Exploring Race and Income Heterogeneity in the Effects of State Merit Aid Loss Among Four-Year College Entrants

Authors:

Kristen M. Cummings

KC Deane

Brian P. McCall

Stephen L. DesJardins

The Journal of Higher Education. 93:6, 873-900.

March 2022

To cite this article:

Kristen M. Cummings, KC Deane, Brian P. McCall & Stephen L. DesJardins (2022) Exploring Race and Income Heterogeneity in the Effects of State Merit Aid Loss Among Four-Year College Entrants, The Journal of Higher Education, 93:6, 873-900, DOI: 10.1080/00221546.2022.2042155

To link to this article:

https://doi.org/10.1080/00221546.2022.2042155

Acknowledgments: We thank the Tennessee Higher Education Commission (THEC) for providing access to data used in this study. We also thank Dominique Baker, Rajashri Chakrabarti, Laura Perna, and members of the DesJardins-McCall research team for helpful comments, as well as attendees of the AEFP 2020 conference, ASHE 2020 conference, CIERS seminar at the University of Michigan, and the SHEEO Public Investment in Higher Education webinar series. Cummings and Deane gratefully acknowledge the funding support of the Institute of Education Sciences (IES), U.S. Department of Education under Grants R305B150011, R305B200011, and R305B150012. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of THEC or IES. Any errors are our own.

The Journal of Higher Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uhej20

Exploring Race and Income Heterogeneity in the Effects of State Merit Aid Loss Among Four-Year College Entrants

Kristen M. Cummings, KC Deane, Brian P. McCall & Stephen L. DesJardins

To cite this article: Kristen M. Cummings, KC Deane, Brian P. McCall & Stephen L. DesJardins (2022) Exploring Race and Income Heterogeneity in the Effects of State Merit Aid Loss Among Four-Year College Entrants, The Journal of Higher Education, 93:6, 873-900, DOI: 10.1080/00221546.2022.2042155

To link to this article: https://doi.org/10.1080/00221546.2022.2042155

Published online: 15 Mar 2022.

THE JOURNAL OF HIGHER EDUCATION
2022, VOL. 93, NO. 6, 873–900
https://doi.org/10.1080/00221546.2022.2042155

Routledge
Taylor & Francis Group

Exploring Race and Income Heterogeneity in the Effects of State Merit Aid Loss Among Four-Year College Entrants

Kristen M. Cummings, KC Deane, Brian P. McCall, and Stephen L. DesJardins

Center for the Study of Higher and Postsecondary Education, University of Michigan, Ann Arbor, Michigan, USA

ABSTRACT

Despite the robust literature on the effects of financial aid, the effects of financial aid loss remain largely understudied. We employ a regression discontinuity design, leveraging a minimum GPA scholarship renewal threshold, to examine the effect of losing state merit aid eligibility on college student stop-out, transfer, and bachelor's degree completion. We estimate the effects of GPA-based eligibility loss for low- and higher-income students, Black and White students, and the interactions of the two separately. Using longitudinal state administrative data from academic years 2011 to 2014 on four cohorts of students attending public four-year colleges, we find evidence that the effects of eligibility loss at the first renewal checkpoint differ among the subgroups of interest. Losing eligibility for a state merit aid scholarship at the first renewal checkpoint leads to increased stop-out among higher-income White students and increased transfer to a community college among low-income Black students. State merit aid loss has no statistically significant effect on on-time bachelor's degree completion but causes a decrease in the probability of 150% time bachelor's degree completion for Black students. We close with implications for policy and research.

ARTICLE HISTORY
Received 5 February 2021
Accepted 11 February 2022

KEYWORDS
Higher education; financial aid; state policy; education equity; state merit aid programs; regression discontinuity; heterogeneous treatment effects

Introduction

Financial aid is a tool institutions and states can leverage to advance college access and completion goals. The merit aid program at the center of this study, the Tennessee HOPE scholarship (TN HOPE), provides scholarships to more than 30,000 first-time freshmen annually. Students can receive TN HOPE for five years, provided they satisfy a GPA requirement when they reach each of four attempted-credit based checkpoints. However, nearly half (42%) of recipients lose their scholarship before their second year of college (Tennessee, Higher Education Commission & Tennessee, Student Assistance Corporation (THEC & TSAC), 2017), with losses disproportionate among Black students, regardless of income, and low-income White students. This paper diagnoses the extent to which merit aid eligibility loss, due to not meeting established GPA requirements, differentially affects students across racial and income

CONTACT Kristen M. Cummings kristenz@umich.edu

subgroups. The estimation of heterogeneous treatment effects, alongside the more commonly-evaluated average treatment effect, builds on the modest literature that evaluates overall effects of merit aid loss.

Critics of statewide merit aid programs note that the scholarships disproportionately provide money to White middle- and high-income students who are likely to attend college regardless of the aid, rather than lower-income students for whom additional aid might induce college enrollment (Heller & Rasmussen, 2002; Jones, 2014). More recent studies indicate that merit aid affects *where* students enroll rather than *whether* students enroll in higher education, with low-income students driving these substitution effects (Bruce & Carruthers, 2014; Cohodes & Goodman, 2014). These patterns of disproportionate receipt across student populations are likewise present in Tennessee. Relative to the population of public high school graduates in the state, Black and low-income students are underrepresented as initial TN HOPE recipients[1] and overrepresented among those who lose eligibility at the first checkpoint (Table 1). The disparity in initial scholarship receipt reflects a host of systemic inequities in the K-12 system—including school funding (EdBuild, 2019) and participation in rigorous coursework (Rodriguez & McGuire, 2019)—that likely contribute to fewer students meeting the academic eligibility criteria for a merit scholarship. We investigate the disparity in scholarship loss by observing the effects of such loss separately for low- and higher-income students, Black and White students, and the interactions of the two.

Table 1. Characteristics of four-year students who reach first checkpoint with HOPE.

	Reach first checkpoint with HOPE	Lose at first checkpoint	Keep at first checkpoint
Number	43,786	16,235	27,551
Female	55%	49%	59%
Race/Ethnicity			
White	72%	64%	76%
Black	17%	24%	12%
Latinx	3%	3%	3%
Asian or Pacific Islander	2%	2%	3%
Multiracial	3%	4%	3%
Unknown	3%	3%	3%
Pell Eligible	43%	51%	38%
From Distressed or At-Risk County	14%	15%	14%
From Rural County	22%	23%	21%
Highest Parental Education			
Less than HS	1%	1%	1%
HS Diploma	26%	31%	24%
College	70%	65%	73%
Other/Unknown	3%	3%	2%
ASPIRE recipient	30%	38%	26%
GAMS recipient	7%	1%	10%

Due to rounding, totals may not equal 100%. Racial categories totaling <1% of the student population (American Indian and Alaskan Native) are excluded from table.
Source: Authors' analyses of THEC datasets

We employ a regression discontinuity design to estimate whether, and if so how, the effects of TN HOPE eligibility loss at the first checkpoint due to insufficient GPA (hereafter eligibility loss) differ for the subgroups of interest. Our study reinforces and builds upon prior work investigating the effects of TN HOPE loss (Carruthers & Özek, 2016) by leveraging changes in data collection protocols to estimate the effect of HOPE loss on additional outcomes and by racial/ethnic and income-based subgroups We examine overall and heterogeneous effects for three student-level outcomes: stop-out, cross-sector transfer, and bachelor's degree completion (100% time and 150% time). Because the effects of loss may differ across sectors, the analyses focus on examining the effects of scholarship loss for students enrolled in four-year institutions at the first checkpoint. We find evidence that losing eligibility for TN HOPE increases stop-out among higher-income White students and causes an increase in cross-sector transfer among low-income Black students. Eligibility loss has no statistically significant effect on the probability of on-time bachelor's degree completion but causes a large decrease in the chances of 150% time bachelor's degree completion for Black students.

Background on the Tennessee HOPE program

Modeled heavily after a program in neighboring Georgia, Tennessee launched its statewide merit aid program in 2004. The Tennessee HOPE scholarship is awarded to graduates of Tennessee high schools who enroll in a two- or four-year Tennessee public or private postsecondary institution within sixteen months of high school graduation. To earn TN HOPE, individuals must file the Free Application for Federal Student Aid (FAFSA) and score a 21 or higher on their ACT Composite test (or 1060+ on the SAT) or achieve at least a 3.00 cumulative high school GPA (HSGPA). In addition to this base HOPE scholarship, students can receive one of two additional supplemental awards if they meet specific eligibility criteria: the General Assembly Merit Scholarship (GAMS) for students who earned a HSGPA of at least 3.75 and an ACT score of at least 29 (or SAT score of 1330+), or the Aspire Award (Aspire) for scholarship recipients with an adjusted gross income (AGI) of $36,000 or less. If a student qualifies for both supplements, they receive the more generous Aspire award. During the years covered by our study, students attending four-year colleges could receive up to $2,000 per semester (including summer) in base HOPE, which covered approximately 48% of tuition and fees at four-year public institutions in the 2011–12 academic year. Students eligible for the GAMS or Aspire supplement could receive an additional $500 and $750 per semester, respectively. HOPE is a guaranteed award up to the cost of attendance (COA), meaning that students are eligible to receive refund checks if their HOPE award exceeds tuition and fees but remains below the COA. Students who are enrolled less than full-time receive a prorated award

amount. Among students in our sample, the average HOPE payment during the first checkpoint semester is $2,252 (sd = $350), with less than 1% of recipients receiving less than the maximum award amount.

To maintain eligibility, recipients must meet a minimum cumulative collegiate GPA renewal threshold in the semesters when they reach each of four attempted-credit based checkpoints (24, 48, 72, and 96 credits). At the first checkpoint, the focus of this study, cumulative GPAs below 2.75 trigger scholarship ineligibility. Additionally, scholarship recipients must maintain continuous enrollment at a TN HOPE-eligible postsecondary institution, complete the FAFSA annually and maintain consistent enrollment intensity within each semester (e.g., a student who drops from full-time to half-time enrollment within a semester will lose scholarship eligibility in subsequent semesters). A student who loses eligibility for any of these reasons loses access to both base HOPE and supplemental awards (i.e., GAMS and Aspire). Barring ineligibility, a scholarship recipient can receive the award until they earn a bachelor's degree; attempt 120 credit hours or receive the scholarship for eight full-time equivalent (FTE) semesters (whichever occurs later); or until five years have passed from their date of initial postsecondary enrollment.

Prior literature and conceptual framework

Although a robust literature examines how receiving state-based merit aid influences student behavior and outcomes (among others, see, Dynarski, 2004; Cornwell et al., 2006), fewer scholars have studied the effect of losing merit-based state financial aid after initial college enrollment (Carruthers & Özek, 2016; Henry et al., 2004). Henry et al. (2004) compare the outcomes of Georgia HOPE recipients in the state's public universities who marginally qualified for the scholarship and subsequently lost it with students who never received the award. They find evidence that students who received then lost merit aid accumulated more credits over four years and earned higher cumulative GPAs than their non-recipient peers. The comparison to non-recipients illuminates the potential long-term effects of even short-lived scholarship receipt, but it is less relevant than a direct comparison to recipients who maintained eligibility. Further, the authors' comparison group, constructed by matching students on high school GPA, differed from the treatment group across key characteristics (SAT score, developmental education enrollment, full-time enrollment). This covariate imbalance increases the possibility that significant effects were driven in part by differences between the treatment and control groups that extend beyond HOPE receipt. An improvement to this methodological limitation, Carruthers and Özek (2016) employ a regression discontinuity (RD) design to compare outcomes between TN HOPE recipients who do and do not maintain scholarship eligibility. Their results suggest that four-year students

who lose the HOPE scholarship substitute some of their lost scholarship dollars with work, evidenced by small increases in the likelihood of at least half-time employment and in biannual earnings in the semester following loss. The authors find no statistically significant effect of scholarship loss on persistence, credit accumulation, or on-time bachelor's degree completion among four-year scholarship recipients.

Our study engages with and furthers Carruthers and Özek's (2016) study in two ways. First, we re-estimate local average treatment effects using more recent data (post-2010) that benefits from a change in THEC's reporting requirements. Whereas Carruthers and Özek could not estimate next-semester outcomes because cumulative GPA for a given semester was recorded in the subsequent semester - meaning students who stopped out in the semester following the first checkpoint are not included in their sample - post-2010 data requires no such conditioning. The analytic consequence of this reporting change is that we can examine the effects of HOPE loss on immediate behavioral outcomes, such as next-semester stop-out and cross-sector transfer.

Second, additional cohorts available in our data provide increased statistical power relative to this earlier work and render the evaluation ofeffect variations for racial and income-based subgroups possible. Carruthersand Özek's study focuses primarily on the average effects of HOPE loss across the full sample ofscholarship recipients and, although they do report subgroup analyses brokenout by institutional sector (two-/four-year enrollees), family income(above/below $60k), and ACT score (above/below 22), the latter two subgroupspool together two- and four-year enrollees. Their sector-level findings suggestthat losing HOPE has differential effects on students enrolled at two- andfour-year colleges, suggestive of the need to estimate any sub-group effects separatelyby sector as well. We focus here on estimating these sub-group effects for-four-year students by race/ethnicity and, like Carruthers and Özek, family-income. By expanding upon the investigation begun in Carruthers and Özek'sstudy, we contribute to the field's continued understanding of the differentialimpacts of scholarship loss.

We turn next to two student-level mechanisms through which we posit scholarship loss may influence stop-out, transfer, and degree completion for four-year students. This conceptual discussion preserves the financial mechanism emphasized in the existing literature (Carruthers & Özek, 2016; Dynarski, 2004) and incorporates an additional mechanism related to self-efficacy. Whether and how these mechanisms manifest likely differs depending on individual characteristics that influence one's experience of the academic environment—specifically, race and/or family income. Existing research on both mechanisms suggests that Black students and low-income students,

compared to White students and higher-income students, are more likely to respond to scholarship loss by stopping out or transferring to a two-year institution and by graduating at lower rates.

Mechanism 1: Shifts in financial costs

The first mechanism through which loss of scholarship eligibility affects outcomes is a shift in the financial cost of college. This mechanism correlates with the notion of price sensitivity, which measures the sensitivity of consumer demand to changes in price (DesJardins & Bell, 2006). Not all consumers react identically to changes in price, and not all students have the same sensitivity to increases in the net price of attendance. In response to the sudden increase in the amount of money required to remain enrolled, a price-sensitive student who loses scholarship eligibility may make enrollment decisions (e.g., stop out or transfer) that stabilize short run costs but negatively affect long-term outcomes (e.g., degree attainment).

Variation in price sensitivity across the income distribution is intuitive: Low-income students are more responsive to changes in price than their high-income peers (Kim, 2012); according to one study, the effect of aid on matriculation is three times larger for low- versus higher-income students (Hurwitz, 2012). Fully subsidized ("free") tuition at two-year colleges induces low-income students to substitute away from the four-year into the two-year sector (Carruthers & Fox, 2016), whereas state merit aid programs tend to induce enrollment substitution in the opposite direction (Bruce & Carruthers, 2014). When considered by race, White students are relatively insensitive to changes in price compared to Black students (Allen & Wolniak, 2019; Kim, 2012). These differences between Black and White students are present in studies of changes in the availability of state need- and merit-based aid (Kim, 2012), receipt of the Pell Grant (Chen & DesJardins, 2010), and increases in tuition prices (Allen & Wolniak, 2019).

An individual's level of price sensitivity stems from a host of factors besides income. Variation in individuals' expectations about the return on investment (ROI) to earning a college degree (Goldrick-Rab et al., 2009) or differential preferences for work in place of schooling (Carruthers & Özek, 2016) may also impact price sensitivity across individuals. Furthermore, the degree to which expected ROI and work preferences influence an individual's assessment of what is "affordable" may reflect differences in how students perceive and experience the benefits of higher education. A 2020 survey of Americans' views on higher education found that Black respondents were less confident than White respondents in the payoff from attending college and more confident in the job opportunities available to someone who foregoes college (Fishman et al., 2020).That Black respondents display reservations about the payoff of higher education is understandable, given that

Black students are disproportionately impacted by student loan debt (Houle & Addo, 2019), attend institutions with lower graduation rates (Monarrez & Washington, 2020), and the Black-White wage gap is present at all education levels and increasing in many of them (Gould, 2020). These findings suggest that Black students may respond more strongly than their White peers to the change in price resulting from scholarship loss.

Mechanism 2: Shifts in self-efficacy

The second mechanism through which loss may affect the focal outcomes is a shift in an individual's academic self-efficacy. The concept of self-efficacy is embedded in Bandura's (1997) Social Cognitive Theory, which posits that an individual's motivation is influenced by their traits, their beliefs about their expectations and capabilities (e.g., self-efficacy), and the environment in which they are embedded. An individual's self-efficacy is derived from four primary sources: academic success over time, individual comparisons to peers or role models, verbal and/or social affirmations of an individual's capabilities, and emotional and physiological states (Usher & Pajares, 2008). Self-efficacy and academic performance share a documented positive relationship (Robbins et al., 2004).

Considered through the lens of self-efficacy and its sources, receipt of a merit-based scholarship may serve as a symbol of academic ability or signal students' long-term academic success, rather than be simply a form of financial assistance. Each time a student succeeds (or fails) at "an academic task, they interpret and evaluate the results obtained" and revise their beliefs accordingly (Usher & Pajares, 2008, p. 752). For this reason, eligibility loss may lead to an immediate decrease in self-efficacy as the individual seeks to realign their self-efficacy given a new signal of their academic performance.

Whether scholarship loss negatively affects self-efficacy and, indirectly, enrollment is likely to vary across individuals. Some individuals may be predisposed to interpret the new information as a more meaningful signal of ability than others and, because of this, make decisions that reflect this self-efficacy revision (e.g., stop-out or transfer). Furthermore, baseline levels of self-efficacy appear to vary by race (Gloria & Hird, 1999), and increases in self-efficacy may have a smaller positive effect on achievement for Black students than for White students in some environments (Cheema & Kitsantas, 2014). Environmental factors, such as a school's racial climate and the presence (or notable absence) of peers with a similar racial identity, may contribute to the strength of the relationship between self-efficacy and academic performance (see, for example, Walton

& Cohen, 2007). For these reasons, the effect of eligibility loss may manifest differently—and possibly more strongly—for students from a population that is under-represented at their institution.

Study methods

Data

We use data from the Tennessee Higher Education Commission and Tennessee Student Aid Corporation (THEC/TSAC) that includes all first-time freshmen who first enrolled at a Tennessee public two- or four-year postsecondary institution in fall 2010–11 or later with a HOPE scholarship (N = 150,568). This dataset is a semester-level college enrollment file that includes details about a student's grade point average (GPA), credits attempted and earned, as well as background information about their sex, race/ethnicity, age, and the high school from which they graduated. To this dataset we add by-semester Tennessee scholarship payment records, annual FAFSA records, and degree records from Tennessee public institutions through summer 2017.

We use data from fall 2010–11 and later, rather than data from the launch of the scholarship in 2004, due to a change in THEC's reporting requirements that occurred in 2010. From 2010–11 onward, the cumulative GPA information for a given student record corresponds to that student's cumulative GPA *at the end of the* semester in question. Before fall 2010–11, as was the case in Carruthers and Özek's (2016) assessment of scholarship loss, institutions reported the cumulative GPA as of the end of the previous semester in which a student was enrolled. Using data from this previous reporting scheme would lead to biased results because it is not possible to observe the cumulative GPA in the last semester enrolled for students who subsequently stop out.

The sample is restricted in several ways (Table 2). Students who earned more than 24 HOPE-eligible credits by the end of their first semester are excluded because this level of credit accrual suggests that they reached the

Table 2. Creation of analytical sample.

Reason for dropping	Lost	# Remaining
Starting Sample (Students who we first observe as first-time freshmen w/ HOPE)	-	150,568
Have 24+ cumulative credits by the end of 1st semester	51	150,517
Does not start in a fall semester	3,750	146,767
Lose scholarship before the first checkpoint	19,860	126,907
Students enrolled at four-year institution during first checkpoint semester		85,774
We observe < 100% time BA degree completion	41,766	44,008
Students who hit the first checkpoint semester during summer	42	43,966
Students who were part-time in first checkpoint semester	179	43,787
Missing lottery GPA in first checkpoint semester	1	43,786
Final Analytical Sample of four-year students at first checkpoint	-	**43,786**

Students can lose their scholarship before the first checkpoint due to failure to meet one of the non-GPA criteria.
Source: Authors' analyses of THEC administrative data

first scholarship checkpoint during their first semester (excludes 51 individuals). Next, the sample is limited to students who began college during a fall semester (excludes 3,750 individuals). Because we are interested in the effects of scholarship loss at the first checkpoint, the sample is limited to students who make it to the first renewal checkpoint (24 credit hours) with their scholarship intact (excludes 19,860 individuals) and who are enrolled at a four-year public institution (excludes 41,133 individuals). The sample is also limited to entering cohorts for whom we have enough years of data to observe on-time (four-year) bachelor's degree completion (excludes 41,766 individuals; to observe 150% time, we further limit the sample to only two cohorts of students). Students who reach the first renewal checkpoint during a summer term or who are enrolled part-time in the first checkpoint semester are removed (excludes 42 students and 178 students, respectively). Finally, students who have a missing GPA value during their first checkpoint semester are removed (one individual). Given these restrictions, the effective sample includes 43,786 students enrolled at a four-year institution when they reached the first TN HOPE renewal checkpoint.

Disaggregating the sample of students who reach the first checkpoint at a four-year institution into those who lose versus maintain GPA eligibility reveals descriptive evidence of race- and income-based inequalities. Aspire recipients are overrepresented among those who lose the scholarship compared to the full sample of students who reach the first checkpoint (38% versus 30%). Furthermore, although 72% of the students who reach the checkpoint identify as White, only 64% of those who lose at the first checkpoint identify as White (Table 1). Black students, in contrast, are overrepresented among the population of students who lose the scholarship at this first checkpoint (24% of those who lose versus 17% of the full sample).

Variables

Treatment

The treatment group includes all students who fail to meet the GPA renewal threshold (2.75) during the first checkpoint semester; students who meet or exceed the threshold serve as the control group. As discussed below, we restrict this sample to students near the cutoff threshold to create a realistic counterfactual. Because HOPE is a guaranteed award that does not crowd out other sources of aid, lost HOPE dollars are not replaced with alternative sources of grant aid. Although we cannot observe in the data how students who remain enrolled compensate for lost HOPE dollars, we expect many do so through (additional) loans or work.

Outcomes

We estimate the effect of losing eligibility for the TN HOPE scholarship on stop-out, cross-sector transfer, and bachelor's degree completion. Stop-out is a dichotomous measure (=1) if a student is not enrolled at any institution in a public Tennessee college in the semester after they hit the checkpoint. Our measure of stop-out includes students who subsequently re-enroll as well as those who permanently drop out.[2] For students who reach a checkpoint during a spring term, stop-out (and transfer) is measured in the following fall semester. Cross-sector transfer is a dichotomous measure (=1) if a student transferred from the four-to the two-year sector in the semester after reaching the checkpoint. Students are categorized as not transferring across sectors if they: stop out based on the above definition,[3] remain enrolled at the same institution, or transfer to another institution within the same sector. Because the enrollments dataset is limited to public postsecondary institutions in Tennessee, we cannot observe whether students no longer enrolled at these institutions have transferred out-of-state or to a private Tennessee institution.

Degree completion is a dichotomous measure corresponding to whether a student earned a bachelor's degree at a public Tennessee college or university within four (100% time) or six years (150% time) of initial postsecondary enrollment. We have data on four-year graduation outcomes for four entering cohorts of students (Fall 2010-Fall 2014) and six-year graduation outcomes for two matriculating cohorts (Fall 2010-Fall 2011).

Covariates and fixed effects

Since the RD design is based on the principle that assignment into the treatment and control groups is as good as random close to the threshold, including covariates in our model is not required to produce unbiased results. However, to improve the precision of the estimates we include several student-level variables including race, sex, whether a student graduated from a high school in an economically distressed or at-risk county, whether they graduated from a rural high school, whether they receive a GAMs or Aspire supplement, whether the student is a Pell Grant recipient, and the highest education level of their parents. These covariates account for individual and environmental factors that may contribute to the relative influence of the two mechanisms discussed above.

Also included are institution and cohort fixed effects; the former captures potential time-invariant cross-institution variation in the effects of loss that may result from students' enrollment at different institutions (e.g., differences in the availability of academic supports). Entering cohort fixed effects captures time-varying differences that may affect student outcomes, such as the Great Recession. We exclude the race covariate in models comparing effects between

Black and White students and exclude the Aspire covariate in models contrasting Aspire and non-Aspire recipients. Both sets of covariates are excluded in models that estimate effects for combined race-income subgroups.

Subgroups

To test the heterogeneous effects of scholarship eligibility loss, we estimate the models separately for each subgroup of interest. For the family income comparisons, students who are eligible to receive a supplemental Aspire grant are compared to those who are not Aspire eligible. This delineation, which occurs at a family AGI of $36,000, aligns with how Tennessee has identified their most financially needy students. The examination of heterogeneous racial effects is restricted to Black and White students due to small sample sizes of the remaining racial groups (i.e., Latinx, American Indian, Asian or Pacific Islander, multiracial),[4] which impose statistical constraints. We decide against pooling these students with Black students into a single "non-White" group, which increases sample size, because doing so is undesirable conceptually and statistically. Conceptually, pooling disregards the variety of environments students from different racial/ethnic backgrounds may experience on a college campus and, therefore, the variety of mechanisms through which a policy may affect their outcomes. Statistically, pooling students masks how these differences may manifest as variation in the size and/or direction of the effects of a policy. To avoid conflating such experiences we elect to focus only on the differential effects for the two largest racial groups in Tennessee, Black and White individuals. We underscore the need to extend this line of inquiry to additional racial/ethnic subgroups in future research since our analytic decision silences their experiences. Lastly, because race and income may interact in meaningful ways, the effects of scholarship loss on Black Aspire students, Black non-Aspire students, White Aspire students, and White non-Aspire are estimated separately.

Empirical strategy

We analyze the impact of losing TN HOPE eligibility due to insufficient GPA on the outcomes of interest using a sharp regression discontinuity design because renewal is contingent on maintaining a cumulative GPA above 2.75.[5] A student is considered treated (i.e., "lost eligibility") if their GPA is below this cutoff. The RD method isolates the causal effect of losing GPA-based scholarship eligibility on student outcomes by comparing the outcomes of students with GPAs just above and below the renewal threshold at the first checkpoint. Because the RD only uses students with GPAs near the cutoff at the renewal checkpoint, the results are local average treatment effects (LATE).

The use of an RD model requires the researcher specify and motivate numerous decisions related to model specification. First, we estimate the models using a non-parametric regression approach (Calonico et al., 2017). Though a parametric approach will generally produce more precise estimates than a non-parametric approach (Gelman & Imbens, 2019), results of a parametric model will be biased if the specification is incorrect. Non-parametric methods are robust to this type of misspecification (Gelman & Imbens, 2019). With respect to bandwidth selection, we utilize data-driven optimal bandwidths (Calonico et al., 2017) that vary for each subgroup and outcome. We also report bias-corrected estimates and robust variance estimators in the RD estimates, providing conservative point estimates rather than conventional estimates and variance estimators. We use triangular kernels, and, as suggested by Gelman and Imbens (2019), local linear estimators. We find similar results using alternative functional forms, including quadratic and cubic slopes rather than linear slopes and using Epanechnikov and uniform rather than triangular kernels (results available upon request). The results are also robust when using a linear parametric specification with samples restricted to various bin widths around the GPA cutoff (see Appendix).

Because we produce estimates for multiple outcomes and several subgroups within each outcome, estimated p-values are adjusted to account for the multiple hypothesis tests conducted. We apply a family-wise error rate (FWER) adjustment, which reduces the probability of reporting a false positive (e.g., a Type I error) by taking the critical value set by the researcher (e.g., $p < .05$) and adjusting the p-values so that the chance of making *any* Type I error across the multiple tests is below that critical value (Porter, 2018). We use the Holm correction, a step-down procedure that helps retain power relative to other multiplicity approaches (Holm, 1979).[6]

Our results should be interpreted as the intent to treat (ITT) effect because not all students comply with their treatment status. The most common form of noncompliance is students who are above the GPA cutoff but lose the scholarship due to non-continuous enrollment.[7] Table 3 summarizes the compliance rates of four-year students at the first checkpoint, which exceed 90% for the overall sample and each subsample. A fuzzy RD set-up, which would account for this imperfect treatment assignment, would require the use of forward-looking data on HOPE payments in the next semester to estimate the first equation in a two-stage set-up. This approach would drop from the second stage regression students who stop-out or transfer to a nonpublic Tennessee institution but who would have received HOPE had they remained enrolled at an eligible institution. Assignment to the treatment group (i.e., lost HOPE) would occur not when eligibility is lost (i.e., at the end of the checkpoint semester), but rather at the start of the following

Table 3. Compliance rates among four-year students at the first checkpoint.

Sample	GPA below threshold, kept scholarship	GPA above threshold, lost scholarship	Compliers
All Students	2%	4%	93%
Income			
Aspire	3%	5%	93%
Non-Aspire	2%	4%	94%
Race			
Black	2%	3%	95%
White	2%	5%	93%
Income x Race			
Aspire/Black	3%	3%	94%
Non-Aspire/Black	2%	3%	95%
Aspire/White	3%	5%	92%
Non-Aspire/White	2%	4%	94%

Calculated based on sample that falls within .3 GPA bandwidth. Due to rounding, totals may not add up to 100%.
Source: Authors' analyses of THEC administrative data

semester. Since factors external to HOPE could impact students' decisions to enroll, using a fuzzy RD approach would invalidate the causal nature of the estimate. That said, we estimated the models using a fuzzy RD setup and found results similar to those produced by the sharp RD approach (see the Robustness Checks sub-section in Results).

For an RD to be valid, several conditions must be met. First, there must be no precise manipulation of the running variable, in this case cumulative GPA at the first checkpoint. Conceptually, it is conceivable that manipulation could occur. The HOPE renewal criteria are publicly available and, ideally, scholarship recipients are aware of these requirements. Students could adopt strategic behaviors, such as enrolling in easier classes or asking professors to award them higher grades, to get their cumulative GPA above the threshold. Though we cannot observe these hypothetical behaviors, we can formally test for manipulation using a local polynomial density estimator (Cattaneo et al., 2018). The results of this test,[8] for the full sample and each subsample, fail to reject the null hypothesis of no difference in the density of students on either side of the cutoff. Figure 1 presents a plot of this relationship for the full sample with no visual indication of a discontinuity in the density of observations at the threshold, leading us to conclude there is no evidence of manipulation of the running variable.

The second condition that must be met for valid inference is that any discontinuities in the outcome at the cutoff must be due to the treatment (i.e., eligibility loss) rather than other differences between the treatment and control groups (e.g., demographic characteristics). We test whether students just above and below the GPA cutoff are similar on observable characteristics by conducting RD estimates of the effect of being below the cutoff on each of the student-level covariates included in our main models. As presented in Table 4, only two estimates are statistically significant: Asian or Pacific Islander students near the GPA renewal threshold are 1.2 percentage points (pp) less likely to lose HOPE eligibility and GAMS recipients near the cutoff

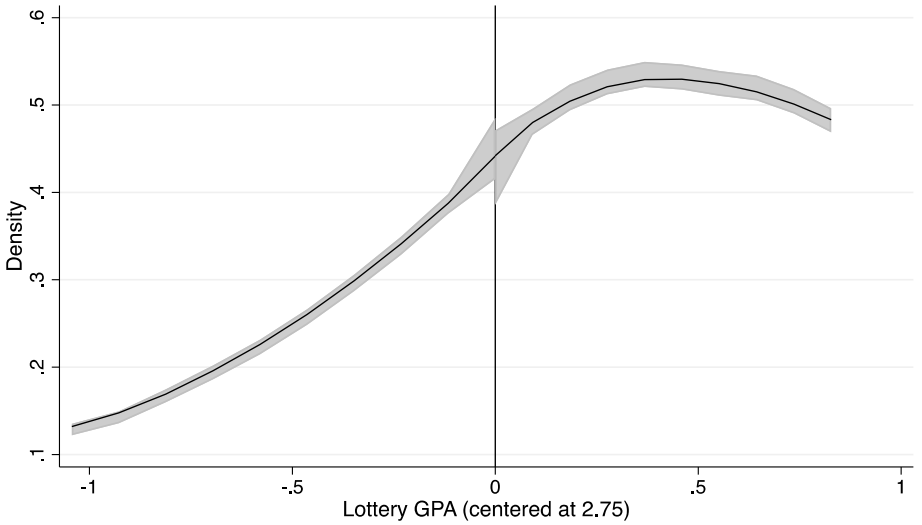

Figure 1. First checkpoint GPA density plot. *Note*: Plot shows no statistical evidence of manipulation of the running variable (robust p-value = 0.377). Results based on data-driven optimal bandwidths. The bandwidth method was comb with triangular kernels and jackknife standard errors. *Source*. Authors' calculations of THEC administrative data

Table 4. Effect of losing HOPE eligibility on student-level observable characteristics.

Outcome	Estimate (SE)	BW	N
Female	0.001 (0.018)	0.584	21,550
White	0.009 (0.017)	0.509	19,016
Black	−0.002 (0.013)	0.595	21,967
Asian or Pacific Islander	**−0.012*** (0.007)	0.297	11,466
Latinx	−0.005 (0.008)	0.317	12,143
Alaskan Native or American Indian	−0.001 (0.002)	0.371	14,122
Multi-Racial	0.009 (0.007)	0.560	20,770
Race Unknown	0.005 (0.006)	0.529	19,652
Student is receiving the Aspire supplement	0.014 (0.020)	0.385	14,689
Student is Pell grant eligible	−0.006 (0.021)	0.392	14,904
Student is receiving the GAMS supplement	**0.017**** (0.008)	0.239	9,103
Graduated HS in distressed or at-risk county	0.015 (0.015)	0.379	14,351
Graduated HS in rural county	0.030 (0.020)	0.288	11,192
Highest parental education level is middle school	−0.001 (0.005)	0.384	14,674
Highest parental education level is high school	0.014 (0.018)	0.465	17,477
Highest parental education level is college or beyond	−0.018 (0.018)	0.500	18,562
Highest parental education level is other or unknown	0.005 (0.006)	0.323	12,370

*p < 0.10, **p < .05, ***p < .01. Significant effects are bolded

Robust standard errors (SE), clustered by postsecondary institution, are included in parentheses. All models use data-driven optimal bandwidth (BW).

Source: Authors' calculations using THEC administrative data

are 1.7pp more likely to lose eligibility. Notably very few students near the cutoff identify as Asian or Pacific Islander or receive a GAMS award (there are < 300 of each student type within a 0.3 GPA of the renewal threshold). To address these differences among students and increase the precision of the estimates, a full set of student-level covariates as well as cohort- and

institution-fixed effects are included. As an additional test of the second condition, we examine whether, in the absence of the renewal requirement, there are any discontinuities in the outcomes. This is done by regressing each outcome on the full set of covariates *but not the treatment variable*, then predicting the probability of each outcome for each GPA value and graphing the results. Figure 2 indicates that in the absence of treatment, there are no discontinuities in the outcomes at the GPA renewal cutoff. Taken together this evidence supports that, aside from their treatment status, the students in the two groups are similar at the cutoff and any effects we find can be attributed to losing HOPE eligibility.

The last condition that must be met is that no other treatments or policy initiatives simultaneously occur at the 2.75 GPA threshold. If there were, we would not be able to attribute the effects only to scholarship eligibility loss. We are not aware of any (e.g., Satisfactory Academic Process has a lower GPA threshold) and do not believe that 2.75 is a meaningful GPA threshold for any programs or policies other than the HOPE scholarship.

Limitations

This study has five primary limitations. First, due to small sample sizes, we cannot estimate effects for students who identify as Latinx, American Indian, Asian or Pacific Islander, or multiracial. Second, our data is limited to Tennessee public postsecondary institutions, meaning we cannot observe if students transfer to and subsequently graduate from private Tennessee or out-of-state colleges. Our original data did include degrees earned at private Tennessee institutions by students with active HOPE scholarships at the time of degree completion, but this number is very small (< 100). We interpret this as plausible evidence that even if students are transferring to those institutions, they stop out before completing a degree. Third, our results represent the causal effect of losing HOPE eligibility at the first checkpoint due to failure to meet the required renewal GPA (2.75). Although this is the primary reason for HOPE loss, some students lose HOPE for other reasons, such as non-continuous enrollment or a within-semester change in enrollment intensity. Our results cannot speak to how losing HOPE eligibility affected these students. Fourth, the analyses are focused on the differential effects of scholarship loss among four-year students, but Tennessee public two-year institutions serve a disproportionate number of the state's Black and lower-income postsecondary students. To fully understand the disparate outcomes of eligibility loss among subgroups, we will examine the differential effects of HOPE loss among community college students in a subsequent paper. Finally, although the methodological design enables the detection of the causal impact of losing eligibility for TN HOPE on student outcomes, we are unable to disentangle the specific mechanism(s) through which aid affects subsequent

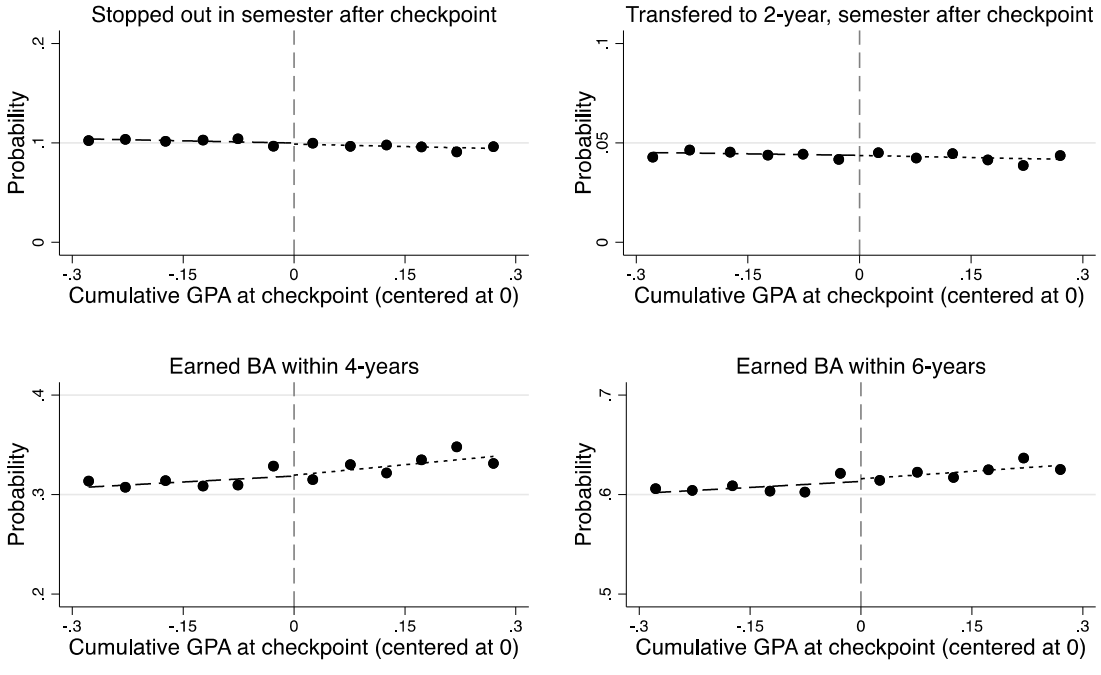

Figure 2. Predicted outcomes in the absence of the HOPE scholarship, by cumulative GPA. *Note*: Plots use data from students with a cumulative TN HOPE GPA near the renewal GPA (within a 0.3 bandwidth). *Source*. Authors' calculations of THEC administrative data

outcomes. Because understanding both components is crucial for responsive policy design, additional research needs to examine the how and why of these effects.

Results

Presented in Table 5 are the effects of losing HOPE eligibility due to insufficient GPA at the first checkpoint across the full sample, as well as the effects for each of the subgroups of interest. Again, the results represent the local average treatment effect and apply to students whose GPAs fall near the 2.75 renewal threshold. Descriptive statistics on the proportion of each subsample who stop-out or transfer to a community college after the first checkpoint, as well as the proportion that earn a bachelor's degree within four and six years, are provided for additional context (Table 6).

Effect on next semester stop-out

In the full sample, those with a GPA below the cutoff have a 2.5 percentage point (pp) higher probability of stopping out in the next semester than students with GPAs above the cutoff. This effect appears to be driven by non-Aspire recipients and White students, who are 3.4pp and 3.3pp, respectively, more likely to stop out upon losing their scholarship. The only statistically significant effect among the income/race subgroups is for non-Aspire White students, who have a 3.3pp higher probability of stop-out after losing TN HOPE. Working within the constraints of their data, Carruthers and Özek (2016) measure whether, conditional on a student enrolling in the semester immediately following the first checkpoint (n+1), the student enrolls the following semester (n+2). Our findings regarding next-semester stop-out are not directly comparable to Carruthers and Özek's (2016) estimation of the effects on persistence. However, when we re-estimate our findings using a definition of persistence that mimics theirs, we find similar statistically non-significant results (-1.5pp and -1.1pp in our and their studies, respectively; results available upon request).

Effect on cross-sector transfer

For the full group of students as well as for the non-Aspire and White subsamples, there is no statistically significant effect of being below the GPA cutoff on next semester cross-sector transfer. In contrast, losing TN HOPE causes a 3.2pp increase in the probability of transferring to a community college for Aspire students and a 4.9pp increase for Black students.

Table 5. Effect of losing HOPE eligibility at the first checkpoint.

	Stop-out (next semester)			Transfer to 2-year (next semester)			100% Time BA			150% Time BA		
	Effect	BW	N	Effect	BW	N	Effect	BW	N	Effect	BW	N
All Students	**0.025****	0.287	10,952	−0.001	0.226	8,647	0.007	0.418	15,657	−0.016	0.323	5,612
	(0.010)			(0.008)			(0.019)			(0.033)		
Income												
Aspire	0.017	0.304	3,070	**0.032*****	0.491	4,726	−0.071	0.241	2,434	−0.034	0.249	1,127
	(0.019)			(0.008)			(0.035)			(0.063)		
Non-Aspire	**0.034*****	0.303	7,040	−0.011	0.186	4,415	0.028	0.322	7,476	−0.027	0.294	3,101
	(0.009)			(0.011)			(0.022)			(0.045)		
Race												
Black	0.019	0.368	2,511	**0.049*****	0.317	2,158	−0.039	0.249	1,641	**−0.155*****	0.352	1,164
	(0.014)			(0.012)			(0.027)			(0.045)		
White	**0.033****	0.313	8,277	**−0**.018	0.190	5,190	0.007	0.402	10,518	0.003	0.297	3,525
	(0.012)			(0.010)			(0.019)			(0.034)		
Income x Race												
Aspire/Black	−0.012	0.287	1,092	**0.048*****	0.311	1,192	−0.070	0.382	1,446	0.008	0.288	535
	(0.022)			(0.012)			(0.038)			(0.105)		
Non-Aspire/Black	0.027	0.495	1,433	0.050	0.364	1,094	−0.036	0.297	887	−0.319	0.223	327
	(0.019)			(0.021)			(0.057)			(0.130)		
Aspire/White	0.029	0.338	2,093	0.022	0.372	2,284	−0.084	0.271	1,732	−0.084	0.280	777
	(0.031)			(0.014)			(0.043)			(0.081)		
Non-Aspire/White	**0.033****	0.303	6,143	−0.020	0.177	3,593	0.036	0.288	5,887	0.003	0.414	3,640
	(0.011)			(0.010)			(0.020)			(0.026)		

*p < 0.10, **p < .05, ***p < .01 where tests are Holm adjusted for multiple comparisons. Significant effects are bolded.

"All Students" includes students within all racial/ethnic categories, but all additional analyses are restricted to Black and White students to better compare subgroup results. Robust standard errors (SE), clustered by postsecondary institution, are included in parentheses. All models include student-level controls and institutional and cohort fixed-effects and use the data-driven optimal bandwidth (BW).

Source: Authors' calculations using THEC administrative data

Table 6. Rates of occurrence of each outcome for students near the GPA cutoff at the first checkpoint.

	N		Stop-out		Transfer to 2-year		100% Time BA		N		150% Time BA	
	Below	Above	Below	Above	Below	Above	Below	Above	Below	Above	Below	Above
All Students	4,864	6,637	6%	3%	5%	2%	19%	30%	2,219	3,085	56%	69%
Income												
Aspire	1,379	1,719	8%	3%	5%	2%	12%	23%	806	609	47%	65%
Non-Aspire	2,918	4,135	6%	3%	5%	3%	23%	33%	1,298	1,862	59%	71%
Race												
Black	967	1,117	5%	2%	4%	1%	14%	22%	456	558	51%	66%
White	3,330	4,737	7%	3%	5%	3%	21%	32%	1,451	2,110	57%	70%
Income x Race												
Aspire/Black	539	641	5%	2%	3%	0%	10%	20%	258	316	51%	64%
Non-Aspire/Black	428	476	5%	1%	5%	1%	18%	25%	198	242	51%	68%
Aspire/White	840	1,078	10%	4%	6%	3%	14%	24%	351	490	44%	65%
Non-Aspire/White	2,490	3,659	6%	3%	5%	3%	23%	34%	1,100	1,620	61%	72%

"All Students" includes students within all racial/ethnic categories, but all additional analyses are restricted to Black and White students to better compare subgroup results. Due to rounding, totals may not equal 100%. Results are based on students whose checkpoint cumulative GPA falls within 0.3 of the renewal threshold.
Source: Authors' calculations using THEC administrative data

Examining the income/race subgroup results, the statistically significant effects are concentrated among Aspire Black students, who experience a 4.8pp increase in the probability of cross-sector transfer.

Effect on earning a bachelor's degree

There are no statistically significant effects of TN HOPE eligibility loss on the probability of on-time degree completion for the full sample of students, or among any of the subgroups. This finding aligns with Carruthers and Özek's (2016) null results for 100% time degree completion among four-ye . Notably, the point estimates for Aspire students, both Black and White, are large and negative (i.e., −7pp to −8.4pp) but imprecisely estimated (as evidenced by large standard errors).

Losing scholarship eligibility at the first checkpoint has a statistically significant negative effect only on Black students' probability of earning a bachelor's degree within 150% time of matriculation. Black students who lose eligibility at the first checkpoint have bachelor's degree completion probabilities that are 15.5pp lower than Black students who maintain eligibility. There are no statistically significant effects on 150% time bachelor's degree completion for the remaining subgroups or the full sample, but again we see large but imprecisely estimated (i.e., not statistically significant) point estimates for two groups, non-Aspire Black students and Aspire White students (−31.9pp and −8.4pp).

Robustness checks

We test the robustness of our results using a series of alternative specifications and find evidence that our results are robust to internal validity threats. We first estimate our models using fixed bandwidths ranging from 0.1 to 0.5, rather than the data-driven optimal bandwidths used in our main specification. Overall, the results using alternative bandwidths (7see Appendix, Table A2) are fairly consistent with the main results. However, the optimal bandwidths may understate the effects on Black students for some outcomes, such as stop-out and 150% time bachelor's completion for non-Aspire Black students and 100% time bachelor's completion for Aspire Black students.

We also reestimate our models using a fuzzy RD design, wherein the first stage provides conditional estimates of the probability of scholarship payment in the semester following the checkpoint, and the second stage estimates whether payment status predicts the outcomes of interest (Table 8see Appendix, Table A3). The statistically significant effects found in the preferred sharp RD specification generally hold, with effect magnitudes increasing slightly in most cases. The effect on 150% time bachelor's completion for non-

Aspire Black students, which was large but imprecise using the sharp RD, became larger (−38.7) and statistically significant ($p < .01$). These results suggest that our preferred sharp RD model may underestimate the true effect of HOPE loss due to this model's inclusion of noncompliers. We present the sharp RD results as our main specification because we believe that this conservative estimate is preferable to a fuzzy RD model that may overestimate the true effect. Under the fuzzy RD set-up, we would drop from the second stage any students who were above the GPA threshold and stopped out because they do not receive a payment in the next semester, even though these students would likely have received a payment had they chosen to enroll. This likely leads to overestimating the effects of HOPE loss, because we keep in the second stage regression students who stop out but have GPAs below the renewal threshold. Importantly, because the results are similar under the sharp and fuzzy RD models, our decision to use the sharp RD results as our preferred specification does not have major implications on our findings and takeaways.

We also conducted falsification tests to examine whether there are any statistically significant effects found at discontinuities other than the 2.75 GPA renewal (see Appendix, Table A4). Using false GPA cutoffs at 2.5 and 3.0 GPA, we find only two statistically significant effects. For the 2.5 GPA discontinuity, we find a statistically significant effect on 150% time graduation among the full sample (10.4pp) and White students (11.7pp). Notably, neither of these findings align with effects found in our main specification. The general lack of statistically significant effects found at these alternative discontinuities adds confidence that the results found at the 2.75 GPA threshold are driven by loss of HOPE eligibility. The results of this series of sensitivity checks, alongside the bias-corrected estimates, robust variance estimators used, as well as the Holm p-value adjustments, provides evidence that our results are robust to internal validity threats.

Discussion and implications for policy

The results indicate differential effects of HOPE eligibility loss between subgroups, demonstrating the importance of disaggregating data and exploring heterogeneous treatment effects. Because the vast majority of students in the sample are White and non-Aspire recipients, the overall treatment effects are driven by the effects for these students and mask the differential effects for Black and Aspire students. For example, the null effect on cross-sector transfer observed for the full sample masks statistically significant effects for Black students and Aspire recipients. For the remainder of this section, we focus on the heterogeneous effects across subgroups.

Immediate responses to scholarship eligibility loss are mixed across groups. White students experience an increase in stop-out, with this effect statistically significant among non-Aspire White students. The adverse effects on cross-

sector transfer are concentrated on Black students, with the effect statistically significant for Aspire Black students. These short run effects of HOPE eligibility loss suggest the need for targeted outreach and support for HOPE recipients before reaching the first checkpoint.

The effect of HOPE eligibility loss on stop-out being statistically significant only for non-Aspire White students contrasts with what we expected to find. We hypothesized that the largest effects of HOPE loss would be experienced by Black and low-income students, but we observed the opposite. The statistically significant finding among only *non-Aspire* White students suggests that the effect cannot be solely explained by a financial cost mechanism, wherein the loss of aid leads price-sensitive students to shift enrollment. Furthermore, the next-semester effects we observe for non-Aspire White students do not follow through to effects on graduation. Notably, our definition of stop-out considers those who stop out and never return (e.g., drop-out) *and* those who stop out temporarily. Although we cannot observe return to school for students who transfer to a nonpublic Tennessee institution, we can observe that 42% of all students (and 44% of non-Aspire White students) who stop out in the semester immediately following the first checkpoint subsequently (within four years of initial college entry) re-enroll at either a two- or four-year public TN institution. Future research that explores the factors that lead some of the students who lose HOPE eligibility and stop-out only to subsequently re-enroll, while others permanently drop-out, would help inform policymakers and institutional stakeholders how to best support these students.

The relative sample size of our various subgroups may influence our ability to detect statistically significant effects. Non-Aspire White students are by far the largest income-race subgroup, enabling us to more precisely estimate effects relative to our other subgroups. Conversely, the much smaller number of both Aspire and non-Aspire Black students in our sample leads to a risk of having insufficient power to be able to detect a true effect for these groups when one exists. As it relates to the next semester stop-out outcome, this means that while we can say with confidence that losing HOPE eligibility causes an increase in next semester stop-out among non-Aspire White students, we do not discount the possibility that an effect also exists for other subgroups. In fact, the magnitude and direction of effect for Aspire White and non-Aspire Black students are similar to the effect for non-Aspire White students but have larger standard errors and are not statistically significant. The stop-out effects of scholarship loss are not statistically different between non-Aspire White students and any other race-income subgroup.

In contrast, we pick up an effect on cross-sector transfer for Aspire Black students even with a small sample size. This concentrated effect on transfer among Aspire Black students could reflect both a financial mechanism *and* a self-efficacy mechanism. With respect to self-efficacy, we theorize that the

effect of eligibility loss on self-efficacy may manifest more strongly for students from a population that is under-represented at the institution. All but one public four-year institution in Tennessee is a predominately white institution (PWI), meaning that Black students who attend these institutions do so alongside noticeably fewer peers, faculty, and staff with same-group racial identities. In the presence of a smaller community of same-group individuals, a Black student at a PWI may read more into the signal received from eligibility loss and, as a result, experience a larger negative effect. It is possible that disaggregating effects by institution would reveal that the effect of eligibility loss on stop-out and transfer is smaller for Black students at the state's one public four-year HBCU.

There are no full sample or sub-group effects of losing HOPE eligibility at the first checkpoint on graduation, but loss at this time does have a large negative effect on 150% time bachelor's degree completion for Black students. There are large and negative, although not statistically significant, point estimates on the effect of eligibility loss on the probability of on-time degree completion for both Black and White Aspire recipients and on the probability of 150% time degree completion for Aspire White recipients. Although these results are imprecisely estimated, they suggest that Aspire recipients, both Black and White, may experience adverse effects of HOPE eligibility loss on degree completion.

These effects, both the statistically significant effect for 150% time completion and the suggestive effects for 100% time graduation, are larger than the effects on the next semester outcomes. This indicates that the behavioral response to eligibility loss is not immediate for many students but is delayed by a semester or more. Students who lose eligibility may enroll in the subsequent semester or multiple semesters, but the compounded lost dollars over time may eventually induce cross-sector transfer or stop-out. Future research should explore the potential compounding effect of loss by considering whether the effects on graduation differ for students who lose and regain their merit aid versus those who lose and never regain.[9] If there is a positive graduation effect among the former group, efforts to graduate more students could target academic support toward helping students regain HOPE eligibility after loss.

Overall, the detrimental effects of scholarship loss are most heavily concentrated among Black students. Non-Aspire White students make up the majority of students who lose eligibility at the first checkpoint but, whereas this subgroup experiences a negative effect of eligibility loss on next semester stop-out, we see no effects on the longer-term outcome of bachelor's degree completion. A major policy implication of these results is that targeting additional resources and interventions toward Black students—the group most impacted by loss—would likely yield the most meaningful improvements in both rates of preliminary loss *and* the subsequent outcomes affected by loss. Lower rates of loss and higher degree attainment levels would each improve

the efficiency of HOPE dollars spent. Returning to the student-level mechanisms proposed—shifts in financial cost and in self-efficacy—the above discussion suggests that the effects of loss result from more than just price sensitivity. The non-financial mechanism of self-efficacy was suggested acknowledging that the research design does not enable us to discern whether, and the extent to which, it contributes to the observed effects.

Future research

There are several possible extensions of this work. First, qualitative research could investigate how students think about the checkpoints before the checkpoint semester and after either renewing the scholarship or losing eligibility. This work would inform our understanding of which student-level mechanisms underlie the observed effects and why student responses differ between subgroups. Second, more work is needed to theorize on the institution as a mediating factor that shapes program administration and affects students' responses to loss.

As it relates to program administration, individuals seeking services can experience onerous administrative burdens that influence their ability to access the desired service (Herd & Moynihan, 2020). These burdens—which come in the form of learning, compliance, and psychological costs to the individual (Moynihan et al., 2014)—result from formal rules and informal practices that guide program administration (Brodkin & Majmundar, 2010). Specifically, embedded in our discussion of TN HOPE is an assumption that scholarship recipients are familiar with renewal timing and requirements. But financial aid is a notoriously complex system to navigate (Dynarski & Scott-Clayton, 2013), and it seems plausible that renewal information is inconsistently communicated to students across and within TN HOPE-eligible institutions. Variation in institutions' informal communication practices—when and how eligibility criteria are communicated in advance of checkpoints, and when students are informed of eligibility loss—could result in differential effects of loss across institutions.

Second, how Black students respond to scholarship loss could stem, in part, from their experiences with the campus's racial climate. Black students who enroll at predominantly white institutions (PWIs) may experience racism and racial bias when navigating campus, in social interactions, and in the classroom (Harper & Hurtado, 2007). A student who has already had a negative experience on campus due to the racial climate might more readily respond to scholarship loss by departing the institution. Here again, the effects of loss may therefore differ not only across sub-groups but across institutions.

Policy design requires thorough and creative consideration of the myriad ways in which financial aid programs interact with students' college experiences. Although student-level mechanisms certainly contribute to what we

observe, we cannot ignore the responsibility of institutions in administering state financial aid programs and cultivating a welcoming racial climate. Future work can better illuminate both types of mechanisms and can likewise lead policy scholars and policymakers to consider aspects of financial aid program design that extend beyond award amount and renewal thresholds.

Notes

1. In 2018, Black students made up 23% of public TN high school graduates but only 13% of first time HOPE recipients. See Tennessee graduation cohort data at https://www.tn.gov/education/data/data-downloads.html and the 2018 Tennessee Education Lottery Scholarship Program Annual Report at https://www.tn.gov/thec/research/tn-hope-scholarship-program.html.
2. Among students in our sample who stop-out in the semester immediately following the first checkpoint, 42% subsequently (within four years of initial college entry) re-enroll at either a two- or four-year public TN institution.
3. Although there is a recently developed method for conducting an RD with categorical outcome variables (Xu, 2017), which could be used for a multinomial set-up that examines next semester transfer, stop-out, and persistence simultaneously, this estimation method is not widely available. However, when Xu applies this method to data from Lindo et al. (2010), they find results similar to Lindo et al. (2010), who use methods similar to ours.
4. The race variable in our dataset includes "Hispanic/Latinx," instead of asking separately about Hispanic/Latinx ethnicity identification.
5. We use the *rdrobust* Stata command for all RD analyses (Calonico et al., 2017) which employs local polynomial RD point estimators and selects the optimal bandwidth for the calculation.
6. For additional insights into the multiple testing problem (MTP) see, Porter (2018) and Schochet (2008).
7. Other, less prevalent, forms of noncompliance include students above the cutoff who lose their scholarship due to a within-semester change in enrollment intensity and students below the cutoff who maintain scholarship eligibility. Although an insufficient GPA should automatically trigger scholarship loss, we observe a small number of students for whom this is not the case.
8. The robust T statistic for the full sample test is −0.884 ($p = 0.377$). Results are based on data-driven optimal bandwidths. The bandwidth method was comb with triangular kernels and jackknife standard errors. Results are robust to alternative bandwidths as well as alternative kernel functions and variance-covariance matrix estimator procedures. We find no statistical evidence of manipulation of the running variable in any of our subgroup samples (results available upon request).
9. Students who lose HOPE eligibility for GPA reasons are able to regain eligibility one time by increasing their GPA to the minimum requirement by the next checkpoint (while meeting all non-GPA criteria).

Acknowledgments

We thank the Tennessee Higher Education Commission (THEC) for providing access to data used in this study. We also thank Dominique Baker, Rajashri Chakrabarti, Laura Perna, and members of the DesJardins-McCall research team for helpful comments, as well as attendees of

the AEFP 2020 conference, ASHE 2020 conference, CIERS seminar at the University of Michigan, and the SHEEO Public Investment in Higher Education webinar series. Cummings and Deane gratefully acknowledge the funding support of the Institute of Education Sciences (IES), U.S. Department of Education under Grants R305B150011 and R305B150012. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of THEC or IES. Any errors are our own.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Allen, D., & Wolniak, G. C. (2019). Exploring the effects of tuition increases on racial/ethnic diversity at public colleges and universities. *Research in Higher Education*, *60*(1), 18–43. https://doi.org/10.1007/s11162-018-9502-6

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Brodkin, E. Z., & Majmundar, M. (2010). Administrative exclusion: Organizations and the hidden costs of welfare claiming. *Journal of Public Administration Research and Theory*, *20* (4), 827–848. https://doi.org/10.1093/jopart/mup046

Bruce, D. J., & Carruthers, C. K. (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics*, *81* , 30–44. https://doi.org/10.1016/j.jue.2014.01.006

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17*(2), 372–404. https://doi.org/10.1177/1536867X1701700208

Carruthers, C. K., & Fox, W. F. (2016). Aid for all: College coaching, financial aid, and post-secondary persistence in Tennessee. *Economics of Education Review*, *51* , 97–112. https://doi.org/10.1016/j.econedurev.2015.06.001

Carruthers, C. K., & Özek, U. (2016). Losing HOPE: Financial aid and the line between college and work. *Economics of Education Review*, *53* , 1–15. https://doi.org/10.1016/j.econedurev.2016.03.014

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261. https://doi.org/10.1177/1536867X1801800115

Cheema, J. R., & Kitsantas, A. (2014). Influences of disciplinary classroom climate on high school student self-efficacy and mathematics achievement: A look at gender and racial–ethnic differences. *International Journal of Science and Mathematics Education*, *12*(5), 1261–1279. https://doi.org/10.1007/s10763-013-9454-4

Chen, R., & DesJardins, S. L. (2010). Investigating the impact of financial aid on student dropout risks: Racial and ethnic differences. *The Journal of Higher Education*, *81*(2), 179–208. https://doi.org/10.1080/00221546.2010.11779048

Cohodes, S. R., & Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy. *American Economic Journal. Applied Economics*, *6*(4), 251–285. http://dx.doi.org/10.1257/app.6.4.251

Cornwell, C., Mustard, D. B., & Sridhar, D. J. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE program. *Journal of Labor Economics*, *24*(4), 761–786. https://doi.org/10.1086/506485

DesJardins, S. L., & Bell, A. (2006). Using economic concepts to inform enrollment management. *New Directions for Institutional Research*, *2006*(132), 59–74. https://doi.org/10.1002/ir.196

Dynarski, S. (2004). *The new merit aid. In college choices: The economics of where to go, when to go, and how to pay for it* (pp. 63–100). University of Chicago Press.

Dynarski, S., & Scott-Clayton, J. (2013). *Financial aid policy: Lessons from research* (NBER Working Paper No. w18710).

EdBuild. (2019). *$23 billion*. https://edbuild.org/content/23-billion/full-report.pdf

Fishman, R., Nguyen, S., & Francisco, M. (2020). *Varying degrees 2020: New America's fourth annual survey on higher education* New America. https://www.newamerica.org/education-policy/reports/varying-degrees-2020/

Gelman, A., & Imbens, G. (2019). What high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics*, *37*(3), 447–456. https://doi.org/10.1080/07350015.2017.1366909

Gloria, A. M., & Hird, J. S. (1999). Influences of ethnic and nonethnic variables on the career decision-making self-efficacy of college students. *The Career Development Quarterly*, *48*(2), 157–174. https://doi.org/10.1002/j.2161-0045.1999.tb00282.x

Goldrick-Rab, S., Harris, D. N., & Trostel, P. A. (2009). Why financial aid matters (or does not) for college success: Toward a new interdisciplinary perspective. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 1–45). Dordrecht: Springer.

Gould, E. (2020) *State of working America wages 2019*. Economic Policy Institute. https://www.epi.org/publication/swa-wages-2019/

Harper, S. R., & Hurtado, S. (2007). Nine themes in campus racial climates and implications for institutional transformation. *New Directions for Student Services*, *2007*(120), 7–24. https://doi.org/10.1002/ss.254

Heller, D. E., & Rasmussen, C. J. (2002). Merit scholarships and college access: Evidence from Florida and Michigan. In D. E. Heller & P. Marin (Eds.), *Who should we help? The negative social consequences of merit scholarships* (pp. 25–40). The Civil Rights Project at Harvard University.

Henry, G. T., Rubenstein, R., & Bugler, D. T. (2004). Is HOPE enough? Impacts of receiving and losing merit-based financial aid. *Educational Policy*, *18*(5), 686–709. https://doi.org/10.1177/0895904804269098

Herd, P., & Moynihan, D. (2020). Administrative burdens in health policy. *Journal of Health and Human Services Administration*, *43*(1), 3–16. https://doi.org/10.37808/jhhsa.43.1.2

Holm. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70 .

Houle, J. N., & Addo, F. R. (2019). Racial disparities in student debt and the reproduction of the fragile black middle class. *Sociology of Race and Ethnicity*, *5*(4), 562–577. https://doi.org/10.1177/2332649218790989

Hurwitz, M. (2012). The impact of institutional grant aid on college choice. *Educational Evaluation and Policy Analysis*, *34*(3), 344–363. https://doi.org/10.3102/0162373712448957

Jones, C. (2014, December 11). *The problem with merit aid*. The Education Trust. https://edtrust.org/the-equity-line/the-problem-with-merit-aid

Kim, J. (2012). Exploring the relationship between state financial aid policy and postsecondary enrollment choices: A focus on income and race differences. *Research in Higher Education*, *53*(2), 123–151. https://doi.org/10.1007/s11162-011-9244-1

Lindo, J. M., Sanders, N. J., & Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal. Applied Economics*, *2*(2), 95–117. https://doi.org/10.1257/app.2.2.95

Monarrez, T., & Washington, K. (2020, June). *Racial and ethnic representation in postsecondary education* (Research Report). Urban Institute. https://www.urban.org/sites/default/files/publication/102375/racial-and-ethnic-representation-in-postsecondary-education_1.pdf

Moynihan, D., Herd, P., & Harvey, H. (2014). Administrative burden: Learning, psychological, and compliance costs in citizen-state interactions. *Journal of Public Administration Research and Theory*, *25*(1), 43–69. https://doi.org/10.1093/jopart/muu009

Porter, K. E. (2018). Statistical power in evaluations that investigate effects on multiple outcomes: A guide for researchers. *Journal of Research on Educational Effectiveness*, *11*(2), 267–295. https://doi.org/10.1080/19345747.2017.1342887

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261. https://doi.org/10.1037/0033-2909.130.2.261

Rodriguez, A., & McGuire, K. M. (2019). More classes, more access? Understanding the effects of course offerings on black-white gaps in advanced placement course-taking. *The Review of Higher Education*, *42*(2), 641–679. https://doi.org/10.1353/rhe.2019.0010

Schochet, P. Z. (2008). *Technical methods report: Guidelines for multiple testing in impact evaluations* (NCEE 2008–4018). National Center for Education Evaluation and Regional Assistance, Institute for Education Sciences, U.S. Department of Education.

Tennessee Higher Education Commission & Tennessee Student Assistance Corporation (THEC & TSAC). (2017). https://www.tn.gov/content/dam/tn/thec/bureau/research/other-research/hope/TELS%20Report_Fall%202017_Suppressed.pdf

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796. https://doi.org/10.3102/0034654308321456

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82. https://doi.org/10.1037/0022-3514.92.1.82

Xu, K. L. (2017). Regression discontinuity with categorical outcomes. *Journal of Econometrics*, *201*(1), 1–18. https://doi.org/10.1016/j.jeconom.2017.07.004