descriptive statistics for the full sample of students, the sample of FIU applicants, and a sample of ‘marginal’ FIU applicants who have grades within 0.3 grade points above or below the cutoff for admission to any SUS institution. I describe the admissions process and the computation of cutoffs in greater detail below. The students in the full sample are predominantly minority, and nearly half receive free or reduced-price lunch at some point during high school. FIU applicants are less likely to be white, more likely to be Hispanic, and slightly more likely to have received free or reduced-price lunch than other students. FIU applicants perform better in high school coursework than other students, but performance for marginal FIU applicants was closer to that of full-sample students than that of other applicants. FIU applicants performed worse on the SAT than the average in-sample student who took the test, and marginal FIU applicants performed substantially worse than the full

⁵The counties are Dade, Broward, Hillsborough, Orange, Polk, Santa Rosa, Charlotte, Putnam, Martin, Highlands, Calhoun, Jefferson, Gulf, Franklin, and Hamilton.

group of applicants (952 to 844 for combined Reading and Math scores).

Nearly all FIU applicants graduated from high school, and 60 percent attended an SUS campus the year following their 12th grade year. 50 percent of marginal applicants attended an SUS campus that year. 35 percent of all applicants and 23 percent of marginal applicants graduated from an SUS institution within six years. This is consistent with a six-year graduation rate of slightly over fifty percent for all applicants and slightly under 50 percent for marginal applicants. 18 percent of all FIU applicants and 24 percent of marginal applicants attended a community college full time the year following their twelfth grade year, but only 8 percent of all applicants and 7 percent of marginal applicants received an associates degree within three years.

The second three columns of Table 1 compare marginal FIU applicants to students who attend an SUS school the year after they graduate from high school and to students who attend a community college within three years after they graduate from high school. The idea is to compare students on the academic margin between college and community college to average students on either side of that threshold. I abstract away from geographic variation in demographics and academic outcomes by restricting the comparison sample to students from the Miami-Dade school district, where most FIU applicants attended high school. SUS attendees are more likely to be white and less likely to receive free or reduced-price lunch than students in the other two groups. The mean unweighted high school GPAs for SUS attendees was 3.0, compared to 2.69 for marginal FIU applicants and 2.58 for CC attendees. Similarly, SUS attendees score an average of 0.66 standard deviations above the mean on the mandatory FCAT assessment test, while marginal FIU applicants scored 0.10 standard deviations above the mean and CC attendees scored 0.08 standard deviations *below* the mean. Academically marginal college applicants lag behind the average college student in terms of academic performance, but out-pace the average community college student.

Table 1: Sample description

	Full sample			Miami only		
	All	FIU	Marg. FIU	SUS	CC	Marg. FIU
White	0.39	0.17	0.15	0.21	0.12	0.12
Black	0.27	0.25	0.33	0.25	0.23	0.30
Hispanic	0.29	0.51	0.48	0.48	0.62	0.54
Male	0.49	0.38	0.37	0.40	0.41	0.36
F/R Lunch	0.43	0.46	0.50	0.40	0.54	0.53
Miami-Dade	0.33	0.75	0.76	1.00	1.00	1.00
HS GPA	2.64	2.94	2.72	3.00	2.58	2.69
FCAT	0.02	0.48	0.10	0.66	-0.08	0.10
FIU GPA	N/A	3.33	3.00	N/A	N/A	2.99
SAT	974	952	844	1011	N/A	842
HS Grad	0.77	0.97	0.96	0.97	0.91	0.97
SUS next yr	0.16	0.60	0.50	1.00	0.02	0.49
Non-FL coll	0.06	0.03	0.03	0.02	0.02	0.03
FL priv. coll	0.04	0.08	0.06	0.03	0.03	0.06
SUS BA in 6 yrs	0.12	0.35	0.23	0.55	0.11	0.22
CC next yr	0.14	0.18	0.24	0.02	1.00	0.25
AA in 3 yrs	0.04	0.08	0.07	0.04	0.17	0.07
	411198	30109	9423	20464	20881	7154

Note: Observations are at the student level. Full sample is the population of students in fifteen Florida school districts over graduating cohorts in the 1995-2003 period. FIU denotes the FIU applicants sample. This is subset of students who apply to FIU during their twelfth-grade year. Marginal applicants are the subset of FIU applicants with applicant GPAs that differ by less than 0.3 grade points from SAT-determined cutoff values for admission. SUS denotes the set of students who attend any SUS campus the first year after they graduate from high school. CC denotes the set of students who attend any Florida community college within three years of high school graduation. GPAs are unweighted and computed at the high school level, while applicant GPAs are computed by FIU according to its own weighting regime. SAT scores reflect students' highest reading scores combined with their highest math scores. FCAT scores are the means of standardized reading and math scores. Reference period for 'in 6 years' or 'next year' is the 12th grade year. Fields are marked N/A when a field is not systematically calculated for the selected students.

Table 2 displays labor force participation and mean earnings statistics for the full sample, FIU applicants, and marginal applicants. Individuals are designated labor force participants if their UI records show non-zero annual earnings. Rates of in-state labor force participation are high in the years immediately following high school. For FIU applicants, participation rises from 75 percent in the first year after high school to 78

percent in the fifth year after high school. They then decline slowly, reaching 62 percent by the 14th year after high school. Labor force participation rates for marginal applicants are several percentage points higher, but follow a similar trend. Labor force participation rates for the full sample are lower throughout. Average earnings (conditional on participation) trend upward in the years after high school, starting at roughly \$7,000 in first year post high school and reaching \$46,000 for all FIU applicants in the 14th year after high school. Earnings are slightly lower for marginal applicants and significantly lower for the full population.⁶

Table 2: Labor force participation

Yrs. post HS	Full sample			Applicants			Marginal applicants		
	LF	Earn	N	LF	Earn	N	LF	Earn	N
1	0.72	7564	428936	0.75	6780	30251	0.79	7178	9478
2	0.71	9710	428936	0.76	8922	30251	0.80	9509	9478
3	0.70	11569	428936	0.76	10829	30251	0.79	11552	9478
4	0.69	13450	428936	0.77	13053	30251	0.79	13716	9478
5	0.68	16144	428936	0.78	17230	30251	0.79	16895	9478
6	0.65	18900	428936	0.75	21778	30251	0.78	20647	9478
7	0.65	21205	352514	0.76	25483	24431	0.78	24158	8029
8	0.63	23327	352514	0.73	29178	24431	0.76	27751	8029
9	0.61	25555	287815	0.72	32579	19114	0.75	30146	6433
10	0.60	27549	227884	0.71	35367	14140	0.73	34707	4868
11	0.58	29341	170236	0.68	38622	9738	0.70	37173	3326
12	0.58	30082	108317	0.68	40017	5544	0.71	37300	1899
13	0.55	31927	108317	0.65	44133	5544	0.67	43703	1899
14	0.51	33890	56978	0.62	46332	2790	0.65	45041	989

Note: Observations are at the student-year level. Full sample is the population of students in fifteen Florida school districts over graduating cohorts in the 1995-2003 period. FIU applicants sample is the subset of these students who apply to FIU during their twelfth-grade year. Marginal applicants are the subset of FIU applicants with applicant GPAs that differ by less than 0.2 grade points from SAT-determined cutoff values for admission.

Several aspects of this table are worth commenting on. First, the decline in sample size with time results from the ‘aging out’ process— I observe 13.5 years of post-high school earnings for the 1996 graduating cohort, but only 5.5 years for the 2004 graduating

⁶Here and following, dollar amounts refer to 2005 dollars.

cohort.⁷ Second, the decline in labor force participation over time is likely the result of migration out of the state of Florida. This will create problems for causal inference using regression discontinuity if the choice to leave the state is related to status relative to the cutoff for marginal students. I discuss this in greater detail below. Third, rates of labor force participation and average earnings for FIU applicants and marginal applicants in the first years after high school seem quite high given that roughly three quarters of these students attend a state university or are full time community college students. This is consistent with the idea that FIU applicants are drawn largely from low-SES groups and must work while in school.

6 College and admissions rule description

6.1 The college in question

Located in Miami, FIU is a large university, and it enrolls many students from traditionally disadvantaged groups. Table 3 describes key attributes of FIU for the 2000-2001 academic year, which falls roughly in the middle of the time frame considered in this paper. In that year, FIU enrolled nearly 24,000 degree-seeking undergraduates, of whom the majority (13,300) were women. 12,975 enrollees— a majority— were Hispanic, and an additional 3,390 were black. 9,546 students attended school on a part-time basis. 42 percent of students received financial aid of some form. Conditional on receipt, the mean value of aid was \$5,163.

For entering freshmen, the 25th percentile SAT score (for both Math and Verbal sections) was 510 and the 75th percentile SAT score was 590. Thus most students in the middle 50 percent of the SAT distribution had scores that slightly exceeded the national average SAT scores for college bound senior in the 1999-2000 school year; these scores were 505 for reading and 514 for math.⁸ Students had mean high school GPAs of 3.46 on a standard four-point scale. Six-year graduation rates for the admitted freshmen in this cohort were roughly 49 percent.⁹ This is somewhat lower than the system-wide six-year

⁷Earnings data is reported quarterly and then aggregated into academic year units. I observe only the first two quarters of the 2009-2010 academic year (fall 2009 and winter 2010). I convert half-year earnings observations for this year to full-year earnings observations by multiplying by 1.82, the ratio of half-year to full-year earnings for non-zero half-year earners in 2008-2009. Section 6 reports results for alternate imputation processes.

⁸NCES Digest of Education Statistics 2006, Table 133.

⁹FIU did not report graduation rates for the Fall 2000 entering cohort in their Common Data Set submis-

SUS graduation rate, which was approximately 62 percent at the time.¹⁰

Table 3: Characteristics of Florida International University, AY 2000-2001

Enrollment		Academics	
Total	23591	SATM: 25th ptile	510
Men	10283	SATM: 75th ptile	590
Women	13308	SATV: 25th ptile	510
PT	9546	SATV: 75th ptile	590
FT	14045	HS GPA	3.46
Black	3390	Grad. rate	0.49
Hispanic	12975		
Applications		Costs	
Total Apps	5891	In-State tuition	2242
Total Acc.	3176	Out-of-state tuition	9580
Total Enroll	2563	Room+Fees	4398
		Pct. Rec. FA	0.42
		Avg. FA value	5163

Source: FIU Common Data Set submissions 2000-2001 and 2007-2008. <http://opir.fiu.edu/cds.htm>. Enrollment data refers to degree-seeking students only. Academic characteristics are for degree-seeking first-time-enrollee freshmen. Six year graduation rates are computed for Fall 2001 entering cohort; graduation rates for the Fall 1999 entering cohort were 0.48. Applications data is for Fall 2000 entrants. Tuition and financial aid are reported in nominal terms. The percentage of students receiving aid includes only full-time undergraduates.

Admissions to FIU were relatively competitive. The bottom left panel of Table 3 presents basic admissions statistics for the Fall 2000 entering cohort. Of 5,891 applicants for that cohort, 3,176 (54 percent) were admitted. Of admitted students, 2,563 enrolled, for strikingly high yield rate of 80 percent. For comparison, the yield rate for the Harvard class of 2012 was approximately 76 percent.¹¹ As I will discuss below, one reason for the high yield rate is that many admitted students had few other options for four-year college attendance.

FIU admissions decisions result from a process that combines a set of statewide require-

sion. Graduation rates for the Fall 1999 and Fall 2001 entering cohorts were 48 and 49 percent, respectively.

¹⁰State University System of Florida 2009 Annual Report. http://www.flbog.org/about/_doc/budget/Volume-I-Published.pdf

¹¹See the Harvard Common Data Set submission at http://www.provost.harvard.edu/institutional_research/common_data_set.php.

ments for admission to any SUS campus with considerable latitude for campus-specific discretion. State regulations establish a set of baseline grade and test score requirements for standard admission. Table 4 describes the ‘sliding scale’ of grade requirements for given levels of test scores. Students with combined SAT scores (math and verbal) of less than 970 must have a high school GPA of least 3.0, while students who 1140 or above on their SATs need a GPA of only 2.0 for standard admission. In practice, there are few students with SAT scores above 970 who have marginal grades, so it is most often the 3.0 grade cutoff that binds. The third column of Table 2 reports the fraction of FIU applicants whose grades are less than 0.3 grade points above or below their SAT-contingent cutoff score that belong to each cutoff group. 72 percent of these marginal applicants have SAT scores below 970, and therefore face a grade cutoff of 3.0. An additional 19 percent did not take the SATs. Some of these students may have taken the ACT, for which there exists a similar crosswalk between scores and required grades. I assign these students a grade cutoff of 3.0, on the assumption that, like SAT takers, ACT takers will generally not score highly enough to reduce the cutoff from this value. I display results with and without SAT non-takers throughout the paper; the choice to include or exclude these students does not affect my findings.

Table 4: Florida SUS admissions rules

SAT	Required GPA	Fraction of marg. applicants
1140	2	0
1110	2.1	0
1090	2.2	0
1060	2.3	0.01
1030	2.4	0.01
1010	2.5	0.01
1000	2.6	0.01
990	2.7	0.01
980	2.8	0.02
970	2.9	0.02
<970	3	0.72
Did not take	3	0.19

Source: Florida Administrative Rule 6C-6.002. Sample: Marginal applicants are defined as all FIU applicants with FIU-computed GPAs within 0.3 grade points of their individual-specific cutoff GPA, computed using SAT scores. N=9,478. 1,829 students who did not take the SATs are assigned to the lowest score group.

Students with grades below the cutoff value may still be admitted, but only through a ‘student profile assessment’ that considers factors like family background, high school quality, and special talents. The number of students admitted through profile assessment is limited to 10 percent of total system wide admissions.¹² Nor does meeting the cutoff values guarantee admission to any SUS campus. State regulations explicitly grant campuses the power to ‘adopt... rules [that] increase standards for eligibility for admission.’

Individual campuses may also exercise discretion in GPA calculations. Rather than using a single high-school computed GPA for admissions to all SUS campuses, each campus admissions office takes transcript data and computes its own ‘admissions GPA.’ State regulations stipulate a broad set of courses that admissions GPAs must include, but evidence from applicants to multiple SUS campuses indicates that, while highly correlated, GPA calculations vary substantially across campuses.

Table 5 examines a sample of students who applied to both FIU and Florida State University (FSU), the SUS campus with which FIU had the largest number of same-year cross-applicants in the applications dataset. Panel A reports mean unweighted high school GPAs, FIU application GPAs, and FSU application GPAs for the set of 6,627 cross-applicants. The mean high school GPA for this group is 3.00, compared to a mean FIU GPA of 3.43 and a mean FSU GPA of 3.22. Clearly neither formula maps directly to unweighted grades computed by high schools, and the formula FIU uses to compute admissions GPAs from high school transcripts is more generous than the formula used by FSU. In fact, as shown in Table B1, FIU GPAs are highly correlated with but systematically higher than GPAs computed by other major SUS campuses.

The relative generosity of FIU GPAs has direct consequences for the status of applicants relative to their required grade cutoffs. Panel B of Table 5 displays counts of position relative to the cutoff for marginal FIU applicants who also applied to FSU. Of the 423 marginal FIU applicants whose FSU grades surpassed the required cutoff, all but 10 also surpassed the FIU cutoff. But of the 1,226 applicants who had grades satisfying the FIU cutoff, nearly two thirds (813) did not satisfy the FSU cutoff value. Panel C presents parallel results for admissions. Of the 162 marginal FIU applicants admitted to FSU, all but 21 were also admitted to FIU. But of the the 1,319 marginal applicants admitted to FIU, only 12 percent (141) were admitted to FSU.

¹²Source: Florida Administrative Rule 6C-6.002. Notably, race, gender, and country of origin are excluded from profile assessments.

There are two key points here. First, there is heterogeneity in the way SUS campuses evaluate identical high school transcripts. This means that even if students know their own GPAs and are aware of SUS grade cutoff rules, it will be difficult for them to manipulate their high school grades so as to land just above, rather than just below, the cutoff value at a particular school. Similarly, even if high schools report weighted grades directly to students, it is unclear which weights they would choose to use. Second, FIU does not grade as harshly as other SUS campuses. This means that some students may satisfy 'system wide' eligibility requirements *only* at FIU. If there is a cost associated with non-standard admissions, FIU may choose to admit more academically marginal students than other campuses, since it can do so via the standard admissions process. This helps explain why policy compliers who do not attend FIU typically do not attend any SUS campus.

Table 5: FIU and FSU admissions GPAs for joint applicants

Panel A: GPA means and SDs for all joint applicants

	Mean	SD
HS GPA	3	0.39
FIU GPA	3.43	0.5
FSU GPA	3.22	0.62

N=6627

Panel B: Status relative to grade cutoffs for marginal applicants

	FSU=1	FSU=0
FIU=1	413	813
FIU=0	10	537

N=1773

Panel C: Admissions for marginal applicants

	FSU=1	FSU=0
FIU=1	141	1,178
FIU=0	21	555

N=1773

Panel A: Sample consists of all students who applied to both FIU and FSU for the year following their senior year in high school. HS GPAs are unweighted cumulative GPAs provided by high schools. FIU and FSU GPAs are university-computed and taken from applications data. Panel B: Sample consists of students who applied to both FIU and FSU for the year following their senior year and had FIU GPAs within 0.3 grade points of their individual-specific admissions cutoff. Cell values are counts of student-level observations.

7 Results

7.1 First stage

Students' status relative to the grading threshold has a strong effect on college admissions and tertiary education choice. Table 6 shows local linear regression estimates of the effect of threshold-crossing on FIU admission, FIU attendance, SUS attendance and

intensity of SUS participation. Labels correspond to the notation in equation (1), so that g_i is $grade_i - c_i$ and Z_i is an indicator variable for threshold-crossing. The coefficient on Z_i represents the estimated discontinuity size. Students just above the threshold were 16.7 percentage points more likely to be admitted to FIU than students with grades just below the threshold. This is a large effect in the context of an overall admissions rate of roughly 50 percent, and it is statistically significant at the 0.1 percent level. As shown in Figure 1, a graphical analysis confirms that this discontinuity is large relative to overall admissions rates.

Table 6: Effect of threshold-crossing on SUS attendance and progress

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Admitted	Attend FIU	Attend SUS	YRS23	YRS25	CRED23	CRED25
g_i	0.835*** (0.104)	0.457*** (0.0928)	0.597*** (0.102)	1.915*** (0.405)	2.178*** (0.545)	3.810*** (0.893)	4.431*** (1.129)
$g_i \times Z_i$	-0.622*** (0.117)	-0.379*** (0.110)	-0.395*** (0.118)	-0.841 (0.473)	-1.083 (0.634)	-0.612 (1.078)	-0.886 (1.348)
Z_i	0.167*** (0.0236)	0.0709** (0.0221)	0.0877*** (0.0237)	0.344*** (0.0937)	0.415*** (0.126)	0.712*** (0.213)	0.755** (0.267)
Observations	9478	9478	9478	9478	8029	9478	8029

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of SUS admissions, attendance, and progress variables on grades relative to the cutoff value for marginal applicants. The YRS23 and YRS25 variables are years of SUS attendance by age 23 and 25, respectively. CRED23 and CRED25 are total SUS credits (divided by 10) by age 23 and 25. YRS25 and CRED25 specifications drop 2003 high school graduates from the sample. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. SUS credits are reported in units of ten credits. Standard errors are heteroskedasticity-robust.

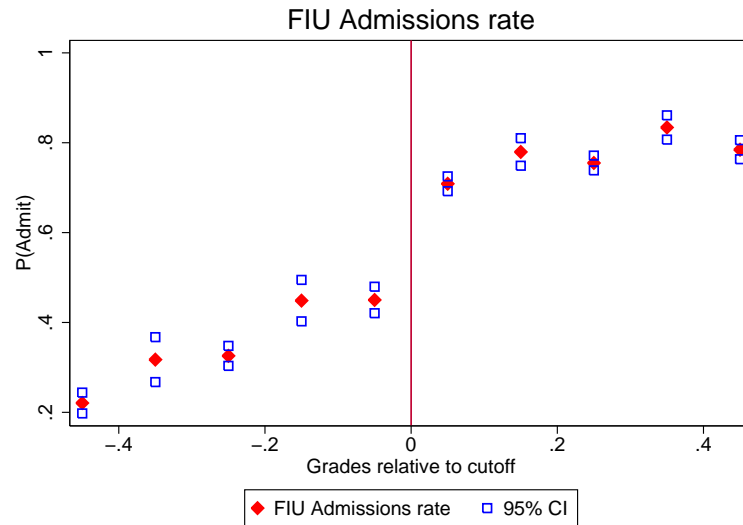


Figure 1: Admission probability

Note: Mean probabilities of college admission by grade relative to cut-off. Means are computed within 0.1 grade-point windows with the lower bound closed and the upper bound open; x-axis values on the graph correspond to interval lower bounds.

Threshold crossing also raised rates of college attendance. Students with grades just above the threshold were 8.8 percentage points more likely to enroll in the SUS in the year immediately following their senior year of high school. FIU enrollment accounts for the bulk of this increase— applicants just above the threshold were 7.1 percentage points more than those just below. This is consistent with evidence that crossing the FIU threshold is a relatively weak predictor of crossing the threshold for admission to other campuses. Both of these results are significant at the 1 percent level.

How does threshold-crossing affect educational attainment? I observe SUS enrollment and course taking through the 2008-2009 school year, the fifth school year after 2003-2004 high school graduates completed high school. Many students remain enrolled in college more than five years after high school. I therefore consider two sets of SUS attainment variables. The first set counts years of enrollment and total credits through five years after high school graduation, or approximately age 23. The second counts years of enrollment and total credits through seven years after high school graduation, or approximately age 25. When using the second set I will drop the 2003 cohort from analysis. Let YRS23 and YRS25 denote total years of SUS attendance by age 23 and 25,

respectively, and CRED23 and CRED25 denote total SUS credits (divided by 10) by ages 23 and 25.

Students induced to attend an SUS campus by threshold-crossing typically spent a significant amount of time in the SUS system. Students just above the threshold were enrolled in the SUS system for 0.344 more years by age 23 and 0.415 more years by age 25 than students just below. Under the assumption that increases in years enrolled and credits earned accrue entirely to students who enroll immediately after high school,¹³ the estimates indicate that, on average, compliers spent 3.9 ($.344 / .0877$) additional years in college by age 23 and 4.7 ($.415 / .0877$) additional years in college by age 25. Dividing estimates from the course credit specifications by the estimates from the years of college specifications indicates that compliers earn about 21 credits per year of enrollment in the first five years post high school, and about 18 credits per year of enrollment in the first seven years post high school. These numbers are consistent with the observation that an average credit load for an FIU student is about 20 credits per year. They also confirm that, at the very least, academically marginal students tend not to drop out of school quickly.

Table 7 shows the effect of threshold-crossing on 5-, 6-, and 7-year graduation probabilities. The effect of threshold crossing on five-year graduation rates is approximately zero. However, the probability of graduation in six or seven years rises by roughly 5.5 percent upon threshold crossing. This implies a six-year graduation rate of $5.5 / 8.77 = 0.63$ for academically marginal students—higher than but not statistically distinguishable from the 49 percent graduation rate for the 2000 entering cohort as a whole. Marginal applicants are unlikely to complete college within the standard four-year timeframe, but eventually many do graduate.

¹³This assumption is empirically valid. Regressions with first-time SUS enrollment in the second year after high school graduation as the dependent variable return zero coefficients on threshold crossing.

Table 7: Effect of threshold-crossing on SUS graduation

	(1)	(2)	(3)
	BA in 5 yrs.	BA in 6 yrs.	BA in 7 yrs.
g_i	0.176** (0.0606)	0.0216 (0.0872)	0.0711 (0.0947)
$g_i \times Z_i$	0.0124 (0.0751)	0.234* (0.105)	0.176 (0.113)
Z_i	0.00726 (0.0149)	0.0548** (0.0203)	0.0542* (0.0220)
Constant	0.125*** (0.0130)	0.170*** (0.0178)	0.225*** (0.0195)
Observations	9478	8029	8029

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of graduation dummies on grades relative to the cutoff value for marginal applicants. 6- and 7-year specifications drop 2003 high school graduates from the sample. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. Standard errors are heteroskedasticity-robust.

For many compliers, the choice to enroll in the SUS is the choice not to enroll in community college. Table 8 shows the effect of threshold crossing on immediate community college attendance, years spent enrolled in community college, and community college graduation. Threshold crossing renders students about 5 percent less likely to enroll in community college full time at any point in the three years after graduation. This estimate is significantly different from zero at the five percent level. Students who would otherwise enroll in community college thus account for about 58 percent (0.0511/0.0877) of threshold-crossers. Threshold crossers spend an average of 0.136 fewer years in community college by age 21 and 0.173 fewer by age 23. Interestingly, there is no corresponding effect for community college graduation—the effect of threshold crossing on rates of associate's degree or vocational certificate attainment five years after high school are approximately zero, as reported in columns three and four. Though many students who are not admitted to the SUS system choose to attend community college, very few graduate.

Columns five and six report estimates of the effects of threshold-crossing on other educational choice probabilities. The dependent variables in these regressions, attendance

at an out-of-state college or attendance at a in-state private college, are taken from surveys of high school seniors conducted by the Department of Education, and are available for 85 percent of marginal applicants. Threshold crossing has no effect on either probability. This is consistent with the observation that students near the admissions eligibility cutoff are very unlikely to attend private or out-of-state colleges.

Tables B2 through B4 display duplicate the estimates from Tables 6 through 8, adding controls for demographics and dropping students who did not take the SATs. As expected in an RD design, point estimates of the size of the discontinuity exhibit negligible changes.

Table 8: Effect of threshold-crossing on CC attendance and progress

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	CC within 3	CCYRS21	CCYRS23	AA in 5	VC in 5	Non-FL	FL priv.
g_i	-0.185 (0.101)	-0.657** (0.248)	-0.720* (0.323)	-0.0504 (0.0821)	-0.00646 (0.0201)	-0.00595 (0.0410)	-0.0926 (0.0597)
$g_i \times Z_i$	-0.0158 (0.118)	0.424 (0.279)	0.296 (0.364)	0.131 (0.0941)	0.0115 (0.0227)	-0.0200 (0.0469)	0.119 (0.0668)
Z_i	-0.0511* (0.0232)	-0.136* (0.0546)	-0.173* (0.0711)	-0.00870 (0.0183)	-0.00112 (0.00450)	-0.0000942 (0.00929)	-0.00366 (0.0127)
Constant	0.642*** (0.0208)	0.850*** (0.0499)	1.076*** (0.0651)	0.167*** (0.0166)	0.0102* (0.00410)	0.0336*** (0.00836)	0.0529*** (0.0117)
Observations	9478	9478	8029	9478	9478	8029	8029

Standard errors in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of SUS admissions, attendance, and progress variables on grades relative to the cutoff value for marginal applicants. CCYRS21 and CCYRS23 count years in which a student attend community college full time by 3 and 5 years after high school, respectively. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. Columns 3, 6, and 7 drop the 2003 cohort due to data limitations. Standard errors are heteroskedasticity-robust.

7.2 The returns to schooling

Table 9 presents OLS and IV estimates of the returns to schooling by age 23. Controls include grades, race dummies, free lunch status, own SAT score, a cubic polynomial in potential experience,¹⁴ and cohort dummies.¹⁵ To limit the impact of labor force partic-

¹⁴Potential experience is defined as years since 12th grade minus years of tertiary education. Years of tertiary education include any year in which a student attends a state university or attends community college full-time.

¹⁵Coefficients on the cohort dummies are not reported.

ipation choices on the analysis, I include only observations in which individuals earn at least \$8,000 (roughly 35 hours of work for 48 weeks at \$5/hour). I also drop a very small number of observations in which individuals earn more than \$200,000.¹⁶ Standard errors are clustered at the individual level. I include only student-year observations for which potential experience is greater than or equal to one.

Column 1 displays standard OLS Mincer regressions for the full population of students. These estimates are generally consistent with the existing literature on the returns to education. A year of SUS enrollment by age 23 raises earnings by 9.1 percent. Earnings are increasing in grades, higher for men, and lower for black students or students who receive free lunch. Earnings are higher for Hispanic students than for non-black, non-Hispanic students; this finding is non-standard and is likely related to the high proportion of Hispanics in the Florida labor market. Columns two and three report equivalent estimates for the sample of FIU applicants and the sample of marginal FIU applicants. For these groups the returns to a year of schooling are roughly 8 percent; signs and magnitudes of other estimated coefficients are similar to those for the full-sample specification; significance levels decrease with sample size.

Column 4 presents 2SLS estimates of the return to education for marginal applicants, instrumenting for years of schooling with threshold crossing. The IV estimate of the return to a year of SUS attendance is 8.8 percent, nearly identical to the OLS estimate for the full sample. This estimate is significantly different from zero at the 10 percent level. One reason for the large standard errors in this regression is the difficulty of simultaneously identifying two slopes and an intercept shift at the cutoff point within a narrow grade range. Though it is important to verify that the regression discontinuity design is robust to the inclusion of flexible slope controls, if one believes the regression discontinuity design is valid and the grade window is narrow enough, there is no need to control for slopes in grades at all. Put another way, point estimates of the return to schooling should not vary with the inclusion or exclusion of slopes in grades. One indication that this may be the case here is the fact that estimates of both slopes effectively zero.

¹⁶In the main sample, there are 551 such observations out of 547,069 total observations.

Table 9: Mincer earnings regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS: FS	OLS: Appl.	OLS: Marg.	IV: Marg.	IV: Marg.	IV: Marg.
YRS23	0.0907*** (0.000715)	0.0803*** (0.00161)	0.0769*** (0.00278)	0.0884* (0.0535)	0.0866*** (0.0117)	0.0744*** (0.0174)
YRS23×FR						0.0427* (0.0236)
YRS23×Black						-0.0218 (0.0227)
YRS23×Male						0.0110 (0.0259)
GPA	0.0468* (0.0254)	0.0582 (0.0812)				
GPA ²	0.0139*** (0.00433)	0.0168 (0.0137)				
g_i			-0.0223 (0.0640)	-0.0536 (0.163)		
$g_i \times Z_i$			0.0351 (0.107)	0.0577 (0.153)		
Pot. Exp.	0.122*** (0.00204)	0.117*** (0.00541)	0.121*** (0.00886)	0.123*** (0.0131)	0.121*** (0.00822)	0.123*** (0.00841)
Pot. Exp. ²	-0.00746*** (0.000389)	-0.00694*** (0.00110)	-0.00728*** (0.00176)	-0.00719*** (0.00181)	-0.00722*** (0.00158)	-0.00725*** (0.00159)
Pot. Exp. ³	0.000202*** (0.0000210)	0.000214*** (0.0000632)	0.000176* (0.000100)	0.000175* (0.000100)	0.000193** (0.0000890)	0.000177* (0.0000909)
Black	-0.0700*** (0.00355)	-0.0519*** (0.00952)	-0.0913*** (0.0165)	-0.0965*** (0.0296)	-0.105*** (0.0145)	-0.0658 (0.0434)
Hisp.	0.0560*** (0.00317)	0.0513*** (0.00806)	0.0367** (0.0149)	0.0353** (0.0165)	0.0279** (0.0136)	0.0304** (0.0138)
Male	0.0844*** (0.00264)	0.0684*** (0.00676)	0.0604*** (0.0117)	0.0620*** (0.0139)	0.0590*** (0.0103)	0.0357 (0.0481)
F/R lunch	-0.0167*** (0.00298)	-0.00349 (0.00659)	-0.00368 (0.0113)	-0.000873 (0.0174)	-0.00771 (0.0109)	-0.0836* (0.0430)
SAT	-0.0000221** (0.00000864)	-0.0000302 (0.0000230)	-0.0000538 (0.0000529)	-0.0000963 (0.000203)		
Constant	9.189*** (0.0381)	9.160*** (0.122)	9.679*** (0.0510)	9.504*** (0.0895)	9.450*** (0.0299)	9.480*** (0.0403)
Observations	547045	86280	28577	28577	36029	36029

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear regressions of log earnings on years of SUS attendance, potential experience, and student covariates. Sample includes only workers defined as full-time (those who make >\$8,000). For full sample and applicant sample, GPA and GPA_2 are high-school computed un-weighted GPAs. g_i is an admissions-computed GPA less the SAT-score specific cutoff value. Z_i is a dummy equal to one if $g_i > 0$. Potential experience is equal to $age - schooling - 5$, where schooling includes years of SUS attendance as well as years of full-time CC attendance. See Appendix A for more detail. Standard errors are heteroskedasticity robust and clustered at the individual level.

Column 5 tests this argument by dropping controls for grades. I also drop controls for SAT score, expanding the sample to include students who did not take the test. The point estimate of the returns to a year of SUS attendance remains effectively unchanged, falling by a statistically negligible 0.002 to 0.087. The standard errors drop by a factor of more than four, so that the coefficient estimate is now significant at the one percent level. I interpret the invariance of the point estimate to the elimination of controls as evidence that the regression discontinuity is valid, and I will exploit the gains in precision in my analysis of heterogeneity in returns.

Column 6 allows returns to college to vary across three demographic categories: free lunch students, black students, and male students. There is no evidence of heterogeneity in returns for black students or male students. For free lunch students, however, each additional year of college yields earnings gains that are 4.3 percent larger than those for other students. The estimate is statistically significant at the ten percent level. These findings validate prediction one from the theoretical model: free lunch students who take up the offer of college admission realize larger returns than other students who do so.

7.3 Labor force participation while in college

I now test the model prediction that free lunch students work more while in college than students from wealthier backgrounds. Table 10 presents results from local linear regressions of labor force participation variables on status relative to the grade cut-off, controlling for experience (defined here as years after the 12th grade year) using a cubic polynomial. Only post-12th-grade student-year observations are included, and standard errors are clustered at the student level. Columns one and two consider the effect of threshold crossing on a dummy variable for any labor force participation (i.e., positive earnings). Column one restricts the effect of threshold-crossing on labor force participation to be an intercept shift that is the same for all levels of experience. The point estimate for this effect is near zero and insignificant. One might imagine that an average shift size of zero masks a negative effect immediately after high school, when above-threshold students are more likely to be in college, and positive effect at higher levels of experience. Column two allows for two shifts; one for the first five years of experience, and another for more than five years of experience. Both effects are still zero, which indicates that a) increased probability of college attendance in the years following high school is not associated with an extensive-margin decrease in labor force

participation, and b) increased probability of college attendance is not associated with differing post-college probability of labor force participation.

Column three investigates the intensive margin of labor force participation. The specification is identical to column 2, but the dependent variable is an indicator equal to one if an individual makes more than \$8,000 dollars in a given year. I refer to individuals who earn at least this much as full time labor force participants. Here again, effects are small in magnitude and statistically insignificant. In the population as a whole, threshold crossing does not have a significant impact on labor force participation.

Average effects, however, mask substantial heterogeneity in the early effects of threshold-crossing. As shown in column four, threshold crossing reduces the probability of full time labor force participation for non-free lunch students less than six years out of high school by 3.8 percentage points. This effect is statistically significant at the 5 percent level. For free lunch students the equivalent effect is slightly positive and not statistically significant. This is consistent with prediction two from the model of college choice and labor supply under credit constraints: constrained students granted access to college work more at first than unconstrained students granted the same access. In columns five and six, I directly estimate the effects of SUS attendance on same-year earnings and labor force participation, instrumenting for SUS attendance with threshold-crossing and allowing different effects for free lunch and non-free lunch students. SUS attendance is associated with a \$6,572 reduction in earnings for non-free lunch students, but this estimate is statistically noisy and insignificant. Free lunch students earn an average of \$7,820 more than other students when they attend SUS, and this difference is significant at the one percent level. They are also 24 percent more likely to earn at least \$8,000. This is consistent with model predictions and to be expected given the results in column four. Importantly, this result does not appear to be driven by selection of harder-working free lunch students into the marginal group, as the main effect of free lunch is zero or negative in all specifications. Poor students work more only when they attend college.

Table 10: Labor force participation by status relative to cutoff and age

	(1)	(2)	(3)	(4)	(5)	(6)
	In LF	In LF	In LF FT	In LF FT	Earnings	In LF FT
g_i	-0.0412 (0.0618)	-0.0412 (0.0618)	-0.0381 (0.0711)	-0.0686 (0.0704)	478.0 (4121.6)	0.0138 (0.163)
$g_i \times Z_i$	0.0437 (0.0709)	0.0438 (0.0709)	0.0943 (0.0813)	0.110 (0.0803)	1259.9 (2816.9)	0.0446 (0.110)
Z_i	0.00498 (0.0140)					
PRE $\times Z_i$		0.00395 (0.0140)	-0.0125 (0.0161)			
POST $\times Z_i$		0.00606 (0.0147)	0.00947 (0.0166)			
PRE \times NOFR $\times Z_i$				-0.0380* (0.0173)		
PRE \times FR $\times Z_i$				0.0194 (0.0177)		
POST \times NOFR $\times Z_i$				-0.00309 (0.0179)		
POST \times FR $\times Z_i$				0.0331 (0.0185)		
SUS					-6571.6 (7343.7)	-0.272 (0.289)
SUS*FR					7820.6** (3021.3)	0.240* (0.115)
F/R Lunch				-0.00740 (0.0118)	-2878.0* (1156.3)	-0.0724 (0.0430)
Constant	0.758*** (0.0156)	0.759*** (0.0156)	0.170*** (0.0179)	0.132*** (0.0201)	6556.7 (4314.1)	0.265 (0.170)
Observations	92340	92340	92340	92340	63448	63448

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of dummies for labor force participation on grades relative to the cutoff value for marginal applicants. Observations are at the student-year level. The labor force dummies in (1) and (2) are equal to one if a student reports any earnings in that year. The labor force dummy in (3) is equal to one if a student reports at least \$8,000 in income that year. Columns one through four present OLS results. In column five I instrument for SUS and $SUS * FR$ using threshold-crossing and threshold-crossing interacted with FR . g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. PRE and $POST$ are dummy variables for, respectively, at most five and more than five years of experience. FR and $NOFR$ are free lunch and no free lunch dummies. Additional controls are a cubic term in experience, cohort fixed effects, and in columns four through six, demographic dummies including a free-lunch dummy. Columns five and six use observations seven or fewer years after high school graduation (five for 2003 cohort). Standard errors are heteroskedasticity-robust and clustered at the student level.

8 Robustness of the RD design

8.1 Grade manipulation

The main concern about the internal validity of the RD design is that students who have an unobservably greater incentive to attend FIU may manipulate their grades so that they fall just above the cutoff value. Were this behavior to occur systematically, it would contravene the assumption that the expected returns to SUS attendance are a smooth function of grades near the cutoff value. One standard falsification exercise is to look for discontinuities in the density of FIU grades at the cutoff point. The argument is that if some students manipulate their grades to surpass the threshold, the density of the grade distribution will be higher just above the cutoff than just below.

Unfortunately, this exercise is unhelpful if distributional discontinuities at the cutoff point can be traced to other factors. That is the case here. For most individuals, the relevant cutoff GPA is 3.0. This corresponds to an unweighted 'B' average—a benchmark grade level that teachers and FIU evaluators may be more likely to assign or students more likely to work to obtain for reasons exogenous to the admissions process than other nearby GPAs. The empirical distribution of grades is consistent with this idea. The left panel of Figure 2 shows a histogram of FIU GPAs for all applicants with SAT scores. There is sharp discontinuity at the 3.0 grade level. This could be the result of strategic cutoff-crossing, or of an alternative process related to the 'B' grade. The numerous other jumps and drops in the density at non-cutoff points suggest the latter story may be important.

Looking only at students for whom the 3.0 cutoff is not in effect provides further evidence of this. The right panel of Figure 2 shows a histogram of FIU GPAs for students with cutoff GPAs of less than 3.0. Because these students by definition have higher SAT scores than students with the 3.0 cutoff, the entire grade distribution is shifted to the right. However, there remains a sharp discontinuity at the 3.0 grade level, which cannot be the result of grade manipulation with respect to the admissions cutoff.

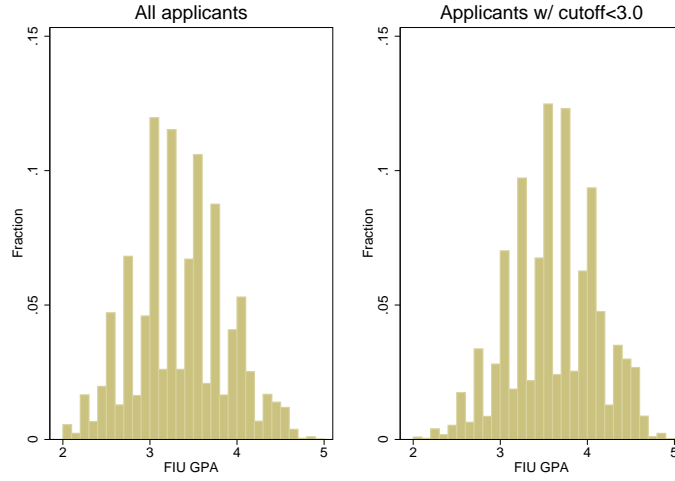


Figure 2: Histogram of FIU GPA

Left panel: histogram of GPAs for the sample of FIU applicants with SAT scores. Bins are discrete, so that each value has its own bin. Right panel: identical to left panel but excluding applicants for whom cutoff GPA is 3.0.

A more informative visual test for grade manipulation in the context of a running variable that may be discontinuously distributed for exogenous reasons is to look for continuity in the ratios of the conditional densities to the unconditional density,

$$\frac{f(g|x)}{f(g)}. \quad (7)$$

$f(g)$ and $f(g|x)$ are the unconditional and conditional densities of g_i , respectively. To understand this test, assume that observable and unobservable wage determinants (x, u_1) have some continuous unconditional distribution $h(x, u_1)$. For the RD design to produce unbiased results, the distribution of wage determinants conditional on grades, $h(x, u_1|g)$, must be continuous in g (Lee and Lemieux, 2009). Via Bayes' rule,

$$h(x, u_1|g) = h(x, u_1) \frac{f(g|x, u_1)}{f(g)} \quad (8)$$

Thus $h(x, u_1|g)$ is continuous if the ratio of the conditional to unconditional densities is continuous. Equation 7 tests this requirement using the observable wage determinants

only. This test is in a sense more direct than looking only at the continuity of $f(g)$, since it focuses specifically on the object that determines the continuity of wage determinants in grades. The intuition is also clear. If discontinuities in the grade distribution are due to a process that is exogenous to the determination of the treatment, discontinuous jumps in the conditional distributions should be matched by discontinuous jumps in the unconditional distribution. The ratio of the two densities should be continuous even if each individual density is not.¹⁷

Figure 3 presents the density ratios described in equation 7 for three different conditioning groups: black students, Hispanic students, and students who receive free or reduced price lunch. Each point represents the ratio of the proportion of observations in the sample of students with the stated characteristic to the proportion of all observations within a 0.1 grade-point bin. Consistent with a valid RD design, each density ratio is continuous around the cutoff value.

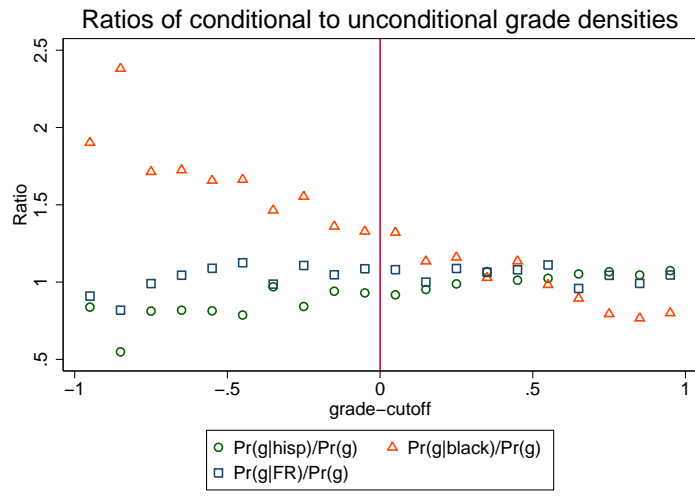


Figure 3: Ratios of conditional to unconditional grade densities

Ratio of the proportion of conditional observations to the proportion of unconditional observations within 0.1 grade-point bins. Bins ordered by grade relative to cutoff. Points are located at the midpoint of the corresponding bin. Sample: FIU applicants with valid SAT scores and demographic data.

¹⁷See Appendix D for a more detailed discussion of the relationship between this continuity assumption and Lee and Lemieux’s ‘imprecise control’ condition for the unbiasedness of RD designs. Appendix D also presents a set of conditions under which RD designs are valid even in the presence of a discontinuously-distributed running variable.

The RD design also passes standard tests for continuity in the means of the X -variables. Table 11 reports estimates of local linear threshold-crossing regressions in which the dependent variables are predetermined demographic characteristics. The coefficient on the threshold-crossing indicator is approximately zero when the dependent variable is sex, lunch status, own SAT score, and school mean 10th-grade FCAT scores for the 2004 cohort.¹⁸ There is thus no indication of selection on the basis of gender, family income, own measured ability, or school quality.

Point estimates of discontinuities for race dummies are larger, though they remain insignificant at the five percent level. The signs on the coefficients are the reverse of what one would expect under the hypothesis of selection into treatment based on earnings levels: students above the cutoff are more likely to be black, while students below the cutoff are more likely to be Hispanic. As shown above, being black is associated with lower earnings in this population of students, while being Hispanic is associated with higher earnings. This suggests that if selection on unobservables mirrors selection on observables, it will bias estimates of returns downward. The more likely hypothesis is that there is little systematic selection on observables and the relatively large coefficients on the race dummies represent statistical noise.

Column six of Table 11 distills this conclusion into a single specification. I construct an earnings-weighted index of the X -variables from columns one through five by regressing log earnings six years post high-school¹⁹ on those variables and computing the predicted earnings value. I then check for discontinuities in this index at the cutoff point. The estimated discontinuity is negative and statistically insignificant at the five percent level.

¹⁸The Florida Comprehensive Achievement Test is an NCLB-mandated reading and math evaluation.

¹⁹The largest value of experience for which I observe every cohort.

Table 11: Local distribution of predetermined variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Black	Hisp.	Male	F/R lunch	SAT	Weighted Index
g_i	-0.372*** (0.104)	0.245* (0.108)	-0.243* (0.107)	-0.121 (0.109)	93.60*** (27.48)	0.0237 (0.0146)
$g_i \times Z_i$	0.167 (0.118)	-0.0620 (0.124)	0.203 (0.122)	0.155 (0.125)	18.28 (30.80)	-0.0133 (0.0164)
Z_i	0.0430 (0.0231)	-0.0370 (0.0244)	-0.00812 (0.0239)	0.00985 (0.0244)	1.972 (6.087)	-0.00533 (0.00321)
Constant	0.293*** (0.0209)	0.499*** (0.0220)	0.363*** (0.0216)	0.478*** (0.0221)	837.9*** (5.604)	10.00*** (0.00294)
Observations	9478	9478	9478	9478	7649	7649

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of student covariates on grades relative to the cutoff value for marginal applicants. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. Standard errors are heteroskedasticity-robust.

8.2 Censoring

Another possible source of bias is the censoring of earnings data. The data only covers in-state earnings, so if students on one side of the cutoff value are more likely to leave the state than students on the other side of the cutoff value, it could confound estimates of earnings equations. One plausible story is that students above the threshold are more likely to attend college, and students who do well in college are more likely to receive lucrative out-of-state job offers. Then high-earning above threshold students would be censored from the data and estimates would be biased down. If selection into out-migration is correlated with free lunch status, estimates of differential returns for free lunch students will also be biased. Fortunately, outmigration does not appear to be correlated with threshold-crossing. Columns two and three of Table 10 (above) show that, in the population of marginal applicants, threshold crossing does not have a statistically significant impact on labor force participation six or more years after high school. Column four shows that these effects are statistically insignificant for both free-lunch and non-free lunch students. Threshold-crossing does not appear to affect long-

term labor force participation outcomes, so there is little reason to believe that censoring will lead to bias in estimates of the returns to schooling.

8.3 Robustness to alternative specifications

I now consider the robustness of my results to changes in variable and sample definitions, and to two simple but possibly confounding theories about the earnings process. Table B5 presents re-estimates the key specifications from Table 11 using years of SUS attendance by age 25 as the independent variable and dropping the 2003 cohort from the sample. Estimates of the returns to SUS attendance drop by roughly one percent in all specifications. This is consistent with the idea that one of the payoffs to completing a year of schooling by age 23 is the option to complete an additional year by age 25. In particular, free lunch students earn about 3.5 percent more than other students per year of age 25 attendance; the p-value for this estimate is 0.102.

Table B6 presents results for a set of alternative IV specifications. Columns 1 and 2 vary the bandwidth used to define marginal students. Column 1 moves the bandwidth in to 0.25 grade points on either side of the cutoff, while column 2 raises it to 0.6 grade points. The first change produces a statistically insignificant reduction in estimated returns, while the second change raises estimated returns slightly to 10.5 percent. Column 3 adds dummies for cutoff groups (which are equivalent to categorical SAT score dummies) and for school districts. The estimated returns do not change. Columns 4 and 5 vary the income cutoff for sample inclusion. In the main specification the cutoff was \$8,000. In column 4 the cutoff is \$1,000. This change raises the point estimate of returns to 12 percent, and also raises the associated standard error. In column 5 the cutoff is \$12,000. This change reduces estimated returns to 7.3 percent. The estimate remains highly significant.

Table B7 reports results for another set of robustness checks. As noted above, only two quarters of earnings data are available for the 2009-2010 academic year. In the main specifications, values for 2009-2010 earnings are imputed by multiplying total earnings for the two available quarters by 1.82, the mean ratio of total earnings to first half earnings in 2008-2009. Columns 1 and 2 show results obtained using alternate imputation methods. Column 1 predicts annual earnings using coefficients from a regression of full-year earnings on first-half earnings in the 2008-2009 year. Column 2 imputes by doubling first half earnings. Neither choice has a meaningful effect on estimates of re-

turns to schooling. Column 3 addresses the same problem via a regression in which half-year earnings are the dependent variable in all years. The resulting changes in the point estimate are not statistically significant.

Column 4 addresses a different critique. FIU is located in Miami. If wage rates in Miami are higher than wage rates in less populated areas of Florida (perhaps at the cost of higher living expenses), and students who move to Miami to attend college are more likely to work there post-college, then the reported estimates of the returns to college might reflect a Miami-specific wage premium. In fact, most FIU applicants are from the Miami-Dade school district, making college-associated migration to Miami less likely to be an issue. Column 4 reports estimates of the returns to college for marginal applicants from the Miami-Dade school district only. These results are again nearly identical to the main results.

Finally, I consider the possibility that the higher returns observed for free lunch recipients result from the additional work experience they accumulate while in college. The expected impact of in-college work experience on post-college earnings is ambiguous: working while in college may provide students with valuable labor market experience, but may also detract from academic effort. The empirical evidence on the effects of work while in college on post-college earnings is also mixed. Light (2001) finds that controlling for work within school reduces estimated returns to schooling by between 25 and 44 percent, while Hotz et al. (2002) find little effect after correcting for selection. In any case, if differential returns for free lunch students could be attributed to this simple mechanism, one would like to know.

Columns 5 and 6 present IV estimates in which I control directly for a quadratic in years of work experience between students' 12th grade year and their last year of college attendance. See Appendix A for a description of how I create these variables. In column 5, I restrict the coefficients on the in-school experience terms to values reported in Light. In column 6, I allow the coefficients to vary freely. Estimates of the main effect of a year of college attendance drop by 20 to 35 percent, similar to estimates reported in Light. However, the differential effect for free lunch recipients is more or less unaffected.

9 Discussion

9.1 Internal rates of return

The results presented here show that the returns to college for students on the academic margin of college attendance are high, and very similar to the returns to college for the broader college student population. To place estimates of returns in context it is helpful to conduct a back-of-the-envelope calculation of the internal rate of return on the investment in schooling for a typical student. Consider a student who attends a state university campus for s years. Assume (conservatively) that the student does not work at all while on attending school. The internal rate of return r is defined as the rate of interest that equates the present value of the net income stream generated for attending s years of college and with the equivalent quantity associated with zero years of college attendance. That is, r is chosen so that

$$\begin{aligned} & \sum_{t=s+1}^{T+s} \left(\frac{1}{1+r} \right)^t E[Y_{it}|X_i = x, g_i = g, S_i = s] - \sum_{t=1}^s \left(\frac{1}{1+r} \right)^t E[C_{it}] \\ & - \sum_{t=1}^T \left(\frac{1}{1+r} \right)^t E[Y_{it}|X_i = x, g_i = g, S_i = 0] = 0 \end{aligned} \quad (9)$$

C_{it} are the direct costs of college attendance. The no-work assumption will tend to bias estimates down, since evidence indicates that most students work while in college. I further assume that all students have careers of length T .

Recall that from equation 6,

$$\ln Y_{it} = X_i' \delta_1 + f(t) + \gamma_1^1 g_i + \gamma_2^1 Z_i g_i + \rho S_i + u_{1i}.$$

Assume further that when g_i is near the cutoff value, u_{1i} is independent of the included covariates and is normally distributed with variance σ^2 .²⁰ Then it is easy to show that

²⁰The RD design already guarantees that u_{1i} is uncorrelated with the independent variables and has zero mean.

$$E[Y_{it}|X_i = x, g_i = g, S_i = s] = \exp(x'\delta_1 + f(t) + \gamma_1^1 g + \gamma_2^1 Z_i g + \rho s) \times \exp\left(\frac{1}{2}\sigma^2\right) \quad (10)$$

For a given value of r , estimates of the first and last terms of equation 5 can be computed by plugging in sample estimates of the parameters in equation 6. Direct costs C_{it} can be computed with reasonable accuracy using published student cost records.

I compute the internal rate of return for a ‘typical’ student: a Hispanic male free-lunch recipient in the 2000 graduating high school cohort who scored 842 (the sample mean for marginal students) on the SAT. I focus on students with grades exactly at the cutoff ($g_i = 0$). I consider the net return to $s = 5$ years of college (the rounded mean number of years obtained by policy compliers), over $T = 40$ year careers. I use estimates from column three of table B-5 for the returns to college through age 25 because the longer time frame better reflects total educational attainment. Because the experience variable has limited support in my data, I replace my estimated experience profiles with quadratic profiles from Heckman et al. (2003).²¹ I assume that this individual paid the average level of full-time tuition, fees, and room and board expenses in the SUS as a whole.²² The cost calculation procedure is conservative because it assumes that those who do not attend college do not pay for room and board, and because SUS average expenses are generally slightly higher than FIU expenses.

I find that the internal rate of return on investment is $r = 0.074$. This value is only slightly smaller than the estimated effect of schooling on earnings, $\rho = 0.081$. This indicates that the direct cost of school attendance and the opportunity cost of forgoing earnings while in school are relatively small compared to the earnings gains that accrue to students who are able to attend school. These findings are similar to estimates of internal rates of return reported elsewhere in the literature. For instance, using data from the NLSY, Belzil and Hansen (2002) estimate the internal rate of return as 0.068 for students attending low-cost colleges. Their result reflects an average over a broad distribution of student types. The results presented here indicate that college is a similarly profitable investment for students at the lower academic bound of admission.

²¹I use reported estimates for white males in the 1990 Census. The linear term is 0.1301 and the quadratic term is -0.0023.

²²Tuition, fees, and room and board values for the academic years 2000-2001 through 2004-2005 are taken from the following NCES tables: 313 (2002), 314 (2004), 320 (2006). Values are inflated to 2005 dollars using the PCE for the spring of the academic year.

They also suggest that the social returns to admitting marginal students to the SUS are quite large even if the state subsidizes students' costs to some extent. Under the assumptions that a) the social returns to college attendance are identical to private returns, and b) the state incurs an additional cost equal in present value to the direct cost paid by the student, the internal social rate of return to a year of college attendance is $r = 0.067$, well above a standard market interest rate of five to six percent.

9.2 Conclusions and directions for future work

I have shown that students at the academic margin of admittance to four-year college realize large returns if they are able to attend, and have provided evidence that at least some of these students face binding credit constraints. These findings imply that straightforward policy changes could benefit marginal students in a number of ways. First, a reduction in the cutoff GPA required for admission to the Florida SUS would enable many Florida students to make human capital investments with high private returns. Even college applicants on the low end of the ability distribution appear ready to benefit from public higher education. For these students, it is enrollment constraints and not preparation constraints that bind. Second, at least some students at this academic margin are credit constrained and would benefit from expanded need-based aid programs or from a reduction in the standards for merit aid eligibility. Florida has a large state-level merit aid program (the Florida Bright Futures scholarship), but the SAT score cutoff for eligibility to the lowest non-vocational tier of the program was 970 for the students studied here. Only 10 percent of SAT-takers in my sample achieved a score this high.

A corollary of these findings is that community college does not appear to be an academically or economically rewarding pathway for students at the margin of attending a four year college. This does not imply that community college does not help other groups of students: the average community college student has a weaker academic background than the average four-year marginal student, and may benefit from different kinds of coursework. But it does indicate that community college does little to help better-prepared students capitalize on their substantial academic potential. Four-year marginal students face the risk of becoming 'caught in the middle' between four-year institutions that do not consistently admit them and two-year institutions that do not help them

Of course, the welfare effects of increasing SUS capacity or providing additional merit aid depend not on private returns for marginal students, but on the social returns to education for all students. The social returns to education may be substantially higher than the private returns, but expanding access to public higher education could have system-wide costs if, for instance, it requires a drop in per-student funding. My findings on the high internal rate of return for marginal students suggest that even if the social returns to education are no higher than the private returns, social costs would have to be substantially higher than private costs to bring returns on investment down to a market rate of five to six percent.

The main weakness of my empirical work is the lack of exogenous variation in credit constraints. Though students who receive free lunch are presumably more likely to face credit constraints than students who do not, they may differ from non-free lunch students in other ways as well. My findings are consistent with any hypothesis under which an unobservable correlate of free lunch status raises work while in college and the returns to college. Available evidence contraindicates some intuitive stories. For instance, say that reaching the margin of college admittance presents a greater challenge for free lunch students than for other students. Free lunch students at that academic level would then be harder workers on average than non-free lunch students, which could affect both labor supply while in college and post-college returns. However, this story suggests that, contrary to what is observed in the data, free lunch students at the margin would out-work and out-earn other students even if they did not attend college. But other stories about cross-group differences in the distribution of unobservables are fundamentally impossible to rule out in this way.

One promising pathway for future work is to combine an admissions cutoff like this one with a merit aid cutoff based on some other measure of student academic achievement, like test scores. The intersection of of aid and admissions eligibility cutoffs would create exogenous variation in both college access and credit constraints for students at those margins. Florida itself has a large state-level merit aid program with eligibility based in part on SAT scores, but students at the admissions margin almost universally have SAT scores that are too low to qualify. However, a number of other states have cutoff based merit aid programs that may interact more favorably with admissions standards. See Dynarski (2004) for a state-by-state summary.

References

- Acemoglu, D. and D. Autor**, "Skills, tasks and technologies: Implications for employment and earnings," *Handbook of Labor Economics*, 2011, 4b.
- Angrist, J.D. and G.W. Imbens**, "Two-stage least squares estimation of average causal effects in models with variable treatment intensity," *Journal of the American Statistical Association*, 1995, 90 (430), 431–442.
- Attewell, Paul A., David E. Lavin, Domina Thurston, and Tania Levey**, "New Evidence on College Remediation," *Journal of Higher Education*, 2006, 77 (5), 886–924.
- Belley, P. and L. Lochner**, "The changing role of family income and ability in determining educational achievement," NBER working paper 13527 2007.
- Belzil, Christian and Jorgen Hansen**, "Unobserved Ability and the Return to Schooling," *Econometrica*, 2002, 70 (5), 2075–2091.
- Cameron, S.V. and C. Taber**, "Estimation of educational borrowing constraints using returns to schooling," *Journal of Political Economy*, 2004, 112 (1), 132–182.
- Card, D.**, "The causal effect of education on earnings," *Handbooks in Economics*, 1999, 5 (3), 1801–1864.
- Carneiro, P., J.J. Heckman, and E.J. Vytlacil**, "Estimating marginal returns to education," NBER working paper 16474 2010.
- Carneiro, Pedro and James J. Heckman**, "The Evidence on Credit Constraints in Postsecondary Schooling," *Economic Journal*, 2002, 112 (482), 705–734.
- Deming, David and Susan Dynarski**, "Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor," NBER Working Papers 15387, National Bureau of Economic Research, Inc 2009.
- der Klaauw, W. Van**, "Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach," *International Economic Review*, 2002, 43 (4), 1249–1287.
- Dynarski, S.M.**, "The behavioral and distributional implications of aid for college," *American Economic Review*, 2002, 92 (2), 279–285.

- , “The new merit aid,” in C. Hoxby, ed., *College choices: The economics of where to go, when to go, and how to pay for it*, Chicago: University of Chicago Press, 2004, pp. 63–100.
- , “Building the Stock of College-Educated Labor,” *Journal of Human Resources*, 2008, 43 (3), 576–610.
- Goldin, C., L.F. Katz, and L. Center**, “The race between education and technology: The evolution of US educational wage differentials, 1890 to 2005,” *NBER Working Paper*, 2007.
- Goldin, C.D. and L.F. Katz**, *The race between education and technology*, Cambridge, MA: Harvard University Press, 2008.
- Heckman, James J., Lance Lochner, and Petra E. Todd**, “Fifty Years of Mincer Earnings Regressions,” NBER Working Papers 9732, National Bureau of Economic Research, Inc 2003.
- Hoekstra, M.**, “The effect of attending the flagship state university on earnings: A discontinuity-based approach,” *The Review of Economics and Statistics*, 2009, 91 (4), 717–724.
- Hotz, V. Joseph, Lixin Colin Xu, Marta Tienda, and Avner Ahituv**, “Are There Returns To The Wages Of Young Men From Working While In School?,” *The Review of Economics and Statistics*, 2002, 84 (2), 221–236.
- Imbens, G.W. and T. Lemieux**, “Regression discontinuity designs: A guide to practice,” *Journal of Econometrics*, 2008, 142 (2), 615–635.
- Jacobson, Louis, Robert John LaLonde, and Daniel G. Sullivan**, “Estimating the returns to community college schooling for displaced workers,” *Journal of Econometrics*, 2005, 125 (1-2), 271–304.
- Johnson, Matthew T.**, “Borrowing Constraints, College Enrollment, and Delayed Entry,” *Mimeo*, 2010.
- Kane, Thomas J.**, “Public Intervention in Post-Secondary Education,” in Erik Hanushek and F. Welch, eds., *Erik Hanushek and F. Welch, eds.*, 1 ed., Vol. 2, Elsevier, 2006, chapter 23, pp. 1369–1401.
- Kane, Thomas J and Cecilia Elena Rouse**, “Labor-Market Returns to Two- and Four-

- Year College," *American Economic Review*, 1995, 85 (3), 600–614.
- Kane, T.J.**, "A quasi-experimental estimate of the impact of financial aid on college-going," *NBER Working Paper 9703*, 2003.
- Katz, L.F. and K.M. Murphy**, "Changes in relative wages, 1963-1987: Supply and demand factors," *The Quarterly Journal of Economics*, 1992, pp. 35–78.
- Keane, M.P. and K.I. Wolpin**, "The effect of parental transfers and borrowing constraints on educational attainment," *International Economic Review*, 2001, 42 (4), 1051–1103.
- Lee, D. and T. Lemieux**, "Regression discontinuity designs in economics," NBER working paper 14273 2009.
- Light, Audrey**, "In-School Work Experience and the Returns to Schooling," *Journal of Labor Economics*, 2001, 19 (1), 65–93.
- Meghir, C. and S.G. Rivkin**, "Econometric methods for research in education," NBER working paper 16003 2010.
- Roderick, M., J. Nagaoka, and V. Coca**, "College readiness for all: The challenge for urban high schools," *The Future of Children*, 2009, pp. 185–210.
- Stinebrickner, R. and T. Stinebrickner**, "The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study," *The American Economic Review*, 2008, 98 (5), 2163–2184.

Appendix

A Data description

This appendix describes the construction of variables used in the analysis.

1. *F/R lunch*. The FLDOE reports free or reduced lunch status for each year in which a student is enrolled in a public school. This variable is equal to one if a student ever reports receiving free or reduced price lunch.
2. *Miami-Dade*. This variable is equal to one if the student was part of a 12th grade

cohort in a high school in the Miami-Dade school district.

3. *HS GPA*. This variable is the high-school reported cumulative unweighted GPA as of a students' last enrollment.
4. *FIU GPA*. FIU computes GPAs as part of the admissions process. This variable is taken from their records.
5. *SAT score*. SAT scores used here reflect maximum combined reading and math scores, possibly over multiple attempts.
6. *SUS next*. A dummy variable equal to one if students attend any SUS campus the academic year following their 12th grade year. This variable is based on enrollment data from all SUS institutions.
7. *YRS23*. The number of academic years within a student's first five years after high school graduation (i.e., by the age of roughly 23) in which a student reports being enrolled as a full-time SUS student for at least one academic term.
8. *YRS25*. The number of academic years within a student's first seven years after high school graduation (i.e., by the age of roughly 25) in which a student reports being enrolled as a full-time SUS student for at least one academic term. Available only for pre-2003 cohorts.
9. *Non-FL college*. The FLDOE collects survey data from Florida students about post-high school plans in their 12th grade year. This variable is a dummy combining all responses that indicate the intent to attend any type of tertiary institution outside of Florida. This variable is not available for the 2004 graduation cohort.
10. *FL private college*. As above, but for in-state private colleges.
11. *SUS BA in 6 years*. This variable is generated using degree records for all public tertiary institutions. It is coded as a one if students receive a BA from any institution within six years after their 12th grade year. Available only for pre-2003 cohorts.
12. *CC next*. Equal to one if a student attends a community college full time at any point in the year following the 12th grade year. Taken from CC enrollment records.
13. *CC within 3*. Equal to one if a student attends a community college full or part time at any point in the three years following the 12th grade year. Taken from CC

enrollment records.

14. *AA in 5 years*. This dummy variable is equal to one if students receive any type of two-year degree from any public tertiary institution within three years after the 12th grade year. It is compiled using degree records from all public tertiary institutions.
15. *Labor force participation*. A dummy equal to one for any individual who reports positive earnings in any quarter of an academic year. Taken from FL UI records.
16. *Full time labor force participation*. As above, but equal to one only for individuals who report earnings of at least \$8,000 in the academic year.
17. *Earnings*. AY earnings from quarterly FL UI records. Deflated to 2005 dollars using the quarterly PCECTPI.
18. *Experience*. Years since the 12th grade year.
19. *Potential experience*. Experience minus total years of full-time education within the first seven years after high school (or five years post high school for the 2003 cohort) minus five.
20. *School FCAT*. The mean combined 2004 reading and math score for 10th graders in an individuals' 12th grade school.
21. *SX*. Labor market experience while in school. To obtain this value I take year total earnings and divide by \$6/hour, an approximation of mean wage for low-wage workers. I cap hours at 2,000 within each year. I then create SX by summing over all years between the senior year of high school and the year in which potential experience is equal to zero, and dividing the result by 2,000.

B Additional tables and figures

Table B-1: Common applicant GPA regressions

	FIU	UCF	UF	USF	UNF	FAU	FSU
FIU	1	0.79	0.82	0.8	0.79	0.83	0.76
	0	0.84	1	0.78	0.82	0.65	0.96
	30251	5285	5785	4422	941	4705	6627
UCF	1.14	1	1.02	0.97	0.97	1.02	0.96
	-0.64	0	0.24	0.1	0.13	-0.1	0.19
	5285	31899	12067	10567	2660	4546	13650
UF	1.08	0.88	1	0.9	0.87	0.92	0.9
	-0.69	0.12	0	0.07	0.2	-0.06	0.08
	5785	12067	35907	8416	1537	2655	15805
USF	1.15	0.98	1.03	1	0.98	1	0.96
	-0.67	0.07	0.22	0	0.09	-0.1	0.18
	4422	10567	8416	31534	2494	3769	9573
UNF	1.1	0.93	1	0.92	1	1	0.92
	-0.53	0.2	0.25	0.22	0	-0.08	0.28
	941	2660	1537	2494	5816	1222	2258
FAU	1.07	0.89	0.93	0.91	0.89	1	0.89
	-0.34	0.4	0.54	0.36	0.41	0	0.46
	4705	4546	2655	3769	1222	14119	3619
FSU	1.19	0.98	1.04	1	0.98	1.03	1
	-0.84	0.01	0.13	-0.03	0.05	-0.19	0
	6627	13650	15805	9573	2258	3619	32926

Note: Table displays results from regressions of applicant GPAs for the SUS campuses listed in the rows on the applicant GPAs for the campuses listed in the columns for the population of same-year cross-applicants. Within each cell, the first row is the slope, the second the intercept, the the third the sample size. The sample consists of all same-year cross applicants in the full sample of Florida students. College names are as follows. FIU: Florida International University. UCF: University of Central Florida. UF: University of Florida. USF: University of Southern Florida. UNF: University of Northern Florida. FAU: Florida Atlantic University. FSU: Florida State University.

Table B-2: Effect of threshold-crossing on SUS attendance and progress

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Admitted	Attend FIU	Attend SUS	YRS23	YRS25	CRED23	CRED25
g_i	0.613*** (0.102)	0.0395 (0.0510)	0.441*** (0.110)	1.178** (0.450)	1.327* (0.613)	2.375* (1.028)	2.995* (1.298)
$g_i \times Z_i$	-0.583*** (0.112)	-0.118 (0.0615)	-0.364** (0.126)	-0.451 (0.518)	-0.650 (0.703)	0.210 (1.217)	-0.207 (1.521)
Z_i	0.172*** (0.0224)	0.0239* (0.0122)	0.0906*** (0.0248)	0.410*** (0.101)	0.499*** (0.137)	0.899*** (0.238)	0.944** (0.298)
Observations	7649	7649	7649	7649	6484	7649	6484

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of SUS admissions, attendance, and progress variables on grades relative to the cutoff value for marginal applicants. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. Additional controls for race dummies, lunch status, SAT score, and cohort. Standard errors are heteroskedasticity-robust.

Table B-3: Effect of threshold-crossing on SUS graduation

	(1)	(2)	(3)
	BA in 5 yrs.	BA in 6 yrs.	BA in 7 yrs.
g_i	0.141 (0.0723)	-0.0249 (0.104)	-0.0130 (0.113)
$g_i \times Z_i$	0.0350 (0.0875)	0.255* (0.123)	0.226 (0.132)
Z_i	0.0112 (0.0173)	0.0585* (0.0236)	0.0690** (0.0255)
Observations	7649	6484	6484

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of graduation dummies on grades relative to the cutoff value for marginal applicants. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. Additional controls for race dummies, lunch status, SAT score, and cohort. Standard errors are heteroskedasticity-robust.

Table B-4: Effect of threshold-crossing on other educational outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	CC within 3	CCYRS21	CCYRS23	AA in 5	VC in 5	Non-FL	FL priv.
g_i	-0.0879 (0.112)	-0.456 (0.282)	-0.490 (0.372)	-0.0202 (0.0972)	-0.00647 (0.0250)	-0.00813 (0.0465)	-0.0631 (0.0680)
$g_i \times Z_i$	-0.0456 (0.129)	0.365 (0.313)	0.265 (0.413)	0.0927 (0.109)	0.0120 (0.0276)	-0.0307 (0.0530)	0.112 (0.0760)
Z_i	-0.0578* (0.0253)	-0.167** (0.0610)	-0.215** (0.0805)	-0.0141 (0.0213)	-0.00263 (0.00536)	0.00511 (0.0105)	-0.00747 (0.0144)
Observations	7649	7649	6484	7649	7649	6484	6484

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Results from linear regressions of alternate choice dummies on grades relative to the cutoff value for marginal applicants. g_i is the difference between an applicant's GPA and the cutoff value. Z_i is a dummy variable equal to one for GPAs above the cutoff. Additional controls for race dummies, lunch status, SAT score, and cohort. Standard errors are heteroskedasticity-robust.

Table B-5: Mincer earnings regressions

	(1)	(2)	(3)	(4)
	OLS: FS	OLS: Marg.	IV: Marg.	IV: Marg.
YRS25	0.0816*** (0.000642)	0.0700*** (0.00237)	0.0809*** (0.0112)	0.0685*** (0.0159)
YRS25×FR				0.0352 (0.0215)
YRS25×Black				-0.0108 (0.0210)
YRS25×Male				0.00141 (0.0230)
GPA	0.0556** (0.0269)			
GPA ²	0.0136*** (0.00459)			
g_i		-0.0179 (0.0664)		
$g_i \times Z_i$		0.0554 (0.111)		
Pot. Exp.	0.123*** (0.00214)	0.123*** (0.00914)	0.122*** (0.00837)	0.121*** (0.00836)
Pot. Exp. ²	-0.00704*** (0.000401)	-0.00685*** (0.00180)	-0.00653*** (0.00164)	-0.00622*** (0.00163)
Pot. Exp. ³	0.000177*** (0.0000214)	0.000149 (0.000102)	0.000155* (0.0000911)	0.000115 (0.0000917)
Black	-0.0717*** (0.00377)	-0.0852*** (0.0170)	-0.103*** (0.0151)	-0.0670 (0.0486)
Hisp.	0.0521*** (0.00339)	0.0386** (0.0153)	0.0276* (0.0141)	0.0366*** (0.0136)
Male	0.0858*** (0.00281)	0.0624*** (0.0123)	0.0594*** (0.0108)	0.0470 (0.0530)
F/R lunch	-0.0236*** (0.00321)	-0.00657 (0.0117)	-0.00956 (0.0112)	-0.0867* (0.0483)
SAT	-0.0000148 (0.00000918)	-0.0000298 (0.0000543)		
Constant	9.237*** (0.0403)	9.573*** (0.0512)	9.546*** (0.0347)	9.575*** (0.0437)
Observations	501288	26788	33694	33547

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear regressions of log earnings on years of SUS attendance, potential experience, and student covariates. Sample includes only workers defined as full-time (those who make >\$8,000). For full sample and applicant sample, GPA and GPA_2 are high-school computed un-weighted GPAs. g_i is an admissions-computed GPA less the SAT-score specific cutoff value. Z_i is a dummy equal to one if $g_i > 0$. Potential experience is equal to $age - schooling - 5$, where schooling includes years of SUS attendance as well as years of full-time CC attendance. See appendix for more detail. Sample does not include the 2003 cohort of high school graduates. Standard errors are heteroskedasticity robust and clustered at the individual level.

Table B-6: Mincer earnings regressions

	(1)	(2)	(3)	(4)	(5)
	BW:.25	BW:.6	SAT/DIST dummies	Low cut	High cut
YRS23	0.0832*** (0.0154)	0.105*** (0.00695)	0.0938*** (0.0119)	0.124*** (0.0175)	0.0728*** (0.0107)
Pot. Exp.	0.119*** (0.00941)	0.125*** (0.00646)	0.122*** (0.00822)	0.181*** (0.0120)	0.0940*** (0.00771)
Pot. Exp. ²	-0.00713*** (0.00179)	-0.00761*** (0.00127)	-0.00720*** (0.00158)	-0.0116*** (0.00235)	-0.00503*** (0.00145)
Pot. Exp. ³	0.000190* (0.000101)	0.000221*** (0.0000721)	0.000193** (0.0000888)	0.000325** (0.000131)	0.000116 (0.0000818)
Black	-0.108*** (0.0164)	-0.0737*** (0.0114)	-0.114*** (0.0146)	-0.102*** (0.0213)	-0.102*** (0.0132)
Hisp.	0.0248 (0.0151)	0.0503*** (0.0107)	0.0194 (0.0138)	0.112*** (0.0198)	0.0122 (0.0124)
Male	0.0689*** (0.0115)	0.0546*** (0.00802)	0.0702*** (0.0104)	0.0300** (0.0146)	0.0726*** (0.00947)
F/R lunch	0.000765 (0.0126)	-0.00504 (0.00825)	-0.0112 (0.0109)	0.0218 (0.0159)	-0.0161 (0.00986)
Constant	9.446*** (0.0354)	9.549*** (0.0303)	8.693*** (0.0988)	9.097*** (0.0653)	9.686*** (0.0285)
Observations	28461	59683	36029	42688	32084

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear regressions of log earnings on years of SUS attendance, potential experience, and student covariates. Sample includes only workers defined as full-time (those who make >\$8,000). For full sample and applicant sample, GPA and GPA_2 are high-school computed un-weighted GPAs. g_i is an admissions-computed GPA less the SAT-score specific cutoff value. Z_i is a dummy equal to one if $g_i > 0$. Potential experience is equal to $age - schooling - 5$, where schooling includes years of SUS attendance as well as years of full-time CC attendance. Standard errors are heteroskedasticity robust and clustered at the individual level.

Table B-7: Mincer earnings regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Reg. Pred.	Doubling	Half-year	Miami-Dade only	SX1	SX2
YRS23	0.0879*** (0.0115)	0.0903*** (0.0118)	0.0927*** (0.0126)	0.0841*** (0.0123)	0.0686*** (0.0171)	0.0563*** (0.0180)
Pot. Exp.	0.122*** (0.00796)	0.124*** (0.00820)	0.151*** (0.00995)	0.122*** (0.00915)	0.127*** (0.00835)	0.135*** (0.00806)
Pot. Exp. ²	-0.00697*** (0.00151)	-0.00708*** (0.00158)	-0.00963*** (0.00185)	-0.00777*** (0.00179)	-0.00738*** (0.00158)	-0.00768*** (0.00158)
Pot. Exp. ³	0.000182** (0.0000845)	0.000202** (0.0000889)	0.000285*** (0.000101)	0.000228** (0.000101)	0.000187** (0.0000905)	0.000212** (0.0000903)
Black	-0.104*** (0.0143)	-0.105*** (0.0146)	-0.122*** (0.0155)	-0.0890*** (0.0183)	-0.0655 (0.0425)	-0.0622 (0.0421)
Hisp.	0.0273** (0.0134)	0.0286** (0.0137)	0.0280* (0.0144)	0.0457*** (0.0169)	0.0260* (0.0136)	0.0163 (0.0134)
Male	0.0580*** (0.0101)	0.0586*** (0.0103)	0.0610*** (0.0111)	0.0559*** (0.0118)	0.0394 (0.0472)	0.0480 (0.0465)
F/R lunch	-0.00803 (0.0107)	-0.00844 (0.0109)	-0.00553 (0.0118)	-0.0179 (0.0123)	-0.0796* (0.0422)	-0.0709* (0.0415)
YRS23*FR					0.0405* (0.0231)	0.0358 (0.0228)
YRS23*Black					-0.0200 (0.0223)	-0.0175 (0.0224)
YRS23*Male					0.00989 (0.0254)	0.00727 (0.0250)
SX					0.03	0.106*** (0.0113)
SX ²					-0.17	-0.722*** (0.135)
Constant	9.481*** (0.0292)	9.466*** (0.0300)	8.695*** (0.0329)	9.622*** (0.0505)	9.439*** (0.0395)	9.337*** (0.0365)
Observations	36029	36029	35536	28075	36029	36029

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear regressions of log earnings on years of SUS attendance, potential experience, and student covariates. Sample includes only workers defined as full-time (those who make $\geq \$8,000$). g_i is an admissions-computed GPA less the SAT-score specific cutoff value. Z_i is a dummy equal to one if $g_i > 0$. Potential experience is equal to $age - schooling - 5$, where schooling includes years of SUS attendance as well as years of full-time CC attendance. SX is years of labor market experience while in school. Coefficients on the SX terms in column five are restricted to the stated values. Standard errors are heteroskedasticity robust and clustered at the individual level.

C Proofs from section 3

C.1 Proof of proposition 1

Proof. By the envelope theorem, we have

$$\frac{dV_C(\theta, R_C)}{d\theta} = \beta U'(c_2^C) Y_C'(\theta) > 0. \quad (\text{C.1})$$

Y_{HS} does not vary with θ , so $V_{HS}(\theta, R_{HS})$ is constant in θ . Since at every value of R_C some agents choose to attend college and some do not, there exist values of θ, θ' and θ'' , such that $V_C(\theta', R_C) > V_{HS}(\theta', R_{HS})$ and $V_C(\theta'', R_C) < V_{HS}(\theta'', R_{HS})$. Since $V_C(\theta, R_C)$ slopes strictly upward in θ while $V_{HS}(\theta, R_{HS})$ is flat, it is clear that a) $\theta' > \theta''$, and b) there exists precisely one $\bar{\theta} \in (\theta'', \theta')$ such that $V_C(\bar{\theta}, R_C) = V_{HS}(\bar{\theta}, R_{HS})$. Agents with $\theta > \bar{\theta}$ choose to attend college, and agents with $\theta < \bar{\theta}$ do not.

How does $\bar{\theta}$ vary with R_C ? Again via the envelope theorem,

$$\frac{dV_C(\theta, R_C)}{dR_C} = \beta (wl^c - c_1^c - \tau) U'(c_2^C) < 0 \quad (\text{C.2})$$

Since college students are net borrowers, raising the interest rate lowers their lifetime utility. Lifetime utility for students who do not attend college does not vary with R_C . Thus the threshold value of θ at which lifetime college utility surpasses lifetime high school utility, which I will now term $\bar{\theta}(R_C)$, rises with R_C .

Corollary 1 follows directly from this result.

$$E[Y_C(\theta) - Y_{HS} | R_C, s = C] = E[Y_C(\theta) - Y_{HS}(\theta) | \theta > \bar{\theta}(R_C)] \quad (\text{C.3})$$

increases with $\bar{\theta}(R_C)$ and therefore with R_C .

□

C.2 Proof of proposition 2

Proof. Due to the strict concavity of u and U and the strict convexity of v , the necessary conditions for the implicit function theorem hold, and l^C is a continuous function of R^C within a neighborhood of the optimal value. Standard algebra yields the formula

$$\frac{dl^C}{dR_C} = -\frac{w\beta u''(c_1^C) [(1 + R_C)U''(c_2^C)(wl^C - c_1^C - \tau) + U'(C_2^C)]}{w^2\beta(1 + R_C)^2 u''(c_1^C)U''(c_2^C) - u''(c_1^C)v''(l^C) - \beta(1 + R_C)^2 U''(c_2^C)v''(l^C)} \quad (\text{C.4})$$

Under the assumptions on the first and second derivatives of u , U , and v , and the assumption that college attendees are net borrowers, the numerator of the fraction is negative and denominator is positive. The value of the expression is thus positive. □

D Imprecise control and discontinuous grade grouping

This appendix describes the connection between the continuous ratio assumption in section 6.2.1 and the ‘imprecise control’ condition developed in Lee and Lemieux (2009). It also develops a set of conditions under which a discontinuously distributed running variable is consistent with unbiased RD estimation.

Recall that wages are determined by observable covariates X_i and unobservable heterogeneity u_{1i} , distributed continuously according to $h(x, u_1)$. Students may allocate effort in high school based at least in part on these characteristics. Lemieux and Lee’s imprecise control condition requires that students be able to control their grades only up to a continuously distributed error term. This guarantees that the conditional grade distribution

$$f(g|x, u_1) \quad (\text{D.1})$$

is continuous in g . Then, by Bayes’ rule, the conditional distribution of earnings determinants

$$h(x, u_1 | g) = h(x, u_1) \frac{f(g|x, u_1)}{f(g)} \quad (\text{D.2})$$

is also continuous in g , since both the numerator and denominator of the g -dependent fraction are continuous. RD estimation is unbiased.

The imprecise control condition is sufficient to guarantee continuity of wage determinants in g . But because it is possible for the fraction to be continuous even if neither the numerator or denominator is continuous, the condition is not necessary. The test posed in section 6.2.1 relaxes the imprecise control condition by looking at the continuity of the fraction, not its component parts.

It is important to ask whether there exist realistic conditions under which the a running variable would satisfy the ratio test but not the imprecise control condition. The answer to this question is yes. To show this I consider an example in which students can perfectly choose their effort level, but effort level is mapped noisily to observed grades.

Let g_i^* measure academic performance for student i . $g_i^* \in [g_l, g_h]$ is (potentially) controlled by students, and is observed by academic administrators but not by the researcher. g^* is distributed according the probability density function $\gamma(g^*)$. Observed grades g are determined stochastically conditional on g^* with the distribution $f(g|g^*)$, which may be discontinuous in g . The idea that there is a probabilistic mapping between latent performance and grades is analogous to the imprecise control condition of Lemieux and Lee, and consistent with the idea that from the perspective of the student there is a stochastic component to academic evaluations and to GPA computation. I make two assumptions about the mapping from academic performance to observed grades:

1. Conditional on academic performance, the distribution of grades is independent of wage determinants, so that $f(g|g^*) = f(g|g^*, u_{i1} = u, X_i = x)$.
2. $\gamma(g^*|g)$ is bounded and uniformly continuous in g for all g^* .

The first assumption guarantees that mapping from latent to observed performance does not depend on wage determinants. Latent performance may still depend on individual demographics, effort, and the interaction between effort and the environment. So long as one believes that there is some noisy component of grade determination, this

can be thought of as a normalization. The second assumption states that students with similar grades are drawn from similar parts of the latent performance distribution. That is, a discontinuity in the grade distribution does not reflect a discontinuity in the distribution of performance of students achieving that grade. It will be violated if, given two students with nearly identical performance levels, one is much more likely receive a certain observed grade than the other.

Theorem 1. *Under assumptions (1) and (2), the density $h(X_i = x, u_{i1} = u | g_i = g)$ is continuous in g .*

Proof. Applying, in turn, the Law of Iterated Expectations, Bayes' rule, and (1) yields

$$\begin{aligned}
 h(x, u | g) &= \int h(x, u | g, g^*) \gamma(g^* | g) dg^* \\
 &= \int f(g | x, u, g^*) \frac{h(x, u | g^*)}{f(g | g^*)} \gamma(g^* | g) dg^* \\
 &= \int h(x, u | g^*) \gamma(g^* | g) dg^*
 \end{aligned} \tag{D.3}$$

The density of wage determinants conditional on grade g is a weighted average of wage determinants conditional on g^* , with weights determined by the density of g^* given g . Then, since $\gamma(g^* | g)$ is uniformly continuous in g , limits can be passed through the integral, so that

$$\begin{aligned}
 \lim_{h \downarrow 0} h(x, u | g + h) &= \lim_{h \downarrow 0} \int h(x, u | g^*) \gamma(g^* | g + h) dg^* \\
 &= \int \lim_{h \downarrow 0} (h(x, u | g^*) \gamma(g^* | g + h)) dg^* \\
 &= \int (h(x, u | g^*) \gamma(g^* | g)) dg^* \\
 &= \lim_{h \downarrow 0} h(x, u | g - h)
 \end{aligned} \tag{D.4}$$

Therefore $h(x, u | g)$ is continuous. □

It follows from Bayes' Rule and the continuity of $h(x, u)$ that

$$\frac{f(g|x, u)}{f(g)}$$

is also continuous in g , even though the numerator and denominator are not.