I Z A

IZA DP No. 9799

The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies

Ben Ost
Weixiang Pan
Doug Webber

March 2016

The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies

Ben Ost
University of Illinois at Chicago

Weixiang Pan
University of Illinois at Chicago

Doug Webber
Temple University and IZA

ABSTRACT

The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies[*]

We estimate the returns to college using administrative data on college enrollment matched to administrative data on weekly earnings. Utilizing the fact that colleges dismiss low-performing students based on exact GPA cutoffs, we use a regression discontinuity design to estimate the earnings impacts of college. Dismissed students are permitted to apply for readmission, but since relatively few do so, these students end up completing fewer years of school and are approximately 10 percentage points less likely to graduate college. Our estimates suggest that low-performing students (on the margin of college dismissal) derive substantial earnings benefits from college.

Corresponding author:

Doug Webber
Department of Economics
Temple University
1301 Cecil B. Moore Ave
Ritter Annex 883
Philadelphia, PA 19122
USA
E-mail: douglas.webber@temple.edu

Introduction

Increasing the rate of college attendance and graduation has become a major

policy goal at both the local and national level. President Obama has called this an

"economic imperative" and has launched a variety of initiatives aimed at restoring the

United States' lead in the rate of college attendance. [1] While there is little doubt that on

average the economic returns to college are high, recent increases in the amount of

student debt as well as relatively high student loan default rates have prompted many to

question whether college pays off for all students.   Critics of Obama's call for expanded

higher education attendance argue that certain students are simply not "college material"

and low-achieving students may be better off dropping out of school and obtaining more

work experience.[2]

In this study, we provide direct evidence on the returns to college for low-

performing students. A major obstacle to identifying the returns to college is that

students who drop out of school are likely different than students who persist, thus

comparisons of dropouts and completers are likely contaminated by omitted variable bias.

In an ideal experiment, researchers would randomly force certain students to drop out of

school while allowing other, similar, students to persist. Our empirical design mimics

this ideal experiment by exploiting institutional policies that dismiss students based on

their exact grade point average (GPA). We use administrative data from 13 public

universities in the state of Ohio to identify students that are at risk of dismissal based on

---

[1] See President Obama's "American College Promise" or the higher education proposals
from Secretary Clinton/Senator Sanders.
[2] See
http://www.slate.com/articles/life/education/2014/03/college_isn_t_for_everyone
_let_s_stop_pretending_it_is.html

their past academic performance. We match these students to administrative data on Unemployment Insurance (UI) earnings records in order to study the impact of university dismissal on earnings 7-12 years after initial college enrollment. Using a regression discontinuity design, we effectively compare mid-run earnings of students with GPA's which fall just below the dismissal cutoff to those just above the cutoff.

The dismissal policies we leverage generate large discontinuities in academic outcomes, in line with the previous literature, which has examined similar institutional discontinuities (Lindo et al, 2010; Vidal-Fernandez, 2011; Barua and Vidal-Fernandez, 2014). Students just below the cutoff are approximately 16 percentage points less likely to enroll in the following term, complete 0.2 fewer years of school and are 10 percentage points less likely to obtain a BA relative to those above the threshold. These reductions in educational investment correlate with a decrease in weekly earnings such that those just below the GPA cutoff earn approximately 4.5 percent less than those just above the cutoff.

While enrollment in college has increased substantially over the past several decades, persistence within college has declined markedly over the same period (College Board Education Pays 2013). As such, many policy makers view improving persistence among already enrolled students as a key step towards increasing college completion rates. Our study estimates the returns to college among the set of students already enrolled and is thus directly relevant for understanding the value of policies that improve college persistence. Because of the nature of our identification strategy, the estimates we present are directly relevant for low-performing students. While these students are not representative of the typical enrolled student, they are in some respects one of the most

policy-relevant subgroups since they are likely to be impacted by policy initiatives aimed at improving persistence.

The primary threat to identification in our context is the possibility that students finely manipulate their GPA in order to fall just above dismissal thresholds. As we discuss later in the paper, this is plausible because dismissal threshold are publicly known and students could potentially alter their effort in order to manipulate their GPA. That said, empirically, we fine little evidence of sorting around the critical threshold. While we cannot definitively rule out manipulation, we view it as a very unlikely explanation for our results. Students just below and just above the threshold are nearly identical in terms of their observable characteristics and yet they have dramatically different college persistence rates, college completion rates, and future earnings. For sorting on unobservables to explain our results, this sorting would need to be simultaneously large in magnitude and completely uncorrelated with observable characteristics.

Though our primary focus is estimating the reduced form benefits to persisting in college, we also provide evidence on the magnitude of so called "sheepskin" effects of earning a college degree. To study this question, we first note that the impact of dismissal on academic outcomes differs depending on the student's standing at the time of the dismissal. For example, a student that is dismissed in her 4th year of college has already completed most of her coursework, whereas a student dismissed in her 2nd year of college still has many years of coursework remaining. Both students may be prevented from obtaining the BA, but the impact on total credits earned is larger for the student dismissed in her 2nd year. Using admittedly strong assumptions regarding the homogeneity in the returns to schooling for different complying populations, we are able

to use this variation to separately identify the BA effect from the years of schooling

effect. While the assumptions necessary to identify this model are fairly strong, they are

considerably weaker than those made by past work (e.g. education is treated as

exogenous conditional on observable characteristics) studying sheepskin effects

(Hungerford and Solon, 1987; Kane and Rouse, 1995; Jaeger and Page, 1996; Liu et al,

2015). We find suggestive evidence that college provides a large sheepskin effect,

though our estimates are imprecisely estimated which prevents us from drawing strong

conclusions.

Literature Review

There is a large literature seeking to understand the economic benefits to college.

Since this literature is far too large to adequately summarize here, see Card (1999) for an

excellent review.

Although researchers have used many different identification strategies to

estimate the returns to schooling in general (e.g. twins, changing compulsory schooling

laws, quarter of birth) studies that specifically consider the returns to college utilize a

more limited set of strategies. With a few exceptions (discussed below) most estimates

of the return to college rely on relatively strong assumptions. One strand of this literature

relies on the assumption that, conditional on observables, college enrollment and

completion are independent of potential outcomes (see Walker and Zhu, 2008 or Webber,

2014 for recent examples). Even when the researcher has access to highly detailed

information about individuals' aptitude and preferences, such as in studies which utilize

the National Longitudinal Survey of Youth database, the argument for a causal

interpretation of the estimated coefficient associated with education still rests on strong

and inherently untestable assumptions. A second strand of the literature uses distance to the nearest college as an instrument for college attendance (see Card, 1995 or Kane and Rouse, 1995). Though arguably less restrictive than assuming that education is exogenous conditional on observables, the distance based IV similarly requires fairly strong assumptions since it rules out the possibility that colleges directly alter the surrounding area.

Our study is most closely related to a recent set of studies which use college admission thresholds and a regression discontinuity design[3] to estimate the benefit of college admission. Several recent studies use data from various international contexts and with the exception of Ockert (2010) these studies find a substantial earnings benefit to gaining admission to a more selective university or course of study (Anelli 2015; Canaan and Mouganie 2015; Hastings, Neilson and Zimmerman 2013; Kirkebøen, Leuven, and Mogstad 2014). In the US context, Hoekstra (2009) studies the earnings effect of admission to a selective flagship public university (as opposed to a less selective public university). Zimmerman (2014) studies the impact of gaining admission to a low-ranked public university on future earnings. Both of these studies find substantial benefit to gaining admission to the university they study.

We view our study as complementary to this set of studies in that they estimate the returns to college admission whereas we estimate the returns to college persistence among the set of admitted students. Our study is most closely related to Zimmerman (2014) because like his study, the students we study are relatively low performing. That said, the students he studies are on the margin of college admission as opposed to being

---

[3] A similar strategy has been used to study academic outcomes , see Van der Klaauw (2002) or Kane (2003) for excellent examples.

on the margin of being dismissed from college. In his setting, college admission requires a 3.0 high school GPA while in our setting college dismissal occurs when a student repeatedly performs below the dismissal threshold, set at 2.0 or lower. While it is difficult to directly compare high school and college GPAs, we suspect that the set of marginal students we study are less qualified for college compared to the set of students studied in past work. This is based on the fact that marginal students in our study *actually* perform poorly during college whereas past work focuses on students that are *predicted* to perform poorly based on their high school record.

Academic Dismissal Policies

Nearly all US universities use some form of a probation—dismissal policy. Though exact policies vary across universities, in most cases, students on probation receive some academic counseling and are informed that if they fail to raise their GPA in subsequent terms, they will be dismissed from school. Typically, students on probation are allowed between 1 and 3 semesters to rehabilitate their cumulative GPA. In our sample, most schools allow dismissed students to apply for re-entry to the university but require that students be away from the university for 1 calendar year before being considered for re-admittance.

Though all public colleges in Ohio utilize the general probation – dismissal policy described above, the specific parameters of these policies vary across schools. Appendix A provides details of the policies used at each public 4-year institution in Ohio. The modal policy places students on probation if their cumulative GPA falls below some critical threshold and dismisses students if they do not raise their GPA above the threshold by the subsequent term. In most cases, students can appeal the dismissal

decision and based on conversations with administrators, these appeals are often times successful. For students who are actually dismissed, the modal waiting period before reapplying is 1 year and readmission is typically not guaranteed.

We use data on the probation—dismissal policy at each school to identify the group of students at risk of dismissal. For example, at Akron University since students get 1 term to rehabilitate their GPA, a student with a cumulative GPA below 2.0 in term t-1 is at risk of dismissal following term t. If his cumulative GPA does not rise above 2.0 in term t, he faces the possibility of dismissal. For students that we designate as at-risk of dismissal, their future enrollment is a potentially discontinuous function of their GPA whereas enrollment should be a smooth function of GPA for students not currently at risk of dismissal. As such, we focus our regression discontinuity analysis on students who are currently at risk of dismissal. For each student that is at risk of dismissal, we define the running variable based on the policies in place at his/her school. In some cases, the cutoffs differ according to the cumulative number of credits a student has previously earned and our running variable, GPA relative to the institution's dismissal threshold, is normalized to reflect these different cutoffs for different students.

There are several reasons why students who fail to meet the GPA requirement while at risk of dismissal will not actually be dismissed. First, we code probation and dismissal policies at the university level, but in practice, these policies can be more or less lenient by major or college. Since our data do not include any information on actual probation or dismissal status, this means that in some cases we identify students as at risk of dismissal when they were never actually on probation to begin with. Second, all schools in Ohio include provisions through which administrators can grant exceptions

due to personal problems such as illness or other adverse shocks. Based on conversations with administrators, this discretion is used frequently and is particularly likely in cases where students are close to the required threshold and for students in their first year of college. Importantly, when administrators grant exceptions to the dismissal policy, they do this by allowing students to enroll despite their low GPAs – not by artificially adding GPA points to help the student rise above the threshold. This distinction is important because it suggests that administrator discretion will contribute to a weaker first-stage discontinuity, but it will not bias our estimates. Finally, as discussed in appendix A, our threshold data may be inaccurate for certain schools because we may not have uncovered all policy changes over time. This will further weaken our first-stage, but should not bias our second stage results. We provide additional discussion of this issue in appendix A.

Data

We use administrative data that links college transcripts for every Ohio public university student to UI earnings records in the state of Ohio. The transcript data span the academic years starting in 2000-2010 while the UI earnings data includes weekly earnings from 2003-2012. These data are made available to researchers by the Ohio Educational Research Data Center (OERDC) and include data from the Ohio Workforce Data Quality Initiative (OWDQI).

The wage and employer data come from the Ohio Department of Job and Family Services and include Quarterly Census of Employment and Work (QCEW) enterprise level data as well as worker-level quarterly earnings data. The higher education data includes the universe of 2- and 4-year public college enrollment in Ohio, though in this study, we focus on students whose first enrollment was at a 4-year institution. With the

exception of federal workers and the self-employed, the UI data covers the universe of workers in Ohio. For some specifications, it is necessary to match quarterly earnings from the UI data to the academic calendar. We match the fall term to 4th quarter (October-December) and the spring term to the 2nd quarter (April-June). For schools on the quarter system, we match the winter term to the 1st quarter. We exclude summer terms from the enrollment analysis but include summers when calculating annual earnings and when calculating total credits earned.

While these administrative data are in many ways uniquely well suited to answer the research questions of this study, they have several key limitations. First, we cannot observe enrollment at private institutions or at public institutions outside of the state of Ohio. This will contribute to our finding relatively larger educational impacts since students that transfer out of state will be miscategorized as dropouts. That said, we are able to track students if they transfer to any of the 38 2- or 4-year colleges in the state of Ohio and these institutions represent more than 75% of students in the state. Second, we cannot distinguish between unemployment, lack of labor force participation, federal employment, self-employment and leaving the state of Ohio.[4] Since we cannot distinguish between these outcomes we focus our analysis on earnings conditional on having a UI earnings record. We later test for whether differences in the probability of missing earnings data can explain our results and find no evidence that it can.

---

[4] Based on our calculations using the 2012 American Community Survey (2012 was chosen to match our study period), for individuals without earnings in our sample's age range with at least some college experience there is a 56% probability that they left the state over the past year, an 8% chance that they are self-employed, 4% chance they are employed by the federal government, and a 32% chance that they truly have no income.

In addition to restricting the sample to students who are at risk of dismissal, we impose several other sample restrictions. First, we restrict the sample to students who begin at one of the 13 4-year public universities that we study. Second, we focus our attention on students who attempt at least 10 credits in the critical term. Third, we restrict the UI earnings data to payments of at least $500 per quarter and we focus on each worker's primary employer for each quarter as measured by total quarterly pay from each employer.[5] Since we cannot observe long-run outcomes for the recent cohorts of students, we focus all our analyses on students that initially enrolled in college at least 6 years prior to the end of our data. This allows us track all students in our analysis sample for at least 6 years in order to measure outcomes such as total credits completed and 6-year graduation rates. For the mid-run earnings analyses, we can observe earnings as far as 12 years after the term of first enrollment, but this is only the case for a single cohort in our data. Since the UI earnings data go further than the higher education data, we are able to observe earnings 7 years after initial enrollment for all students for whom we observe 6-year academic outcomes.

In order to measure earnings at the latest possible data for as many students as possible, our preferred earnings measure is average weekly earnings in the latest year in which an individual is observed in the data. For the majority of individuals, this is their earnings during 2012, but some individuals are missing earnings for this year since it is

---

[5] We drop small quarterly payments because the UI data include any payment from a firm to an individual, even in cases where that payment would not constitute what we normally think of as a job (e.g. legal payment, consulting service, etc.). For related reasons, we also only use an individual's "dominant job" (highest earning) in earnings calculations in the event they work for multiple employers in the same quarter. These restrictions are standard when using UI data, see Webber (2015) as an example and for further citations.

increasingly likely to have missing earnings records in later years of the data.[6] For

individuals missing earnings in 2012, we use their average weekly earnings from earlier

periods if they exist. Since we expect the short-run and mid-run impact of enrollment to

go in opposite directions, we focus only on earnings measured at least 7 years after initial

college enrollment. For example, if a student has earnings in their 3rd year after first

enrolling but never appears in the UI data again, we do not use their 3rd year earnings. For

the 2004 cohort, our earnings measure uses earnings mostly from 2012, but fills in

missing observations using 2011 if necessary. For the 2000 cohort, it measures earnings

mostly in 2012, but fills in missing observations using 2007-2011 if necessary.

Relative to simply using 2012 earnings as the dependent variable, our preferred

earnings measure has the advantage of increasing power and also minimizing potential

biases caused by sample attrition. That said, the time at which earnings are measured

depends on choices made by the individual, so the measurement timing could be

endogenous. In practice, however, there is no discontinuity in the time at which earnings

is measured, and all results are unchanged whether or not we control for the age at which

earnings are measured.[7]

Table 1 shows descriptive statistics for the overall sample and for the restricted

sample of students that are at risk of being dismissed based on their historical GPAs. In

that table, probation is an indicator for whether a student's GPA falls below the probation

threshold, as opposed to being a measure that comes directly from the data. Students at

risk of probation are quite different than the general population as they are much more

---

[6] This pattern of increasing missing with age is consistent with that found in other studies
using UI earnings records (e.g. Zimmerman (2014) finds similar rates of attrition).
[7] The results are also similar if we simply drop those missing earnings in 2012.

likely to be male and more likely to be black. Not surprisingly, these students have much

lower graduation rates and complete far fewer credits on average. The mean age at

earnings measured is 28, which is similar to other recent studies using administrative

earnings and education data (e.g. Andrews et al. 2014, Zimmerman 2014).

Descriptive Analysis

Given that relatively few studies have utilized data that include both

administrative quarterly earnings and administrative enrollment information, we begin by

describing the earnings trajectory of college graduates and dropouts in our data. This

analysis is in the spirit of Mincer's (1974) seminal analysis, however, rather than relying

on repeated cross sections measured after schooling is completed, we use panel data

including earnings while in school.

Figure 1 shows average log weekly earnings vs quarters since first college

enrollment, split according to whether the individual eventually obtains a BA. As

discussed in the data section, we show earnings conditional on employment since we

cannot identify true non-employment. The difference in the earnings of those that earn

and do not earn a BA likely reflects a combination of the causal impact of a BA, selection

patterns with respect to who decides to complete a BA and sample selection into

employment. Nevertheless, Figure 1 illustrates several interesting patterns.

First, among individuals who eventually obtain a BA, we see a very clear seasonal

pattern in the first 4 years in which earnings are highest during the 3rd quarter and lowest

during the 1st quarter. Though quarters do not perfectly line up with academic terms, the

3rd quarter covers summer vacation for most students. A similar (though weaker) pattern

of seasonality is observed for individuals that do not obtain a BA. This pattern lessens

after the first year as these students begin to drop out of college. Second, Figure 1 shows that while dropouts out-earn those with a BA in the 2nd, 3rd and 4th years, the earnings of dropouts and completers rise at a similar rate over the first 16 quarters, so this gap does not increase over time. Finally, Figure 1 shows that BA completers out earn dropouts almost immediately after the 16th quarter. Since the vast majority of these students complete the BA in the 16th quarter or later, this suggests that college graduates are able to overtake non-graduates shortly after graduating. Importantly, by the 28th quarter following initial enrollment, there already exists a substantial earnings differential between graduates and non-graduates suggesting that year 7 earnings is likely not too early in the life course to be looking for possible earnings effects of a college education. Though we can observe earnings as late as 12 years after college for the 2000 cohort, it is useful to know that the earnings gap appears as early as 7 years after initial enrollment since we can observe earnings 7 years after enrollment for all of the cohorts for which we have long-run academic outcomes.

Empirical model

Our analysis uses a regression discontinuity design based on enrollment discontinuities caused by academic dismissal policies. We limit our attention to students who are at risk of dismissal based on their past GPAs. For this group, we construct our running variable, C_{it}, as

$$C_{it} = GPA_{it} -\ cutoff_{it} \qquad (1)$$

If we perfectly identify at risk students and universities strictly enforce dismissal policies then all students with $C_{it} < 0$ would be dismissed following term t and none of these students would enroll in term t+1. In practice, not all students with GPA's below

the threshold will be dismissed because both the probation and dismissal policies are not rigidly enforced. Hence, our identification strategy can be categorized as a fuzzy (rather than a sharp) regression discontinuity design. Since our data do not include information on whether each student is dismissed, it is not possible to estimate the relationship between our running variable and dismissal. That said, we do observe whether a student re-enrolls in the following term and ultimately, we are more interested in estimating the impact of college enrollment as opposed to college dismissal.

While it is possible to estimate the enrollment effect separately by every term, in order to increase power, we stack the data by term and thereby estimate the average effect across all terms. In the heterogeneity section, we show analyses split according to the timing of the dismissal, but this analysis generally lacks power, so we focus primarily on the average effect. We estimate a version of Equation 2 on a series of outcomes.

$$Y_{i,t+1} = \alpha C_{i,t} + \beta D_{i,t} + \gamma C_{i,t}D_{i,t} + \varepsilon_{i,t} \qquad (2)$$

$Y_{i,t+1}$ denotes enrollment in period t+1. $C_{i,t}$ is the running variable defined above and $D_{i,t}$ represents an indicator for whether $C_{i,t}$ is below zero or not. We use a two-way cluster method to produce standard errors, clustering both on the individual and on the running variable. We cluster around the individual because some individuals could appear in the regression in multiple terms. Though in our context the running variable is virtually continuous, we also cluster on the running variable based on the recommendation of Card and Lee (2015).

Initial enrollment effect

Our first step is to explore whether the dismissal policies we study actually impact next term enrollment. Figure 2 shows that there is clear evidence of a discontinuity at the

cutoff suggesting that our basic empirical approach is viable. The fact that students who

fall below the cutoff still appear to enroll at fairly high rates indicates either that the

policies are not rigidly enforced or that our measurement of these policies has some error.

From the standpoint of identification, neither of these issues will bias estimates since they

simply weaken the power of our instrument.

Conversations with administrators suggested that first-year students are often

treated more leniently than other students, so we split our first-stage analysis by whether

the student has first-year standing at the beginning of term t. Figure 3a shows that for

students with first-year standing, there does not appear to be any discontinuity in terms of

future enrollment. Figure 3b shows the same specification for students with second-year

or higher standing and for these students, there is a clear discontinuity in enrollment at

the cutoff.[8] Being just below the performance cutoff reduces the probability of

enrollment in t+1 by approximately 15 percentage points.

Given that there is no discontinuity in enrollment for first-year students, we

cannot use our empirical approach to study the impact of college persistence for these

students. As such, for all analyses moving forward, we drop the students with first-year

standing from our sample.   While this sample restriction does not impact internal

validity, it does further limit the external validity since we focus on students who make it

to at least their second year of college.

---

[8] We have also explored whether the discontinuity generally increases with standing and
we did not find this to be the case. The estimated discontinuity is approximately 15
percentage points for all students outside of those with first-year standing.

In addition to showing a clear discontinuity, Figure 3b shows that there appears to be different slopes on either side of the discontinuity. Appendix B provides a discussion of various reasons why the slope may change around the discontinuity.

Specification checks

The primary threat to identification is the possibility that students finely manipulate their GPA in order to land just above the threshold. In our context, thresholds are publicly known and students could feasibly manipulate their GPA by exerting extra effort (or pleading with professors) in order to obtain higher grades. Theoretically, the type of students who may engage in this sort of manipulation may be particularly motivated and as such, they might have better potential outcomes than their less motivated counterparts. While the institutional incentives to manipulate one's GPA exist, there are several reasons why this form of manipulation might not occur in practice. First, we restrict the sample to those already on probation so these students have already failed to meet performance standards set forth by the university. This suggests that the type of students we study are perhaps less likely to respond to performance standards than the typical college student. Second, since appeals of the dismissal decision are frequently granted, students may not perceive the threat of dismissal as something really worth responding to.

While it is useful to consider the likelihood of manipulation from an ex-ante perspective, ultimately, whether manipulation is likely to substantially bias our estimates is an empirical question. We explore the magnitude of potential manipulation both by examining the density of observations surrounding the performance standard and also by testing for whether covariates move smoothly through the threshold. If manipulation

occurs, we expect to see too few observations just to the left of the threshold and too many observations just to the right of the threshold. To the extent that this manipulation is correlated with students' characteristics, the manipulation will also generate discontinuities in predetermined characteristics.

Figure 4a shows the density of the running variable around the threshold. Figure 4a shows clear evidence that students are much more likely to obtain a GPA exactly at the cutoff than any other GPA. Though this could be indicative of manipulation, another possibility is that GPAs cluster at certain numbers mechanically. As noted in past work (Zimmerman 2014, Barreca et al 2015) GPAs tend to spike at whole numbers because the number of combinations that result in whole number GPAs are far larger than the number of combinations that result in decimal GPAs. Since the most common GPA cutoff is 2.0, if students are more likely to obtain exactly a 2.0 mechanically, this could feasibly explain the large spike in the running variable at zero. To illustrate this point, in Figure 4b we show the histogram of GPAs around 1 point above the dismissal threshold for students who are not at risk of dismissal. These students have no particular incentive to manipulate their GPA to be above the 1-point mark, but Figure 4b shows a clear spike, just as was observed in Figure 4a. Though we do not view the spike at the discontinuity necessarily as evidence of manipulation, this large spike does make it difficult to assess whether students sort around the threshold based solely on the histogram. Furthermore, even if there is no manipulation, as noted by Barreca et al. (2015) this type of spike has the potential to bias estimates and so we follow their recommendation and estimate both the typical regression discontinuity and the "donut" RD in which those with a running variable exactly at zero are excluded from the estimation.   The intuition in the donut RD

is that even in the presence of a spike at 2.0, it will still be the case that those with a 2.01 are a reasonable counterfactual to those with a 1.99.

In order to further explore the likelihood that students finely manipulate the running variable, we examine whether there are discontinuities in observable predetermined characteristics. In addition to testing for smoothness in each of the individual student characteristics, we also construct 4 composite variables in which we predict our outcomes using all of the covariates and test for smoothness in the predicted outcome. Considering whether there are discontinuities in predicted outcomes is particularly useful because these measures show not only whether there are covariate imbalances, but also whether and in what direction any imbalances may bias estimates.

Figures 5a-5l show our estimates of the discontinuities in each covariate along with the predicted outcomes. Table 2 also displays these same estimates. On the whole, there is no evidence of discontinuities in the covariates. Importantly, all of the predicted outcomes and key covariates (such as log weekly earnings in period t-1) are smooth through the threshold. As discussed in appendix B, there is a clear non-monotonic relationship between certain covariates and the running variable. Most notably, credits-earned at time t-1 shows an inverse V shape with the apex near the cutoff. Since the relationship between credits-earned and the running variable is smooth at the cutoff, this will not create bias, however, it can lead to sharp slope changes right at the cutoff which can contribute to imprecision of the RD estimates. As such, in all of our analysis below, we condition on credits earned at time t-1.

Taken together, these results suggest that while we cannot definitively rule out the possibility of manipulation, it is unlikely that manipulation is sufficiently strong so as to

be able to explain away the results presented below. In the robustness section, we provide further evidence on this point by showing that our estimates are virtually unchanged regardless of whether or not we control for covariates and are also robust to using a "donut" RD by excluding observations exactly at the cutoff (Barreca et al 2015).

Results

       We present all of our core results in figure form using a fixed bandwidth of 0.5 grade points. Tables 3 and 4 show these same estimates in table form and examines the robustness of each result to changes in the bandwidth, adding student observable characteristics, or using a "donut" RD. We begin by showing the reduced form impact of falling below the cutoff on mid-run earnings and then explore several key academic variables that likely contribute to this reduced form relationship.

Impact on Earnings

       Figure 6a shows the reduced form relationship between the running variable and mid-run log weekly earnings as discussed in the data section. We find that falling below the cutoff decreases weekly earnings by approximately 5.6 percent. The slopes on each side of the discontinuity are not statistically distinguishable from zero. While, we cannot differentiate between earnings gains from increased hours and earnings gains from increased wages, it is worth nothing that unlike many studies using administrative UI records, we have weekly rather than quarterly earnings and so these estimates cannot be driven by weeks worked.

       While our data cover earnings for a majority of the sample, earnings are not observed for students who are self-employed or leave the state. As such, it is possible that some of the earnings effect that we estimate is explainable by differential attrition.

Although we find little evidence of differential attrition above and below the critical GPA cutoff (see Figure 6b), it is possible that even with perfectly balanced attrition rates, our estimates could be biased if the form of attrition differs between the treatment and control groups. For example, suppose college simultaneously increases each person's propensity to move out of the state and also increases their employment probability by the same amount. In other words, if treatment influences attrition non-monotonically, then balanced attrition does not eliminate attrition bias.

An alternative test of whether missing earnings data is likely to bias our estimates is to examine whether we continue to have balanced covariates in the sample for which we can observe earnings in the later period. Column 2 of Table 2 shows the estimated discontinuities in covariates around the critical threshold when the sample is restricted to the set of students with non-missing mid-run earnings. We find no evidence of discontinuities in any of the covariates. The smoothness of the covariates in the sample with non-missing mid-run earnings suggests that sample selection does not cause those just to the left of the cutoff to be quite different than those just to the right of the cutoff.[9]

Academic Outcomes

Though we have already established that dismissal reduces enrollment in the short-run, there are several reasons why the impact on an outcome such as BA receipt is not a mechanical function of the disenrollment effect. First, it is possible that the policy only dismisses students that were going to drop out eventually, and thus the policy would have no impact on completion. Second, students who are dismissed are permitted to

_____

[9] While we view the balanced covariates as the strongest evidence regarding sample selection, we have also explored bounding estimates using the set identification strategy presented in Lee (2009). Given that there is little evidence of differential attrition to begin with, it is no surprise that our results are robust to this bounding exercise.

apply for reentry to the university after their dismissal and so being dismissed itself does not directly preclude earning a degree at that institution.

To explore the long-run academic impacts empirically, we examine how falling below the threshold impacts total credits earned and the probability of BA receipt. In addition to studying these outcomes at the initial school, we measure these outcomes at other public 2 and 4- year colleges in Ohio. Theoretically, we expect that a student's educational investment likely falls at the institution from which the student was dismissed, but his educational investments might rise at other schools.

Figure 7 shows the impact of falling below the cutoff on the total credits earned at one's starting institution. We find that students who fall below the cutoff earn approximately 5.1 fewer credits at their starting institution, which corresponds to slightly over 1/6th of a full year's load.[10] While credits earned at the starting institution falls, it is possible that dismissal increases credits earned at other schools. To investigate this possibility, we change the outcome to be all the credits earned by a student at any public Ohio college outside of their starting institution. We find a relatively small impact (less than one credit at a different institution) in this model (see Figure 8) and in our robustness analysis we find that the estimate is sensitive to bandwidth choice.

Figure 9 shows the impact of falling below the cutoff on the probability of earning a BA from the initial college of enrollment. Falling below the cutoff reduces the probability of graduating by approximately 11 percentage points. Rerunning this analysis

---

[10] Credits are measured based on semester-credit equivalents so the typical school requires 120 credits to graduate. We follow Kane and Rouse (1995) and define a year of education as 30 earned credits so that the typical college degree is 4 years of credits. This definition is consistent with the treatment of educational measurement in most labor studies where researchers assume that individuals with a BA have 16 years of education, even if that BA took more than 4 years to complete.

for institutions other than the one a student initially enrolled in shows no evidence of a discontinuity (see Figure 10).

While we have no data on enrollment at private schools or public universities outside of Ohio, the fact that we see relatively little movement of students across public institutions within Ohio is suggestive that falling below the cutoff has a real impact on the total educational investment of the students in our sample. Since Ohio's higher education system is dominated by public institutions, we suspect that if a large fraction of students were obtaining degrees from other schools as a result of the dismissal, we would observe this effect at public institutions. The 2-year institutions in particular would likely be an attractive institution to transfer to after failing out of a 4-year public institution, but we see no evidence of an increase in AA degrees (see Figure 11). That said, we cannot definitively rule out the possibility that our estimates of the impact on educational investments are too large because students substitute towards private institutions. To the extent that we are overestimating the impact on educational investment, we will be underestimating the earnings returns to education.

Robustness

For each outcome, we show how the estimates vary according to the bandwidth chosen, whether or not we include covariates, and whether or not we estimate the donut RD suggested by Barrecca et al (2015).[11] Table 3 shows the robustness of the results for academic outcomes while Table 4 shows the robustness of the results for the labor market outcomes. In Table 3, the estimates appear to be generally invariant to whether or not

---

[11] Based on the recommendation of Gelman and Imbens (2014), we focus on the local linear models. In our context, though, the local linear and polynomial models yield similar estimates.

covariates are included providing further suggestive evidence that differences in

unobservables between those on either side of the cutoff are unlikely to explain our

results. The results are similarly robust to implementing the donut RD suggesting that

the mechanical heaping at 2.0 is unlikely to explain our results. Broadly, the estimates are

also robust to bandwidth choices, though the smaller bandwidths generally yield

somewhat smaller estimates and larger standard errors. Our visual inspection of the 0.75

bandwidths makes clear that a linear fit does not approximate the data well with this

larger bandwidth, but the fitted lines for the 0.5 and 0.25 bandwidths both appear to

match the figures fairly well.

Table 4 shows that the estimated impact on mid-run earnings is similar across

various specifications. The earnings estimates are similar across different bandwidth

choices and vary only modestly depending on whether controls are included or exact

zeroes are excluded. The second panel of Table 4 shows that after conditioning on

covariates or estimating a donut RD, the insignificant "missing earnings" discontinuity

documented in Figure 6b becomes statistically significant. Though not as robust across

specifications, based on the statistically significant estimates, falling below the cutoff

reduces the probability of having missing earnings data by approximately 1.5-2.5 percent

depending on the specification. While this differential attrition has the potential to bias

our estimates, it is worth emphasizing that the magnitude of the differential attrition is

quite small. In order for 2 percent monotonic differential attrition to explain away our

earnings estimates, it would be necessary for the potential earnings outcomes of the

differential missings to be approximately 250% larger than their observed counterparts.

As noted earlier, Table 2 shows that there are no observable differences in the sample that

have observable mid-run earnings making it even less plausible that differential attrition substantially biases our estimated earnings effects.

Though the smaller bandwidth increases the likelihood that the linear fit on either side of the threshold is approximately correct, it also increases the probability that the slope of the lines over-fit noisy data. As such, it is not a priori obvious whether the smaller or larger bandwidth is preferable. We use the 0.5 bandwidth as our preferred specification moving forward, though the results are qualitatively similar for other choices. Furthermore, the 0.5 bandwidth yields more conservative IV estimates since it tends to estimate similar wage effects and larger educational effects. We include covariates to improve precision and we use the donut RD specification because it is found to outperform the standard RD specification in contexts with spikes in the running variable density. (Barrecca et al 2015)

IV estimates of the returns to schooling

The preceding section shows the reduced form estimated impact on both educational outcomes and earnings. In this section, we use the discontinuity as an instrument and estimate the impact of years of schooling on earnings. While computationally, the IV estimate is just the ratio of the earnings estimate to the years of schooling estimate, in order to interpret this as a local average treatment effect, we require several standard assumptions. Let Z denote our instrument (falling below the dismissal threshold), Y denote our outcome of interest, X denote years of schooling, and ε denote the error term in the causal equation with Y as the dependent variable. We consider the plausibility of each of these assumptions in turn below.

Assumption 1: $Cov(Z,X) \neq 0$ (relevance condition)

This assumption is testable and as we show earlier, a first-stage relationship exists. We further confirm this by considering the F-statistic on the excluded instrument.

Assumption 2: Cov(Z,ε) = 0 (exclusion restriction).

One potential threat to the exclusion restriction is that dismissal may be psychologically distressing so that in addition to reducing total time in school, it impacts emotional health in a fashion that hurts future earnings. To the extent that it exists at all, we believe this effect is likely small for several reasons. First, we study earnings measured many years after the dismissal so it seems implausible that the direct emotional consequences of dismissal would continue to be important. Second, while dismissal may be emotionally difficult, voluntary dropout due to low performance is likely similarly emotionally difficult in many ways. Voluntary dropouts may leave school with similar feelings of inadequacy as would dismissed students. Finally, while dismissed students are given a negative signal about their academic ability, those that just narrowly avoid dismissal are also likely aware of their generally poor performance.

While none are a violation of the exclusion restriction, it is worth mentioning that there are several different channels through which years of schooling may impact earnings. First, increasing years of schooling likely reduces labor market experience and thus the estimated return to schooling reflects the difference in human capital acquired through school vs work.[12] In this regard, we view lost experience as part of the schooling effect as opposed to something we might want to control for. Second, the impact of increasing years of schooling partly operates through increasing BA receipt. Just as with

---

[12] This conceptualization of the causal impact of education is somewhat different from the standard Mincer wage regression that simultaneously controls for education and experience.

years of employment, an increase in BA receipt does not introduces bias into our estimates of the returns to schooling, but it is important to keep in mind that our estimates include any sheepskin effect averaged over the number of years of college attended.[13]

Assumption 3: The impact of Z on X is either weakly positive or weakly negative for all individuals (monotonicity).

Unlike many IV contexts, the monotonicity assumption is not innocuous in our case because the dismissal may act as a wake-up call for certain students. In other words, it is possible that certain students on the path to dropout could be encouraged to complete more years of schooling by being forced to take a break. Unlike violations of the exclusion restriction, violations of monotonicity do not bias our estimates in a homogenous treatment effects world. The IV estimates become misleading, however, if the returns to schooling are very different for compliers compared to defiers. In particular, if the defiers derive no benefit from schooling, their existence reduces the size of the first-stage but leaves the reduced form unchanged, thereby inflating our IV estimates. On the other hand, if on average the defiers derive a similar benefit from schooling as the compliers, the IV estimates will be unbiased. We discuss the monotonicity assumption in greater detail in Appendix C.

Table 5 shows IV estimates of mid-run earnings for 3 different bandwidths. In each specification, our instrument has an F-statistic above conventional rule-of-thumb thresholds. Focusing on the results corresponding to a bandwidth of 0.5, as we have done previously, we see that an extra year of schooling is predicted to increase earnings by

---

[13] This conceptualization is consistent with the literature since most estimates of the returns to schooling include any sheepskin effects. For example, the average return to a year of college is estimated as the total return to college divided by 4, even though the sheepskin effect accrues entirely in the final year.

nearly 25%. While at first glance, this estimate appears to be much larger than that of past work, this difference is to be expected if sheepskin effects are important. Due to the nature of our identification strategy, the marginal years of education we examine must occur late in an individual's college career (because we focus on students at risk of dismissal in their second year or later). Thus, to the extent that there is a sheepskin effect of earning a BA separate from continuous human capital accumulation, this effect will be averaged out over a smaller number of years in our study than in most other studies, which effectively average the sheepskin effect over 4 years. The fact that our estimate of the return to a year of schooling is larger than many past estimates is suggestive evidence of sheepskin effects as opposed to suggesting that either our estimate is too large or past estimates are too small.

Extensions

Contemporaneous earnings effects of enrollment

Though our main goal is to understand how educational investments impact mid-run earnings, our empirical design also allows us to study the degree to which students forgo current earnings in order to attend college. The classical educational investment model assumes that college students forgo all of their earnings to attend school, but increasingly, the typical college student works substantial hours while enrolled in school (Darolia 2014). As such, the forgone earnings could be substantially smaller than typically assumed.

Figure 12 shows the reduced form relationship between the running variable and log weekly earnings in t+1. We estimate that falling below the cutoff increases immediate weekly earnings by approximately 5.3 percent. Recall, that there was no

discontinuity in log weekly earnings in period t-1, so it is unlikely that this estimate can be explained by differences in the types of students on either side of the cutoff.[14] The change in log weekly earnings can result either from changes in hourly wages or changes in hours worked and our data do not allow us to differentiate between these effects. That said, given that there is little reason to expect that immediate human capital changes substantially as a result of the enrollment decision, we suspect that much of the immediate change in earnings reflects changes in hours worked. However, it is also possible that wages increase in the short-run because students are able to devote more time to employment search and could end up with a better match.

In order to estimate the impact of enrollment on contemporaneous earnings, it is necessary to scale the 5.3 percent estimate by the estimated first-stage impact on enrollment, implying a roughly 41% increase in contemporaneous earnings. To interpret this as the local average treatment effect, we must assume an analogous set of IV assumption as were discussed previously. In the case of enrollment in the next term, the direction of the effect of our instrument is unambiguously negative, and thus the monotonicity assumption is of less concern here.

Heterogeneity

In Table 6, panels A and B, we show the estimated impact of falling below the threshold on long-run academic outcomes for various groups. Panel A shows that there is relatively little heterogeneity across groups in terms of the impact on total credits earned and the standard errors are too large to statistically detect differences across groups.

---

[14] In results not shown, we find that there is also no discontinuity for any of the observable characteristics when restricting the sample to individual with non-missing t+1 earnings.

Panel B similarly shows that there is relatively little heterogeneity in terms of the impact on whether individuals earn a BA. The one exception is that the effect on BA receipt is considerably smaller for students with 2nd year standing compared to students with higher standing at the time of the dismissal.

While the heterogeneity in the BA effect could be the result of a differential response to dismissals in earlier and later terms, we suspect that it mostly reflects heterogeneity in the control group. Recall that for students who are unlikely to graduate in any case, the dismissal leads to an earlier dropout but has no impact on whether or not the student graduates. There are many students with 2nd year standing that are unlikely to ever graduate, but conditional on obtaining 4th year standing, most students are expected to graduate. Although it makes sense that dismissal will have a larger impact on the graduation rates of those likely to graduate, it is somewhat surprising that dismissal seem to derail students so close to graduation. One possibility is that these students perceive that sheepskin effects are small so that the returns to completion are small conditional on almost completing. Alternatively, it is possible that the students we study are only tenuously connected to their academic institutions and are delicately balancing a variety of factors that allow them to stay in school.[15]

In panel C, we show heterogeneity in the IV estimates of the impact of years of schooling on earnings. Large standard errors for the subgroup analysis means that many of the estimates in panel C are statistically insignificant, but the estimates are positive and

---

[15] While it may be difficult to rationalize the lack of persistence in college from a neoclassical perspective, it is worth noting that it is similarly difficult to rationalize why students so close to graduation would be performing so poorly in the first place. Anecdotally, each term approximately 10% of economics majors in our upper level economics courses earn an F and these students earn low grades as a result of poor attendance or failure to attempt exams.

large for all groups. While the large standard errors prevent us from statistically distinguishing across groups, in terms of the point estimates, women and disadvantaged minorities have a larger return than men and non-minorities respectively. Though all the estimates are noisy, we observe much larger estimated returns to a year of schooling for those with 4th year standing at the time of dismissal. The statistical imprecision prevents any strong conclusions, but note that a large estimated return to a year of schooling for those with higher standing is consistent with a large sheepskin effect. This is because late dismissal has a relatively small impact on total credits earned and a large impact on BA completion. In the extreme, for students with 1 day left before graduation, the return to that 1 day of schooling could be extremely large if there are sheepskin effects.

Separating the BA effect from the years of schooling effect

One unique feature of our identification strategy is that the quasi-experimental variation used to identify the returns to education occurs at different points in time for different individuals; this makes it possible to separately identify the BA effect from the years of schooling effect. The intuition is that since there is a larger BA effect and smaller credits earned effect for those with later standing, if we see a larger impact on earnings from early dismissals, this suggests that credits earned are more important whereas if the earnings impact is larger for later dismissals the sheepskin effect is more important. Econometrically, we treat years of schooling and BA completion as both endogenous and our two instruments are whether an individual is below the cutoff and that same variable interacted with standing.

In order to separately identify the impact of BA completion and years of schooling, it is necessary to make two assumptions in addition to the standard IV

assumptions discussed earlier. First, we assume linearity in the years of schooling effect. Kane and Rouse (1995) also impose this assumption since without it the BA effect could simply reflect a sharp spike in the returns to schooling. Second, we assume that the local average treatment effect of the BA and years of schooling is the same across our instruments. In other words, the compliers with lower standing must on average have the same returns to education as the compliers with higher standing. Without this assumption, we would conflate heterogeneity in the returns to years of schooling with the BA effect. Although both of these assumptions are fairly strong, these assumptions are considerably weaker than the assumptions used in past work, which has also assumed that education is exogenous conditional on observables.

Table 7 shows our estimates of the BA and years of schooling effect using 3 different bandwidths. Columns 1, 4 and 7 show the IV estimates whereas the other columns show the first-stage estimates. Though the IV estimates are quite similar across the 3 bandwidths, the smallest bandwidth provides less power so the F-statistic on the years of schooling first-stage is only 6.4. Consistent with the rest of this manuscript, we focus on a a bandwidth of .5 to balance statistical precision and parameter identification. Columns 5 and 6 show that the negative impact of dismissal on BA receipt is larger in magnitude for those with higher standing whereas the negative impact on credits completed is smaller in magnitude for those with higher standing. The IV estimates are generally imprecise, but the point estimates are large and positive for the BA effect and positive (although small) for the years of schooling effect..

Taken at face value, our estimates suggest that a student that completes 4 years of schooling and obtains a degree should obtain approximately a 41 percent total return

(0.363+4*0.0139). It is important to emphasize that these results are likely specific to the sample of low-performing students that we study since we suspect that the human capital benefits of a year of school are likely smaller for students that have very low performance. These students likely skip more classes, are less likely to complete assignments and are more likely to be lost during the classes they do attend.

The large standard errors in Table 7 prevent any strong conclusions, but given that ours is the first study to attempt to separate out the years of schooling effect from the BA effect using credibly identified variation, we still view these estimates as a valuable starting point. The type of data we use in this study is increasingly available in a variety of states, so future research might use a similar empirical methodology combined with a larger sample to potentially obtain more precise estimates.

Conclusion

We show that low-performing students who are dismissed from public 4-year colleges in Ohio suffer substantial earnings losses measured between 7-12 years after college enrollment. Given the large economic returns to persisting in college for this group, a natural question is why so many similar students drop out. We see four broad categories of potential explanations.

First, it is possible that the dropout decision is not really a decision at all and instead is forced upon students by their university. For example, it might be that given the incoming skill level of students combined with a time constraint, many students cannot satisfy the performance standards and are thus prevented from graduating by probation-dismissal policies.

Second, it is possible that the returns to schooling that we identify are larger than for the typical dropout. Our instrument only alters the behavior of students who would otherwise have stayed in school. In other words, the dismissal only binds for students who view the benefits of school as larger than the costs. As such, it is possible that students who voluntarily drop out of school have a lower return to schooling than the compliers in our study. While Oreopolous (2006) finds evidence that the average treatment effect is similar to the local average treatment effect in the context of high school students in the United Kingdom, the returns to college have the potential to be much more heterogeneous.

Third, it is possible that despite large economic benefits, dropouts still view the total costs as larger than the total benefits. For example, dropouts could have relatively high discount rates such that the upfront psychic and economic costs of college attendance outweigh the large long-term benefits. Alternatively, if students are risk averse, the certainty of the upfront costs combined with the uncertainty of the long-run benefits could reduce the attractiveness of investing in college. Finally, dropouts may not know the causal economic returns to a college degree given that there remains uncertainty regarding the magnitude of these returns among researchers.

Moreover, even if students view college to have a high rate of return, credit constraints could prevent certain students from persisting (Belley & Lochner, 2007; Goodman, 2010;). Though there exists many different sources of loans for college students in the United States, students may be debt averse (Marx and Turner, 2015), may need to finance familial obligations for which they cannot borrow, or may need more

loans than the capped amounts in order to sufficiently smooth consumption (Lochner and Monge- Naranjo, Forthcoming).

While it is difficult to distinguish between the forces behind the average dropout decision, among the students we study, there are clear external forces contributing to dropout. While this loss of educational investment likely reduces these students' total welfare, their dismissal likely has positive externalities so that the total social welfare impact of the dismissal policies may well be positive. In particular, restricting low-performing students from graduating maintains the signaling value of the degree and provides incentives for all students to succeed. As such, we view our study as uninformative regarding the total impact of dismissal policies. Instead, our study uses these dismissal policies to provide evidence on whether low-achieving students benefit financially from persisting in college. Though we do not assess the costs of attendance faced by these students, the magnitude of our estimates combined with the fact that the returns to college accrue over a lifetime suggests that persisting likely provides a large positive return, even for low-performing students.

References

Andrews, Rodney, Jing Li, and Michael F. Lovenheim. "Heterogeneous paths through college: Detailed patterns and relationships with graduation and earnings." *Economics of Education Review* 42 (2014): 93-108.

Anelli, M. (2015). Returns to Elite College Education: A Quasi Experimental Analysis.

Angrist, Joshua D., and Guido W. Imbens. "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." *Journal of the American statistical Association* 90.430 (1995): 431-442.

Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell. "Heaping-Induced Bias in Regression-Discontinuity Designs." *Economic Inquiry* 54.1 (2016): 268-293.

Baum, Sandy, Jennifer Ma, and Kathleen Payea. "Education Pays 2013."The College Board (2013).

Barua, Rashmi, and Marian Vidal-Fernandez. "No Pass No Drive: Education and Allocation of Time." *Journal of Human Capital* 8.4 (2014): 399-431.

Belley, Philippe, and Lance Lochner. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital* 1.1 (2007): 37-89.

Canaan, S., & Mouganie, P. (2014). Quality of higher education and earnings: Regression discontinuity evidence from the French Baccalaureate. Available at SSRN 2518067.

Card, David "*Using geographic variation in college proximity to estimate the return to schooling*" in Aspects of Labour Economics: Essays in Honour of John Vanderkamp, edited by Louis Christofides, E. Kenneth Grant and Robert Swindinsky. University of Toronto Press. (1995)

Card, David. "The causal effect of education on earnings." Handbook of Labor Economics. Volume 3 (1999): 1801-1863.

Card, David, Andrew Johnston, Pauline Leung, Alexandre Mas, and Zhuan Pei.. "Inference on Causal Effects in a Generalized Regression Kink Design." (2015): IZA Discussion Paper 8757.

Card, David, and Thomas Lemieux. "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis." *Quarterly Journal of Economics* 116.2 (2001).

Darolia, Rajeev. "Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students." *Economics of Education Review* 38 (2014): 38-50.

Fiorini, Mario, et al. Monotonically Hopeless? Monotonicity in IV and fuzzy RD designs. Tech. rep, 2013.

Gerard, Francois, Miikka Rokkanen, and Christoph Rothe. Partial Identification in Regression Discontinuity Designs with Manipulated Running Variables. working paper, Columbia University, 2015.

Goodman, Joshua. "Skills, schools, and credit constraints: Evidence from Massachusetts." *Education* 5.1 (2010): 36-53.

Hungerford, Thomas, and Gary Solon. "Sheepskin effects in the returns to education." *The review of economics and statistics* (1987): 175-177.

Jaeger, David A., and Marianne E. Page. "Degrees matter: New evidence on sheepskin effects in the returns to education." *The review of economics and statistics* (1996): 733-740.

Kane, Thomas J. *A quasi-experimental estimate of the impact of financial aid on college-going*. No. w9703. National Bureau of Economic Research, 2003.

Kane, Thomas J., and Cecilia Elena Rouse. "Labor-market returns to two-and four-year college." *The American Economic Review* (1995): 600-614.

Kirkebøen, L., Leuven, E., & Mogstad, M. "Field of study, earnings, and self-selection" *Quarterly Journal of Economics*. Forthcoming.

Lee, David S. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76.3 (2009): 1071-1102.

Lindo, Jason M., Nicholas J. Sanders, and Philip Oreopoulos. "Ability, Gender, and Performance Standards: Evidence from Academic Probation." *American Economic Journal: Applied Economics* 2.2 (2010): 95-117.

Liu, Vivian YT, Clive R. Belfield, and Madeline J. Trimble. "The medium-term labor market returns to community college awards: Evidence from North Carolina." *Economics of Education Review* 44 (2015): 42-55.

Marx, Benjamin M., and Lesley J. Turner. *Borrowing Trouble? Student Loans, the Cost of Borrowing, and Implications for the Effectiveness of Need-Based Grant Aid*. No. w20850. National Bureau of Economic Research, 2015.

Mincer, Jacob A. "Age and Experience Profiles of earnings." *Schooling, experience, and earnings*. NBER, 1974. 64-82.

Lochner, Lance and Alexander Monge-Naranjo, "Student Loans and Repayment: Theory, Evidence and Policy", in E. Hanushek, S. Machin, and L. Woessmann (eds.), Handbook of the Economics of Education, vol. 5, Amsterdam: Elsevier Science, forthcoming.

Van der Klaauw, Wilbert. "Estimating the effect of financial aid offers on college enrollment: A Regression–Discontinuity Approach*." *International Economic Review* 43.4 (2002): 1249-1287.

Vidal-Fernández, Marian. "The effect of minimum academic requirements to participate in sports on high school graduation." *The BE Journal of Economic Analysis & Policy* 11.1 (2011).

Walker, Ian, and Yu Zhu. "The College Wage Premium and the Expansion of Higher Education in the UK*." *The Scandinavian Journal of Economics* 110.4 (2008): 695-709.

Webber, Douglas A. "The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors." *Labour Economics* 28 (2014): 14-23.

Webber, Douglas A. "Firm market power and the earnings distribution." *Labour Economics* 35 (2015): : 123-134.

Zimmerman, Seth. "The returns to college admission for academically marginal students." *Journal of Labor* Economics 32.4 (2014): 711-754.

Figure 1: Average log weekly earnings by BA receipt status
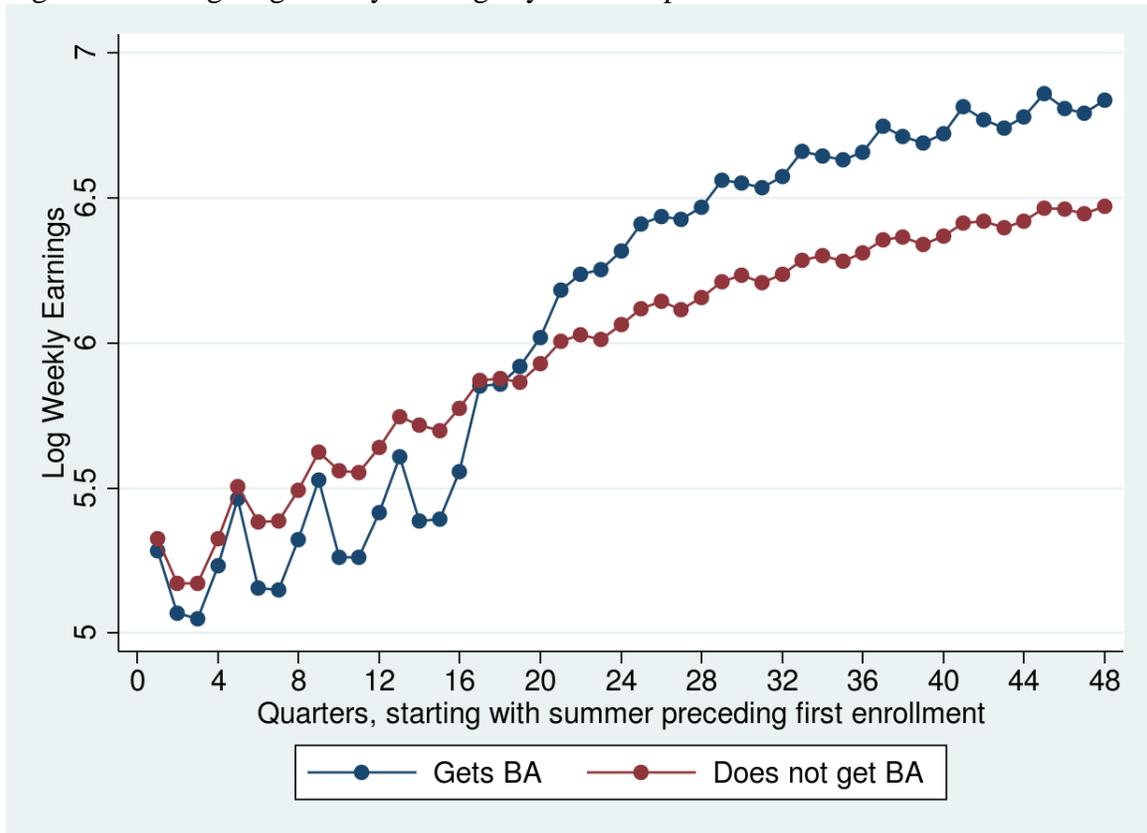
Note: This figure depicts the average log weekly earnings for all Ohio workers who we observe attending a public 4-year university in the state of Ohio during the sample period.

Figure 2: Enrollment discontinuity (full sample)
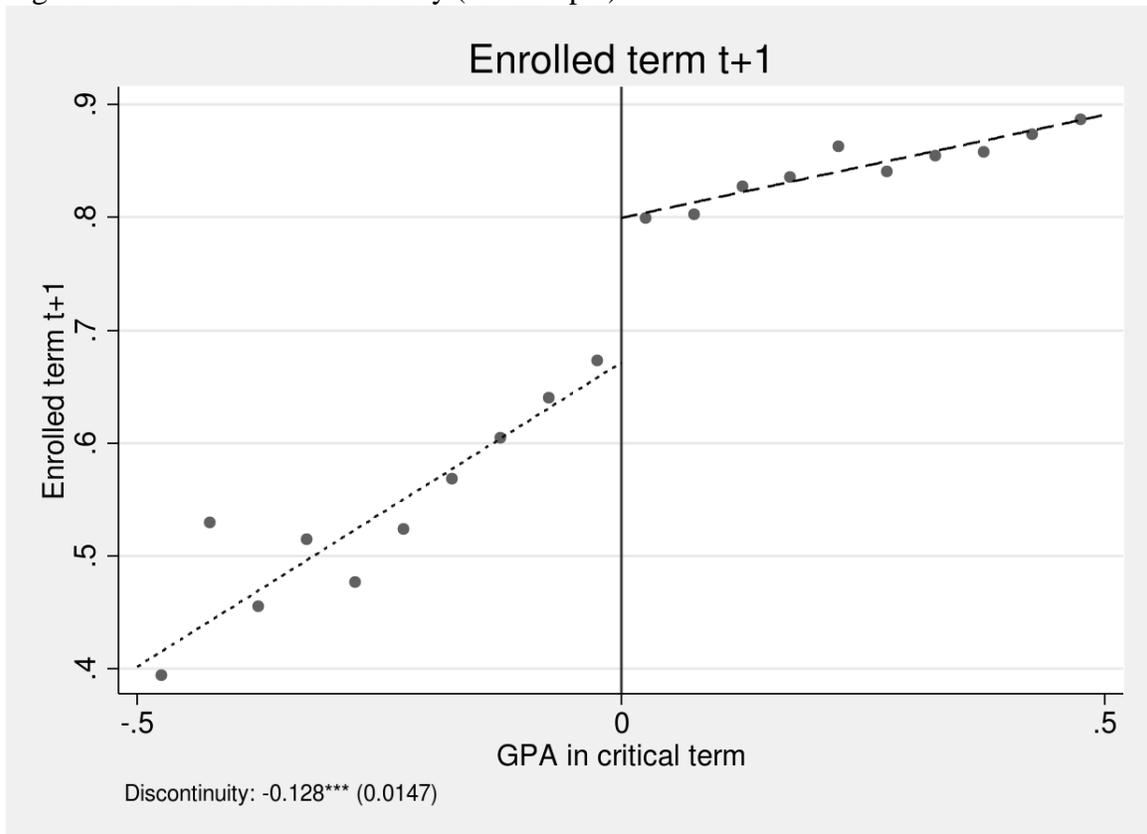
Note: This figure depicts the probability of enrollment in the subsequent term, by GPA in term t, for students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 3a: Enrollment discontinuity (first-year standing)
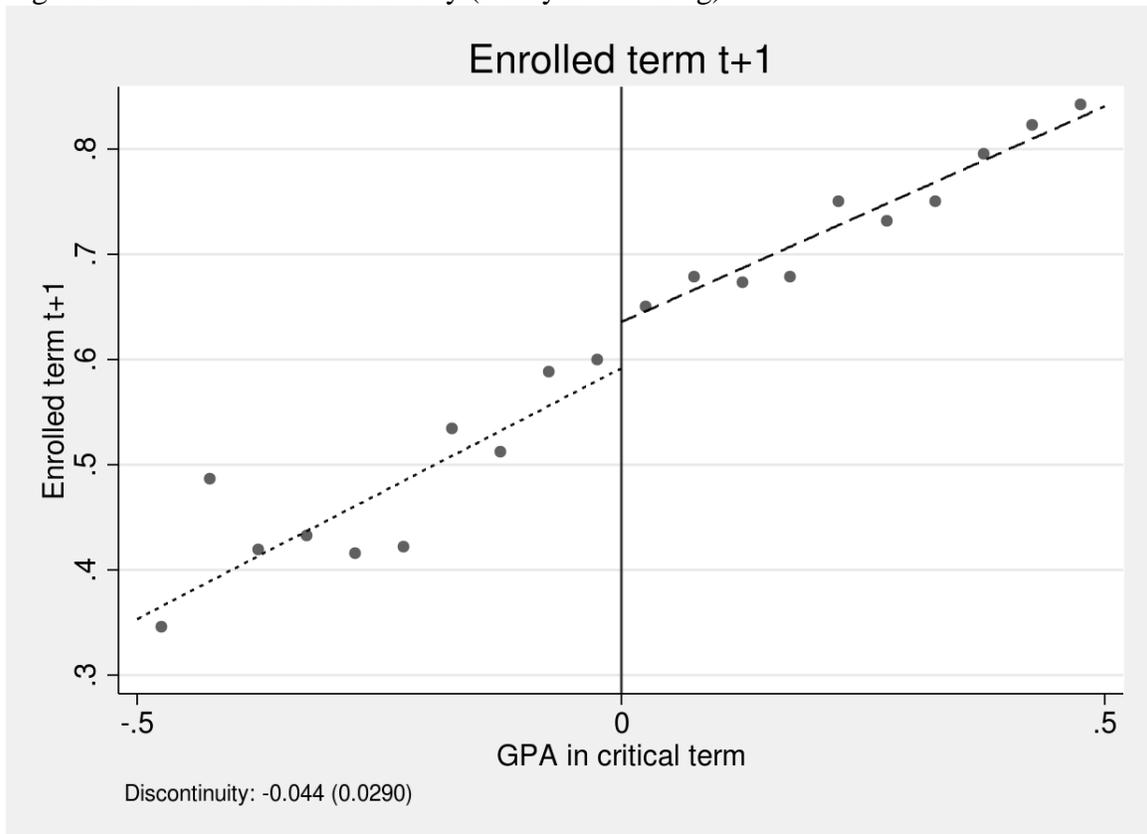
Note: This figure depicts the probability of enrollment in the subsequent term for students with first-year standing, by GPA in term t, for students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 3b: Enrollment discontinuity (second-year or higher standing)
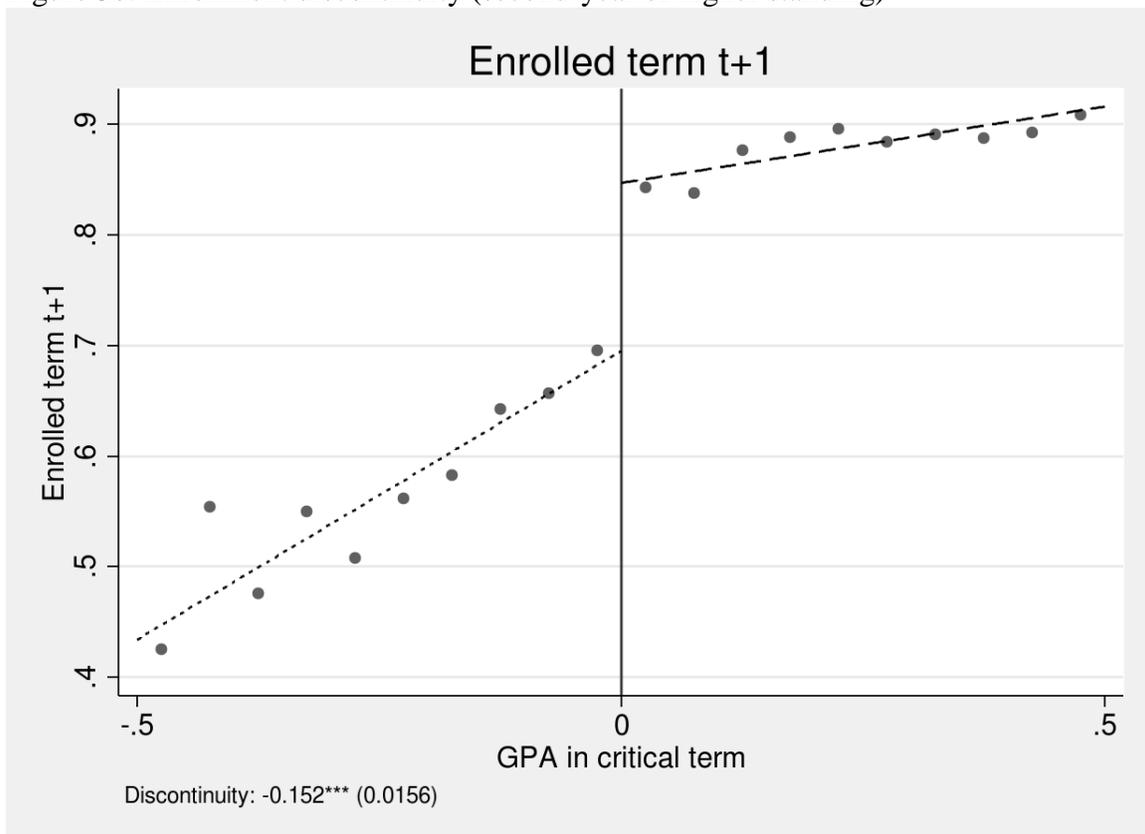
Note: This figure depicts the probability of enrollment in the subsequent term for students with second-year or higher standing, by GPA in term t, for students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 4a: GPA histogram
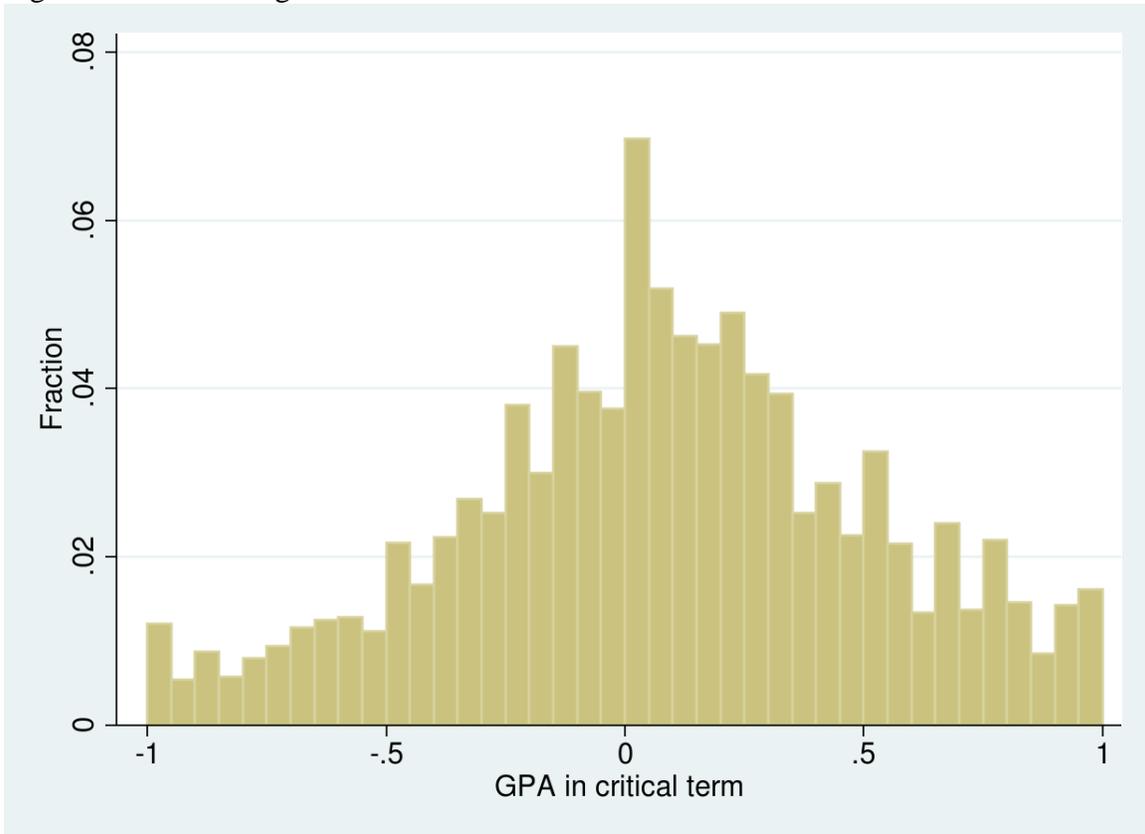
Note: This figure plots the distribution of GPA's for students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 4b: GPA histogram
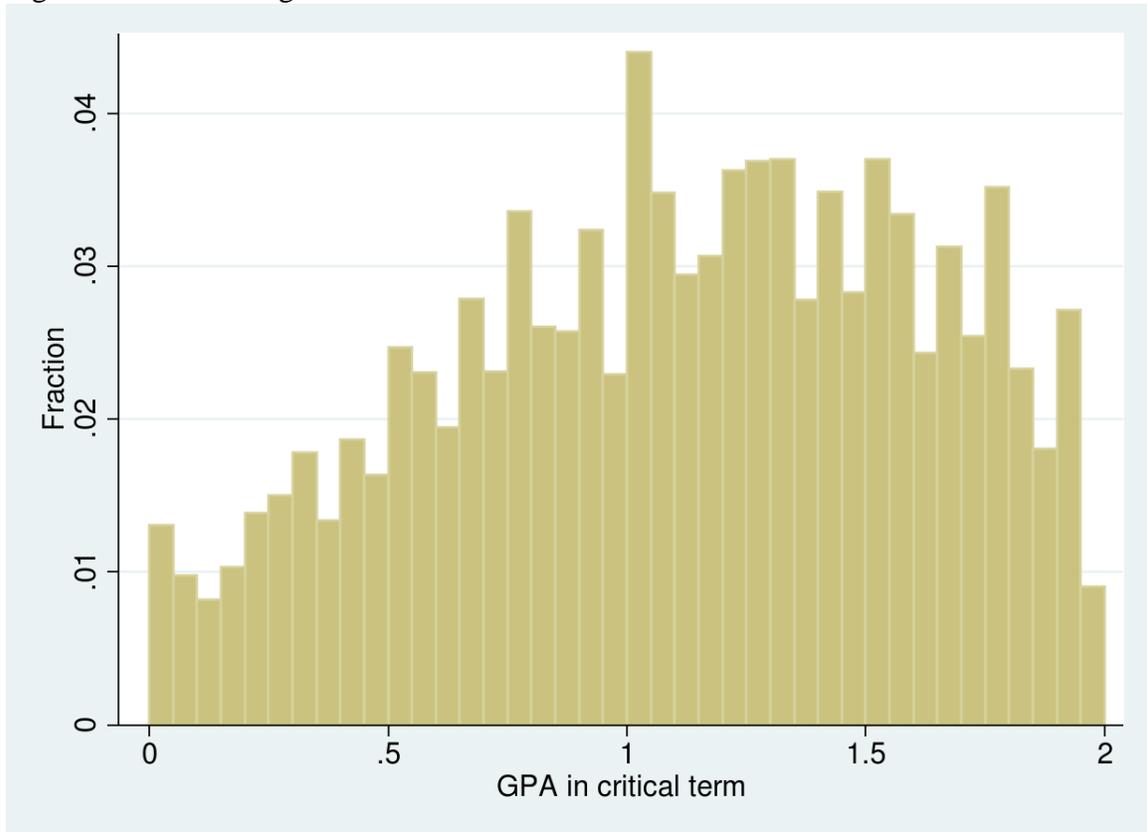
Note: This figure plots the distribution of GPA's for all students (including those never
on academic probation). The running variable is normalized to be 0 at the dismissal
threshold so all the students in the above figure are above the dismissal threshold. The
spike at 1 (and at various other points with no practical importance) is meant to illustrate
that the nature of how GPA's are calculated leads to heaping, often at integers.

Figures 5a-5l: Tests of covariate balance

Female

Discontinuity: -0.026 (0.0189)

Black

Discontinuity: -0.001 (0.0130)

Age at college entry

Discontinuity: -0.051 (0.0628)

Term GPA in term 1

Discontinuity: 0.025 (0.0313)
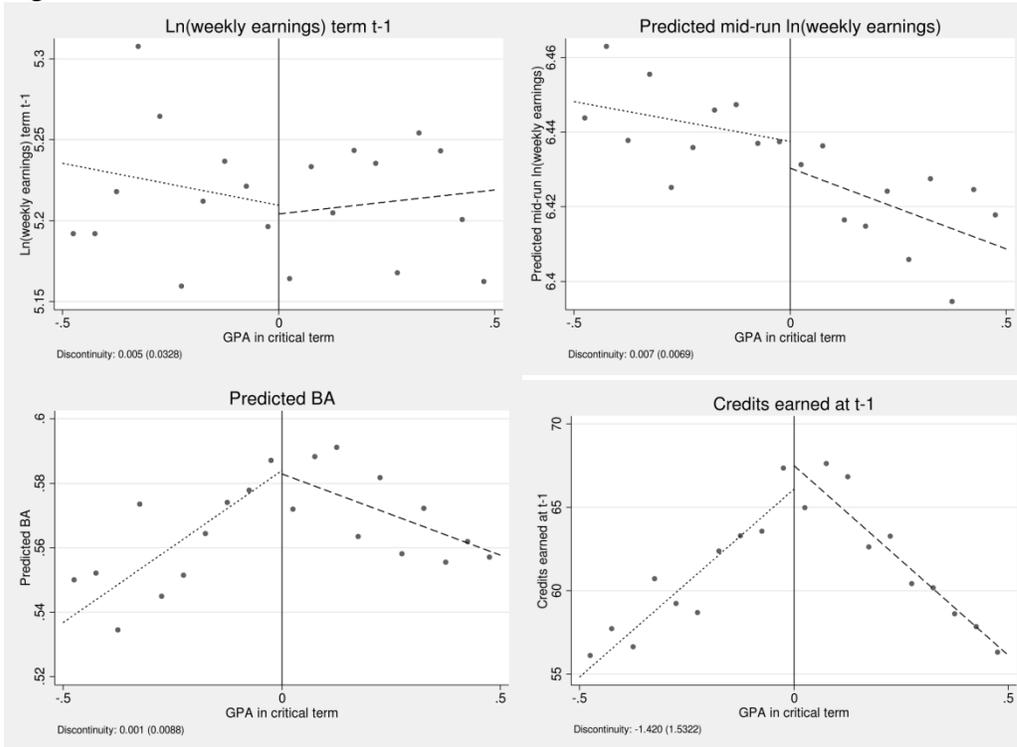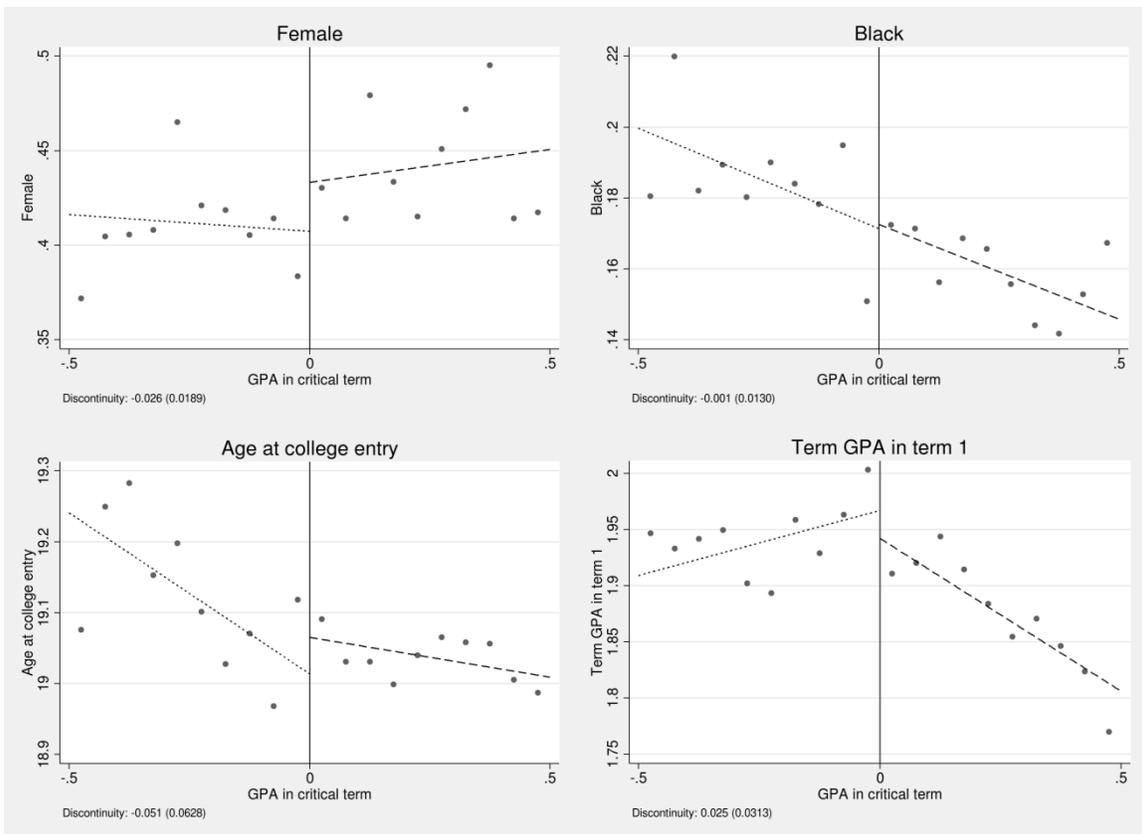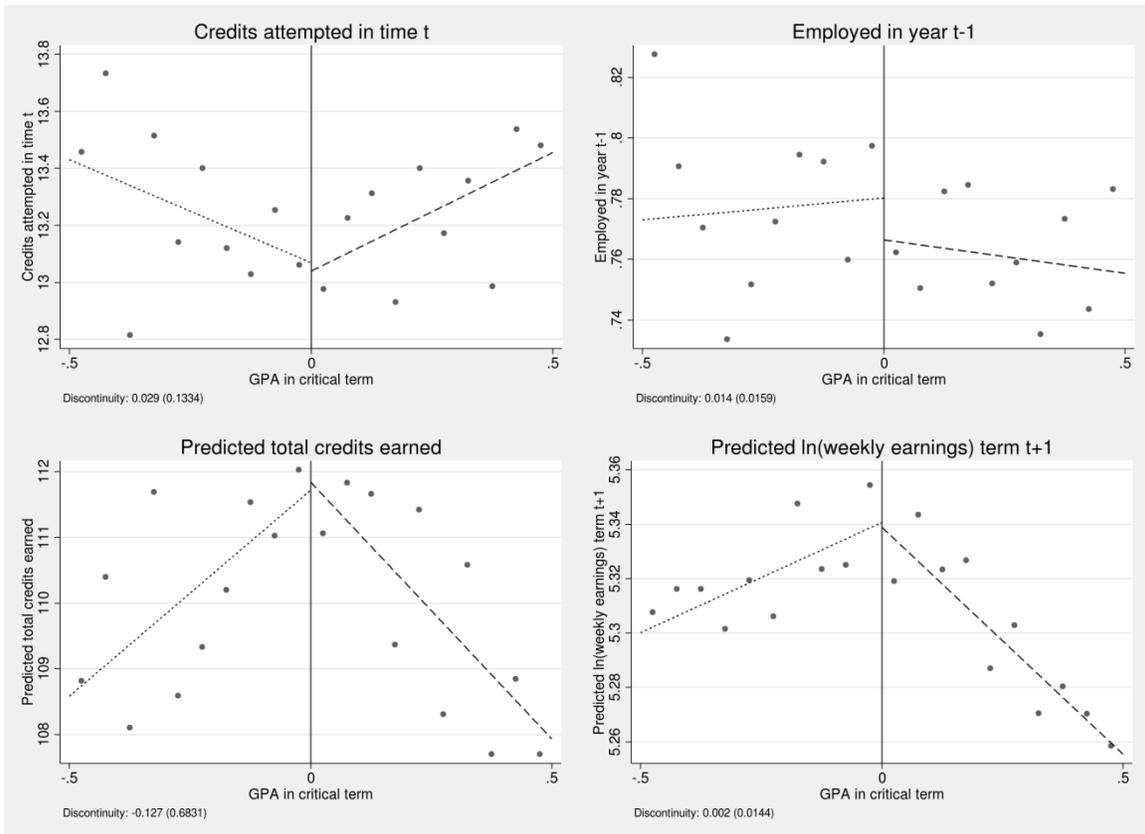
Note: The above figures provide tests of covariate balance around our discontinuity threshold.

Figure 6a: Discontinuity in mid-run earnings
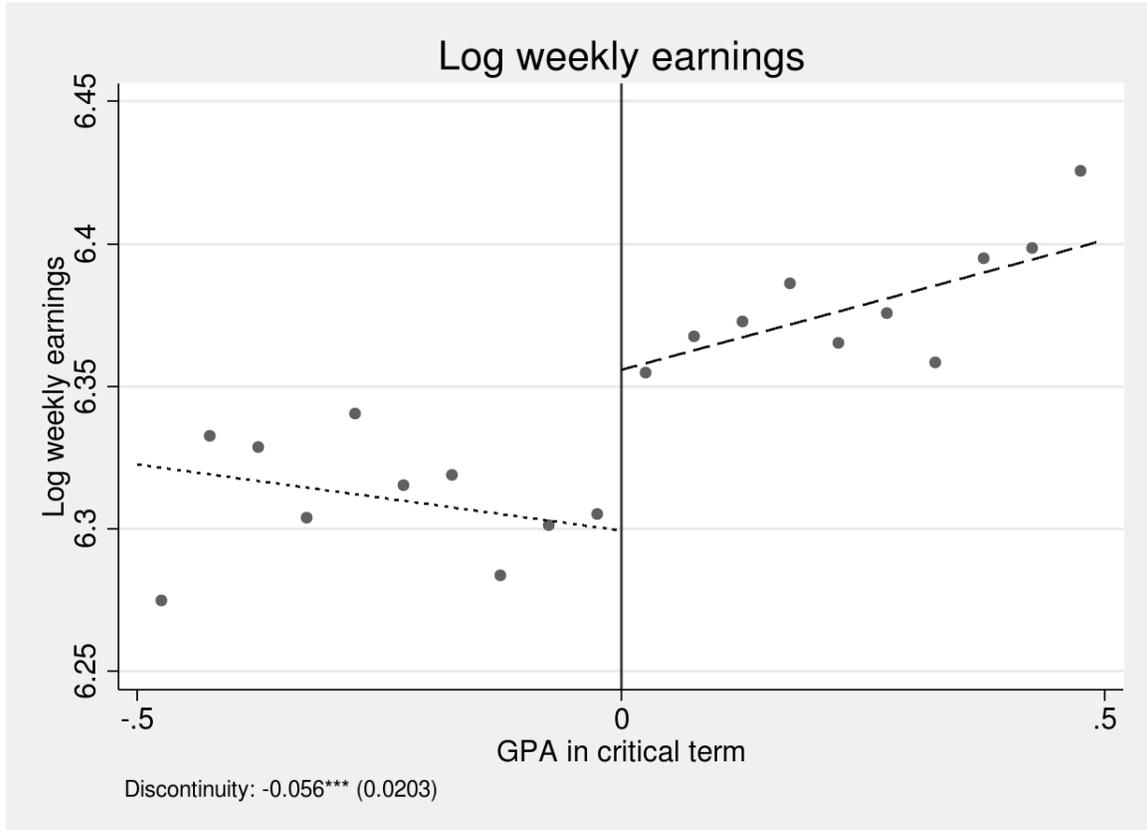
Note: This figure plots average log weekly earnings measured 7-12 years after initial enrollment in college (see the data section for a more detailed description of the dependent variable). The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 6b: Differential attrition around GPA threshold

Missing earnings year 7-12

Discontinuity: -0.017 (0.0143)
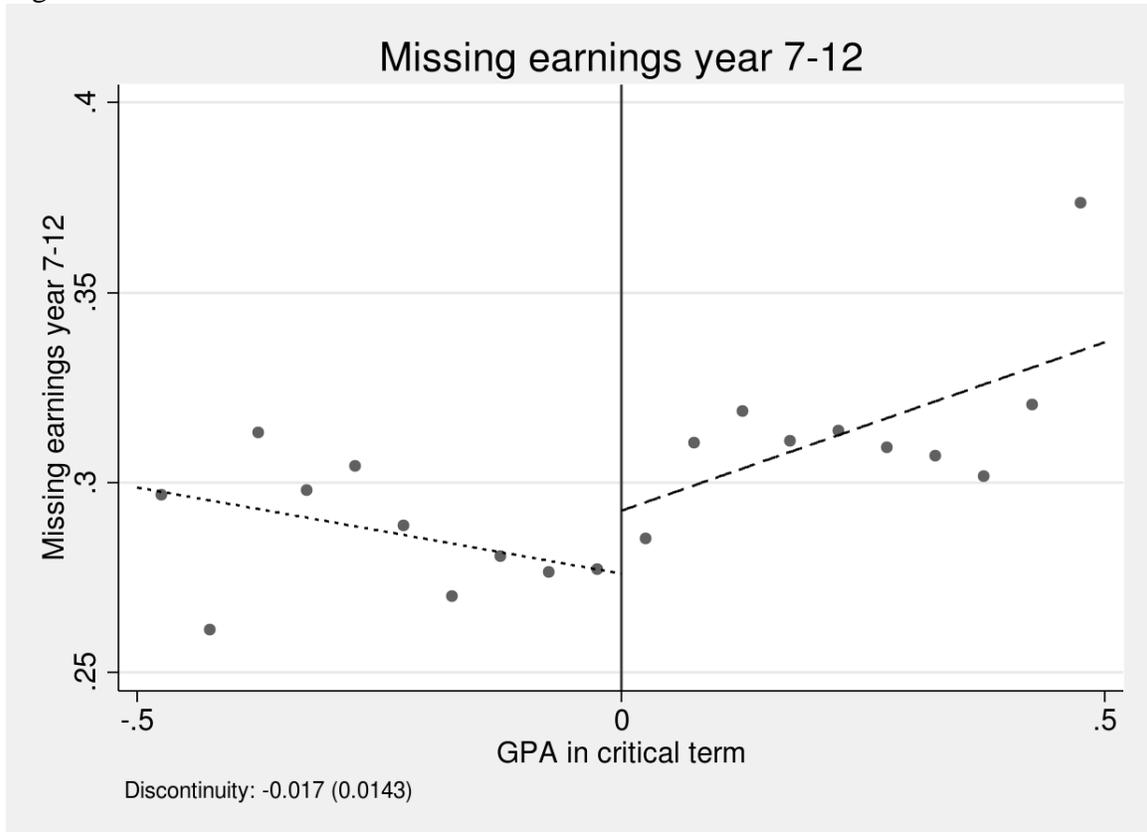
Note: This figure plots the likelihood of missing earnings 7-12 years after initial college enrollment. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 7: Discontinuity in total credits earned

Total Credits Earned

Discontinuity: -4.850*** (1.1438)
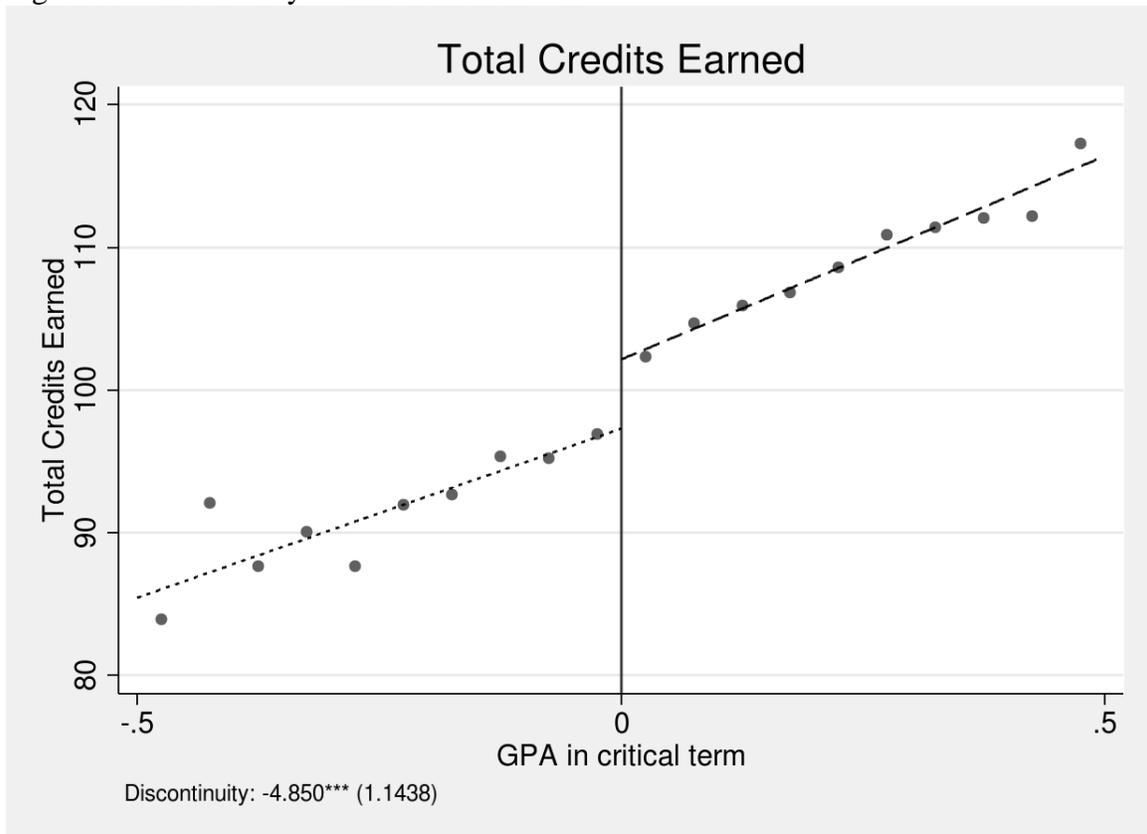
Note: This figure plots average total credits earned in college by GPA relative to the dismissal threshold at their institution. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 8: Discontinuity in credits earned outside of starting institution
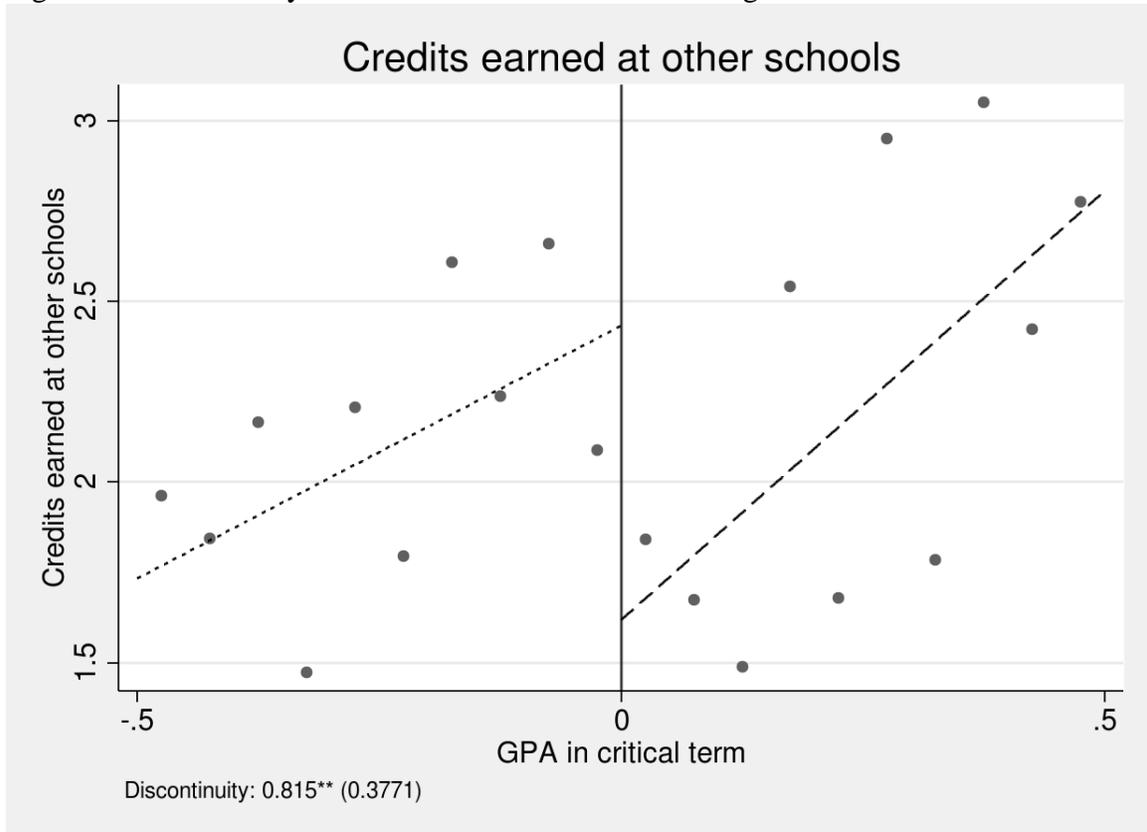
Note: This figure plots the total credits earned by an individual outside of their starting institution by GPA relative to the dismissal threshold at their institution. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 9: Discontinuity in BA receipt from starting institution

BA from starting institution

Discontinuity: -0.108*** (0.0146)
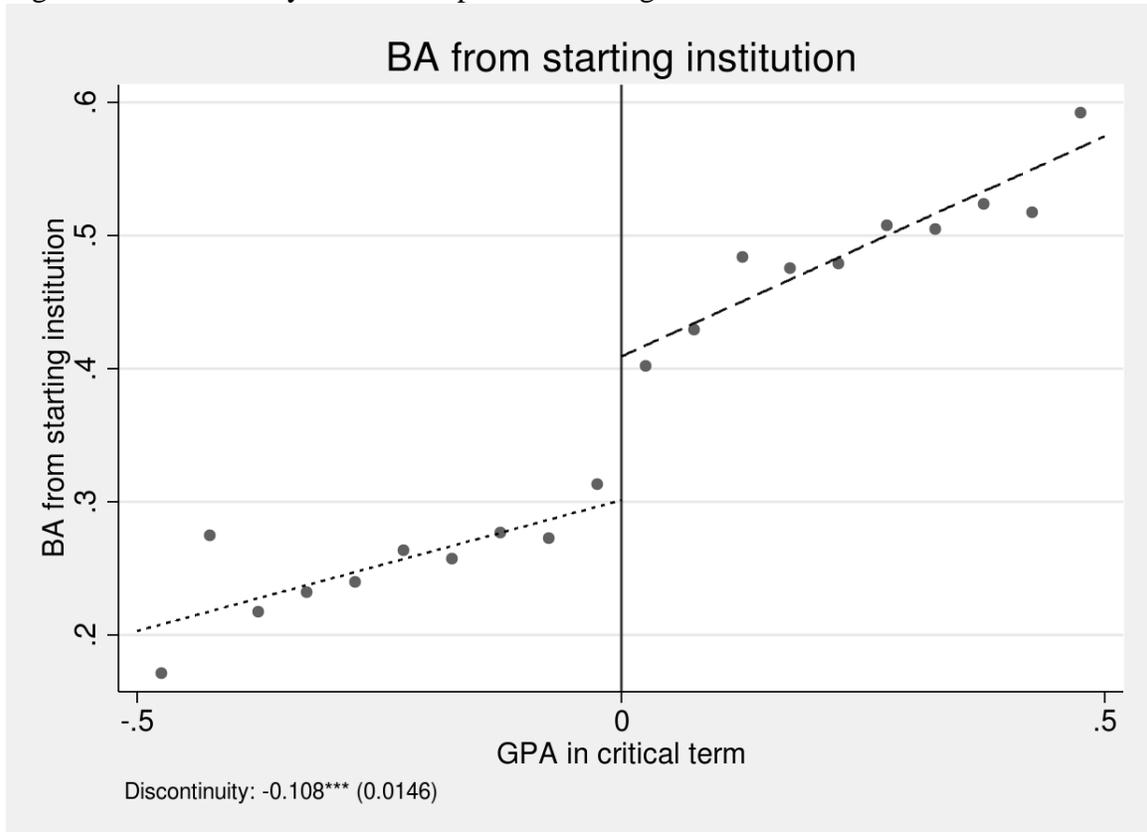
Note: This figure plots the average probability of completing a college degree by GPA relative to the dismissal threshold at their institution. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 10: Discontinuity in BA Degree receipt from a different institution

BA earned at other schools

Discontinuity: -0.001 (0.0034)
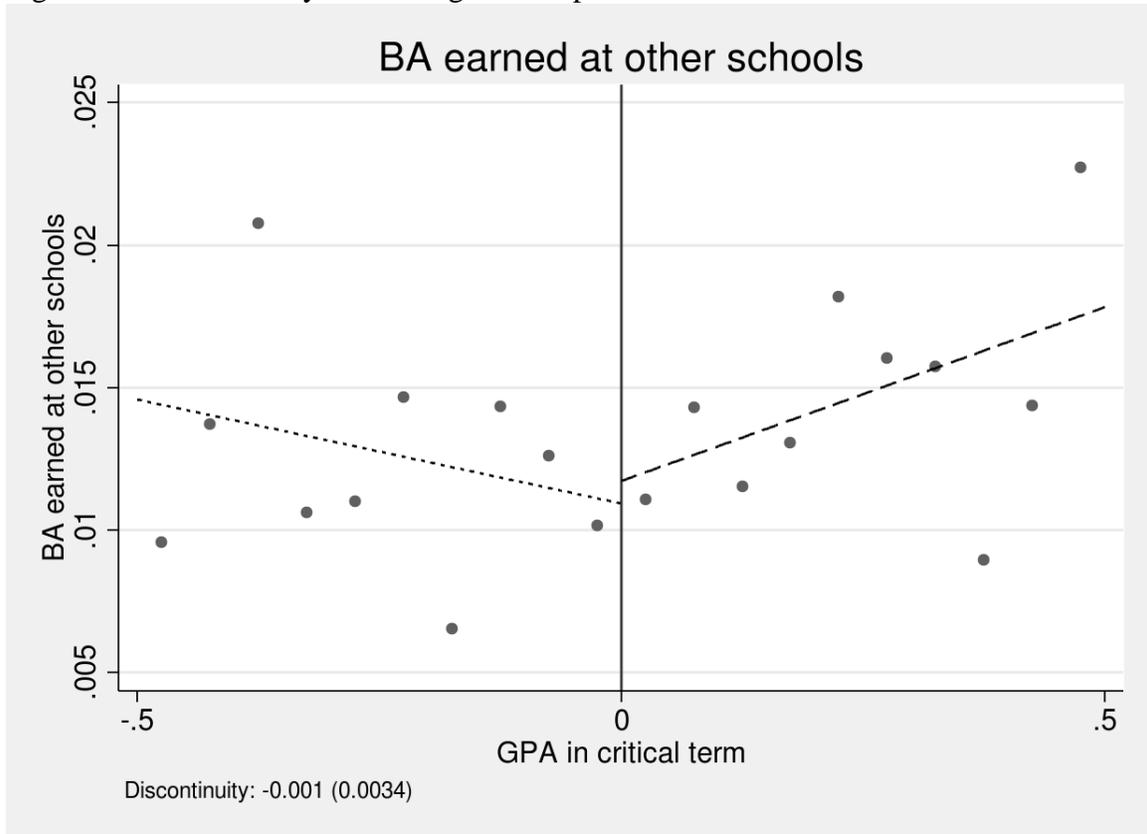
Note: This figure plots the likelihood of ever receiving an BA degree from a n institution other than an individual's starting institution by GPA relative to the dismissal threshold at their institution. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 11: Discontinuity in AA Degree receipt

## AA from any Ohio public

Discontinuity: -0.010 (0.0081)
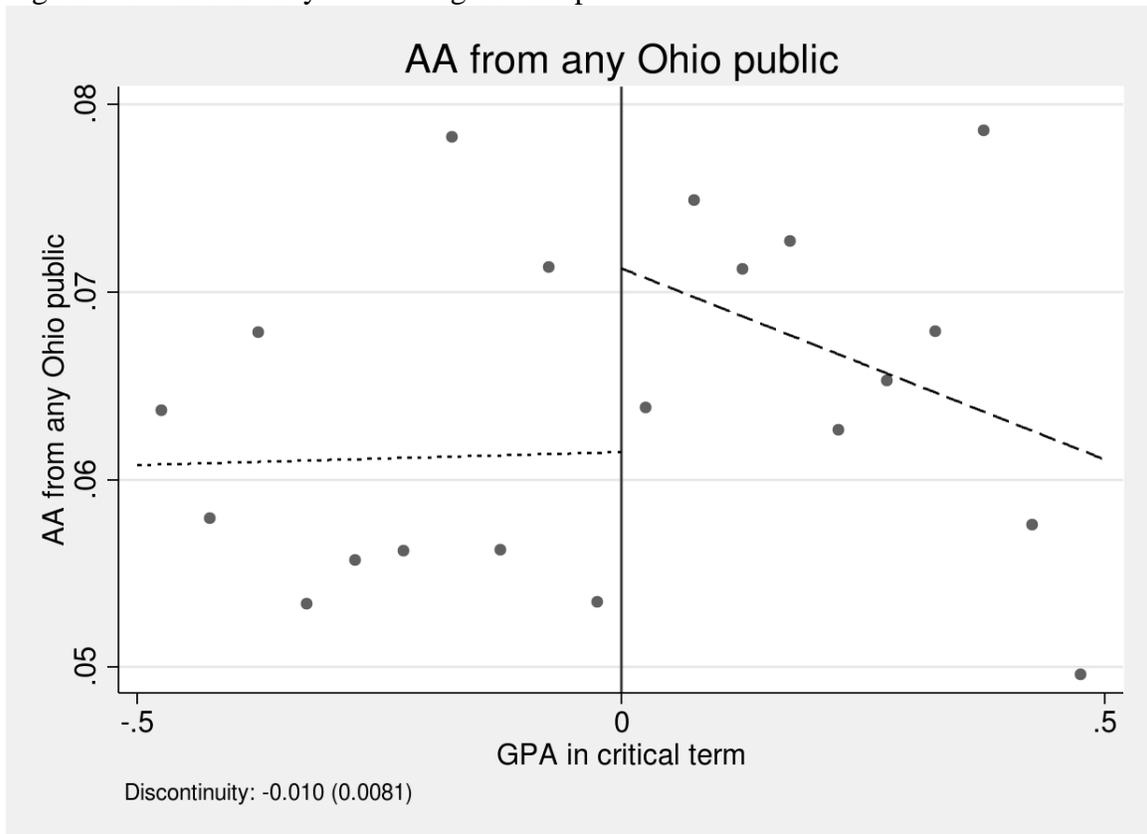
Note: This figure plots the likelihood of ever receiving an AA degree from a 2-year university by GPA relative to the dismissal threshold at their institution. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Figure 12: Discontinuity in contemporaneous earnings
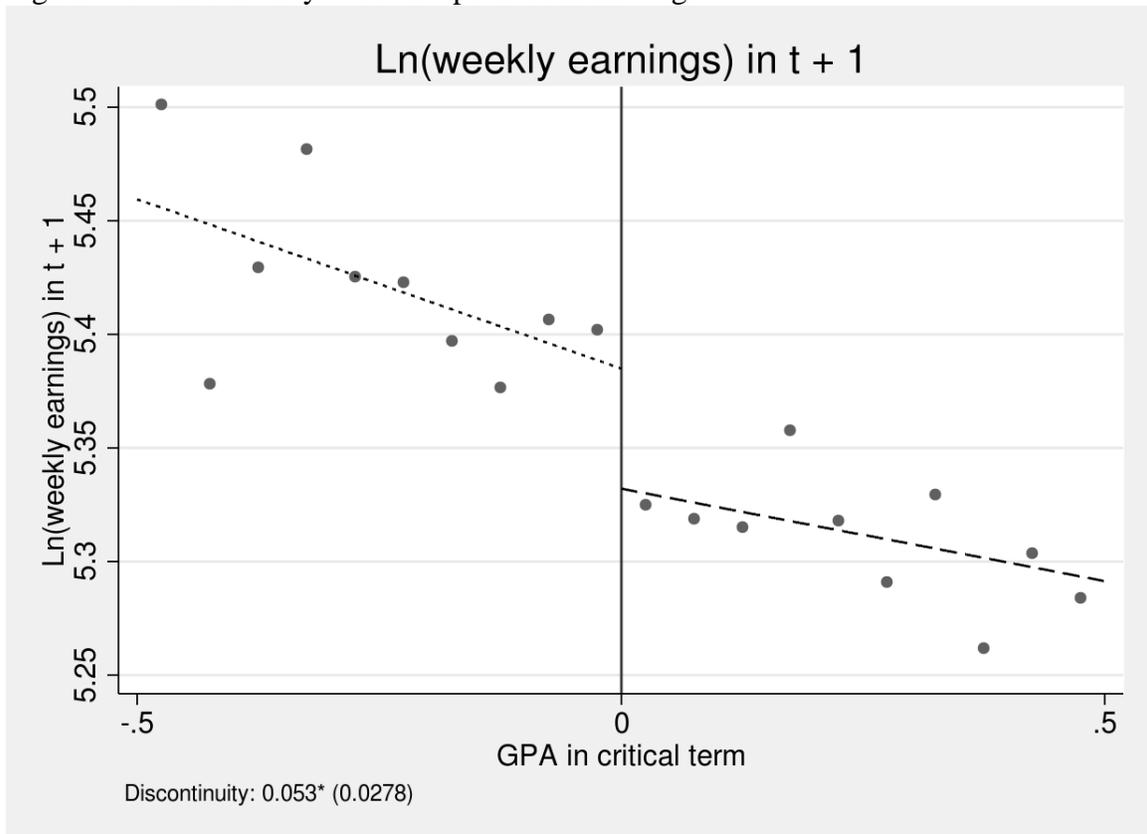
Note: This figure plots average log weekly earnings measured one term following the probation term by GPA relative to the dismissal threshold at their institution. The sample includes only students who are predicted to be on academic probation (based on the assignment rules collected by the authors). The running variable is normalized to be 0 at the dismissal threshold.

Table 1: Summary Statistics

Variables	Full Student Sample	At-Risk Sample
Ever on probation	0.24	1.00
Ever at risk of dismissal	0.10	1.00
Ever dismissed	0.06	0.53
Female	0.55	0.45
Black	0.08	0.17
Age at college entry	19.14	19.02
BA from starting institution	0.57	0.27
BA from any Ohio public institution	0.60	0.29
AA from any Ohio public institution	0.06	0.06
Total credits earned from starting institution	105.08	76.12
Total credits earned from any institution	109.32	79.97
Years to degree	4.17	4.62
Missing weekly earnings	0.38	0.32
Weekly earnings	805.85	647.09
Age when weekly earnings measured	28.80	28.70
Observations	218,030	21,664

Note: The full student sample is comprised of all students for whom we observe their initial enrollment in a public Ohio institution. The at-risk sample is comprised of only those students who are predicted to be on academic probation based on the rules for each institution collected by the authors.

Table 2: Tests of covariate balance

Sample	All at-risk students	Only students with future mid-run earnings
Credits earned at t-1	-1.420	-1.159
	(1.532)	(1.764)
Female	-0.0259	-0.0135
	(0.0189)	(0.0234)
Black	-0.00121	0.0121
	(0.0130)	(0.0147)
Age at college entry	-0.0514	-0.0333
	(0.0628)	(0.0660)
Term GPA in term 1	0.0247	0.0183
	(0.0313)	(0.0373)
Credits attempted in time t	0.0294	-0.0135
	(0.133)	(0.142)
Ln(weekly earnings) term t-1	0.00541	-0.0141
	(0.0328)	(0.0353)
Employed in year t-1	0.0138	0.0166
	(0.0159)	(0.0152)
Predicted mid-run ln(weekly earnings)	0.00713	0.00496
	(0.00693)	(0.00742)
Predicted BA	0.000992	0.00196
	(0.00884)	(0.0100)
Predicted total credits earned	-0.127	-0.155
	(0.683)	(0.742)
Predicted ln(weekly earnings) term t+1	0.00169	-0.00111
	(0.0144)	(0.0158)
Observations	14,071	10,228

Note: The at-risk sample is comprised of only those students who are predicted to be on academic probation based on the rules for each institution collected by the authors. Mid-run earnings are measured 7-12 years after initial enrollment in college, and is described further in the data section.

Table 3: RD Estimates of Educational Outcomes

Bandwidth	.25	.25	.25	.50	.50	.50	.75	.75	.75
RD Estimator	Standard	Standard	Donut	Standard	Standard	Donut	Standard	Standard	Donut
Variable									
Total credits earned	-4.355***	-4.137***	-4.504***	-4.854***	-5.332***	-5.680***	-7.163***	-8.090***	-8.466***
	(1.449)	(1.244)	(1.336)	(1.141)	(1.026)	(1.057)	(0.962)	(0.917)	(0.906)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	9163	9163	8711	14618	14618	14166	18045	18045	17593
R-Squared	0.282	0.328	0.331	0.291	0.328	0.329	0.295	0.327	0.328
BA from starting institution	-0.0989***	-0.107***	-0.103***	-0.108***	-0.117***	-0.117***	-0.134***	-0.142***	-0.144***
	(0.0183)	(0.0174)	(0.0196)	(0.0147)	(0.0135)	(0.0146)	(0.0133)	(0.0121)	(0.0126)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	9157	9157	8705	14606	14606	14154	18026	18026	17574
R-Squared	0.0926	0.123	0.124	0.109	0.136	0.138	0.115	0.143	0.144
Credits earned from other schools	0.596	0.517	0.493	0.816**	0.764**	0.846**	0.621*	0.710**	0.780**
	(0.487)	(0.451)	(0.503)	(0.377)	(0.363)	(0.387)	(0.345)	(0.337)	(0.351)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	9163	9163	8711	14618	14618	14166	18045	18045	17593
R-Squared	0.00500	0.0159	0.0167	0.00505	0.0153	0.0158	0.00489	0.0161	0.0165
BA earned from other schools	-0.000713	-0.000181	-0.00223	-0.000773	-0.000245	-0.00126	-0.000115	0.000707	0.000190
	(0.00449)	(0.00432)	(0.00472)	(0.00336)	(0.00327)	(0.00335)	(0.00292)	(0.00287)	(0.00297)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	9157	9157	8705	14606	14606	14154	18026	18026	17574
R-Squared	0.00246	0.0107	0.0116	0.00242	0.0106	0.0110	0.00219	0.0104	0.0107
AA earned from any Ohio public	-0.00741	-0.00322	-0.0107	-0.00982	-0.00641	-0.0107	-0.00697	-0.00353	-0.00670
	(0.00993)	(0.0104)	(0.00999)	(0.00810)	(0.00795)	(0.00774)	(0.00694)	(0.00675)	(0.00669)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	9163	9163	8711	14618	14618	14166	18045	18045	17593
R-Squared	0.00864	0.0330	0.0346	0.00744	0.0283	0.0292	0.00656	0.0268	0.0273

Note: The above table presents estimates of the discontinuity for a variety of educational outcomes around the critical GPA threshold. In addition to the controls noted in equation (2), all specifications include controls for: sex, race, age, first term GPA, employment status in ter t-1, cumulative credit hours earned and attempted in term t-1, institution fixed effects and year fixed effects. Standard errors are clustered around both the individual and running variable (GPA).

Table 4

Bandwidth	.25	.25	.25	.50	.50	.50	.75	.75	.75
RD Estimator	Standard	Standard	Donut	Standard	Standard	Donut	Standard	Standard	Donut
Variable									
Mid-run log weekly earnings	-0.0660**	-0.0445*	-0.0497*	-0.0563***	-0.0453**	-0.0485**	-0.0538***	-0.0506***	-0.0530***
	(0.0259)	(0.0237)	(0.0260)	(0.0203)	(0.0187)	(0.0200)	(0.0184)	(0.0174)	(0.0183)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	6448	6448	6132	10228	10228	9912	12533	12533	12217
R-Squared	0.00966	0.139	0.141	0.00951	0.133	0.134	0.0110	0.134	0.134
Missing earnings	-0.0117	-0.0238	-0.0187	-0.0165	-0.0265*	-0.0249*	-0.0168	-0.0252**	-0.0244*
	(0.0179)	(0.0176)	(0.0189)	(0.0143)	(0.0137)	(0.0145)	(0.0129)	(0.0120)	(0.0126)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	9163	9163	8711	14618	14618	14166	18045	18045	17593
R-Squared	0.00192	0.0826	0.0812	0.00291	0.0808	0.0799	0.00388	0.0824	0.0818

Note: The above table presents estimates of the discontinuity in mid-run (7-12 years after initial college enrollment) log weekly earnings and missing earnings around the critical GPA threshold. In addition to the controls noted in equation (2), all specifications include controls for: sex, race, age, first term GPA, employment status in ter t-1, cumulative credit hours earned and attempted in term t-1, age at which earnings are measured, institution fixed effects and year fixed effects.   Standard errors are clustered around both the individual and running variable (GPA).

Table 5: IV Estimates of the returns to a year of college education

Variable			
Credits/30	0.311*	0.246**	0.165***
	(0.173)	(0.108)	(0.0591)
Bandwidth	.25	.50	.75
F-statistic	10.56	24.83	77.77
Observations	6252	10150	12533

Note: The above table presents instrumental variables estimates of the returns to a year of education on future mid-run earnings. The discontinuity in college credits earned generated from institutional dismissal policies serves as the instrument in the above specifications. In addition to the controls noted in equation (2), all specifications include controls for: sex, race, age, first term GPA, employment status in term t-1, cumulative credit hours earned and attempted in term t-1, institution fixed effects and year fixed effects. Standard errors are clustered around both the individual and running variable (GPA).

Table 6 Heterogeneity

Panel A

Credits/30	-5.680***	-6.759***	-4.935***	-6.433***	-4.328***	-7.617***	-5.201***	-6.736***	-6.877***	-4.105***
	(1.057)	(1.607)	(1.341)	(1.431)	(1.427)	(2.274)	(1.184)	(1.610)	(1.652)	(1.434)
Sample	All	Women	Men	Selective	Nonselective	Black/Hispanic	Non-Black/Hispanic	2nd year	3rd year	4th+
N	14166	5814	8325	8493	5673	2709	11457	8139	3681	2346
R2	0.329	0.338	0.329	0.322	0.355	0.355	0.325	0.167	0.157	0.263

Panel B

BA Degree	-0.117***	-0.126***	-0.113***	-0.0985***	-0.135***	-0.117***	-0.117***	-0.0762***	-0.114***	-0.210***
	(0.0146)	(0.0234)	(0.0191)	(0.0181)	(0.0226)	(0.0336)	(0.0163)	(0.0189)	(0.0292)	(0.0361)
Sample	All	Women	Men	Selective	Nonselective	Black/Hispanic	Non-Black/Hispanic	2nd year	3rd year	4th+
N	14154	5809	8318	8489	5665	2707	11447	8139	3680	2335
R2	0.138	0.148	0.135	0.131	0.163	0.145	0.138	0.100	0.100	0.127

Panel C

Earnings	0.247**	0.339**	0.223	0.235**	0.289	0.264	0.207	0.115	0.357*	0.524
	(0.109)	(0.164)	(0.153)	(0.120)	(0.256)	(0.192)	(0.127)	(0.107)	(0.214)	(0.356)
Sample	All	Women	Men	Selective	Nonselective	Black/Hispanic	Non-Black/Hispanic	2nd year	3rd year	4th+
F-Stat	24.91	11.22	12.81	19.44	4.9	6.99	17.6	18.35	8.72	5.91
N	9912	4291	5621	5915	3997	1853	8059	5778	2591	1543

Note: The above table presents RD estimates separately by various demographic and educational factors. In addition to the controls noted in equation (2), all specifications include controls for: sex, race, age, first term GPA, employment status in term t-1, cumulative credit hours earned and attempted in term t-1, institution fixed effects and year fixed effects. Standard errors are clustered around both the individual and running variable.

Table 7: Separating credits earned from BA effect

Bandwidth	.25	.25	.25	.50	.50	.50	.75	.75	.75
Dependent Variable	Log(earnings)	BA	Credits/30	Log(earnings)	BA	Credits/30	Log(earnings)	BA	Credits/30
Variable									
BA	0.345			0.363			0.347		
	(0.541)			(0.325)			(0.259)		
Credits divided by 30	0.0587			0.0139			0.00412		
	(0.320)			(0.181)			(0.118)		
Below Cutoff		-0.0635	-0.367**		-0.0134	-0.393***		-0.0516	-0.594***
		(0.0764)	(0.172)		(0.0569)	(0.128)		(0.0492)	(0.111)
Standing*Below Cutoff		-0.0178	0.0659		-0.0421**	0.0593		-0.0374**	0.0997***
		(0.0279)	(0.0528)		(0.0210)	(0.0399)		(0.0183)	(0.0346)
Observations	6129	6129	6132	9908	9908	9912	12209	12209	12217
R-Squared	0.140	0.120	0.297	0.145	0.132	0.300	0.149	0.140	0.302
F-Statistic		11.37	6.4		26.41	17.57		50.09	44.56
Specification		First	First		First	First		First	First
	IV	Stage	Stage	IV	Stage	Stage	IV	Stage	Stage

Note: The above table presents IV estimates of both the return to a year of college and college degree completion. Separate instruments for each educational measure are generated from the fact that students are on academic probation at different points in time, meaning that the interaction between standing and the discontinuity effectively creates multiple experiments. In addition to the controls noted in equation (2), all specifications include controls for: sex, race, age, first term GPA, employment status in term t-1, cumulative credit hours earned and attempted in term t-1, institution fixed effects and year fixed effects. Standard errors are clustered around both the individual and running variable.

Appendix A: Academic Probation Policies

Below we describe the dismissal policies in place in 2011-2012 for each Ohio public 4-year school. We code policies at the institution level which likely introduces some measurement error since standards can differ across departments within an institution. For each school, we spoke to registrar officials to determine the institutions policies and asked about historical changes. While we were not told about any historical policy changes, it is possible that some institutions changed their policy over time in ways that we are not aware of.[16] To the extent that we miscode some policies, this weakens the strength of our first stage, but it will not bias estimates.

Akron University

Students with a cumulative GPA below 2.0 are placed on probation. Students are dismissed after two terms of probation. Dismissed students are not eligible to register for credit work, but they can enroll in noncredit work. Readmission is possible only with specific approval from the Dean. For students that are readmitted, the length of time they must wait before readmission is case specific. Students may appeal both the initial probation placement and the dismissal decision.

Bowling Green State University

The academic standards that a student must achieve depend on the number of credits that student has earned. Students with 1st, 2nd, 3rd and 4th year standing face cumulative GPA cutoffs of 1.5, 1.6, 1.8 and 1.9 respectively. Students are placed on probation if their cumulative GPA falls below the cutoff and are dismissed if their cumulative GPA falls below the cutoff after having been on probation. There are also academic warnings based on higher cutoffs, but these warnings carry no actual consequences. Dismissed students cannot enroll in any courses at BGSU for 1 academic year. Students may appeal both the initial probation placement and the dismissal decision.

University of Cincinnati

Students are placed on probation only if they have attempted at least 30 credit hours and have a cumulative GPA below 2.0. Once on probation, students must maintain a term

---

ailure to turn in assignments.

[16] In addition to discussing the possibility of policy changes with registrar personnel, we searched through historical student handbooks and used the data to empirically search for changes in the size of the enrollment discontinuity over time for each school. The empirical investigation is based on the idea that if a school changed its policy, the enrollment discontinuity should appear only after that policy change. We did not find any indication of changing dismissal policies empirically, which corroborates our findings based on historical student handbooks and conversations with the registrars.

GPA above 2.0 or they are dismissed. Since students are dismissed based on term GPA, we use term GPA as the running variable for students at University of Cincinnati as well as other schools that use the same sort of policy (e.g. Central State University). The academic dismissal is for one academic year and prevents students from enrolling in any classes at the University of Cincinnati. According to the registrar website, "Readmission after the mandatory suspension period is not automatic. Dismissed students must petition for readmission after serving the suspension period. The dismissed student must submit an application for readmission supported by a letter documenting the reasons for the previous academic difficulties". Students may appeal the dismissal decision but they may not appeal the initial probation placement.

Cleveland State University

Students are placed on probation if their cumulative GPA falls below 2.0 and they are not in their first term. Students are also placed on probation if they fail to complete 67% of their courses, but we do not have access to exact course grades so we only classify probationary status based on the GPA cutoff. This misclassification mostly affects students who withdraw from several courses since their GPA may be well above 2.0 but they can still be placed on probation. Students are dismissed if they do not raise their cumulative GPA above a 2.0 in the term following their probation, though there is an appeal process. Dismissed students cannot enroll for 2 consecutive terms after the dismissal. According to the registrar website, "readmission is a lot like applying to college, though in this case information is sought about what the student has done to make him- or herself academically successful upon return.". The registrar also states that while it is possible to gain admission to another university following dismissal, "Many universities are becoming less willing to accept students who have been suspended from another university, even if they are only applying to be part-time students."

Central State University

The academic standards that a student must achieve depend on the number of credits that student has earned. Students with between 1-20 credits earned must achieve a 1.7. Students with between 20-40 credits earned must achieve a 1.8. Students with between 40-60 credits earned must achieve a 1.9. Students with over 60 credits earned must achieve a 2.0.   Students are placed on probation if their cumulative GPA falls below the cutoff and are dismissed if they do not maintain at least 2.0 term GPA while on probation. Dismissed students must leave for at least 1 term following the dismissal. After the term away, students must reapply to the university in order to gain readmission.

Kent State University

The academic standards that a student must achieve depend on the number of credits that student has earned. Students with between 16-30 credits earned must achieve a 1.5. Students with between 30-60 credits earned must achieve a 1.7. Students with between 60-90 credits earned must achieve a 1.8. Students with over 90 credits earned must

achieve a 1.9. Dismissed students must leave for at least 12 months following the dismissal. After the time away, students must reapply to the university in order to gain readmission. According to the registrar "Reinstatement after dismissal from Kent State University is neither automatic nor guaranteed. A student may be reinstated only if the student provides convincing evidence of probable academic success if permitted to return to the university. A dismissed student who has previously accumulated a substantial number of credit hours and/or an excessively low GPA should expect that reinstatement is not likely to be approved." The registrar also specifically stated that the provost has full discretion with dismissals and students can appeal.

Miami University

Students are placed on probation only if they have attempted at least 16 credit hours and have a cumulative GPA below 2.0. Students with fewer than 16 credits attempted are given an academic warning which has no specific consequences. Once on probation, students must maintain a term GPA above 2.0 or they are dismissed. Students are only dismissed if they have already attempted 30 credit hours. The academic dismissal is for 2 consecutive semesters and prevents students from enrolling in any classes at the Miami University. While students are guaranteed re-entry following the dismissal, students are not permitted to transfer any credits earned from other institutions during the dismissal period. Students may appeal the dismissal decision.

Ohio State University

Relative to the other public schools in Ohio, Ohio State has a much more flexible dismissal policy that does not emphasize sharp GPA thresholds as heavily. Students with a cumulative GPA below 2.0 are generally placed onto academic probation. To avoid dismissal, students must perform adequately in the preceding term where the definition of adequate is determined based on individual student review by academic advisors and an evaluation board. While we do not have data on the exact manner through which every academic advisor judges adequate performance in the previous term, the administrator we spoke with stated that dismissal is unlikely for a student that is very close to the cumulative GPA threshold or for students who earn above a 2.0 in the most recent term. As such, we use term GPA and a threshold of 2.0 following our approach from Cincinnati University. Given the flexibility of Ohio State's policies, we also explored simply excluding Ohio State from the analysis and the results are robust to this restriction. Dismissed students can apply for readmission after a case-specific waiting period, but readmission is not guaranteed.

Ohio University

Students are placed on probation if their cumulative GPA falls below a 2.0. So long as they continue to make progress towards their degree, students can stay on academic probation for a maximum of 4 terms. At the end of the 4th term on probation, students are dismissed and there is no appeal process. After dismissal, students must leave for at least 12 months and then can reapply. Readmission is not guaranteed and in some cases,

students must wait for longer than 12 months to be allowed to reapply. In 2010-2011 there were major changes to the probation-dismissal policies at Ohio University but none of our sample is impacted by these changes.

Shawnee State University

The academic standards that a student must achieve depend on the number of credits that student has attempted. Students with between 0-28 credits attempted must achieve a 1.0. Students with between 28-37 credits attempted must achieve a 1.3. Students with between 37-44 credits attempted must achieve a 1.55. Students with between 44-51 credits attempted must achieve a 1.8. Students with between 51-58 credits attempted must achieve a 1.9. Students with over 58 credits attempted must achieve a 2.0. Students are placed on probation if their cumulative GPA falls below the cutoff and are dismissed if they do not maintain at least 2.0 GPA while on probation. Students must leave for at least 12 months following the dismissal. After the time away, students can appeal their dismissal, though lost financial aid cannot be reinstated through appeal.

University of Toledo

The academic standards that a student must achieve depend on the number of credits that student has attempted. Students with between 10-20 credits attempted must achieve a 1.0. Students with between 20-30 credits attempted must achieve a 1.5. Students with between 30-40 credits attempted must achieve a 1.7. Students with between 40-50 credits attempted must achieve a 1.8. Students with between 50-60 credits attempted must achieve a 1.9. Students with over 60 credits attempted must achieve a 2.0.   Students are placed on probation if their cumulative GPA falls below the cutoff and are dismissed if they do not raise their cumulative GPA to above the threshold while on probation. Students must leave for at least 1 semester following the dismissal. During the suspension period, students can work to remove incompletes from their transcript but course work taken at other educational institutions during the suspension periods is not accepted as transfer credit. As such, students are effectively barred from any for credit educational investment if they want to return to the University of Toledo.

Wright State University

Students are placed on probation only if they have attempted at least 12 credit hours and have a cumulative GPA below 2.0. Students are dismissed if their cumulative GPA drops below the 2.0 while on probation. Dismissal is at the discretion of the chief academic officer and these officers will consider the student's progress towards meeting degree requirements. After dismissal, students are barred from applying for readmission for 1 full year. After the waiting period, students must apply for readmission and this those that are readmitted may be subject to special requirements. Students that are granted readmission may also petition using the "Fresh Start Rule" to have some low grades earned prior to the dismissal stricken from their cumulative GPA calculation.

Youngstown State University

The academic standards that a student must achieve depend on the number of credits that student has attempted. Students with between 0-32 credits earned must achieve a 1.75. Students with above 30 credits earned must achieve a 2.0. Students are given an academic warning the first time their cumulative GPA falls below the threshold. Students are placed on probation if their cumulative GPA falls below the cutoff a second time and are dismissed if they do not raise their cumulative GPA to above the threshold while on probation. Students must leave for at least 1 semester following the dismissal. Students must petition the dean to be reinstated and reinstatement following the dismissal is not guaranteed. Students that are granted readmission may be subject to certain terms and conditions.

Appendix B

  This appendix discusses several potential reasons why the slope might change at the discontinuity. As discussed below, the slope change could be policy driven and if this were the only reason for the slope change, the regression-kink design would be appropriate; however, there are several other forces leading to a change in slope, most of which are expected to be correlated with observable and unobservable student characteristics.

  First, because we focus on students whose t-1 cumulative GPA is below the cutoff, this generates a particular form of sample selection that could explain the changing slope. In particular, upper-year students in our sample are unlikely to have a t-1 cumulative GPA very far below the threshold since upper-year students with very low GPAs are likely to dropout prior to enrolling in the at-risk term since they know that they are unlikely to be able to raise their GPA in a single term. In addition to being less likely to be far below the threshold, upper year students are also less likely to have a term t GPA very far above the threshold. This is because by construction all students have t-1 GPAs below the cutoff and it is difficult for upper-year students to raise their cumulative GPA substantially. As we show in the specification checks section, there is an inverse V shaped relationship between the running variable and past credits earned with the apex at the cutoff. While this V shaped relationship is continuous at the cutoff, it makes it difficult to use the regression kink design in practice since the change in slope is partly driven by changes in the slope of observable characteristics at the threshold.

  In addition to a changing slope caused by academic standing, our sample selection generates a non-linear relationship between the running variable and t-1 cumulative GPA and the curvature of this relationship is mechanically strongest near the cutoff. The intuition for this non-linear relationship is that the probability of being included in the sample decreases as cumulative GPA in term t increases. To illustrate how exactly the sample selection generates the nonlinearity, we use a simple simulation where for each student we generate two randomly assigned term GPAs (centered around zero) and then use these artificial GPAs to calculate cumulative GPAs for each of the two terms. We then construct an outcome variable, Y, that is simply equal to the cumulative GPA in term 2 plus some noise. Appendix Figure B1 plots Y vs second term cumulative GPA and shows the linear relationship with a slope of one. Appendix Figure B2 creates the same figure, but first drops observations in which the first-term GPA is below zero. Figure B2 shows that this sample restriction leads to a non-linear relationship between cumulative GPA and the outcome variable. The relationship is approximately linear well below the threshold since students with very low cumulative GPAs in term 2 are unlikely to be affected by the sample restriction. Well above the threshold, the relationship becomes fairly flat because the sample selection nearly perfectly balanced out the underlying relationship between the outcome and cumulative GPA in term 2. While there is a changing slope near the cutoff, there is visually no discontinuous jump created by the sample selection and the change in slope itself is smooth. While the smoothly changing slope near the cutoff does not theoretically bias the RK estimates, it makes it difficult to

precisely identify point estimates since it is necessary to dramatically reduce the bandwidth to eliminate the mechanical curvature. In practice, we do not estimate RK models because we lack sufficient data density right at the cutoff.

While all of the above points complicate an RK analysis, the primary RD analysis is mostly unaffected by these concerns because all of these concerns still result in continuous potential outcomes through the cutoff. The one complication is that curvature in the relationship between the running variable and various outcomes can lead to biased estimates to the extent that the linear extrapolation on either side of the cutoff does not accurately fit the data. As such, we carefully examine the figures relating outcomes to the running variable to ensure that our RD estimates appear to capture true discontinuities and not simply a changing slope. Furthermore, we have confirmed that our estimates are robust to fitting quadratic or cubic trends on either side of the threshold.

Appendix Figure 1
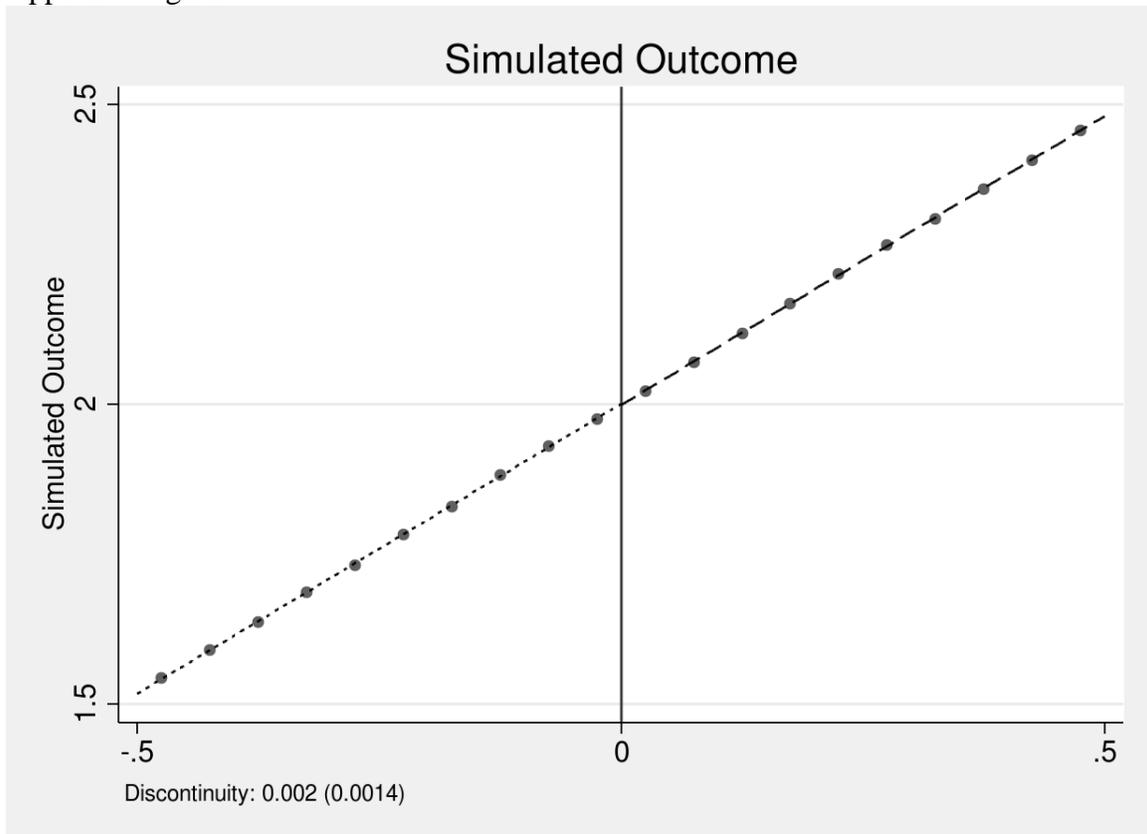
Note: As described in Appendix B, the two simulations in Appendix Figures 1 and 2 are meant to illustrate how we might expect the slope of various outcomes measured above and below the GPA threshold to differ for mechanical reasons.

Appendix Figure 2
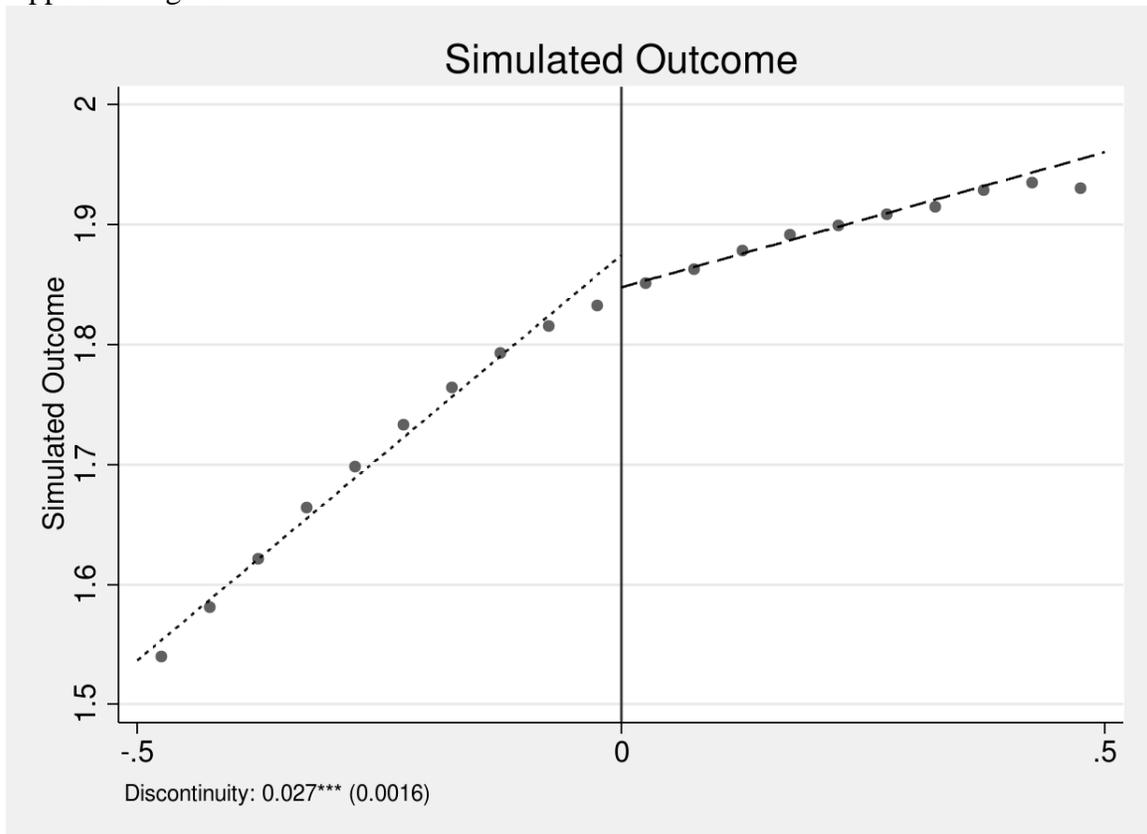
Note: As described in Appendix B, the two simulations in Appendix Figures 1 and 2 are meant to illustrate how we might expect the slope of various outcomes measured above and below the GPA threshold to differ for mechanical reasons.

Appendix C

In this appendix we provide a brief discussion of the monotonicity assumption necessary for an IV approach to yield unbiased estimates. While this assumption is often overlooked in many studies, it is not innocuous in our setting (as is discussed in the main text).

Angrist, et al. (1996, p.451), derives the bias from a violation of the monotonicity assumption to be equal to

$$\frac{P_i(defier)}{P_i(complier) - P_i(defier)}[E(Y_i(1) - Y_i(0)|defier) - E(Y_i(1) - Y_i(0)|complier)]$$

Since the direction of the bias depends on the relative returns to schooling for defiers vs compliers, it is useful to consider from a theoretical perspective which group is likely to derive larger returns from educational investments. In our context, the compliers are individuals who would have completed more years of schooling but did not as a result of the dismissal. Compliers intended to complete additional postsecondary schooling, and thus they expected to obtain substantial benefits from that schooling. That said, being dismissed for one year was sufficient to deter compliers from completing their intended amount of education. As such, the group of compliers likely has larger returns than those who drop out voluntarily but perhaps smaller returns than the typical graduate. The defiers in our context are students who would normally not complete many years of school, but end up persisting because of the dismissal. As such, we suspect that the returns to education for these students cannot be too large (since without the dismissal they would, by definition, not invest heavily in their education) but the returns are large enough so that they eventually do complete more years of schooling. Thus, from an optimal human capital investment point of view, one might expect that the defiers and

compliers are fairly similar in terms of their estimated return to schooling. Naturally, the preceding discussion is a bit ad-hoc and we are aware of no empirical evidence that would shed light on the relative rates of return to these two types of students.

The bias equation above is likely to be small when 1) the proportion of defiers in the population is small, and/or 2) when there is a large first stage effect. The first condition seems intuitively likely, since it is difficult to imagine that there are a large number of individuals who choose to obtain more education because they were dismissed from college previously, although this is not directly testable. The second condition, a large first-stage impact, effectively means the compliers outnumber the defiers to such a degree that the estimated coefficient is identified primarily, if not entirely, from the compliers. This condition is given strong support from the results presented in Table 3.

In order to provide further evidence that the monotonicity assumption is not violated in our setting, we also perform a test suggested by Angrist and Imbens (1995). We examine the cdf of completed credit hours for treated students and find that it first-order stochastically dominates the corresponding cdf for nontreated students. Angrist and Imbens (1995) first presented this as a test for the monotonicity assumption in the case where the endogenous variable (credit hours in our case) is multi-valued. Fiorini et al (2013) later argues that first order stochastic dominance is a necessary, rather than a sufficient condition for monotonicity, but is still strong suggestive evidence that the monotonicity assumption holds.